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

First revision

Guidance from your Editor

Please submit by **6 Feb 2020** for the benefit of the authors .

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

Author notes
Have you read the author notes on the guidance page?

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.

1 Tracked changes manuscript(s)
1 Rebuttal letter(s)
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 using environmental DNA metabarcoding

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 C Campus, 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

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. We here adopted the MiFish pipeline to know if the ~~system~~ is reliable to estimate fish biodiversity in ~~the~~ Korean rivers. Total 125 unique haplotypes and 73 confirmed fish species were identified from 16 water samples collected from a single survey of four Korean rivers (Hyeongsan, Taehwa, Seomjin, and Nakdong) indicating MiFish pipeline is a useful tool to estimate the fish biodiversity with relatively low cost and labors. However, low 12S sequences of endemic species in the database and low resolution of MiFish region for differentiating several taxa should be upgraded for their wide use. Among the four rivers, the highest species richness was identified in Seomjin river (52 species), followed by Taehwa river (42 species), Hyeongsan river (40 species). Nakdong river (26 species) showed the lowest species richness and endemic species numbers presumably due to its metropolitan location and anthropogenic impacts such as dams or weirs ~~there.~~ We were also able to ~~know~~ that five exotic species (*Carassius cuvieri, Cyprinus carpio, Cyprinus megalophthalmus, Lepomis macrochirus,* and *Micropterus salmoides*) are widely distributed in all surveyed rivers, ~~which would be problematic in the Korean river ecosystem~~. These findings strongly support the idea that the eDNA metabarcoding technique ~~would be~~ one of the cost-effective and scientific tools in the management and conservation of fish resources among Korean rivers.

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

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. We here adopted the MiFish pipeline to know if the system is reliable to estimate fish biodiversity in the Korean rivers. Total 125 unique haplotypes and 73 confirmed fish species were identified from 16 water samples collected from a single survey of four Korean rivers (Hyeongsan, Taehwa, Seomjin, and Nakdong) indicating MiFish pipeline is a useful tool to estimate the fish biodiversity with relatively low cost and labors. However, low 12S sequences of endemic species in the database and low resolution of MiFish region for differentiating several taxa should be upgraded for their wide use. Among the four rivers, the highest species richness was identified in Seomjin river (52 species), followed by Taehwa river (42 species), Hyeongsan river (40 species). Nakdong river (26 species) showed the lowest species richness and endemic species numbers presumably due to its metropolitan location and anthropogenic impacts such as dams or weirs there. We were also able to know that five exotic species (*Carassius cuvieri, Cyprinus carpio, Cyprinus megalophthalmus, Lepomis macrochirus,* and *Micropterus salmoides*) are widely distributed in all surveyed rivers, which would be problematic in the Korean river ecosystem. These findings strongly support the idea that the eDNA metabarcoding technique would be one of the cost-effective and scientific tools in the management and conservation of fish resources among Korean rivers.

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

INTRODUCTION

Fish communities have been considered as ~~one of the good~~ bioindicators of ecosystem status due to their vulnerability to environmental or anthropogenic stresses such as pollution, climate changes, or other disturbances ~~in habitats~~ (Dudgeon, 2010). Traditional monitoring methods for fish biodiversity, which have relied on the direct capture or observations of specimens, are often costly and time-consuming 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 fish biodiversity and demonstrated a potential to improve the traditional ~~methods a~~ cost-effective way (Foote et al., 2012; Kelly et al., 2017; Kelly et al., 2014; Shaw et al., 2016; Stoeckle et al., 2017; Yamamoto et al., 2017). This technique is sensitive as to allow the identification of rarely identified species (Pilliod et al., 2013), invasive species (Ardura et al., 2015; Cai et al., 2017; Clusa et al., 2017; Dejean et al., 2012; Klymus et al., 2017; Takahara et al., 2013; Williams et al., 2018) or migratory species (Gustavson et al., 2015; Pont et al., 2018; Yamamoto et al., 2016; Yamanaka and Minamoto, 2016).

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. Three fish-specific universal primer sets are currently reported; two sets for 12S rRNA regions, Eco Primers (Riaz et al., 2011) and MiFish (Miya et al., 2015b), and one for 16S rRNA region (Shaw et al., 2016). Among them, ~~MiFish primer~~ set demonstrated its reliability for eDNA metabarcoding analysis of fish biodiversity both in ~~seawater~~ (Ushio et al., 2017; Yamamoto et al., 2017) and ~~freshwater~~ (Sato et al., 2018). More recently, the web-based MiFish pipeline in MitoFish was publically open (http://mitofish.aori.u-tokyo.ac.jp/mifish/), which considerably boost-up the way of fish biodiversity analysis by eDNA metabarcoding 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 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 by 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, which have been isolated for long times without exchanging genetic information with ~~those other~~ habitats (Seehausen and Wagner, 2014), direct application of MiFish platform may produce a high amount of ~~the~~ 'unidentified' regional species. ~~Besides,~~ ~~the~~ relatively much lower amount of MiFish sequence data (12S region) is currently deposited compared with those of 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 the accurate fish biodiversity analysis using eDNA metabarcoding.

In this study, we firstly employed eDNA metabarcoding analysis of water samples collected from four rivers using MiFish primer ~~set to know freshwater~~ fish biodiversity in Korea. After that, we analyzed the haplotypes obtained by the MiFish pipeline to ~~know~~ 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 a 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 river, Taehwa river, Seomjin river, and Nakdong river, which are four large rivers in the southern part of the Korean peninsula (Fig.1 and Table 1). 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). One liter of water sample was collected at each station with disposable plastic bottles. After collecting water, the bottles were immediately stored in ~~the~~ icebox ~~until brought~~ to the laboratory for filtration. Water temperature and salinity were measured 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 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.

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

In order to assess the fish biodiversity, amplicon libraries of partial 12S rRNA region by the MiFish universal primer sets were constructed (Miya et al., 2015a). 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

The MiSeq raw reads paired by open-source software (Python 2.7) with the specific script (Zhang, 2015), then uploaded the paired sequences to the web-based MiFish pipeline (http://mitofish.aori.u-tokyo. ac.jp/mifish/). In MiFish pipeline, the low-quality tail of reads (QV ≤ 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 ≥ 99 %, then the sequence was ascertained ~~as species~~. If the sequence identity from 97 % to 99 %, the sequence was ascertained ~~as a genus~~, ~~and the~~ sequences ~~having~~ 97 % to 95 % identity (putative genera) to the GenBank database were assigned as ‘unidentified’ genera. The habitat distribution of each species was assessed on the FishBase website (https://www.fishbase.org/). ~~The~~ ~~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 total species in an ecosystem (Gray, 2000; Magurran, 1988). The H' index and the heat map clustering analysis were enumerated ~~by using~~ the PRIMER® software v7 (Clarke and Gorley, 2015).

RESULTS

Physico-chemical 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 difference (5.4 ˚C) in temperature from the upstream (HS1) to the downstream (HS4), whereas lowest levels of temperature variation were observed in Seomjin river (0.8 ˚C) and Nakdong river (1.5 ˚C). 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 Hyeongsan river. Salinity level increased from upstream to downstream in all rivers sampled, except for the Nakdong river, where an artificial dam has been constructed to block water from the ocean (Table 1).

Analysis of fish haplotypes obtained by the MiFish pipeline

The reliability of the MiFish pipeline (http://mitofish.aori.u-tokyo.ac.jp/mifish/workflows) for the biodiversity assessment of fish species inhabiting the sampled 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. Among the 238 haplotypes, ~~we found~~ 125 unique ~~haplotypes~~, which were identified using the phylogenetic tree analysis by the MEGA7 ~~program~~ (Kumar et al., 2016) with Maximum likelihood algorithm (Fig. 2-5). A total of 2,241,130 reads (98.26 %) were assigned to 73 ~~confirmed~~ species, 46 genera and 13 families of the Teleostei at 99 % as cutoff identity. The remaining 39,720 reads (49 haplotypes), which showed less than 99 % identity, were further assigned into 11 genera and 8 unidentified genera (Table 3). A total of 34,755 reads (1.50 %) with low identity (below 95 %) to the GenBank database were discarded from further ~~analysis,~~ The highest species number was identified in the family Cyprinidae (35), followed by Gobiidae (11), Cobitidae (8), and the remaining (19) are from other families of the Teleostei. Among them, the highest species numbers (4 species) were identified in the genus *Acheilognathus,* followed by *Carassius, Misgurnus, Tridentiger*, and *Squalidus* ~~with 3 species each genus~~ (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 99 % ~~and higher~~ ~~in~~ 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 Korean haplotype of *Hemibarbus labeo* (GenBank Number: DQ347953) and the Japanese haplotype of *Hemibarbus maculatus* (LC146032), respectively. Among four endemic species in the genus *Hemibarbus, Hemibarbus labeo* and *Hemibarbus 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 the Taehwa river (TH1) showed 97 % and 95 % identity to *Hemibarbus longirostris* (LC049889), respectively, which suggests that those three haplotypes may be either *Hemibarbus longirostris* or *Hemibarbus mylodon* (Fig. 2). Since *Hemibarbus mylodon* is an endangered freshwater species, which has been exclusively identified in Han and Geum rivers and further study should be conducted.

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

Fishes of the subfamily Acheilognathinae, commonly known as bitterlings, deposit eggs in the gill cavities of freshwater mussels (Kitamura, 2007; Kitamura et al., 2012). About 60 fish species of bitterlings are ~~currently found~~ in the genera *Acheilognathus*, *Tanakia*, and *Rhodeus* (Arai, 1988). ~~We here identified~~ *Acheilognathus intermedia,* ~~*Acheilognathus*~~ *macropterus,* ~~*Acheilognathus*~~ *majusculus,* ~~*Acheilognathus*~~ *rhombeus, Rhodeus suigensis,* ~~*Rhodeus*~~ *uyekii, Tanakia somjinensis,* and ~~*Tanakia*~~ *signifier* ~~with~~ higher than 99 % ~~sequence identity to~~ the database. Three haplotypes from the Seomjin river showed a 99 % sequence identity to the respective Korean haplotype of *Acheilognathus intermedia* (EF483933), *Tanakia somjinensis* (FJ515921), and *Tanakia signifier* (EF483930). Among them, *Tanakia somjinensis* and ~~*Tanakia*~~ *signifier* are endemic to Korea (Kim and Park, 2002). One haplotype from the Taehwa river (TH3) showed a 100 % identity to the Korean haplotype of *Rhynchocypris semotilus* (KT748874). This species is currently categorized as critically endangered in the Red Data Book of endangered fishes in Korea (Ko et al., 2011).

Two species are currently known in the genus *Sarcocheilichthys* in Korea, ~~*Sarcocheilichthys*~~ *nigripinnis morii* and ~~*Sarcocheilichthys*~~ *variegatus wakiyae* (Kim and Park, 2002). Two haplotypes from the Seomjin river (SJ2) and the Hyeongsan river (HS2) showed 100 % and 97 % sequence identity to the Korean haplotype of ~~*Sarcocheilichthys*~~ *variegatus wakiyae* (GenBank Number: KU301744). One haplotype from the Hyeongsan river (HS2) showed 100 % and 99.43 % sequence identity to the Japanese haplotype of *Sarcocheilichthys soldatovi* (LC146036) and the Korean haplotype of *Sarcocheilichthys nigripinnis morii* (AP017653) respectively. However, ~~*Sarcocheilichthys*~~ *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.

Gobiidae

We identified 16 haplotypes of the family Gobiidae, which represent 7 genera and 11 species (Fig. 3). Five haplotypes were identified in the genus *Tridentiger*, which represents five known species ~~in~~ the genus ~~*Tridentiger*~~ recorded in Korea (Kim et al., 2005). One haplotype from the Taehwa river (TH4) showed a 100 % identity with the Korean haplotype of *Tridentiger obscurus* (GenBank Number: KT601092). One haplotype from the Hyeongsan river (HS4) showed 100 % identity to the Japanese haplotype of *Tridentiger trigonocephalus* (GenBank Number: LC385175) and another haplotype from Seomjin river (SJ3) showed 100 % identity with the Korean haplotype of ~~*Tridentiger*~~ *trigonocephalus* (GenBank Number: KM030481). According to the phylogenetic tree recovered, the ~~*Tridentiger*~~ *trigonocephalus* haplotype from ~~that of~~ 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 Korean (KM030471) and Japanese (LC049760) ~~haplotype~~ of *Rhinogobius brunneus* with 100 % identity, whereas the other ~~one~~ haplotype showed 100 % identity (KM030475) to the Korean haplotype of *Rhinogobius giurinus.* Two haplotypes of *Gymnogobius* sp. from the Taehwa river and Hyeongsan river showed a 98 % sequence identity to *Gymnogobius 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.

Cobitidae

Sixteen species in five genera of the family Cobitidae are currently reported in Korean rivers (Kim, 2009). A total of 18 haplotypes, which represent five genera ~~in~~ the family, were identified herein (Fig. 4). Two haplotypes in the genus *Cobitis* from the Seomjin river were most closely related to the Japanese haplotype of ~~*Cobitis*~~ *tetralineata* (LC146139) with 100 % and 99 % identity, respectively. Two haplotypes from the Taehwa river showed 98 % and 97 % identity to ~~*Corbitis*~~ *hankugensis* (LC146140). Two species ~~in~~ the genus *Misgurnus* are currently reported from ~~the~~ Korean waters, ~~*Misgurnus*~~ *mizolepis* and ~~*Misgurnus*~~ *anguillicaudatus* (Kim, 2009). Interestingly, two phylogenetically distinct clades ~~in~~ *M. anguillicaudatus* were identified ~~by~~ the phylogenetic analysis (Fig. 4). One of them was grouped with the Chinese haplotype of *Misgurnus bipartitus* (KF562047), while the other one ~~was~~ clustered with the Korean haplotype of *Misgurnus mizolepis* (AP017654). *Misgurnus bipartitus* is currently reported as endemic to China, and sequence data of Korean freshwater fishes in GenBank data should be reexamined.

Two haplotypes from the Hyeongsan river (HS1; KJ699181) and the Taehwa river (TH4; KM186182) showed 100 % identity with the distantly located Chinese haplotypes of *Paramisgurnus dabryanus* (Fig. 4). This species is regarded as endemic to China, and *P. dabryanus* is often imported to Korea mixed with *Misgurnus anguillicaudatus* due to their ~~morphological~~ similarity. ~~The previous study showed~~ that there are several geographically different populations of *P. dabryanus* ~~(Shimizu and Takagi, 2010)~~, and ~~those~~ two haplotypes ~~indicated~~ that *P. dabryanus* had been imported from various locations of China. One haplotype from the Taehwa river (TH1) showed a 100 % sequence identity to the Korean haplotype of *Niwaella multifaciata* (EU670806), while another one from the Hyeongsan river (HS1) showed lower (96 %) identity to *Niwaella* sp. ~~So,~~ further ~~study~~ should be conducted to confirm the ~~haplotype~~ ~~of the genus~~ in the Hyeongsan river.

Other families of the Teleostei

Besides the three main families of the Teleostei identified in this study, 27 additional haplotypes, representing 19 species belonging to14 genera and 11 families were also identified, in the following families: Bagridae (5), Mugilidae (4), Anguillidae (1), Centrarchidae (3), Channidae (1), Clupeidae (2), Odontobutidae (3), Pleuronectidae (1), Siluridae (3), Sinipercidae (3), and Amblycipitidae (1). All the haplotypes ~~in~~ the family Bagridae were clearly identified, which included *Pseudobargrus ussuriensis*, *~~Pseudobargrus~~ koreanus*, *Tachysurrus nitidus*, and *~~Tachysurus~~ fulvidraco* (Fig. 5). Two species of *Silurus* are currently known in ~~the~~ Korean waters, *~~Silurus~~ microdorsalis*, and *~~Silurus~~ asotus* (Park and Kim, 1994). One haplotype from the Taehwa river (TH1) showed a 99 % sequence identity with the Korean haplotype of *~~Silurus~~ microdorsalis* (GenBank Number: KT350610), whereas another haplotype from the Seomjin river (SJ1) showed lower identity (96 %) with *~~Silurus~~ microdorsalis* (KT350610).

One haplotype of the Amblycipitidae from the Seomjin river showed 97 % and 96 % identity to the Chinese haplotype of *Liobagrus styani* (KX096605) and the Korean haplotype of *Liobagrus mediadiposalis* (KR075136), respectively. This result ~~indicated~~ that haplotypes in the family should be supplemented for their accurate identification. Three species of *Odontobutis* are currently known in Korea*,* *~~Odontobutis~~ interrupta*, *~~Odontobutis~~ platycephala*, and *~~Odontobutis~~ obscura* (Kim et al., 2005). Two of them (*O. interrupta* and *O. platycephala*) were identified in this study. Two haplotypes ~~in~~ genus *Coreoperca* showed 100 % and 97 % sequence identity to the Korean haplotype of *Coreoperca herzi* (KR075132). Since two species of *Coreoperca* are reported as endemic to the Korean peninsula (Kim et al., 2005), the second haplotype is most likely *~~Coreoperca~~ kawamebari*, but further study should be conducted for confirmation of this haplotype. Two invasive species of the family Centrarchidae, the Bluegill (*Lepomis macrochirus*), and the Largemouth bass (*Micropterus salmoides*) were also identified in this study. Those two species are endemic to North America but were introduced in the Korean peninsula for aquaculture purposes without considering the impacts on the ecosystem.

Fish biodiversity in the four rivers

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

The highest number of species was recorded in the Seomjin river (52 species), followed by the Taehwa river (42 species), Hyeongsan river (40 species), and Nakdong river (26 species). A total of 17 species were exclusively recorded ~~in~~ Seomjin river~~, which include~~ *Cobitis tetralineata, Squalidus gracilis, 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. By contrast, five species ~~from~~ Taehwa River~~:~~ *Pseudogobius masago, Mugilogobius abei, Acanthogobius lactipes, Rhynchocypris semotilus,* and *Silurus microdorsalis*, followed by four species ~~from~~ Nakdong River~~;~~ *Tachysurus nitidus, Rhinogobius giurinus, Pseudobagrus ussuriensis,* and *Plagiognathops microlepis* were identified, respectively. Only three species, including *Squalidus multimaculatus, Sarcocheilichthys soldatovi,* and *Nipponocypris koreanus* were exclusively detected in the Hyeongsan river (Fig. 6).

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

Clustering analysis

In order to ~~know~~ the correlation between the fish assemblage and sample stations, ~~we conducted~~ a heat-map analysis with ~~30~~ most abundant ~~species using Primer~~ (Clarke and Gorley, 2015). The result ~~clearly demonstrated~~ species distribution according to different sampling stations from upstream to downstream (Fig. 8). In upstream (Station 1 and 2), dominant species are *Zacco platypus*, *Odontobutis interrupta, ~~Odontobutis~~ platycephala, Nipponocypris temminckii, Rhynchocypris lagowskii, Misgurnus mizolepis, Coreoperca herzi, Acheilognathus intermedia,* and *Tanakia signifier.* In station 3, the dominant species are *Pseudorasbora parva, Gymnogobius breunigii, Rhinogobius giurinus, ~~Rhinogobius~~ brunneus,* and *Mugil cephalus,* whereas in its downstream (Station 4), *Tridentiger obscurus, ~~Tridentiger~~ trigonocephalus, Konosirus punctatus, Mugil cephalus, Anguilla japonica, Planiliza haematocheila* were identified as the dominant species, all of which are either euryhaline or anadromous (https://www.fishbase.org). This result ~~indicated~~ that salinity is one of the essential factors to determine the fish assemblage at the downstream of ~~the~~ rivers.

DISCUSSION

~~In the present study, we were able to know~~ that eDNA metabarcoding using the MiFish pipeline ~~would be~~ a useful tool for the fish biodiversity ~~analysis~~ ~~which recovered~~ a total of 125 unique haplotypes including at least 73 ~~species only~~ by a single-day survey of 16 sampling stations of ~~the~~ four rivers (Fig. 2-5). According to the “Survey and Evaluation of Aquatic Ecosystem Health (SEAEH)”, a total of 130 freshwater species of fishes were identified from 953 sampling sites in ~~most of the~~ Korean rivers and lakes (Yoon et al., 2012). The numbers of confirmed fish species by eDNA metabarcoding were approximately 56.15 % of those obtained by the year-long conventional survey, and its proportions would be higher ~~considering~~ ‘unidentified’ ~~species~~. This result strongly suggest~~ed~~ ~~that a freshwater fish biodiversity survey in Korea would be possible using~~ eDNA metabarcoding ~~platform~~ with the MiFish ~~pipeline~~ ~~for its incomparable~~ cost and labors ~~compared~~ with ~~conventional~~ morphology-based surveys ~~in Korea~~. Although the methodology in each group may be slightly different, similar conclusions have been ~~drawn from the~~ other studies (Bista et al., 2017; Deiner et al., 2016). This is also adequate for surveying aquatic species ~~inside~~ in protected areas to minimize disturbance of vulnerable communities as well (Fernandez et al., 2018). Notably, most of the rivers in Korea are the primary source for the drinking water in metropolitan cities, and eDNA metabarcoding would be more importantly used for those rivers.

~~Although~~ eDNA metabarcoding analysis using the MiFish pipeline ~~seems to be~~ a useful tool to monitor the biodiversity of freshwater fish, several drawbacks still need to be overcome. First, MiFish sequence data for the endemic species in 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 uploaded to the GenBank database. Besides the lack of sequence data, habitats for freshwater fish species have been fragmented and isolated for a long time, and the intra-~~species~~ genetic distance is generally higher than those ~~for the~~ marine species (Seehausen and Wagner, 2014). Therefore, it is strongly required to establish the haplotype database of the endemic fish species for accurate species identification. Secondly, MiFish primer amplifies the 12S rRNA gene (163-185 bp) region of mitochondrial DNA, which is much smaller than in size as well as lower in sequence variance compared with the typically used COI region (IVANOVA et al., 2007). In fact, the MiFish region was unable to differentiate several closely related marine fish taxa, such ~~as *Sebastes* spp. and *Takifugu* spp.,~~ (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 species ~~in the genus~~ *Carassius* was too low (0.01) to discriminate against one another in the MiFish region (Fig. 2). The supplemented strategy should be designed for those taxa to obtain accurate results.

Although we here analyzed fish biodiversity based on the MiFish pipeline, further study should be made to adopt the quantitative analysis. 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 ~~the river~~ including water dynamics (Deiner and Altermatt, 2014; Jerde et al., 2016; Wilcox et al., 2016) or decaying times with different physical, chemical, or biological factors (Shapiro, 2008). Although several studies about the decaying times of eDNAs in the laboratory and natural conditions (Alvarez et al., 1996; Matsui et al., 2001; Zhu, 2006), ~~it is generally known that the~~ short fragments of DNA are degraded slower than larger ~~ones~~ increasing the probability of detection ~~from~~ the natural environments (Deagle et al., 2006). However, it is still far from establishing reliable methods for the accurate measurement of eDNA in rivers/streams yet, and more data should be accumulated for accurate values. For the quantitative study, ~~the~~ standardized collection methods and pretreatment procedures for the NGS sequencing analysis should be ~~established~~ as well. One of the strongest points in the biodiversity survey by eDNA metabarcoding is ~~a~~ large number of data sets, which would be useful for the statistical analysis compared with ~~the~~ conventional surveys. However, large amounts of data have been produced ~~by~~ different water collection methods, eDNA preparation, sequencing, and bioinformatics analysis platforms ~~in respective~~ research groups in different countries. Therefore, the interconversion of data is currently not possible, and it is required to establish a standard in the overall methodology of eDNA metabarcoding. As one of them, the MiFish pipeline would be a feasible bioinformatic platform for eDNA metabarcoding analyses of fish biodiversity with little modifications and supplementation for the regional application.

~~We here identified~~ the highest species richness ~~in the~~ Seomjin river (3.48) compared with ~~those of~~ the other three rivers: Taehwa river (3.06), Hyeongsan river (2.95), and Nakdong river (2.86). ~~Low species richness~~ in Nakdong, Hyeongsan, and Taehwa ~~river~~ ~~presumably due~~ to the higher anthropogenic ~~effects in~~ ~~these~~ rivers. Like the other Korean rivers, those three rivers run through highly populated metropolitan cities, in which rivers are exposed to various human impacts, which directly or indirectly promote changes in diversity and distribution of freshwater fishes (Finkenbine et al., 2000). In particular, the lowest species richness (2.86) ~~and endemic~~

~~species numbers~~ (only one, *Odontobutis interrupta*) were identified in the Nakdong river ~~along which~~ the highest numbers of constructions and populations exist among the sampled rivers. Lee et al.~~,~~ (2015) reported only two endemic species (*Coreoperca herzi* and *Odontobutis platycephala*) ~~from~~ the Nakdong river by the traditional survey method. On the other hand, eight endemic species including *Coreoleuciscus splendidus, Iksookimia longicorpa, Microphysogobio koreensis, Microphysogobio yaluensis, Odontobutis interrupta, ~~Odontobutis~~ platycephala, Pseudobagrus koreanus,* and *Squalidus gracilis* were identified ~~in~~ Seomjin river, ~~which was~~ similar to ~~the~~ previous results (Jang et al., 2003; Lee et al., 2015). The ~~various~~ constructions along the urbanized watershed, 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). Different from those three rivers, there is no metropolitan city along ~~with~~ the Seomjin river, which is, therefore, less exposed to anthropogenic impacts. As freshwater ecosystems are easily disturbed~~,~~ and ~~it takes a long time to recover compared to other ecosystems~~ (Ricciardi and Rasmussen, 1999), ~~The~~ long-term survey should be conducted to establish the clear correlations between anthropogenic factors and fish assemblage in the Korean rivers.

The eDNA metabarcoding analysis also ~~revealed some~~ exotic ~~species~~ are widely distributed in ~~inland Korean waters~~. We were able to identify at least five exotic ~~fish~~ species, including *Carassius cuvieri, Cyprinus carpio, ~~Cyprinus~~ megalophthalmus, Lepomis macrochirus*, and *Micropterus salmoides* (Table S3). Those exotic species may impact ~~on~~ the native fishes for shelter and spawning sites as well as disturbing the food change preying on the native fishes. In addition, since the species has a high reproductive capacity makes it potential invasive species (Keller & Lake, 2007; Koster *et al*, 2002; Nico & Fuller 2010). Our results also surprisingly revealed that the largemouth bass, *M. salmoides*, and ~~bluegill,~~ *L. macrochirus* are likely ~~to~~

present in all sampled ~~four~~ rivers. ~~As native to eastern North America, those~~ two ~~species~~ were artificially introduced in the ~~1970s~~, as freshwater fish stock without any further ~~consideration~~ of the effects on the freshwater ~~ecosystem~~ in ~~Korea,~~ The species ~~has spread~~ throughout the Korean peninsula ~~competing~~ with the native ~~species and~~ their long-term survey should be conducted (Jang et al., 2002; Yoon et al., 2012). Freshwater ecosystems are much more vulnerable to invasive species causing biodiversity loss and global change (Clavero and García-Berthou, 2005) and the eDNA metabarcoding analysis would be useful to monitor the distribution patterns of the invasive species in Korean rivers.

Collectively, we firstly analyzed the fish biodiversity in four rivers (Nakdong, Hyeongsan, Seomjin, and Taehwa) using the eDNA metabarcoding with the MiFish platform. Our result clearly showed that eDNA metabarcoding is a reliable tool to monitor the fish biodiversity with low cost and labors compared with the traditional survey methods. This method is also useful to monitor the exotic species or rare species with a little adverse effect on the ecosystem in the river. eDNA metabarcoding platform would be much more effective if several issues were upgraded, such as the supplement of the local species data, standardized sample preparations, and quantitative methodologies. As those data accumulate, we would able to obtain better information about the changes in fish assemblage structure in a river caused by various biotic or abiotic factors, including climate change, pollution, or the introduction of foreign species.

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, 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., Christmas, M., 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., Geng, Y., 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, e0177724.

Clarke, K., Gorley, R., 2015. Getting started with PRIMER v7. PRIMER-E: Plymouth, Plymouth Marine Laboratory.

Clavero, M., García-Berthou, E., 2005. Invasive species are a leading cause of animal extinctions. Trends in Ecology & Evolution 20, 110.

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.

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, 11.

Deiner, K., Altermatt, F., 2014. Transport distance of invertebrate environmental DNA in a natural river. PloS one 9, 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, 953-959.

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.

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, 1149-1160.

Foote, A.D., Thomsen, P.F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L.A., Salling, A.B., Galatius, A., Orlando, L., Gilbert, M.T.P., 2012. Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. PloS one 7, e41781.

Gray, J.S., 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.

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.

IVANOVA, N.V., ZEMLAK, T.S., HANNER, R.H., HEBERT, P.D.N., 2007. Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes 7, 544-548.

Jang, M.-H., Lucas, M.C., 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.

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, 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, 8770-8779.

Kelly, R.P., Closek, C.J., O'Donnell, J.L., Kralj, J.E., Shelton, A.O., Samhouri, J.F., 2017. Genetic and manual survey methods yield different and complementary views of an ecosystem. Frontiers in Marine Science 3, 283.

Kelly, R.P., Port, J.A., Yamahara, K.M., Martone, R.G., Lowell, N., Thomsen, P.F., Mach, M.E., Bennett, M., Prahler, E., Caldwell, M.R., 2014. Harnessing DNA to improve environmental management. Science 344, 1455-1456.

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

Kim, I.-S., Park, J.-Y., 2002. Freshwater fishes of Korea. Kyo hak sa.

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

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.

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, 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. Molecular biology and evolution 33, 1870-1874.

Lee, J.W., Yoon, J.D., Kim, J.H., Park, S.H., Baek, S.H., Yu, J.J., Jang, M.H., Min, J.I., 2015. Length-weight relationships for 18 freshwater fish species from the Nakdong River in South Korea. Journal of Applied Ichthyology 31, 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, 353-359.

Magurran, A.E., 1988. Ecological diversity and its measurement. Princeton 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, 95-102.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J.Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Araki, H., 2015a. 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.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J.Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Araki, H., Kondoh, M., Iwasaki, W., 2015b. 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.

Nilsson, C., Reidy, C.A., Dynesius, M., Revenga, C., 2005. Fragmentation and flow regulation of the world's large river systems. Science 308, 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, 105-112.

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., 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.

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.

Ricciardi, A., Rasmussen, J.B., 1999. Extinction rates of North American freshwater fauna. Conservation biology 13, 1220-1222.

Riley, S.P., Busteed, G.T., Kats, L.B., Vandergon, T.L., Lee, L.F., Dagit, R.G., Kerby, J.L., Fisher, R.N., Sauvajot, R.M., 2005. Effects of urbanization on the distribution and abundance of amphibians and invasive species in southern California streams. Conservation Biology 19, 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, 1553-1555.

Seehausen, O., Wagner, C.E., 2014. Speciation in Freshwater Fishes. Annual Review of Ecology, Evolution, and Systematics 45, 621-651.

Shapiro, B., 2008. Engineered polymerases amplify the potential of ancient DNA. Trends in biotechnology 26, 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.

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, e0175186.

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.

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, T.M., McKelvey, K.S., Young, M.K., Sepulveda, A.J., Shepard, B.B., Jane, S.F., Whiteley, A.R., Lowe, W.H., 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, 688-695.

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.

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.

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. Springer, 175-191.

Zhang, Y., 2015. An Introduction to Python and computer programming, An Introduction to Python and Computer Programming. Springer, 1-11.

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, 3231-3238.

Figure 1

Water sample collection sites of four Korean rivers

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

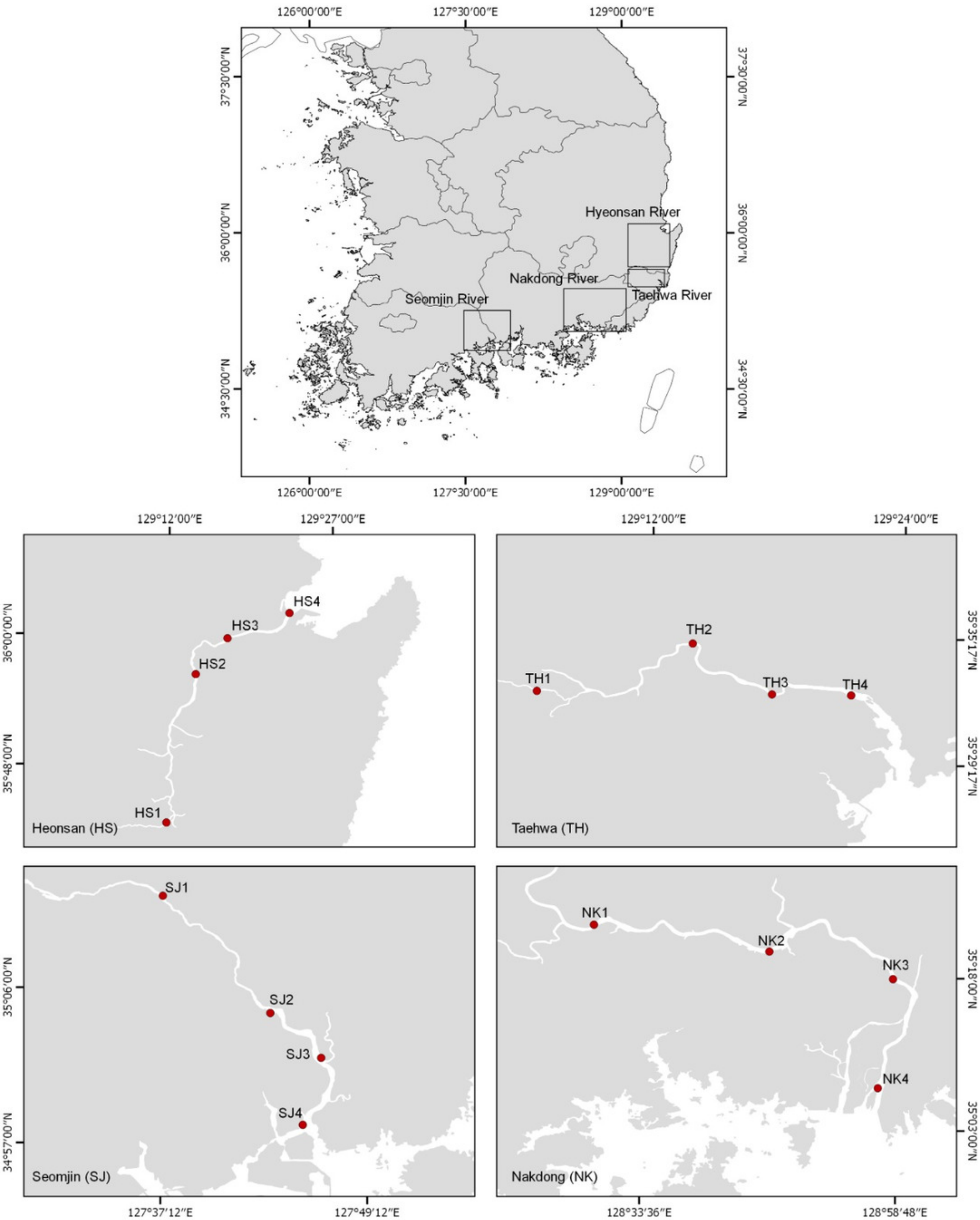
126°00′00″E
127°30′00″E
129°00′00″E
37°30′00″N
36°00′00″N
34°30′00″N
Hyeonsan River
Nakdong River
Taehwa River
Seomjin River
129°12′00″E
129°27′00″E
36°00′00″N
35°48′00″N
HS1
HS2
HS3
HS4
Heonsan (HS)
129°12′00″E
129°24′00″E
35°35′17″N
35°29′17″N
TH1
TH2
TH3
TH4
Taehwa (TH)
35°06′00″N
34°57′00″N
SJ1
SJ2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

Figure 2

Phylogenetic tree of the fish species under the family Cyprinidae

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.

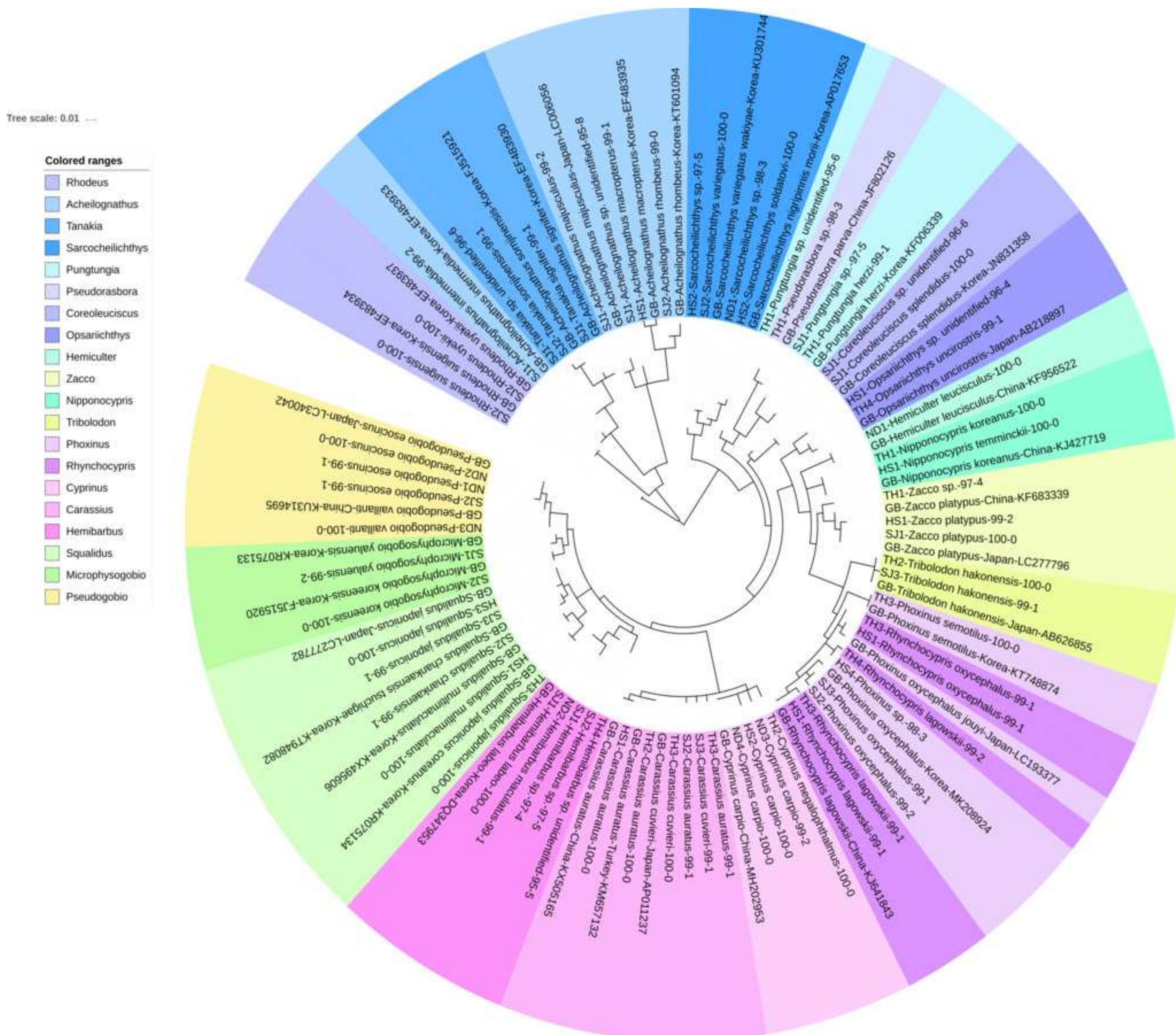

Figure 3

Phylogenetic tree of the fish species under the family Gobiidae

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.

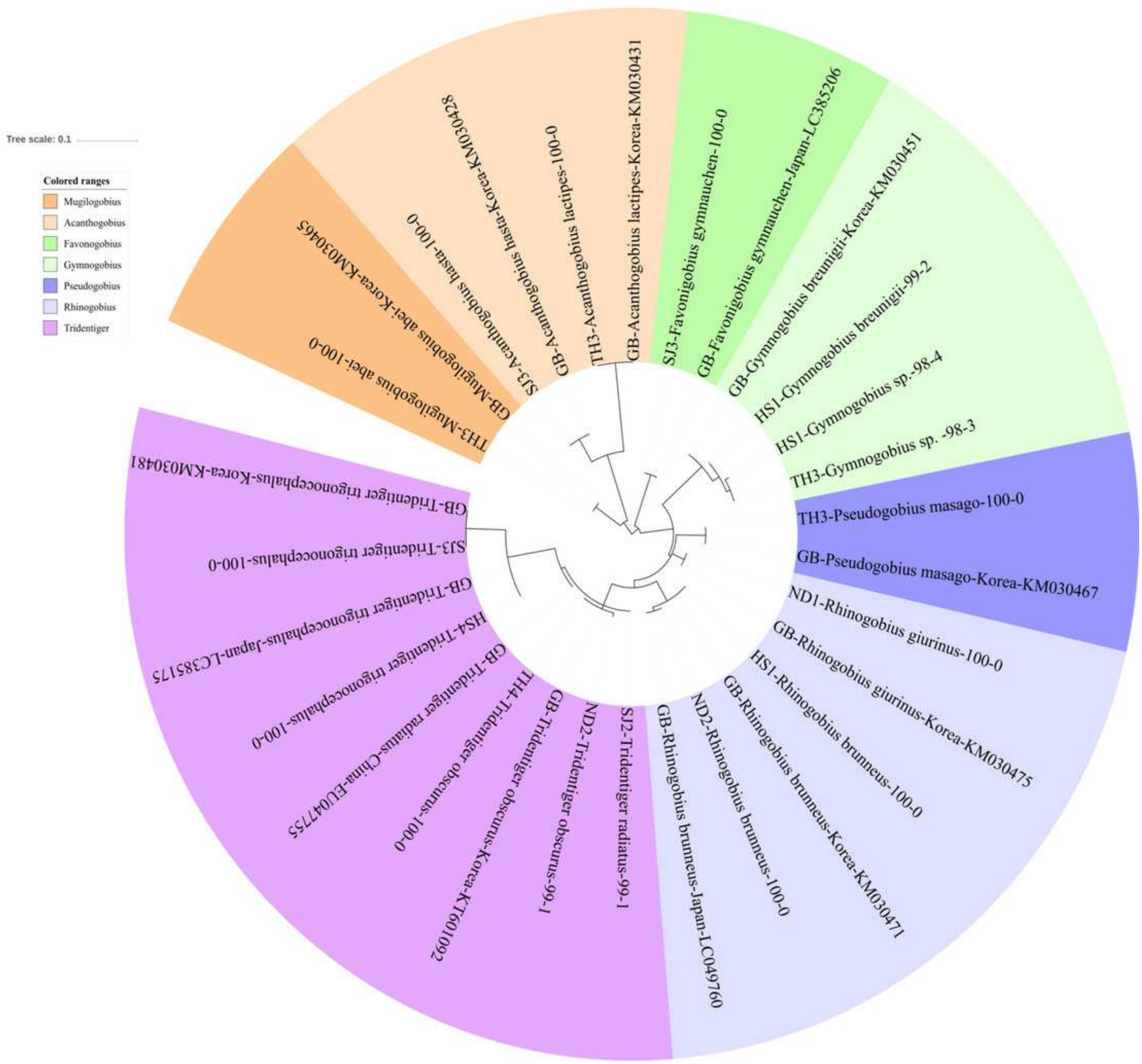

Figure 4

Phylogenetic tree of the fish species under the family Cobitidae

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.

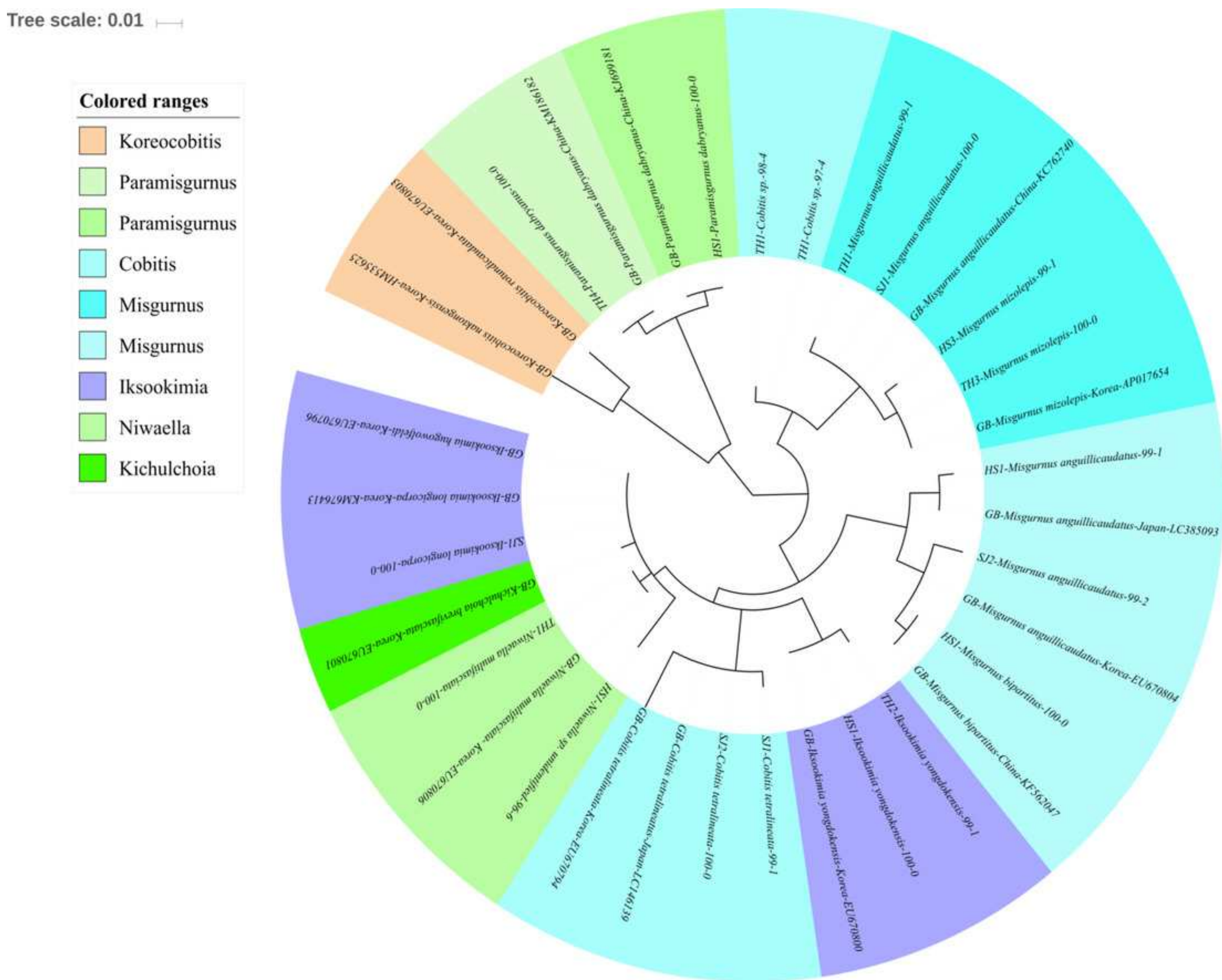

Figure 5

Phylogenetic tree of the fish species under the other families

Figure 5 Phylogenetic tree analysis of fish species under the other families of Teleostei. Phylogenetic tree was constructed by Maximum likelihood (ML) algorithm (MEGA 7.0) under the 1000 replication bootstrap.

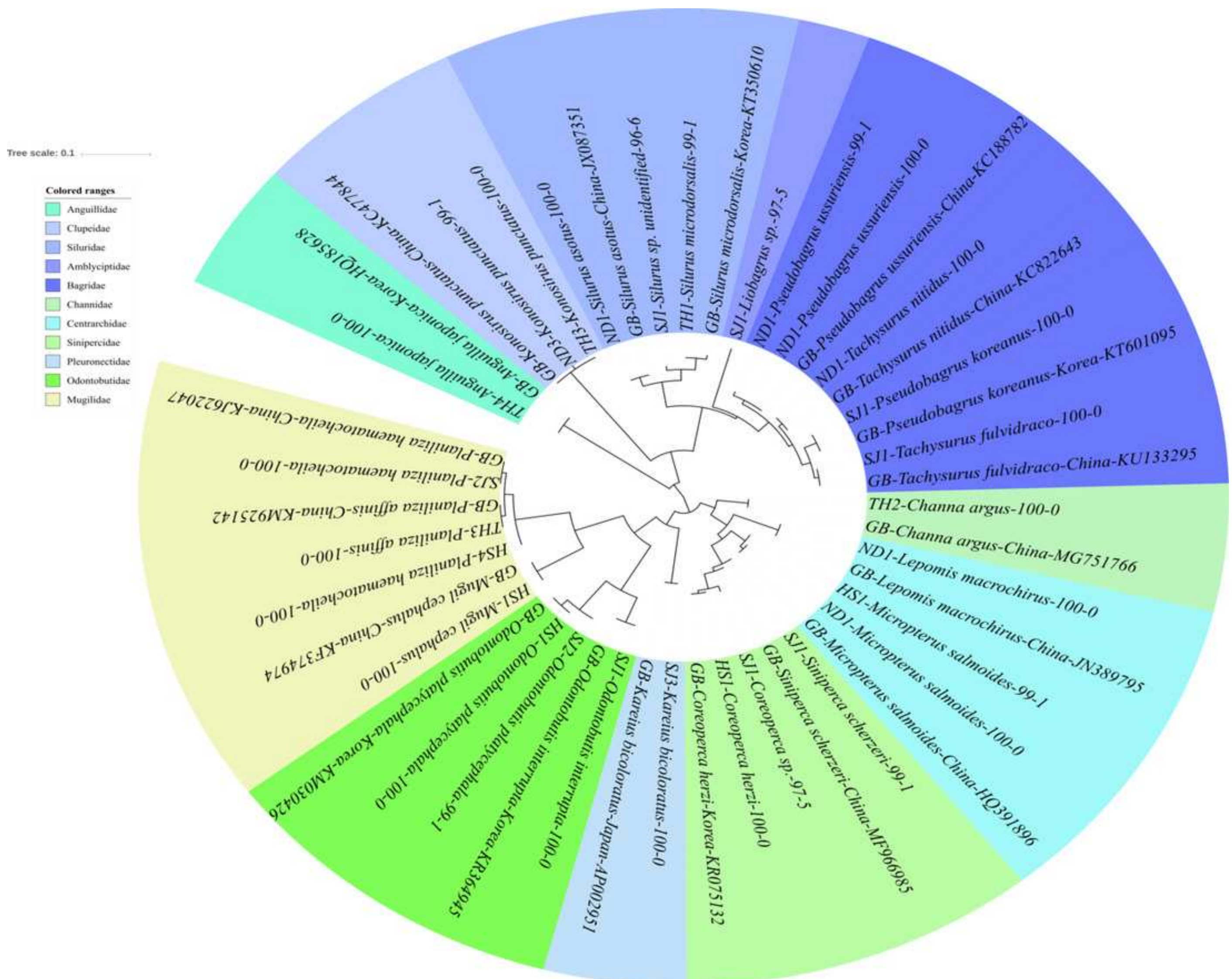

Figure 6

Venn diagram of fish species identified in the four Korean rivers.

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

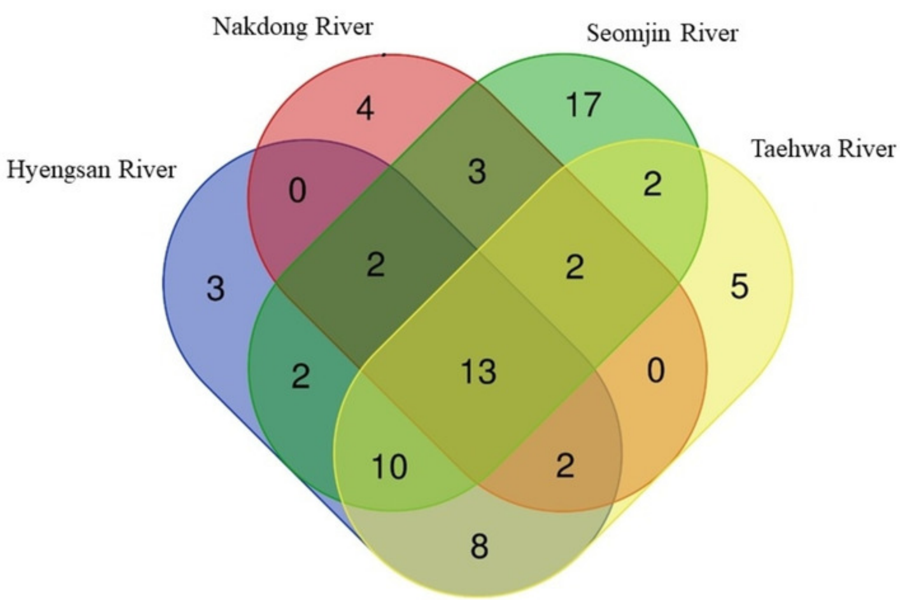

Figure 7

Proportion of families detected from the four Korean rivers

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

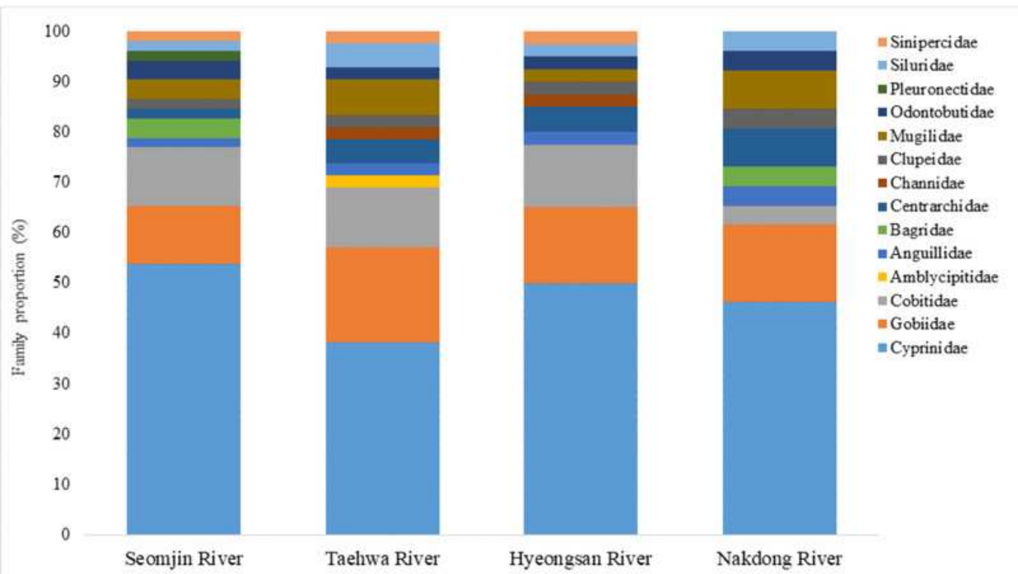

Figure 8

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

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.

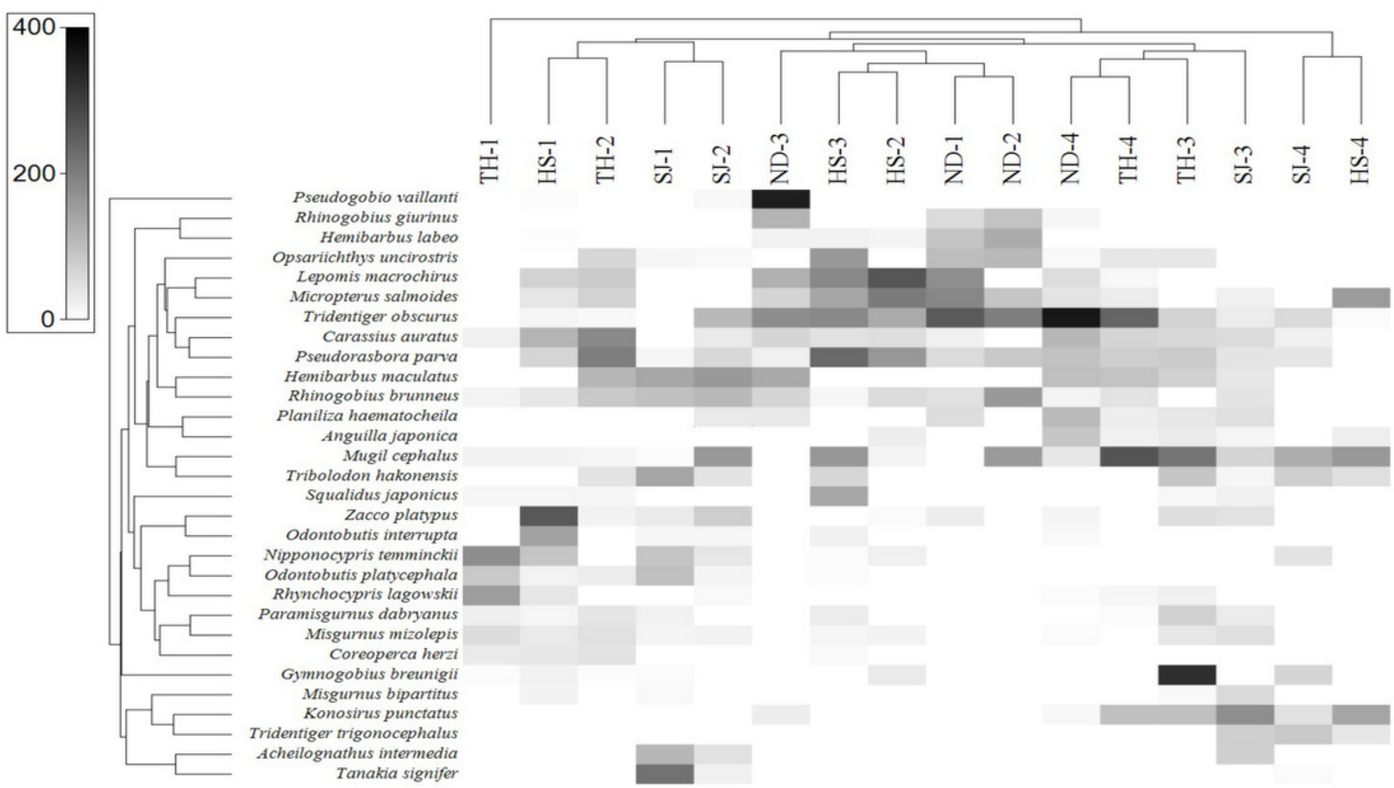

Table 1(on next page)

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

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

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)

Table 2 Taxonomic assignment summary of of the MiSeq reads from four Korean rivers

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

* Final number, after removal of duplicated one in brackets

Table 3(on next page)

Table 3 List of haplotypes of fishes identified by eDNA metabarcoding study in four Korean rivers

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)

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

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	-