Thesis for the Degree of Doctor of Engineering

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

by

Md. Jobaidul Alam

Interdisciplinary Program of Biomedical, Mechanical

and Electrical Engineering

The Graduate School

Pukyong National University

Feruary 21, 2020

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

(한국하천에서의 환경 DNA 메타바코딩 기법을 이용한 어류 생물다양성 평가)

Advisor: Prof. Hyun-Woo Kim

by

Md. Jobaidul Alam

A thesis submitted in partial fulfillment of the requirements
for the degree of

Doctor of Engineering

in Interdisciplinary Program of Biomedical, Mechanical and Electrical Engineering, The graduate School,
Pukyong National University

Feruary 2020

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

A dissertation

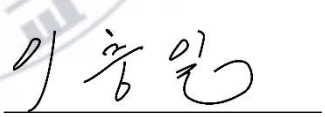
by

Md. Jobaidul Alam

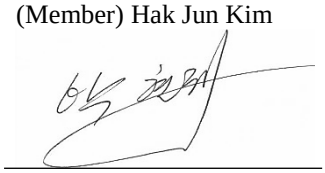
Approved by:

(Chairman) Do-Hyung Kim

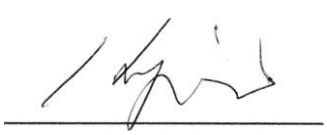
(Member) Hak Jun Kim



(Member) Chung Il Lee



(Member) Hyun Je Park



(Member) Hyun-Woo Kim

Feruary 21, 2020

CONTENTS

Contents …… i

List of Figures …….. iii

List of Tables ……..... **v**

Abstract ……..... **vi**

Chapter 1. General Introduction …… **1**

Chapter 2. Fish biodiversity study in four Korean rivers by the environmental DNA metabarcoding approach ……. 11

2.1. Introduction …….. **12**

2.2. Materials and Methods …… **15**

2.2.1 Sample collection and environmental DNA extraction …. 15

2.2.2 Construction of library and MiSeq sequencing ……... 17

2.2.3 Bioinformatics analysis of NGS data …… 18

2.3 Results ……..... **19**

2.3.1 Physico-chemical parameters …… 19

2.3.2 Analysis of fish haplotypes obtained by the MiFish pipeline …….. 19

2.3.3 Cyprinidae ……. 25

2.3.4 Gobiidae …… 36

2.3.5 Cobitidae …… 38

2.3.6 Other families …… 40

2.3.7 Fish biodiversity in four rivers …….. 43

2.3.8 Salinity and relationship …… 46

2.3.9 Clustering analysis …… 47

2.4 Discussion ……... **48**

2.5 Conclusion …….. **58**

Chapter 3. Fish biodiversity patterns estimated from the environmental DNA metabarcoding surveys in the Suyeong River, Korea …… 59

3.1. Introduction …… **60**

3.2 Materials and Methods …… **62**

3.2.1 Sample collection and DNA Extraction ……... 62

3.2.2 Construction of library and MiSeq sequencing …… 64

3.2.3 Bioinformatics analysis of NGS data …… 66

3.2.4 Statistical analysis for fish biodiversity indices ……. 67

3.3 Results …… 68

3.3.1 Temperature and salinity changes during the survey period in the Suyeong River…… 68

3.3.2 Analysis of the taxonomic assignment and haplotypes … 70

3.3.3 Fish biodiversity in the Suyeong River …….. 78

3.3.4 Salinity and relationship …… 84

3.3.5 Clustering analysis …… 86

3.4 Discussion …….. **89**

3.5 Future directions for the eDNA metabarcoding research….. 100

3.6 Conclusion ……. **101**

IV. References ……. **102**

V. Abstract in Korean languange ……. 119

VI. Acknowledgment …….. **121**

LIST OF FIGURES

Figure 2.2.1	Sample collection sites of the four Korean rivers …….	16
Figure 2.3.2	Phylogenetic tree of Cyprinidae family by maximum likelihood method …….	26
Figure 2.3.3	Phylogenetic tree of Gobiidae family by maximum likelihood method …….	37
Figure 2.3.4	Phylogenetic tree of Cobitidae family by maximum likelihood method …….	39
Figure 2.3.5	Phylogenetic tree of other families by maximum likelihood method …….	41
Figure 2.3.6	A. Venn diagram of fish species identified from the four rivers, and B. Venn diagram of fish species identified at the different stations of the four rivers ...	43
Figure 2.3.8.	Fish community structure at family level in the four Korean rivers ……..	43
Figure 2.3.9	Heat map of top 30 fish species identified in 16 sampling stations of 4 rivers ……..	47
Figure 3.2.1	Environmental DNA sample collection sites of the Suyeong River ……	66
Figure 3.3.2	Water temperature changes in the Suyeong River …..	
Figure 3.3.3	Changes in salinity (psu) during the study period in the Suyeong River ……..	69
Figure 3.3.4	The phylogenetic tree of the identified fish species in the Suyeong River ……..	73
Figure 3.3.5	Shannon-Wiener diversity index in the Suyeong River...	78

Figure 3.3.6 Proportions of detection frequencies with different sample stations in the Suyeong River …….. 80
Figure 3.3.7 Number of fish species identified at the different sampling stations of the Suyeong River …….. 82
Figure 3.3.8 Venn diagram of fish species identified at different sampling sites in the Suyeong River by the eDNA metabarcoding analysis …… 83
Figure 3.3.9 Fish species distribution with salinity measurement in the Suyeong River …… 85
Figure 3.3.10 Heatmap of the eDNA metabarcoding result based on the top 20 species at four sampling stations of the Suyeong River ……..... 86
Figure 3.3.11 A and B. The Bray-Curtis similarity and Non-metric MDS Resemblance of fish species from the different sampling stations …… 87

LIST OF TABLES

Table 2.2.1	List of primer sets used in this study ……	19
Table 2.3.2	Environmental DNA sample collection sites from the four Korean rivers …….	21
Table 2.3.3	Summary of taxonomic assignment of the MiSeq reads from the four rivers eDNA samples ……..	22
Table 2.3.4	Fish species identified from the four rivers by eDNA metabarcoding ……...……	23
Table 2.3.5	Fish haplotypes with the GenBank numbers identified from the eDNA metabarcoding study of the four rivers..	29
Table 2.3.6	Shannon Index (SI) in the four Korean rivers ……	
Table 2.3.7	Genetic distance of species under the family Cyprinidae	51
Table 2.3.8	List of the exotic fish species identified from the four river eDNA water samples ……	57
Table 3.2.1	Environmental DNA sample collection sites with water temperature (^{0}C) and salinity (psu) measurement from the Suyeong River ……	63
Table 3.3.2	Numbers of haplotypes and reads obtained from the Suyeong River by MiSeq sequencing platform …….	72
Table 3.3.3	List of haplotypes identified by the environmental DNA metabarcoding study from the Suyeong River ….……	74
Table 3.3.4	Number of fish species identified in the Suyeong River in the study period ……	81
Table 3.3.5	Analysis of Similarity (ANOSIM) pairwise test by Prime v7	89
Table 3.3.6	Comparison of the taxonmic assignment between MiFish pipeline and FishBase website ……	99

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

Md. Jobaidul Alam

Interdisciplinary Program of Biomedical, Mechanical and Electrical Engineering,

The Graduate School, Pukyong National University

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.

In the second part of this dissertation, we compared the species abundance and seasonal variation for one year surveys in the Suyeong River. For this study, we have collected water samples from August 2017 to June 2018 at the four sampling stations of the Suyeong River, Korea. The MiFish universal primer set was used for eDNA metabarcoding analysis in this river. Here we identified 65 fish haplotypes from four sample collection stations of the Suyeong River, among those haplotypes, the highest 31 were identified from the family Cyprinidae, followed by Gobiidae (8), and the remaining 15 were from the other 11 families. The highest Shannon-Wiener (H') index was found at the station A (2.364), followed by station C (1.186) and station B (1.039), while the lowest was at the station D (0.976). The average Margalef index was also highest at the station A (3.406), followed by station C (3.073), station D (2.462), and the lowest was at the station B (1.963). Among all sampling stations, station A had the highest average species detection rate (17.33), while the lowest (9.33) was at station B. During the study period, the highest average number of species was found in June (24.25), while the lowest (8.5) was in February. We also identified five exotic species (*Carassius cuvieri*, *Cyprinus carpio*, *Micropterus salmoides*, *Lepomis macrochirus*, and *Oreochromis niloticus*) in the Suyeong River, and now they are widely distributed in Korean waters. Our findings suggest that eDNA metabarcoding required less time and taxonomic expertise, and it is better to understand fish distribution and biodiversity in rivers/streams than the traditional survey systems. Although MiFish metabarcoding successfully presented fish species inhabiting in Korean rivers, additional sequence data should be supplemented for better result. The accuracy for detection of endemic species, endangered species, invasive species, and fish distribution by eDNA analysis is possible very effectively to supplement the traditional monitoring approaches in different ecosystems. In conclusion, we expect that these findings would provide useful information for the

effective management or conservation of fish resources in the inland open water of the Korean peninsula.

Chapter 1

General introduction

1.1 Introduction

Fish is the most diverse group of animals and act as an indicator for understanding the fluctuations of an aquatic ecosystem (Hutchings, 2000; Mahapatra et al., 2014; Pauly et al., 2002). Freshwater fishes and their habitats are facing more threats from anthropogenic activities than terrestrial and marine ecosystems due to the construction of various types of dams, industrial, agricultural and household pollution, and overexploitation or overfishing (Revenga et al., 2000; Revenga and Mock, 2000; Thomsen et al., 2012b).

Nowadays, almost all natural habitats are under stress, especially freshwater habitats (lakes, rivers, and wetlands) are the most vulnerable (Revenga and Mock, 2000). During the last three decades, due to the fast-growing of industrialization, the endemic freshwater fish population of South Korea has been reducing greatly for human activities, various types of water pollution, construction of dams, habitat destruction, and the introduction of exotic fish in the natural water bodies (Colautti et al., 2003; Hong et al., 2017; Sato et al., 2010).

Biodiversity is the genetic variation of living organisms, species, and populations; and their complex amalgamation in ecosystems (Mace et al., 2005). Generally, it directs to the variety and variability of life in the Universe;

biodiversity typically measures variation at three levels-genetic diversity, species diversity, and diversity in the ecosystem (Maclaurin and Sterelny, 2008). Its responses according to the fluctuation of the surrounding environment, vindicate actions of habitats to deliver commodities and services to facilitation for the ethnic welfare (e.g., supply of drinking water, nutrient cycling) in the universe (Costanza et al., 1997; Díaz et al., 2006; Hiddink et al., 2008; Hooper et al., 2005).

Fish biodiversity has intrinsic and aesthetic value: many people admired the beautiful colors and decorative body shapes of coral reef fishes and other fish and shellfishes of aquatic habitats. Few amenities of fish biodiversity currently may not be visible, however, it may be revealed in the upcoming days, i.e. some active compounds derived from the aquatic animals will be a source of essential drugs for restraint or treatments of sickness (Hiddink et al., 2008).

In addition, fish biodiversity is monumental for the sake of sustainable use of water resources, including natural and industrial fisheries management. For sustainable fisheries management, catching/harvesting more than one species has better catch stability to comprises odd species fishery (Dulvy et al., 2000; Hilborn et al., 2003). Following the World Summit on Sustainable Development

(2002) and other international agreements, the natural resources of the countries have to manage in systems as protecting/safeguard the biodiversity and resources (Communities, 2009; Hiddink et al., 2008; Nations, 2002).

Fish biodiversity monitoring results may be able to utilize for various purposes. 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. This result can be able to take some effective steps to enhance efforts from the government and improve public conversance. Fish biodiversity monitoring results also useful for reference data (Yoon et al., 2012).

Nowadays, organisms having commercial importance are not only taking into account, but other biological organisms have to figure out as valuable national assets, protection and conservation of these organisms are increasing. Freshwater habitats are the highest threatened habitats (Sala et al., 2000), and freshwater fish is one of the greatest endangered groups of animals after amphibians (Bruton, 1995). Consequently, in protecting freshwater fish species, set up management strategies aiming to conserve endangered or endemic species, figure out the effects of exotic species, and denominate the protected areas are necessary.

Freshwater fish biodiversity monitoring is essential to supply inevitable information/data, and it should be executed by using a standardized process to the river system (Yoon et al., 2012). A river is a rich ecosystem for diverse fishes and other aquatic organisms. It also plays a vital role in drinking water sources, irrigation for agriculture, sources for various kinds of food materials, recreational activities and employment opportunities (Revenga and Mock, 2000). The major population of the world lives near freshwater environments and depends on it; also most inland cities are located very close to a waterway (Moyle and Leidy, 1992).

In 2014, a total 130 freshwater fish species of 28 families from 953 sites were identified in Korea, among them *Zacco platypus* (28.2%) and *Zacco koreanus* (19.3%) were the most abundant species, then 51 endemic, 20 endangered and 4 exotic fish species were also identified (Yoon et al., 2018). The highest freshwater fish diversity (96 species) and lowest (72 species) diversity were found in the Han River and the Yeongsan/Seomjin River system respectively. The *Macropterus salmoides, Lepomis macrochirus*, and *Carassius cuvieri* were found in all river systems without the north Yeongdong, on the contrary, *Oreochromis niloticus* was found only at three sites, may be due to cold water temperatures in winter (Yoon et al., 2018).

The south sub-region rivers of the Korean peninsula (Nakdong, Seomjin, and Yeongsan River) show resembling fish fauna with Japan and illustrate 13

endemic fish species (Jang et al., 2002). In 2009, total of 124 freshwater fish species, belonging to 27 families were found in the four major river systems of the Korean Peninsula. Among those fish species, the most abundant (85.7%) family was Cyprinidae (54 fish species), followed by Cobitidae (15 species) and Gobiidae (12 species) respectively.

Two exotic fish species of Centrarchidae showed a little bit high (1.9%) abundance (Yoon et al., 2012). To know the fish biodiversity in a river, its regular surveys are required. Traditional fish surveys have been mainly dependent on the direct observation of catches using different types of nets, traps, angling, electrofishing, and counting methods (Bonar et al., 2009; Murphy and Willis, 1996). This type of sample collection and data gathering method is not standard and it needs practical and taxonomic expertise, requires extensive and expensive fieldwork (Hopkins and Freckleton, 2002; Wheeler et al., 2004).

There are some visual and hydroacoustic fish survey methods, i.e. camera recordings, video monitoring system and underwater visual census (Edwards and Schindler, 2017; Egerton et al., 2018); but these systems required sophisticated instruments and skilled manpower and also less effective to figure out the biodiversity of a river/stream. In addition, some areas (e.g. the deep sea, river, and the stream) which is difficult to collect the sample or observe the communities properly. Moreover, the early stages of

fish (i.e. larvae) or other invertebrates also difficult to collect and identification (Jones, 2008; Lenat and Resh, 2001; Pfrender et al., 2010). Moreover, it is impossible to identify the damaged or decayed specimen because of the shortage of morphological characteristics, which is the most important key factor for visual identification (Deiner et al., 2013). Therefore, costly surveys are required for ecological and conservation research to study the specific hypothesis and to understand the biodiversity in a given area (Yamamoto et al., 2017). As a consequence, fish biologists and policymakers accepted an ecosystem-friendly, inexpensive and effective management of fisheries resources (Sinclair et al., 2002).

Environmental DNA (eDNA) metabarcoding (detection of multispecies by using degraded DNA from environmental sample) was introduced as an alternative method of direct observation, which may be useful to reduce systematics error to estimate the species richness resulting misidentification, the low detection rate of species and overall complementary approaches of morphology-based identification (Baird and Sweeney, 2011; DeWalt, 2011; Evans and Lamberti, 2017; Valentini et al., 2009).

Environmental DNA is a brand new tool in Ecology and Conservation Biology, this method is going to be an amazing and effective approach in complying to detect the abundance/distribution of species, which was previously imperceptible to us (Pilliod et al., 2013b). A method that collected

DNA from the environment is using to rummage the aquatic animals in Ecological and Biological studies (Bohmann et al., 2014; Valentini et al., 2009). The genetic materials released by organisms into habitats, collected by water filtration, DNA extraction from the filter, DNA sequenced by polymerase chain reaction (PCR) to detect the species.

This approach is suitable to identify the presence of aquatic organisms living in freshwater ecosystems e.g. ponds, lakes, and the lagoon (Dejean et al., 2012; Doi et al., 2015; Ficetola et al., 2008; Jerde et al., 2013; Takahara et al., 2013; Takahara et al., 2012; Thomsen et al., 2012b; Uchii et al., 2016); rivers, and streams (Fukumoto et al., 2015; Goldberg et al., 2011; Ikeda et al., 2016; Jerde et al., 2011; Mahon et al., 2013; Minamoto et al., 2012; Pilliod et al., 2013a; Takahara et al., 2012; Wilcox et al., 2016; Yamanaka and Minamoto, 2016); marine, and coastal habitats (Miya et al., 2015a; Thomsen et al., 2012a; Thomsen et al., 2016; Yamamoto et al., 2017; Yamamoto et al., 2016).

In most cases, this approach has been used to detect invasive fish species (Doi et al., 2015; Jerde et al., 2013; Jerde et al., 2011; Mahon et al., 2013; Takahara et al., 2013; Uchii et al., 2016), endemic/endangered fish species (Doi et al., 2017; Fukumoto et al., 2015; Ikeda et al., 2016; Thomsen et al., 2012b), but a few studies have been conducted to detect the whole fish

biodiversity through the eDNA approach (Doi et al., 2017; Thomsen et al., 2012a; Yamamoto et al., 2017).

In this study, to our knowledge, for the first time in Korea, we used the eDNA metabarcoding approach to rendering the fresh-water fish biodiversity from river water samples. To perceive the freshwater fish biodiversity through environmental DNA analysis, a set of fish-universal primers, MiFish (Miya et al., 2015a) were used, which is compatible with fish eDNA metabarcoding (Yamamoto et al., 2017). MiFish primers amplify the 12S rRNA gene (163-185 bp) of mitochondrial DNA, which is a hypervariable region and contains adequate information for identifying fishes in most cases up to the species level.

A total of 232 fish species from 152 genera and 70 families have been detected from eDNA sample of seawater by MiFish primers using MiSeq Illumina platform (Miya et al., 2015a). These primers detected more than 90% fish species by using eDNA metabarcoding (168 species belongs 14 orders) in an aquarium. Having a short amplicon length, less cross-reactivity, these primers can amplify from decayed/degraded DNA sample, and short amplicons are more competent for MiSeq sequencing (Yamamoto et al., 2017). Moreover, to eDNA, this metabarcoding method is suitable for bulk samples, e.g. net collection samples having the extent of juveniles, fish

larvae/eggs, and damaged or decayed specimens having a few signs for species identification (Miya et al., 2015a).

In this study, we analyzed fish biodiversity patterns by the environmental DNA metabarcoding analysis in five major rivers of the southern territory of the Korean peninsula, i.e. the Hyeongsan River, Taehwa River, Seomjin River, Nakdong River, and the Suyeong River. The physicochemical parameters of water and fish abundance were compared to understand the relationship between the water quality parameters with the fish availability, as well as to understand a healthy ecological condition of the freshwater ecosystem.

Chapter 2

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

2.1 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., 2013b), 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.

2.2 Materials and Methods

2.2.1 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.

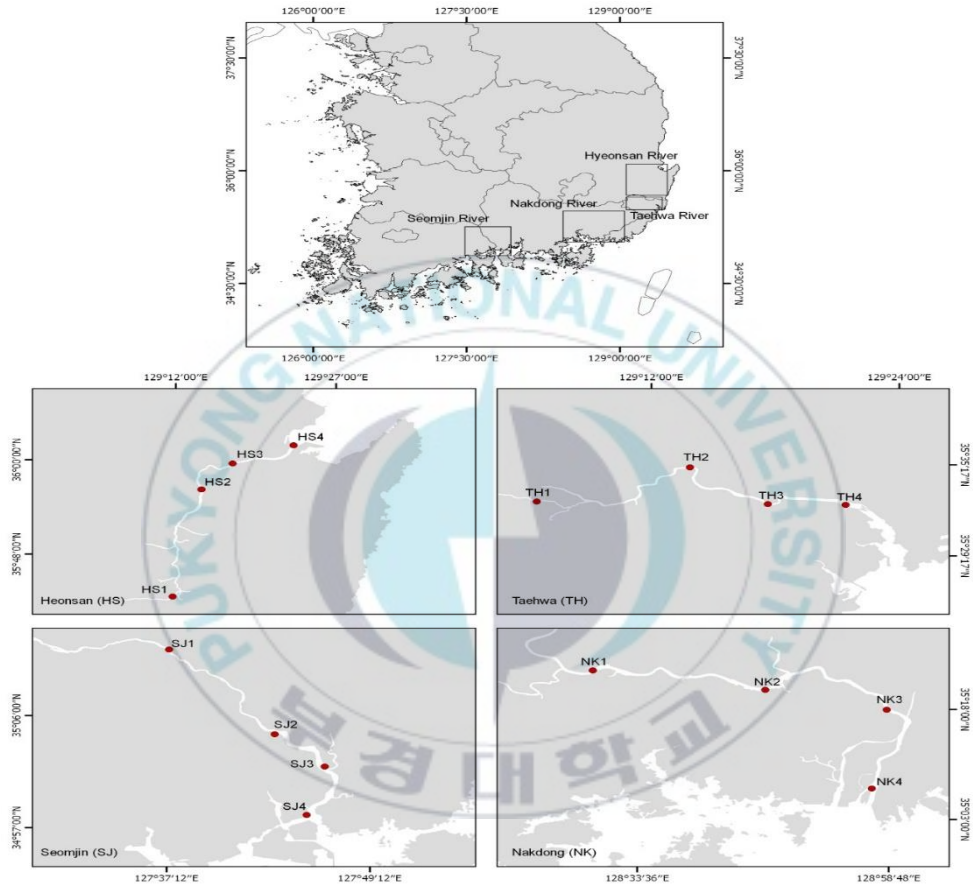


Figure 2.2.1. Sample collection sites of the four Korean rivers

2.2.2 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).

Table 2.2.1 Primer list were used in this study

Primer	Sequences (5’ to 3’)	Target region	Reference
MiFish F	GTCGGTAAAACTCGTGCCAGC	12s rRNA	(Masaki Miya et al., 2015)
MiFish R	GTTTGACCCTAATCTATGGGGTGATAC		

2.2.3 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).

2.3 Results

2.3.1 Physico-chemical parameters

The water temperature of the sample sites ranged from 18.6 ˚C to 24.20 ˚C (Table 2.3.2). 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 2.3.2).

2.3.2 Analysis of the fish haplotypes obtained by the MiFish pipeline

The reliability of MiFish pipeline (http://mitofish.aori.utokyo.ac.jp/mifish/workflows/new) for the biodiversity assessment of fish species inhabiting the sampled rivers was analyzed. From the 2,315,605 raw reads, 2,280,850 merged reads were obtained by the MiFish pipeline showing 98.50% yields (Table 2.3.3). 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 MEGA program v7 (Kumar et al., 2016) with Maximum likelihood algorithm (Fig. 2.3.2-2.3.5).

Table. 2.3.2 Environmental DNA sample collection from the four rivers

River	Date	Station	GPS location	Temp. (˚ C)	Salinity
Hyeongsan	2018.06.11	HS1	N 35° 42' 36", E 129° 11' 42"	18.60	1.00
		HS2	N 35° 56' 14", E 129° 14' 24"	19.50	2.00
		HS3	N 35° 59' 32", E 129° 17' 19"	20.00	3.20
		HS4	N 36o 01' 51", E 129° 23' 01"	24.00	20.20
Taehwa	2018.06.11	TH1	N 35° 32' 52", E 129° 06' 27"	19.40	1.02
		TH2	N 35° 35' 07", E 129° 13' 52"	19.80	2.04
		TH3	N 35° 32' 42", E 129° 17' 38"	22.70	14.02
		TH4	N 35° 32' 39", E 129° 21' 24"	19.20	17.80
Seomjin	2018.06.12	SJ1	N 35° 11' 18", E 127° 37' 21"	24.20	0.15
		SJ2	N 35° 04' 30", E 127° 43' 35"	23.40	2.01
		SJ3	N 35° 01' 54", E 127° 46' 32"	23.00	12.90
		SJ4	N 34° 58' 01", E 127° 45' 28"	23.00	16.80
Nakdong	2018.06.12	ND1	N 35° 23' 19", E 128° 29' 09"	24.00	1.92
		ND2	N 35° 20' 40", E 128° 46' 26"	24.10	2.40
		ND3	N 35° 17' 57", E 128° 58' 37"	23.20	2.78
		ND4	N 35° 07' 13", E 128° 57' 07"	22.50	4.50

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 2.3.3).

Table 2.3.3 Summary of taxonomic assignment of the MiSeq reads from the four 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)

* Final number, after removal of duplicated one in brackets

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 2.3.4).

Table 2.3.4 Fish species identified from the four rivers by the eDNA metabarcoding approach

No.	Species	Family	Identity (%)	GenBank number
1	*Acanthogobius hasta*	Gobiidae	100	KM030428
2	*Acanthogobius lactipes*	Gobiidae	100	KM030431
3	*Acheilognathus intermedia*	Cyprinidae	99	EF483933
4	*Acheilognathus macropterus*	Cyprinidae	99	EF483935
5	*Acheilognathus majusculus*	Cyprinidae	99	LC006056
6	*Acheilognathus rhombeus*	Cyprinidae	99	KT601094
7	*Anguilla japonica*	Anguillidae	100	HQ185628
8	*Carassius auratus*	Cyprinidae	100	KX505165
9	*Carassius cuvieri*	Cyprinidae	100	AP011237
10	*Channa argus*	Channidae	100	MG751766
11	*Cobitis tetralineata*	Cobitidae	100	EU670794
12	*Coreoleuciscus splendidus*	Cyprinidae	100	JN831358
13	*Coreoperca herzi*	Sinipercidae	100	KR075132
14	*Cyprinus carpio*	Cyprinidae	100	KX710076
15	*Cyprinus megalophthalmus*	Cyprinidae	100	KR869143
16	*Favonigobius gymnauchen*	Gobiidae	100	LC385206
17	*Gymnogobius breunigii*	Gobiidae	99	KM030451
18	*Hemibarbus labeo*	Cyprinidae	100	DQ347953
19	*Hemibarbus maculatus*	Cyprinidae	99	NC018534
20	*Hemiculter leucisculus*	Cyprinidae	100	LC340359
21	*Iksookimia longicorpa*	Cobitidae	100	KM676413
22	*Iksookimia yongdokensis*	Cobitidae	100	EU670800
23	*Kareius bicoloratus*	Pleuronectidae	100	AP002951
24	*Konosirus punctatus*	Clupeidae	100	KC477844
25	*Lepomis macrochirus*	Centrarchidae	100	JN389795
26	*Microphysogobio koreensis*	Cyprinidae	100	FJ515920
27	*Microphysogobio yaluensis*	Cyprinidae	99	KR075133
28	*Micropterus salmoides*	Centrarchidae	100	HQ391896
29	*Misgurnus anguillicaudatus*	Cobitidae	100	KC762740
30	*Misgurnus bipartitus*	Cobitidae	100	KF562047
31	*Misgurnus mizolepis*	Cobitidae	100	AP017654
32	*Mugil cephalus*	Mugilidae	100	KF374974
33	*Mugilogobius abei*	Gobiidae	100	KM030465
34	*Nipponocypris koreanus*	Cyprinidae	100	KJ427719
35	*Nipponocypris temminckii*	Cyprinidae	100	AP012116
36	*Niwaella multifasciata*	Cobitidae	100	EU670807

Table 2.3.4 Continued.

No.	Species	Family	Identity (%)	GenBank number
37	*Odontobutis interrupta*	Odontobutidae	100	KR364945
38	*Odontobutis platycephala*	Odontobutidae	100	KM030426
39	*Opsariichthys uncirostris*	Cyprinidae	99	AB218897
40	*Paramisgurnus dabryanus*	Cobitidae	100	KM186182
41	*Phoxinus oxycephalus*	Cyprinidae	99	MK208924
42	*Phoxinus semotilus*	Cyprinidae	100	KT748874
43	*Planiliza affinis*	Mugilidae	100	KM925142
44	*Planiliza haematocheila*	Mugilidae	100	KJ622047
45	*Pseudobagrus koreanus*	Bagridae	100	KT601095
46	*Pseudobagrus ussuriensis*	Bagridae	100	KC188782
47	*Pseudogobio esocinus*	Cyprinidae	100	LC340042
48	*Pseudogobio vaillanti*	Cyprinidae	100	KU314695
49	*Pseudogobius masago*	Gobiidae	100	KM030467
50	*Pungtungia herzi*	Cyprinidae	99	KF006339
51	*Rhinogobius brunneus*	Gobiidae	100	KT601096
52	*Rhinogobius giurinus*	Gobiidae	100	KM030475
53	*Rhodeus suigensis*	Cyprinidae	100	EF483934
54	*Rhodeus uyekii*	Cyprinidae	100	EF483937
55	*Rhynchocypris lagowskii*	Cyprinidae	99	KJ641843
56	*Rhynchocypris oxycephalus*	Cyprinidae	99	LC193377
57	*Sarcocheilichthys soldatovi*	Cyprinidae	100	LC146036
58	*Sarcocheilichthys variegatus*	Cyprinidae	100	KU301744
59	*Silurus asotus*	Siluridae	100	JX087351
60	*Silurus microdorsalis*	Siluridae	99	KT350610
61	*Siniperca scherzeri*	Sinipercidae	100	MF966985
62	*Squalidus chankaensis*	Cyprinidae	100	KT948082
63	*Squalidus japonicus coreanus*	Cyprinidae	100	KR075134
64	*Squalidus multimaculatus*	Cyprinidae	100	KX495606
65	*Tachysurus fulvidraco*	Bagridae	100	KU133295
66	*Tachysurus nitidus*	Bagridae	100	KC822643
67	*Tanakia signifer*	Cyprinidae	99	EF483930
68	*Tanakia somjinensis*	Cyprinidae	99	FJ515921
69	*Tribolodon hakonensis*	Cyprinidae	100	AB626855
70	*Tridentiger obscurus*	Gobiidae	100	KT601092
71	*Tridentiger radiatus*	Gobiidae	99	EU047755
72	*Tridentiger trigonocephalus*	Gobiidae	100	KM030481
73	*Zacco platypus*	Cyprinidae	100	LC277796

2.3.3 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 (Figure 2.3.2). 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* (Figure 2.3.2). Since *Hemibarbus mylodon* is an endangered freshwater species, which has been exclusively identified in Han and Geum rivers (KIM et al., 2007), so further study should be made for confirmation.

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).

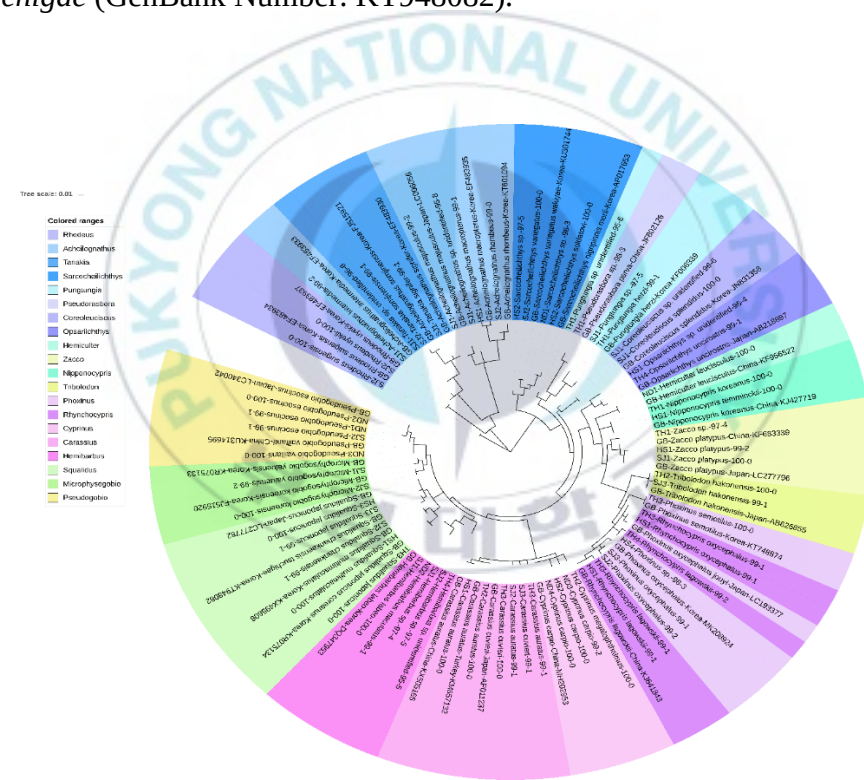


Figure 2.3.2. Phylogenetic tree of Cyprinidae family by maximum likelihood method

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). 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. 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 99% sequence identity to the Korean haplotypes of *Acheilognathus intermedia* (EF483933), *Tanakia somjinensis* (FJ515921), and *Tanakia signifier* (EF483930). One haplotype from Taehwa river (TH3) showed 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 Seomjin river (SJ2) and

Hyeongsan river (HS2) showed 100% and 97% sequence identity to the Korean haplotype of *Sarcocheilichthys variegatus wakiyae* (GenBank Number: KU301744). One haplotype from 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.

Table 2.3.5 Fish haplotypes with the GenBank numbers identified from the eDNA metabarcoding study of the four rivers

No.	Family	ID No.	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	

Table 2.3.5 Continued.

No.	Family	ID No.	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
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	-	-	

Table 2.3.5 Continued.

No.	Family	ID No.	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
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	

Table 2.3.5 Continued.

No.	Family	ID No.	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
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	-	

Table 2.3.5 Continued.

No.	Family	ID No.	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
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	

Table 2.3.5 Continued.

No.	Family	ID No.	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
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*	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	

Table 2.3.5 Continued.

No.	Family	ID No.	Haplotypes	Identity (%)	Korean haplotype	Chinese haplotype	Japanese haplotype	Others
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		

2.3.4 Gobiidae

We identified 16 haplotypes of the family Gobiidae, which represent 7 genera and 11 species (Fig. 2.3.3). 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 Taehwa river (TH4) showed 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. 2.3.3).

All three haplotypes in the genus *Rhinogobius* showed 100 % identity to the GenBank 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 to the Korean haplotype (KM030475) of *Rhinogobius giurinus.* Two haplotypes of *Gymnogobius* sp. from the Taehwa river and Hyeongsan river showed 98 % sequence identity to *Gymnogobius taranetzi* (GenBank Number:

LC385155) (Fig. 2.3.3). 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.

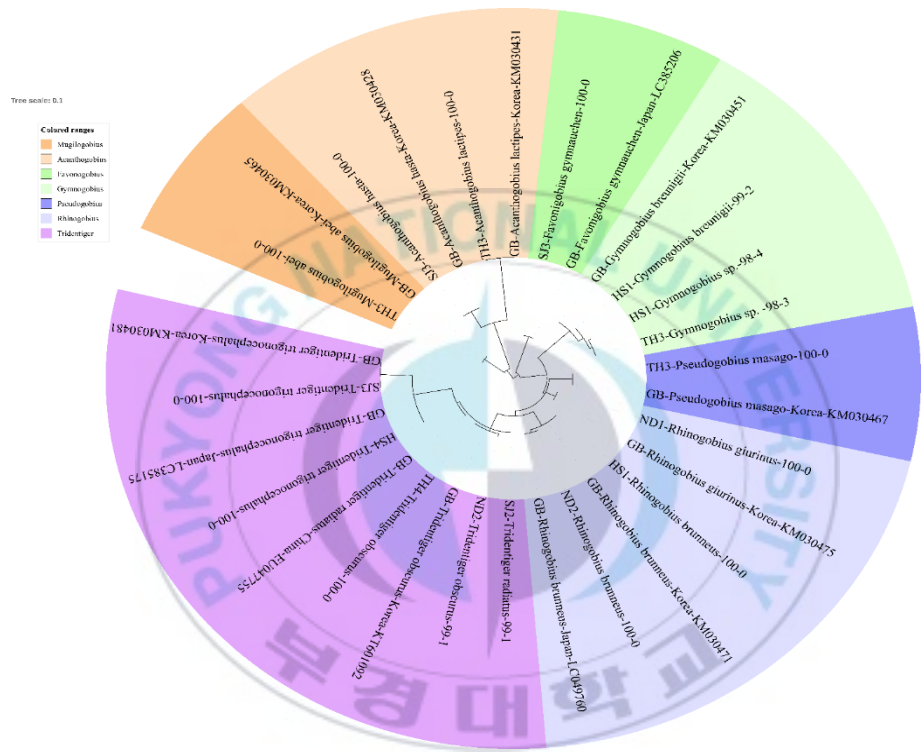


Figure 2.3.3. Phylogenetic tree of Gobiidae family by maximum likelihood method

2.3.5 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. 2.3.4). 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 *Misgurnus anguillicaudatus* were identified by the phylogenetic analysis (Fig. 2.3.4). 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.

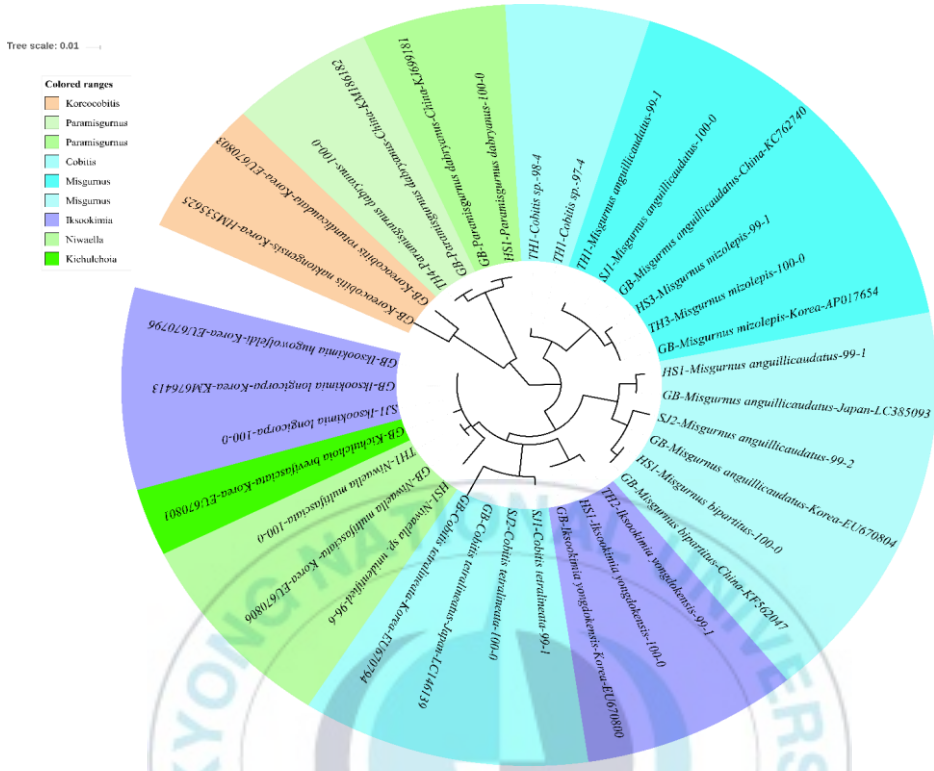


Figure 2.3.4. Phylogenetic tree of Cobitidae family by maximum likelihood method

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. 2.3.4). This species is regarded as endemic to China and *P. dabryanus* is often imported to Korea together 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 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.

2.3.6 Other families

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. 2.3.5). 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). Further studies should be made to identify this haplotype.

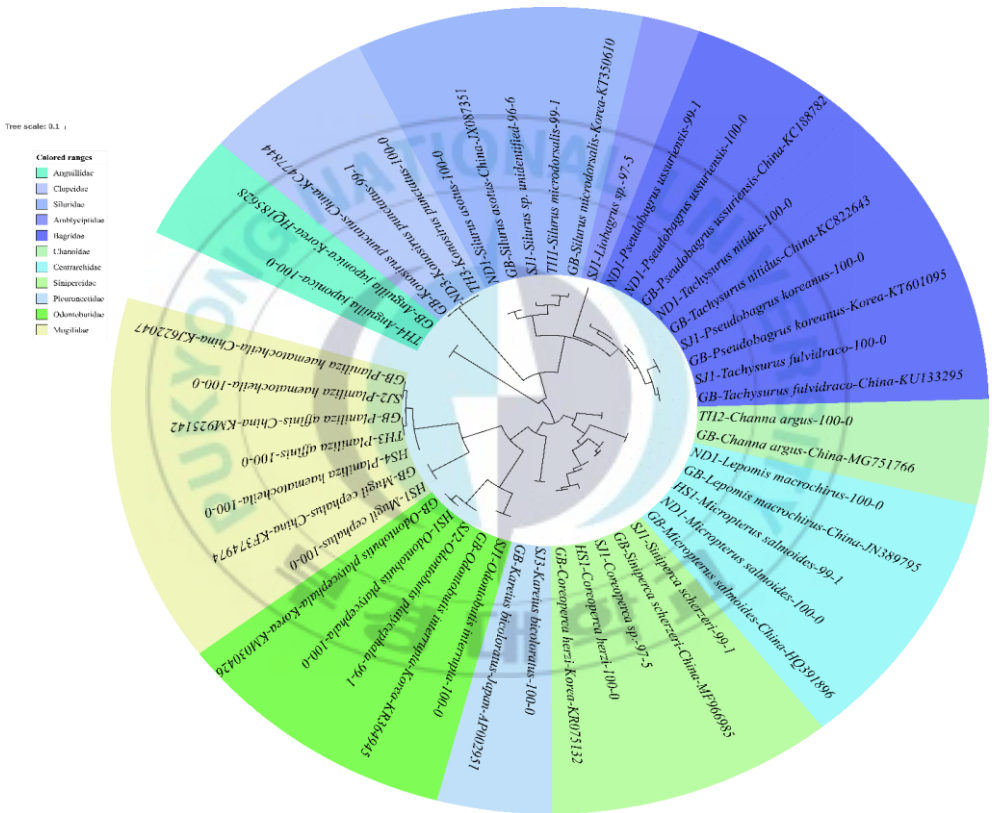


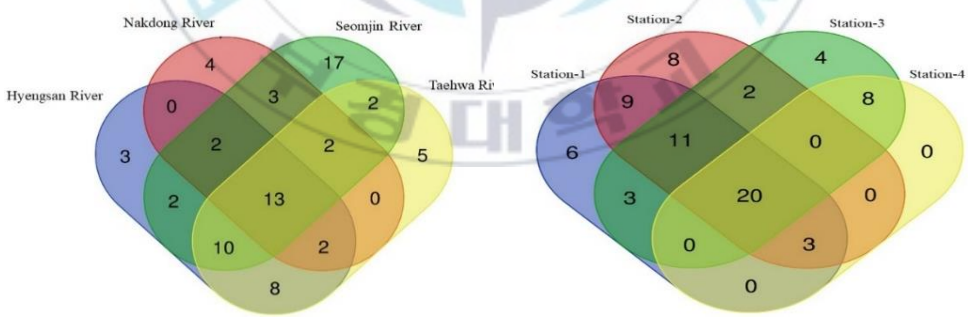
Figure 2.3.5. Phylogenetic tree of other families by maximum likelihood method

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. Five endemic species in the family Amblycipitidae are currently reported in Korea: *Liobagrus andersoni*, *Liobagrus obesus*, *Liobagrus mediadiposalis*, *Liobagrus somjinensis*, and *Liobagrus hyeongsanensis* (Kim and Park, 2002). Their MiFish region/complete mitochondrial DNA sequences should be supplemented to the GenBank database. This result indicates that haplotypes in the family Amblycipitidae 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.

2.3.7 Fish biodiversity in the rivers sampled

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. 2.3.6 A and B).



6. A 6. B

Figure 2.3.6 A. Venn diagram of fish species identified from the four rivers and 2.3.6 B Venn diagram of fish species identified at the different stations of the four rivers.

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. 2.3.7). 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 %) (Fig. 2.3.7).

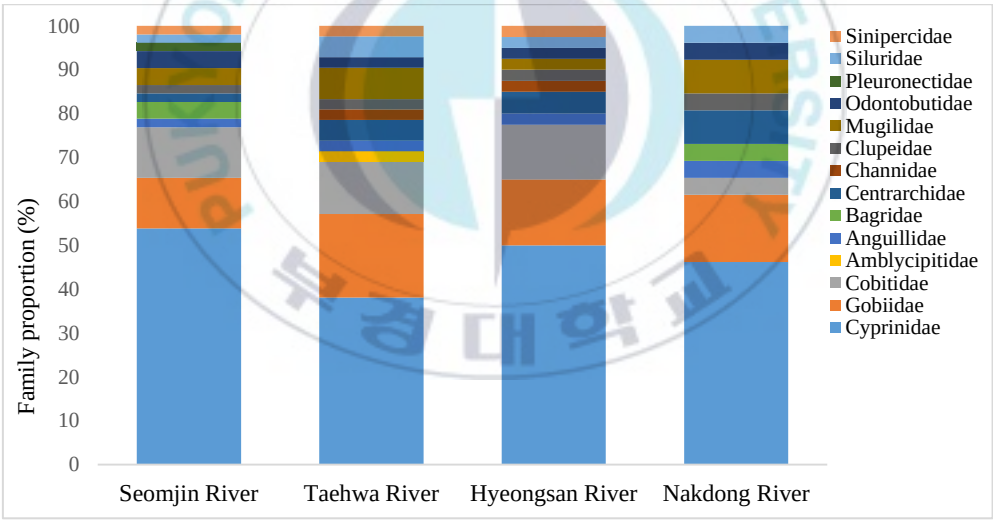


Figure 2.3.7. Fish community structure at family level in four Korean rivers

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. 2.3.6).

The highest Shannon Index (SI) was identified in the Seomjin river (3.48), 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 2.3.6).

Table 2.3.6 Shannon Index (SI) in the four Korean rivers

	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	

2.3.8 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 (Table 2.3.1). 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 (Fig. 2.3.9).

2.3.9 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® software v7 (Clarke and Gorley, 2015). The result clearly demonstrated species distribution according to different sampling stations from upstream to downstream (Fig. 2.3.9).

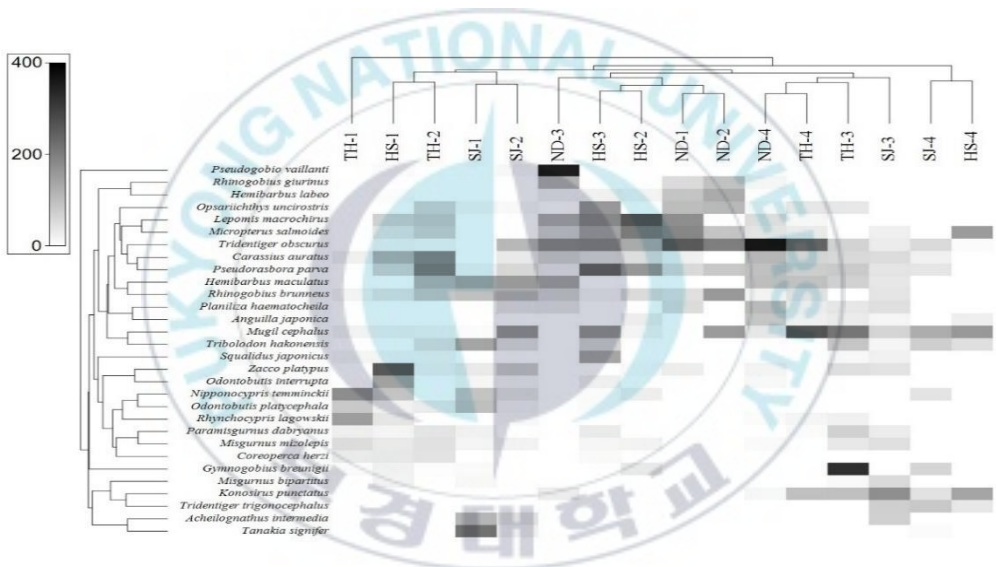


Figure 2.3.9. Heat map of top 30 fish species identified in 16 sampling stations of 4 rivers

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.

2.4 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. 2.3.2-5). 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 suggested 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. 2.3.2 and Table 2.3.7). 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).

Table 2.3.7. Genetic distance of species under the family Cyprinidae

No.	Species														
1	*Carassius auratus* (China KX505165)														
2	*Carassius auratus* (Turkey-KM657132)	0.006													
3	*Carassius auratus* (SJ2 99)	0.018	0.012												
4	*Carassius auratus* (TH2-100)	0.006	0.00	0.012											
5	*Carassius auratus* (TH3-99)	0.006	0.012	0.012	0.012										
6	*Carassius auratus* (HS1-100)	0.00	0.006	0.018	0.006	0.006									
7	*Carassius cuvieri* (Japan-AP011237)	0.018	0.012	0.012	0.012	0.012	0.018								
8	*Carassius cuvieri* (SJ3-99)	0.024	0.018	0.018	0.018	0.018	0.024	0.006							
9	*Carassius cuvieri* (TH3-100)	0.018	0.012	0.012	0.012	0.012	0.018	0.00	0.006						
10	*Carassius gibelio* (China-KX505166)	0.00	0.006	0.018	0.006	0.006	0.00	0.018	0.024	0.018					
11	*Cyprinus carpio* (China-MH202953)	0.018	0.012	0.012	0.012	0.012	0.018	0.012	0.018	0.012	0.018				
12	*Cyprinus carpio* (HS2-100)	0.03	0.024	0.024	0.024	0.024	0.03	0.024	0.03	0.024	0.03	0.012			
13	*Cyprinus carpio* (ND4-100)	0.018	0.012	0.012	0.012	0.012	0.018	0.012	0.018	0.012	0.018	0.00	0.012		
14	*Cyprinus carpio* (ND3-99)	0.03	0.024	0.024	0.024	0.024	0.03	0.024	0.03	0.024	0.03	0.012	0.024	0.012	
15	*Cyprinus megalophthalmus* (TH2-100)	0.03	0.024	0.024	0.024	0.024	0.03	0.024	0.03	0.024	0.03	0.012	0.024	0.012	0.024
16	*Cyprinus megalophthalmus* (China-KR869143)	0.03	0.024	0.024	0.024	0.024	0.03	0.024	0.03	0.024	0.03	0.012	0.024	0.012	0.024	0.0

Table 2.3.7. Continued.

No.	Species														
1	*Acheilognathus intermedia* (Korea-EF483933)														
2	*Acheilognathus intermedia* (SJ1-99)	0.012													
3	*Acheilognathus macropterus* (Korea-EF483935)	0.232	0.223												
4	*Acheilognathus macropterus* (HS1-99)	0.223	0.214	0.018											
5	*Acheilognathus majusculus* (Japan-LC006056)	0.198	0.198	0.127	0.119										
6	*Acheilognathus majusculus* (SJ1-99)	0.198	0.198	0.119	0.112	0.012									
7	*Acheilognathus rhombeus* (Korea-KT601094)	0.251	0.232	0.077	0.07	0.105	0.097								
8	*Acheilognathus rhombeus* (SJ2-99)	0.233	0.215	0.07	0.063	0.084	0.077	0.018							
9	*Acheilognathus chankaensis* (Japan-AB016671)	0.233	0.215	0.105	0.083	0.111	0.104	0.076	0.056						
10	*Acheilognathus koreensis* (Korea-NC013704)	0.09	0.09	0.248	0.239	0.24	0.231	0.287	0.268	0.278					
11	*Acheilognathus yamatsutae* (Korea-NC013712)	0.205	0.205	0.111	0.104	0.07	0.056	0.083	0.063	0.111	0.229				
12	*Acheilognathus signifer* (Korea-EF483930)	0.063	0.063	0.231	0.222	0.205	0.197	0.249	0.232	0.241	0.024	0.204			
13	*Tanakia signifer* (SJ1-99)	0.07	0.07	0.24	0.231	0.214	0.205	0.259	0.241	0.251	0.018	0.213	0.006		
14	*Acheilognathus somjinensis* (Korea-FJ515921)	0.084	0.07	0.223	0.214	0.206	0.198	0.232	0.215	0.215	0.056	0.205	0.031	0.037	
15	*Tanakia somjinensis* (SJ2-99)	0.091	0.077	0.232	0.223	0.215	0.206	0.241	0.224	0.224	0.063	0.214	0.037	0.044	0.006

Table 2.3.7. Continued.

No.	Species			
1	*Nipponocypris koreanus* (China-KJ427719)			
2	*Nipponocypris koreanus* (TH1-100)	0.00		
3	*Nipponocypris temminckii* (Japan-LC468890)	0.011	0.011	
4	*Nipponocypris temminckii* (HS1-100)	0.011	0.011	0.00

Table 2.3.7. Continued.

No.	Species			
1	*Rhodeus uyekii* (Korea-EF483937)			
2	*Rhodeus suigensis (*Korea-EF483934)	0.164		
3	*Rhodeus suigensis* (SJ2-100)	0.164	0.00	
4	Rhodeus uyekii (SJ1-100)	0.00	0.164	0.164

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 2.3.8). 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).

Table 2.3.8 List of the exotic species identified from the four Korean rivers sampled

No.	River	Species Name	Family	Order	Identity (%)
1	Hyeongsan	*Carassius cuvieri*	Cyprinidae	Cypriniformes	100.0
2		*Cyprinus carpio*	Cyprinidae	Cypriniformes	100.0
3		*Cyprinus megalophthalmus*	Cyprinidae	Cypriniformes	100.0
4		*Micropterus salmoides*	Centrarchidae	Centrarchiformes	100.0
5		*Lepomis macrochirus*	Centrarchidae	Centrarchiformes	100.0
6	Nakdong	*Cyprinus carpio*	Cyprinidae	Cypriniformes	100.0
7		*Cyprinus megalophthalmus*	Cyprinidae	Cypriniformes	100.0
8		*Micropterus salmoides*	Centrarchidae	Centrarchiformes	100.0
9		*Lepomis macrochirus*	Centrarchidae	Centrarchiformes	100.0
10	Seomjin	*Carassius cuvieri*	Cyprinidae	Cypriniformes	99.42
11		*Micropterus salmoides*	Centrarchidae	Centrarchiformes	99.41
12	Taehwa	*Carassius cuvieri*	Cyprinidae	Cypriniformes	100.0
13		*Cyprinus carpio*	Cyprinidae	Cypriniformes	100.0
14		*Cyprinus megalophthalmus*	Cyprinidae	Cypriniformes	100.0
15		*Micropterus salmoides*	Centrarchidae	Centrarchiformes	100.0
16		*Lepomis macrochirus*	Centrarchidae	Centrarchiformes	100.0

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.

2.5 Conclusion

The eDNA metabarcoding approach, combined with NGS to identify multiple species is a potential technique to monitor species diversity in aquatic habitat and offer a more precise estimation of biodiversity rather than single or a handful of species surveillance. We firstly analyzed the fish biodiversity in four rivers (Nakdong, Hyeongsan, Seomjin, and Taehwa) using 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.

Chapter 3

Fish biodiversity patterns estimated from the environmental DNA metabarcoding surveys in the Suyeong River, Korea

3.1 Introduction

The composition of fish communities provides the basic information needed for biological conservation and management of a freshwater ecosystem (Elliott et al., 2007). Therefore, regular surveys have been conducted to estimate the fish biodiversity of the many rivers in Korea. Traditional fish surveys have been mainly dependent on the direct observation of catches using different types of nets, traps, angling, electrofishing, and counting methods (Bonar et al., 2009; Murphy and Willis, 1996). Those conventional capture methods often detect limited numbers of fish species thus requiring the repetitive costly surveys with different seasons for the reliable evaluation of the fish communities in a given area (Gabriel et al., 2008). Those methods also often resulted in the destructive effects on the ecosystem as well (Li et al., 2018). As a consequence, both researchers and policymakers would like to adopt a reliable and non-destructive survey system with low cost to manage and conserve the aquatic ecosystem and its resources (Sinclair et al., 2002).

The environmental DNA (eDNA) metabarcoding is becoming a popular tool for the monitoring the fish diversity due to its higher diversity of taxa with lower efforts and times compared with the conventional morphology-based surveys of captured specimens (Baird and Sweeney, 2011; DeWalt, 2011; Evans and Lamberti, 2017; Valentini et al., 2009). It is also Thus, eDNA

metabarcoding analysis is a promising avenue for monitoring the ocean's biodiversity It is a brand new tool in Ecology and Conservation Biology, this method is going to be an amazing and effective approach in complying to detect the abundance/distribution of species, which was previously imperceptible to us (Pilliod et al., 2013b).

Among the most widely used eDNA metabarcoding platforms, to perceive the freshwater fish biodiversity through environmental DNA analysis, a set of fish-universal primers, MiFish (Miya et al., 2015a) were used, which is compatible with fish eDNA metabarcoding (Yamamoto et al., 2017). The MiFish primers amplify the 12S rRNA gene (163-185 bp) of mitochondrial DNA, which is a hypervariable region and contains adequate information for identifying fishes in most cases up to the species level. Having a short amplicon length, less cross-reactivity, these primers can amplify from decayed/degraded DNA sample, and short amplicons are more competent for MiSeq sequencing (Yamamoto et al., 2017). The MiFish primer set demonstrated its reliability for fish biodiversity analysis both in seawater (Ushio et al., 2017; Yamamoto et al., 2017) and freshwater (Sato et al., 2018).

As the second largest river in Busan, the Suyeong River begins from the Hoidong reservoir to the Suyeong bay its total length reaches about 28.6 km. It is an urban stream, where the one-third population of the Busan city lives around this area and transports various types of industrial and residential

wastewater in drainage of the mainstream (Dong–Myung et al., 2013). Water is supplying in the Hoidong reservoir from the Nakdong River since 2008 and this is the main source of water for the Suyeong River. Huge amounts of domestic wastes and industrial sewages are mixing to the Suyeong River, resulting in the water quality deteriorating severely since it runs through the urbanized city center. For this reason, eutrophic algal bloom and pollution in bottom sediments are very common in the Suyeong River (Kim et al., 2014; Park, 1997; WON et al., 1979).

The purpose of this study was to understand the changes in fish biodiversity in different seasons by using environmental DNA analysis in the Suyeong River from August 2017 to June 2018 and to make a correlation between the Physico-chemical parameters of water with fish abundance to understand a healthy ecological condition in this river. This information would be helpful to understand the relationship between the water quality parameter with fish availability in a freshwater ecosystem.

3.2 Materials and Methods

3.2.1 Sample collection and environmental DNA extraction

Water samples (eDNA) were collected with a one-month interval from August 2017 to June 2018 (six times) from the four stations (A, B, C, and D) of the Suyeong River (Figure 3.2.1 and Table 3.2.1).

Table 3.2.1 Environmental DNA sample collection sites with water temperature (^{0}C) and salinity (psu) measurement from the Suyeong River

Month	Station A 35.264935 N 129.114018 E		Station B 35.217105 N 129.118479 E		Station C 35.189336 N 129.114921 E		Station D 35.170747 N 129.124724 E	
	Temp.	Salinity	Temp.	Salinity	Temp.	Salinity	Temp.	Salinity
August 2017	27.50	0.10	27.30	0.10	25.70	10.50	25.80	19.70
October 2017	11.40	0.20	9.80	0.20	13.30	9.80	16.40	19.10
December 2017	6.50	0.20	7.30	0.35	8.50	4.70	13.10	20.00
February 2018	4.00	0.20	7.70	0.40	7.70	5.80	11.00	19.60
April 2018	19.74	0.12	14.64	0.17	19.51	8.49	17.55	18.39
June 2018	21.94	0.14	16.85	0.16	21.75	15.60	22.89	21.79
Average	15.18 ± 9.30	0.16 ± 0.05	13.93 ± 7.58	0.23 ± 0.12	16.08 ± 7.37	9.15 ± 3.88	17.70 ± 5.66	19.76 ± 1.14

Water samples (eDNA) were collected with a one-month interval from August 2017 to June 2018 (six times) from the four stations (A, B, C, and D) of the Suyeong River (Fig. 1 and Table 1). Four sample stations in this river covered from the upstream (Hoidong Reservoir) to the downstream (Close to BEXCO, Busan) of the river. One liter of water sample was collected from each station in disposable plastic bottles by using a plastic bucket with a nylon rope. To prevent the contamination, all equipment (e.g. plastic bottles, buckets, rope, etc.) dipped for at least 10 minutes with 10% commercial bleach. After collecting water, the bottles were immediately stored in ice until brought to the laboratory for filtration. Another one-liter water (underwater 0.5 m.) was collected from each station to measuring the water temperature and salinity with a conductivity meter (CD-4307SD, LUTRON). On the same

day of water collection, one liter of water was filtered with 0.45 μm pore-sized GN-6 membrane (PALL Life sciences, Mexico) and stored the membrane filters at -20˚C until DNA extraction from the filters.

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 membranes 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), and stored at -20˚C.

3.2.2 Construction of library and MiSeq sequencing

The MiFish universal primer set was used to get the sequence of partial 12S rRNA gene (Miya et al., 2015a). The first PCR was performed to amplify the MiFish regions with an overhanging linker sequence for each Nextera XT index (Illumina, USA). The PCR mixture (20 μL) contained 1.0 μL of MiFish (Forward & Reverse) primers (5 pmol 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 EX Taq Hot Start (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. Gel electrophoresis was run with

1.5% agarose gel, then the amplicon with the expected size (250 bp~350 bp) was purified by AccuPrep® Gel Purification Kit (Bioneer, Republic of Korea). 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 24 field negative controls; consequently, the 24 negative controls were discarded from the next analysis. After gel purification, the quality and quantity of the indexed PCR products with the expected size were analyzed by qubit dsDNAHS Assay Kit (Invitrogen, Carlsbad, CA, USA) followed by the sequencing using MiSeq platform (2 X 300 bp).

Figure 3.3.1. Environmental DNA sample collection sites of the Suyeong River

3.2.3 Bioinformatics analysis of NGS data

After getting the NGS raw data, an open-source software (Python 2.7) was used for pairing both the reverse and forward sequences with the specific script (Zhang, 2015), after pairing the NGS raw reads, we uploaded the paired data to the publically open web-based MiFish pipeline (http://mitofish.aori.u-

tokyo.ac.jp/mifish/). In the MiFish pipeline, the raw reads by MiSeq sequencing run FASTQC, the low-quality tail of reads (QV ≤20) were trimmed, and we used the 95% sequence identity option. After a taxonomic assignment 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 sequences were ascertained as species. If the sequence identity from 97% to 98%, the sequence was ascertained as a genus, and the sequences having <97% similarity (putative genera) with the database were assigned as 'unidentified' genera. The habitat distribution for each species was confirmed by the FishBase website (https://www.fishbase.org/).

3.2.4 Statistical analysis for fish biodiversity indices

The alpha biodiversity was measured by the average read number from each sampling station of the Suyeong River. The alpha diversity index, Shannon-Wiener (H') indicates the heterogeneity of species or the richness of total species in an ecosystem (Gray, 2000; Magurran, 1988). The H' index, the Margalef index (d), and the heat map clustering analysis were enumerated by using the PRIMER® software v7(Clarke and Gorley, 2015).

3.3 Results

3.3.1 Temperature and salinity changes in the Suyeong River

The annual water temperatures of four sample stations ranged from 4.0 ˚C to 27.5 ˚C (Table 1). The lowest average water temperature was identified at station B (13.93 ± 7.58 ˚C), while the highest one was identified at station D (17.79 ± 5.66 ˚C). During the survey period, the temperature change was greatest at station A (from 4 ˚C to 27.50 ˚C), while and lowest (11.0 ˚C to 25.80 ˚C) at station D (Table 3.2.1 and Figure 3.3.2). By contrast, average salinity increased from the upstream to downstream. Salinity changes highly in station C (4.70 to 15.60) by the influx of tides while the stable high amount of salinity was identified at station D (18.39 to 21.79) showing the estuarial characteristics. In this river, at the station C and station D has higher salinity level, because of the influx of seawater from the ocean during the high tide, which is the correspondent to the high degree of salinity in the summer season (June and August) of the sampling period (Table 3.2.1 and Figure 3.3.3).

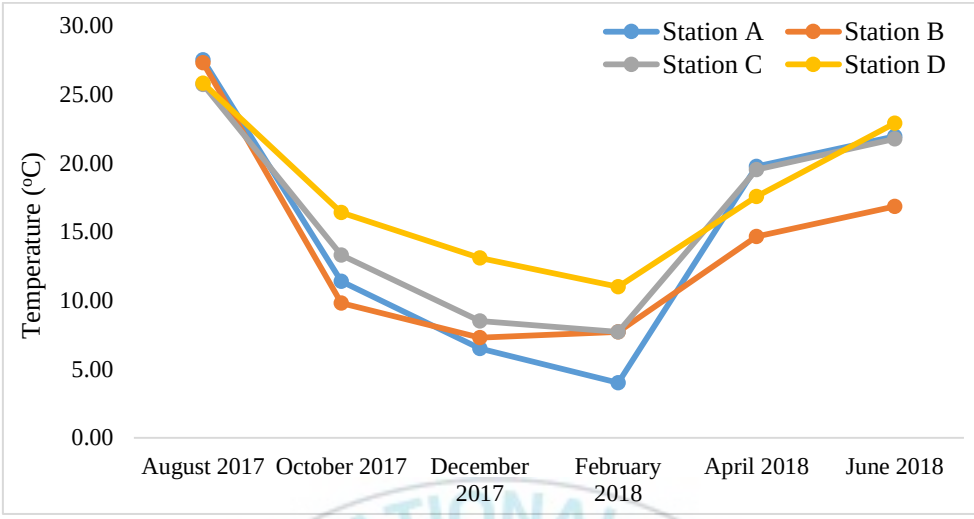


Figure 3.3.2. Water temperature changes in the Suyeong River

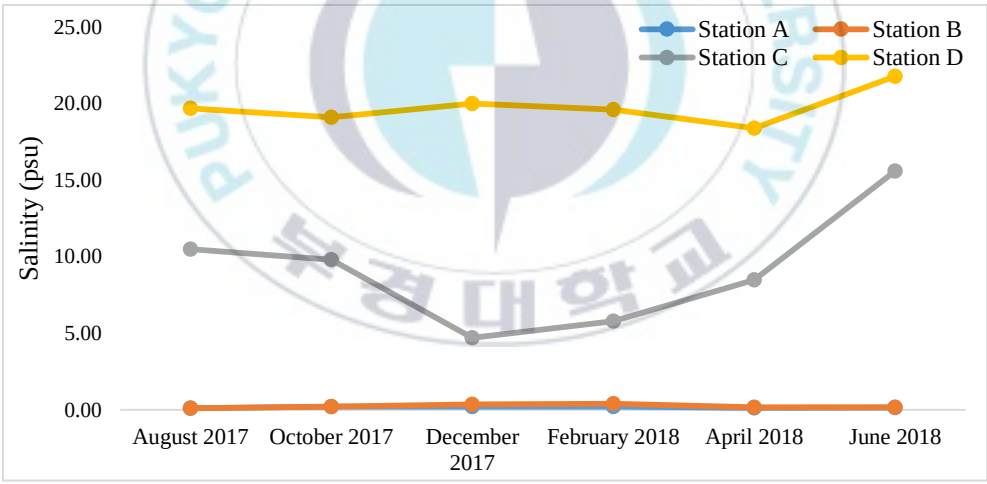


Figure 3.3.3. Changes in salinity (psu) during the study period in the Suyeong River

3.3.2 Analysis of the taxonomic assignment and haplotypes

After clustering and trimming the raw reads (2,696,913) from the MiFish pipeline, a total of 2,657,364 merged reads (98.53%) were clustered and assigned to 391 haplotypes. A total of 2,587,724 merged reads (97.38%) showed more than 99% sequence identity to the database and assigned to 332 haplotypes belonging to 38 confirmed species, and 61,975 merged reads (2.3%) showed 99% to 97% sequence identity and assigned to 36 haplotypes belonging to 8 genera, and the remaining 7,665 merged reads showed 97% to 95% sequence identity assigned to 23 haplotypes belonging to 8 unidentified genera (Table 3.3.2). A total of 39,549 raw reads (1.47%) with lower than 95% sequence identity to the database were discarded from further analyses.

After eliminating the duplicated ones from 391 haplotypes obtained from the four sample stations, 65 haplotypes were finally identified as the representative haplotypes in the Suyeong River. Among them, 49 haplotypes showed a high sequence identity to the database with 99% or higher to the database (https://www.ncbi.nlm.nih.gov/genbank/). Those included 38 confirmed species, 31 genera, 13 families and 8 orders (Table 3.3.3). Among the 65 haplotypes, 16 (24.61%) showed a lower identity (<99% to 97%) to the database (Table 3.3.2). This result indicated that current sequence data for Korean freshwater fish species are good enough to adopt the metabarcoding technique for the fish biodiversity analysis with supplement additional

sequence data. These 16 sequences with low identity to the database were further analyzed to estimate the more precise species numbers. First, three haplotypes showed the high identity to *Hemiculter leucisculus* (LC340359) (https://species.nibr.go.kr/index.do), two species in the genus were currently reported in Korea, *Hemiculter leucisculus* and *Hemiculter eigenmanni*. Since *H. leucisculus* was identified, those may be *H. eigenmanni* but further study should be made (Table 3.3.3).

Five haplotypes of *Squalidus* showed the highest identity to *Squalidus japonicus* (KR075134). Seven species in the genus *Squalidus* are currently reported and sequences in the genus should be supplemented. Finally, three haplotypes in the genus *Zacco* also should be further analyzed, which would be one of three species in the genus (Table 3.3.3). As a result of manual analysis, we further assigned 16 species from 16 haplotypes with low sequence identity (<99 %) to the database. Collectively, a total of 54 species were assigned by Mifish pipeline and manual analysis, which included 31 genera, 13 families and 8 orders from the Suyeong River (Fig. 3.3.4 and Table 3.3.3).

Table 3.3.2 Numbers of haplotypes and reads obtained from the Suyeong River by MiSeq sequencing platform

Station	Description	August 2017	October 2017	December 2017	February 2018	April 2018	June 2018
Station A	Raw reads	101790	245542	7039	40519	197844	234116
	Merged reads	100286	241913	6935	39920	194920	230656
	Representative haplotypes	18	17	15	15	19	26
	Reads (>99 % identity) with haplotypes	97598, 16	241430, 16	6749, 13	38850, 14	189696, 17	224474, 24
	Reads (<99 % identity) with haplotypes	2688, 2	483, 1	186, 2	1070, 1	5224, 2	6182, 2
Station B	Raw reads	11757	152997	1751	40220	229466	213197
	Merged reads	11583	150736	1725	39626	226075	210046
	Representative haplotypes	10	7	6	6	11	19
	Reads (>99 % identity) with haplotypes	11583, 10	146696, 6	1725, 6	38564, 5	226075, 11	204417, 17
	Reads (<99 % identity) with haplotypes	0	4040, 1	0	1062, 1	0	5629, 2
Station C	Raw reads	104798	5811	44983	54654	228615	226777
	Merged reads	103249	5725	44318	53846	225236	223426
	Representative haplotypes	22	7	7	6	19	29
	Reads (>99 % identity) with haplotypes	100482, 17	5725, 7	44318, 7	53846, 6	219200, 18	217438, 27
	Reads (<99 % identity) with haplotypes	2767, 5	0	0	0	6036, 1	5988, 2
Station D	Raw reads	165316	251591	11527	50767	159959	236609
	Merged reads	162873	247873	11357	50017	157595	233112
	Representative haplotypes	16	14	12	11	14	25
	Reads (>99 % identity) with haplotypes	158508, 16	247873, 14	11053, 11	48677, 10	157595, 14	226865, 21
	Reads (<99 % identity) with haplotypes	0	0	304, 1	1340, 1	0	6247, 4

Among 65 haplotypes, the highest 31 haplotypes were identified from the family Cyprinidae, followed by Gobiidae (8), and the remaining 15 were from the other 11 families (Table 3.3.3).

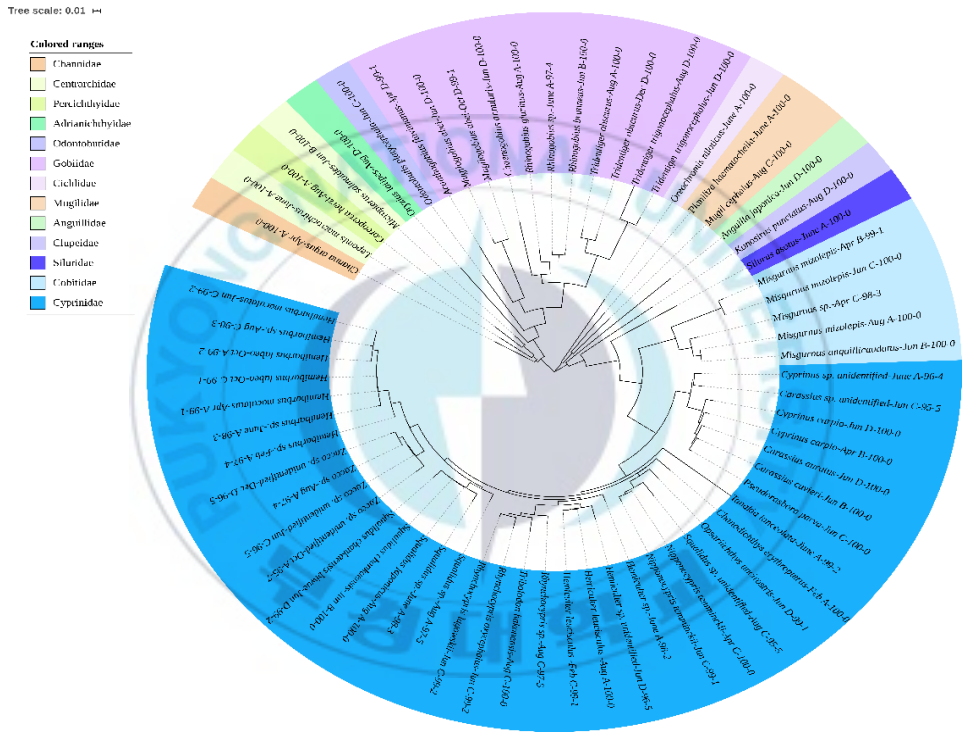


Figure 3.3.4. The phylogenetic tree of the identified fish species in the Suyeong River

Table 3.3.3 List of haplotypes identified by the environmental DNA metabarcoding study from the Suyeong River

No.	Haplotypes ID	Haplotypes name	Family	Identity (%)	GenBank number	Frequency
1	Caraau 1	*Carassius auratus*	Cyprinidae	100	KX505165	24
2	Pseupa 1	*Pseudorasbora parva*	Cyprinidae	100	KJ135626	23
3	Misgan 1	*Misgurnus anguillicaudatus*	Cobitidae	100	KC762740	16
4	Hemile 1	*Hemiculter leucisculus*	Cyprinidae	100	LC340359	13
5	Mugice 1	*Mugil cephalus*	Mugilidae	100	KF374974	12
6	Misgmi 1	*Misgurnus mizolepis*	Cobitidae	100	AP017654	11
7	Squaja 1	*Squalidus japonicus*	Cyprinidae	100	KR075134	11
8	Tridob 1	*Tridentiger obscurus*	Gobiidae	100	KT601092	11
9	Planha 1	*Planiliza haematocheila*	Mugilidae	100	KJ622047	10
10	Rhinbr 1	*Rhinogobius brunneus*	Gobiidae	100	KT601096	10
11	Zacc 1	unidentified	Cyprinidae	96	LC277796	9
12	Cyprca 1	*Cyprinus carpio*	Cyprinidae	100	KX710076	8
13	Hemima 1	*Hemibarbus maculatus*	Cyprinidae	99	LC146032	8
14	Lepoma 1	*Lepomis macrochirus*	Centrarchidae	100	AP005993	8
15	Micrsa 1	*Micropterus salmoides*	Centrarchidae	100	HQ391896	8
16	Rhingi 1	*Rhinogobius giurinus*	Gobiidae	100	KM030475	8
17	Caracu 1	*Carassius cuvieri*	Cyprinidae	100	AP011237	7
18	Konopu 1	*Konosirus punctatus*	Clupeidae	100	KC477844	7
19	Odonpl l	*Odontobutis platycephala*	Odontobutidae	100	KM030426	7
20	Zacc 2	unidentified	Cyprinidae	95	LC277796	7

Table 3.3.3 Continued.

No.	Haplotypes ID	Haplotypes name	Family	Identity (%)	GenBank number	Frequency
21	Chaean 1	*Chaenogobius annularis*	Gobiidae	100	AP014796	6
22	Tribha 1	*Tribolodon hakonensis*	Cyprinidae	100	AB626855	6
23	Tridob 2	*Tridentiger obscurus*	Gobiidae	100	KT601092	6
24	Misgmi 2	*Misgurnus mizolepis*	Cobitidae	100	AP017654	5
25	Mugiab 1	*Mugilogobius abei*	Gobiidae	100	KM030465	5
26	Nippte 1	*Nipponocypris temminckii*	Cyprinidae	100	AP012116	5
27	Anguja 1	*Anguilla japonica*	Anguillidae	100	MH050933	4
28	Chanar 1	*Channa argus*	Channidae	100	MG751766	4
29	Hemima 2	*Hemibarbus maculatus*	Cyprinidae	99	LC146032	4
30	Opsaun 1	*Opsariichthys uncirostris*	Cyprinidae	99	AB218897	4
31	Tridtr 1	*Tridentiger trigonocephalus*	Gobiidae	100	KM030481	4
32	Acanfa 1	*Acanthogobius flavimanus*	Gobiidae	100	LC474211	3
33	Chaner 1	*Chanodichthys erythropterus*	Cyprinidae	100	KJ801524	3
34	Corehe 1	*Coreoperca herzi*	Percichthyidae	100	KR075132	3
35	Cypr 1	unidentified	Cyprinidae	96	KX710076	3
36	Hemila 1	*Hemibarbus labeo*	Cyprinidae	99	DQ347953	3
37	Hemile 2	*Hemiculter leucisculus*	Cyprinidae	99	LC340359	3
38	Hemi 2	unidentified	Cyprinidae	95	LC340359	3
39	Oreoni 1	*Oreochromis niloticus*	Cichlidae	100	GU477626	3
40	Rhynla 1	*Rhynchocypris lagowskii*	Cyprinidae	99	KJ641843	3

Table 3.3.3 Continued.

No.	Haplotypes ID	Haplotypes name	Family	Identity (%)	GenBank number	Frequency
41	Rhynox 1	*Rhynchocypris oxycephalus*	Cyprinidae	99	LC193377	3
42	Siluas 1	*Silurus asotus*	Siluridae	100	NC015806	3
43	Squach 1	*Squalidus chankaensis*	Cyprinidae	100	KT948082	3
43	Squach 1	*Squalidus chankaensis*	Cyprinidae	100	KT948082	3
44	Squach 2	*Squalidus chankaensis*	Cyprinidae	99	KT948082	3
45	Squa 2	*Squalidus* sp.	Cyprinidae	97	KR075134	3
46	Zacc 3	unidentified	Cyprinidae	95	LC277796	3
47	Cara 2	unidentified	Cyprinidae	95	KX505165	2
48	Cyprca 2	*Cyprinus carpio*	Cyprinidae	100	KX710076	2
49	Hemila 2	*Hemibarbus labeo*	Cyprinidae	99	DQ347953	2
50	Hemi 1	*Hemibarbus* sp.	Cyprinidae	98	DQ347953	2
51	Hemi 2	*Hemibarbus* sp.	Cyprinidae	98	DQ347953	2
52	Hemi 1	*Hemiculter* sp.	Cyprinidae	98	LC340359	2
53	Misgmi 3	*Misgurnus mizolepis*	Cobitidae	99	AP017654	2
54	Misg 1	*Misgurnus* sp.	Cobitidae	98	KC762740	2
55	Mugiab 2	*Mugilogobius abei*	Gobiidae	99	KM030465	2
56	Nippte 2	*Nipponocypris temminckii*	Cyprinidae	99	AP012116	2
57	Oryzla 1	*Oryzias latipes*	Adrianichthyidae	100	AP008947	2
58	Rhin1	*Rhinogobius* sp.	Gobiidae	97	KT601096	2
59	Rhyn 1	*Rhynchocypris* sp.	Cyprinidae	97	KJ641843	2

Table 3.3.3 Continued.

No.	Haplotypes ID	Haplotypes name	Family	Identity (%)	GenBank number	Frequency
60	Squa 3	*Squalidus* sp.	Cyprinidae	97	KR075134	2
61	Squa 4	unidentified	Cyprinidae	96	KR075134	2
62	Squa 5	unidentified	Cyprinidae	95	KR075134	2
63	Tanala 1	*Tanakia lanceolata*	Cyprinidae	99	LC458035	2
64	Tridtr 2	*Tridentiger trigonocephalus*	Gobiidae	100	KM030481	2
65	Zaccpl 1	*Zacco platypus*	Cyprinidae	99	LC277796	2

3.3.3 Fish biodiversity in the Suyeong River

Based on the eDNA metabarcoding analysis from the MiFish pipeline, the alpha diversity index, Shannon-Wiener (H') index, Margalef index (d) was analyzed from the four sampling stations in the Suyeong River by using the PRIMER® software v7.

The highest Shannon-Wiener (H') index was found at the station A (2.364), followed by station C (1.186) and station B (1.039), while the lowest was at the station D (0.976). The average Margalef index was also highest at the station A (3.406), followed by station C (3.073), station D (2.462), and the lowest (1.963) was at the station B (Figure 3.3.5).

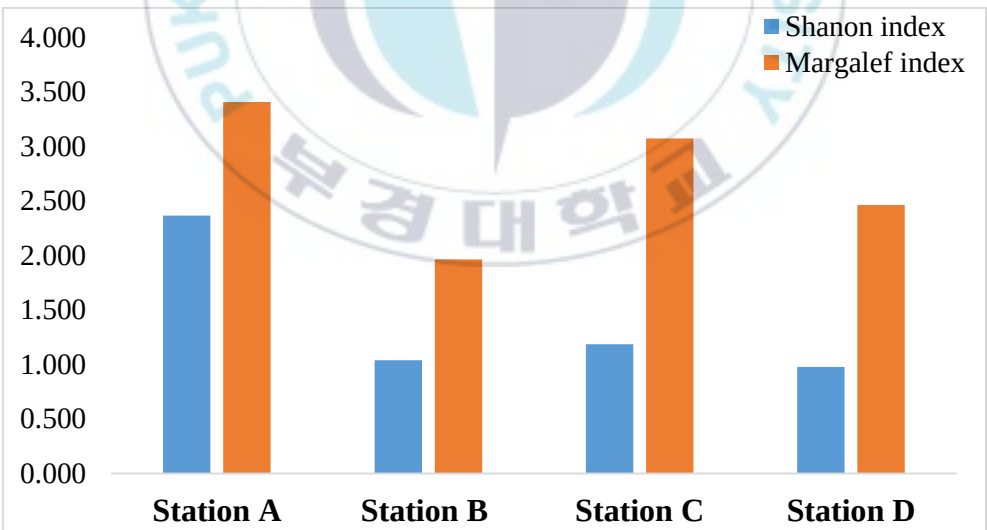


Figure 3.3.5 Shannon-Wiener diversity index in the Suyeong River

The *Carassius auratus* species was detected 24 times in all sampling stations, followed by *Pseudorasbora parva* (23 times), and *Misgurnus anguillicaudatus* (16 times), *Hemiculter leucisculus* (13 times), and *Mugil cephalus* (12 times). The *Misgurnus mizolepis, Squalidus japonicas,* and *Tridentiger obscurus* were found 11 times respectively. A total of 15 haplotypes and 19 haplotypes were found three times and two times respectively, while 32 haplotypes were detected once in the study area, so those haplotypes were discarded from further analysis (Fig. 3.3.6).

A total of 65 fish haplotypes from the four sample collection stations of the Suyeong River, among those haplotypes, the highest 31 were identified from the family Cyprinidae, followed by Gobiidae (8), and the remaining 15 were from the other 11 families.

Haplotypes	Species	StationA	Station B	Station C	Station D
15	*Micropterus salmoides*	75.0	12.5	12.5	0.0
16	*Rhinogobius giurinus*	75.0	12.5	12.5	0.0
45,60,61	*Squalidus* sp.	71.4	0.0	28.6	0.0
11	*Zacco* sp.	66.7	0.0	33.3	0.0
10	*Rhinogobius brunneus*	60.0	10.0	20.0	10.0
14	*Lepomis macrochirus*	50.0	50.0	0.0	0.0
30	*Opsariichthys uncirostris*	50.0	0.0	0.0	50.0
50,51	*Hemibarbus* sp.	50.0	0.0	50.0	0.0
7	*Squalidus japonicus*	45.5	9.1	18.2	27.3
17	*Carassius cuvieri*	42.9	42.9	14.3	0.0
36,49	*Hemibarbus labeo*	40.0	0.0	60.0	0.0
43,44	*Squalidus chankaensis*	40.0	20.0	40.0	0.0
4,37	*Hemiculter leucisculus*	37.5	31.3	25.0	6.3
8,23	*Tridentiger obscurus*	35.3	17.6	11.8	35.3
13,29	*Hemibarbus maculatus*	33.3	33.3	0.0	33.3
19	*Odontobutis platycephala*	28.6	28.6	42.9	0.0
6,24,53	*Misgurnus mizolepis*	26.7	26.7	26.7	20.0
1	*Carassius auratus*	25.0	25.0	25.0	25.0
28	*Channa argus*	25.0	0.0	75.0	0.0
2	*Pseudorasbora parva*	21.7	26.1	26.1	26.1
3	*Misgurnus anguillicaudatus*	18.8	37.5	18.8	25.0
21	*Chaenogobius annularis*	16.7	0.0	50.0	33.3
22	*Tribolodon hakonensis*	16.7	16.7	16.7	50.0
31,64	*Tridentiger trigonocephalus*	16.7	0.0	0.0	83.3
12,48	*Cyprinus carpio*	10.0	40.0	30.0	20.0
5	*Mugil cephalus*	0.0	0.0	50.0	50.0
9	*Planiliza haematocheila*	0.0	0.0	40.0	60.0
18	*Konosirus punctatus*	0.0	0.0	14.3	85.7
20	Unidentified1	0.0	42.9	28.6	28.6
27	*Anguilla japonica*	0.0	0.0	0.0	100.0
25,55	*Mugilogobius abei*	0.0	0.0	14.3	85.7
26,56	*Nipponocypris temminckii*	0.0	0.0	57.1	42.9

Figure 3.3.6 Proportions of detection frequencies with different sample stations in the Suyeong River

Table 3.3.4 Number of fish species identified at the different sampling stations

of the Suyeong River

Month	Station A	Station B	Station C	Station D	Average
August 2017	16	10	17	21	16±4.55
October 2017	13	6	8	18	11.25±5.38
December 2017	12	6	7	13	9.5±3.51
February 2018	11	5	7	11	8.5±3.00
April 2018	16	12	17	16	15.25±2.22
June 2018	26	17	29	25	24.25±5.12
Average	15.67±5.47	9.33±4.63	14.17±8.68	17.33±5.16	

Among all sampling stations, station D had the highest average species detection rate (17.33), while the lowest (9.33) was at station B. During the study period, the highest average number of species was found in June (24.25), while the lowest (8.5) was in February (Figure 3.3.7).

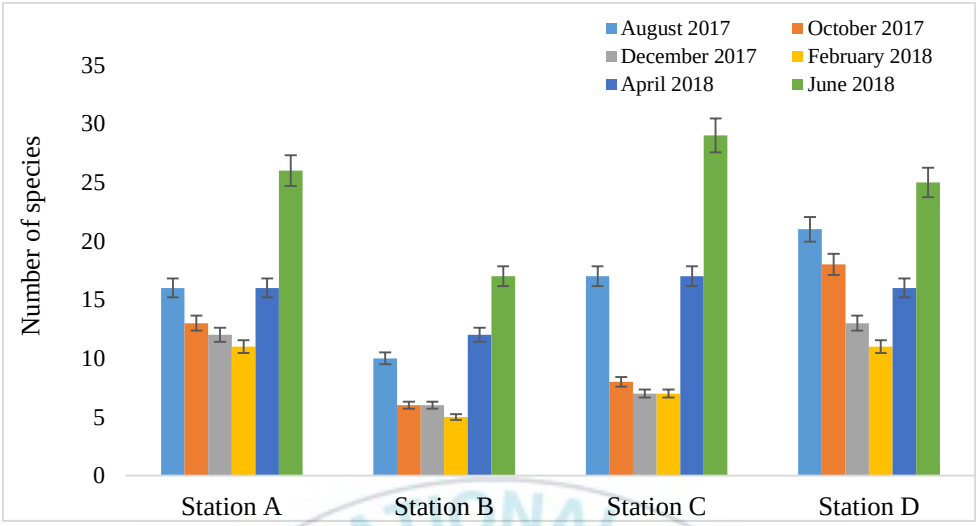


Figure 3.3.7 Number of fish species identified at the different sampling stations of the Suyeong River

Among the 65 haplotypes in the Suyeong River, 12 haplotypes e.g. *Rhinogobius brunneus*, *Squalidus japonicas*, *Misgurnus mizolepis*, *Hemiculter leucisculus*, *Pseudorasbora parva*, *Carassius auratus*, *Zacco* sp., *Cyprinus carpio*, *Tridentiger obscurus*, *Hemibarbus maculatus*, *Tribolodon hakonensis*, and *Misgurnus anguillicaudatus* were present at all the four sampling stations (Figure 3.3.8).

While station A, B, and C have 4 common haplotypes (e.g. *Carassius cuvieri*, *Rhinogobius giurinus*, *Micropterus salmoides*, and *Odontobutis platycephala*). Furthermore, station A and B have 2 common haplotypes (e.g. *Lepomis macrochirus* and *Oreochromis niloticus*), station C and D has 6

common haplotypes (e.g. *Nipponocypris temminckii*, *Mugil cephalus*, *Konosirus punctatus*, *Planiliza haematocheila*, *Mugilogobius abei*, and *Oryzias latipes*). Moreover, station A has 8 unique haplotypes e.g. *Rhinogobius* sp., *Nipponocypris* sp. (unidentified), *Chanodichthys erythropterus*, Coreoperca *herzi*, *Hemibarbus labeo*, *Cyprinus* sp. (unidentified), *Pseudogobio* sp., and *Zacco platypus*. No unique haplotype was found in the station B (Figure 3.3.8).

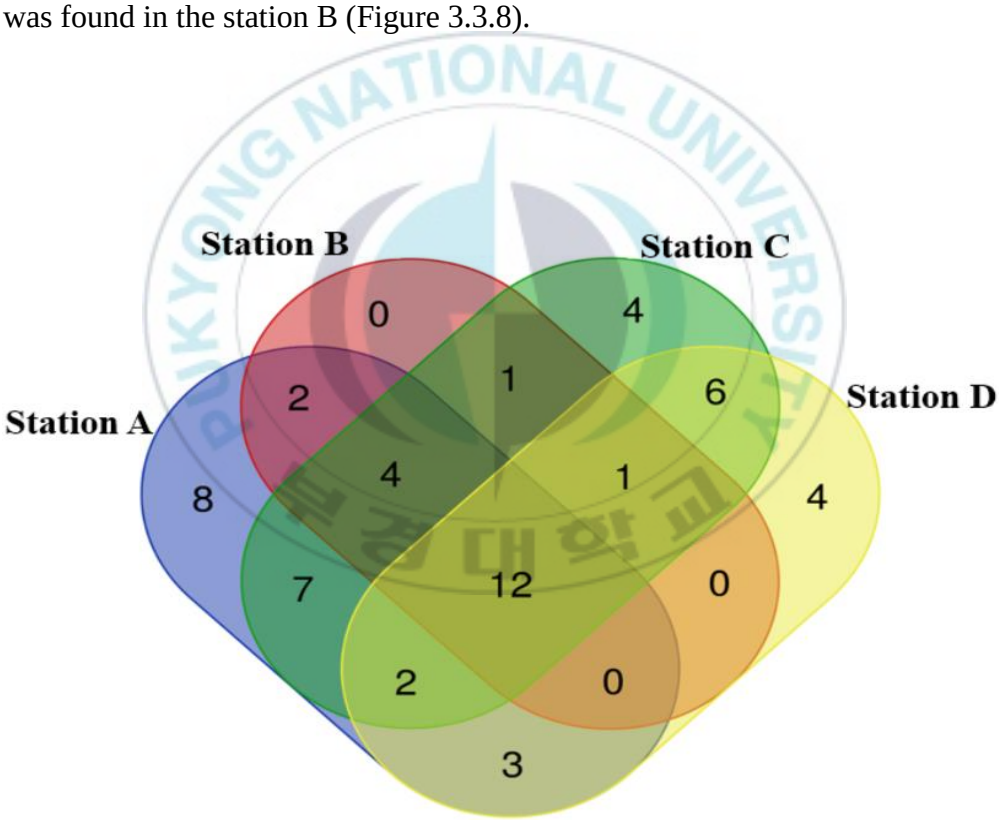


Figure 3.3.8 Venn diagram of fish species identified at different sampling sites in the Suyeong River by the eDNA metabarcoding analysis

In the present study, by the eDNA metabarcoding approach in the Suyeong River from August 2017 to June 2018, the *Mugil cephalus* was dominant (38.13%) followed by *Carassius auratus* (24.27%), and *Pseudorasbora parva* (18.82%). Among four sampling stations, we found that, at the station A, *Carassius auratus* was relatively abundant (29.05%) than *Pseudorasbora parva* (19.06%) and *Carassius cuvieri* (12.53%), while *Pseudorasbora parva* was relatively abundant (43.36%), followed by *Carassius auratus* at station B with 39.87%. At station C, *Mugil cephalus* was dominant (56.85%) than *Carassius auratus* (28.31%), and at station D, *Mugil cephalus* was also dominant (71.31%) than *Pseudorasbora parva* (13.58%).

3.3.4 Salinity and relationship

Salinity was increased from the upstream to the downstream of the Suyeong River, the lower average salinity (0.16 psu) was measured at the upstream (station A), and 0.23 psu at station B, while the highest average salinity (19.75 psu) was found at station D. The fish species distribution with the salinity level was also measured and found that the freshwater fish species distributed at the upstream of rivers (station A and B), while the brackish water fish species distributed at the midstream (station C) and downstream area (station D) of the Suyeong River (Figure 3.3.9).

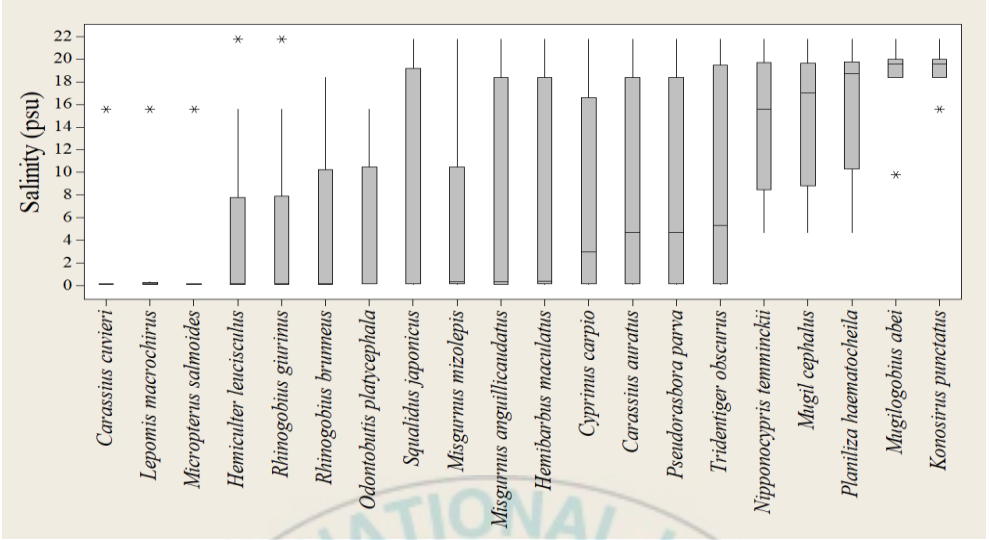


Figure 3.3.9 Fish species distribution with salinity measurement in the Suyeong River

At station A, five fish species e.g. *Tanakia signifier, Coreoperca herzi, Gymnogobius breunigii, Zacco platypus, Pseudogobio vaillanti* were found, where the average salinity range was 0.16 and the salinity range was 0.10 to 0.20 psu. On the other hand, at station D (downstream), seven estuarine fish species e.g. *Chaenogobius gulosus, Anguilla japonica, Acanthopagrus latus, Acanthopagrus schlegelii, Trachurus japonicas, Acanthogobius flavimanus, Rhynchopelates oxyrhynchus* were found, where the average salinity range was 19.76 and salinity range was 18.39 to 21.79 psu (Figure 3.3.9).

3.3.5 Clustering analysis

In this study, we have categorized the sampling stations of the Suyeong River as upstream (station A and B), and downstream (Station C and D). Among the 54 identified fish species in this study, we have taken top 20 species and by using a statistical program (Primer v7), we have drawn a heat map and found that it demonstrated the fish species distribution in different sampling stations (Figure 3.3.10).

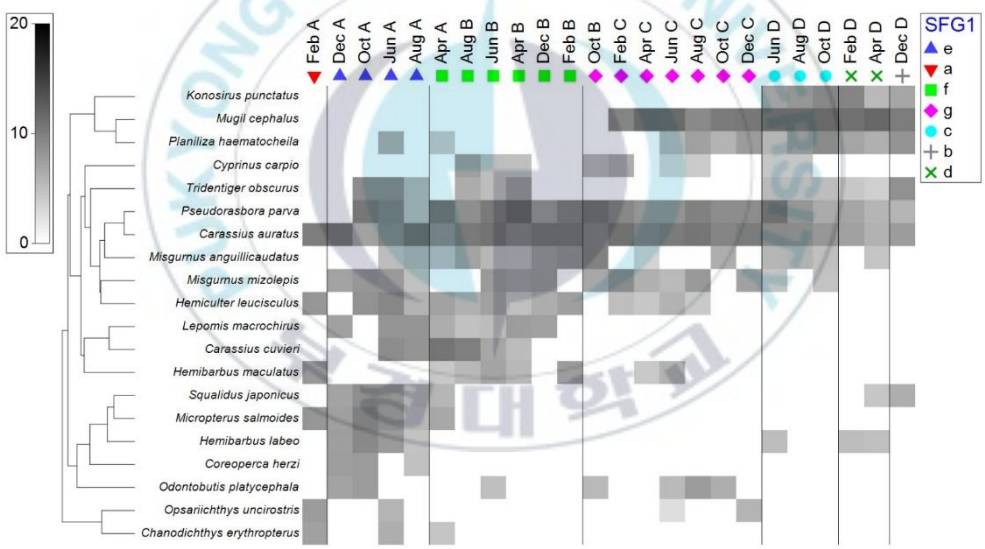


Figure 3.3.10 Heatmap of the eDNA metabarcoding result based on the top 20 species at four sampling stations of the Suyeong River

Based on Bray-Curtis Similarity analysis showed that there is a distinct four cluster of the four sampling stations of the Suyeong River. The upstream sampling stations (stations A and B) were cluster closely and the downstream sampling stations (stations C and D) were clustered closely. One exception was found in February 2018 of the station A was clustered separtely (Figure 3.3.11 A and 3.3.11 B).

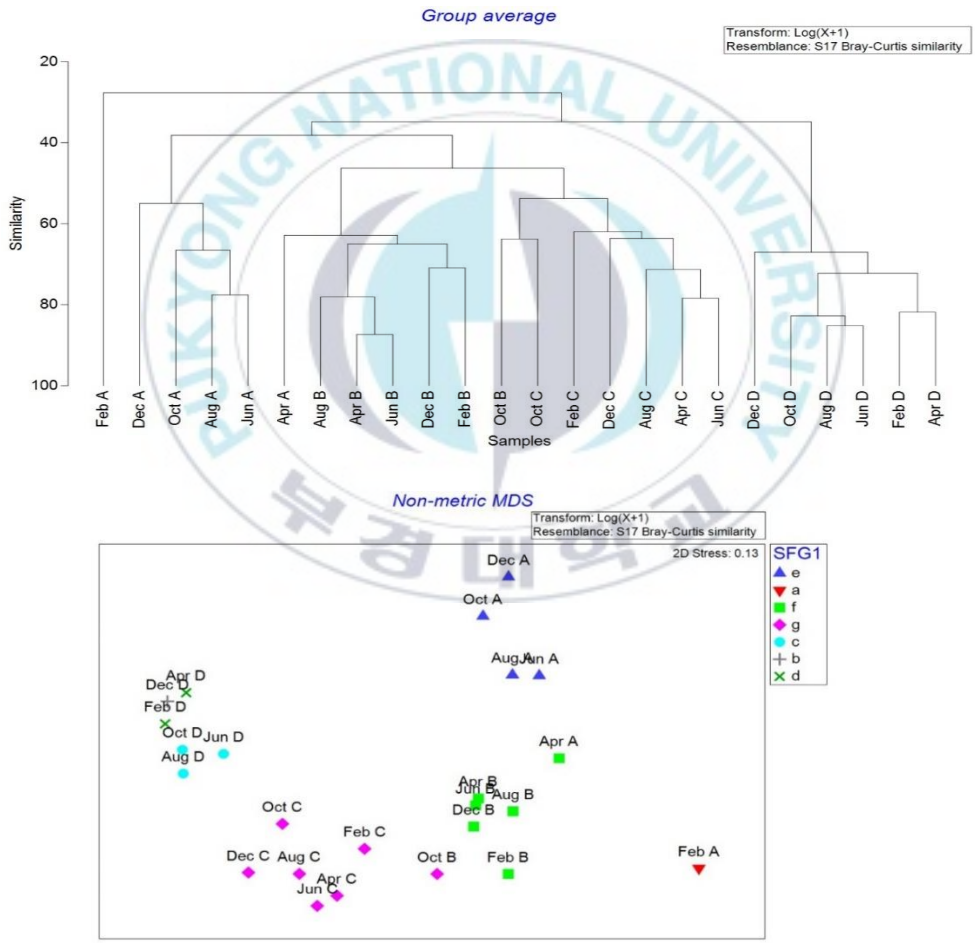


Figure 3.3.11 A and B. The Bray-Curtis similarity and Non-metric MDS Resemblance of fish species from the different sampling stations

In upstream sampling sites (station A and B) of the river, the dominant species were *Zacco platypus*, *Odontobutis interrupta*, *Odontobutis platycephala*, *Nipponocypris temminckii*, *Rhynchocypris lagowskii*, *Misgurnus mizolepis*, *Coreoperca herzi*, *Acheilognathus intermedia*, *and Tanakia signifier* (Figure 3.3.10); these species are non-migratory freshwater species and most of them are endemic in the Korean peninsula (Kim, 1997). At the midstream of the Suyeong River (station C), we found the dominant species were *Pseudorasbora parva*, *Gymnogobius breunigii*, *Rhinogobius giurinus*, *Rhinogobius brunneus*, and *Mugil cephalus*, while at the downstream area (station D), we found *Tridentiger obscurus*, *Tridentiger trigonocephalus*, *Konosirus punctatus*, *Mugil cephalus*, *Anguilla japonica*, and *Planiliza haematocheila;* these all species are estuarine and migratory species (https://www.fishbase.org).

The analysis of similarity (ANOSIM) pairwise test was done by Prime v7 program and demonstrated that the average R statistic vale was 0.81, where the highest (1.0) and the lowest value (0.65) was found at the combination of station B and D, and station A and B respectively (Table 3.3.5). The Bray-Curtis similarity and Non-metric MDS Resemblance also clearly clustered the detected fish species of upstream, midstream and the downstream of the Suyeong River (Figure 3.3.11 A and B).

Table 3.3.5 Analysis of Similarity (ANOSIM) pairwise test by Prime v7

Groups	R Statistic	Significance Level (%)	Actual Permutations
Station A, Station B	0.656	0.2	462
Station A, Station C	0.809	0.2	462
Station A, Station D	0.952	0.2	462
Station B, Station C	0.726	0.2	462
Station B, Station D	1	0.2	462
Station C, Station D	0.715	0.2	462
Average	0.81		

3.4 Discussion

Biodiversity is a measure of the variety of life in an ecosystem. It is determined by both the richness (number of species), and the evenness (comparing the relative abundances of the species). The most diverse community has both high richness and high evenness. In this study, we identified that environmental DNA metabarcoding using the MiFish pipeline is a useful tool for the fish biodiversity analysis in Korean rivers. Due to the short amplified size of the MiFish universal primer, we were able to obtain 54 fish species at the 4 sampling stations of the Suyeong River from August 2017 to June 2018. In 2012, a nationwide survey conducted, Survey and Evaluation of Aquatic Ecosystem Health (SEAEH), total 130 freshwater fish species of 28 families were identified from the 953 sites in Korean waters, among those species, 51 fish species were identified as endemic, 20

endangered and 4 were exotic fish species.

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). Here we identified 31 fish species belonging to the family Cyprinidae (57.41 %), followed by 8 from Gobiidae (14.81 %), and the remaining 15 species (27.78 %) were from other 11 families.

In this study, by the environmental DNA metabarcoding analysis, we found the seasonal variation of fish species abundance in different sampling stations, but the most important factors to be considered about the fish biodiversity study in the Suyeong River was the water salinity variation. For example, we found a high degree of similarity in fish diversity from station C (where Oncheon stream merged with the Suyeong River) and station D (Figure 3.10). This may have come from the influx of seawater to the upstream of the Suyeong River during the high tide, which is correspondent to the high degree of salinity (± 17 psu) in station C in June, which is 2.2 fold higher than in Dec/Feb (7.7 psu).

The major water supply of the Hoidong reservoir (station A), the origin of Suyeong River, comes from the Nakdong River and there is a similar fish

biodiversity structure of the Nakdong river found from our previous study (26 species identified by eDNA metabarcoding on June 2018), where 19 fish species are common in two rivers. By contrast, we also found a strong correlation between two freshwater stations, station A and B (Fig. 3.3.12).

We also identified differences in the fish biodiversity between station A and C, because of station A (Hoidong reservoir), where water is supplying from the Nakdong River, so station A has almost the same fish biodiversity pattern with the Nakdong River. On the other hand, station C is the end of the Oncheon Stream which merged with the Suyeong River, so this station has mixed biodiversity of Oncheon Stream and Suyeong River. Station D is estuarine water, so in this station, we found both freshwater and brackish water fish species.

Lee et al. (2015) reported 18 fish species from the Nakdong River by using traditional survey methods (Lee et al., 2015). Among these 18 fish species, 11 fish species (e.g. *Carassius auratus*, *Carassius cuvieri*, *Chanodichthys erythropterus*, *Coreoperca herzi*, *Cyprinus carpio*, *Hemibarbus labeo*, *Micropterus salmoides*, *Odontobutis platycephala*, *Silurus asotus*, and *Tanakia lanceolata*, *Zacco platypus*) were also identified by this eDNA metabarcoding study.

In this study, by the eDNA metabarcoding approach, we also identified five exotic fish species in the Suyeong River (e.g. *Carassius cuvieri*, *Cyprinus*

carpio, Lepomis macrochirus, *Micropterus salmoides*, and *Oreochromis niloticus*), and now they are widely distributed in Korean waters, which is very alarming and important for the policymakers and the biologists for the better management strategies in the inland Korean waters.

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 from 16 families, they also identified 5 exotic fish species e.g. *Carassius cuvieri, Cyprinus carpio, Micropterus salmoides, Lepomis macrochirus,* and *Oreochromis niloticus* (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).

By using the traditional survey 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 et al., 2014). Since freshwater ecosystems are much more vulnerable to invasive species causing biodiversity loss and global change (Clavero and García-Berthou, 2005). It is highly required to conduct a long-term survey throughout the country in the river systems.

The spatial distribution of environmental DNA in a lotic environment is a little bit difficult to detect species accurately in rivers or streams (Deiner et al., 2016; Jerde et al., 2016). It is tough to confirm species detection related to sampling location, where the Environmental DNA transported from the upstream to the downstream, which may lead to detecting as a 'false positive' (Li et al., 2018). Environmental DNA transported distance was found 50-1450 m in low discharge (Q=2-10 L/s) experimental streams (Jerde et al., 2016; Wilcox et al., 2016) but in high discharge systems (Q>3000 L/s), at the 10 km downstream, eDNA has been detected (Deiner and Altermatt, 2014). Further studies required for better understanding and quantify eDNA transport dynamics in rivers/streams.

As freshwater is lighter than the marine water, the genetic materials of some commercial marine fish or non-native fish used in the residence or restaurant may also be transported to the rivers and offshore areas (Yamamoto et al., 2017). Even though the eDNA metabarcoding approach draws an

apparent structure of the fish community, this transported DNA should be considered as an influential noise. Some statistical approaches are being applied to minimize this potential noise (Lahoz-Monfort et al., 2016; Stoeckle et al., 2017). Sometimes it is also possible that the DNA may come from some predators transfer carcasses or defecation (Merkes et al., 2014). So, consistent and long-term monitoring would increase the reliability of the data from the environmental samples.

The annual phytoplankton succession during the summer and the winter seasons typically happened in the temperate zone (Townsend et al., 1992). In the summer conditions with an abundance of phytoplankton will increase the zooplankton, which supporting the nutrients coming from the Suyeong River. The essential nutrients (Nitrate, Silicate, and Phosphorus) from the river to coastal water is an important factor for phytoplankton growth (Justić et al., 1995; Redfield, 1963). Between freshwater and marine water mixing in the estuarine ecosystem, where nutrients load support high biodiversity in this area. These changes occur around 3-4 months in the year, and when winter comes, it has a declining impact on primary productivity. With the decline in primary productivity, the abundance of tropics on it will also decrease. In the winter season, the food chain forms a new balance with its limited type of phytoplankton (Thompson et al., 2008).

However, several species identified in this study were able to survive in the conditions in both seasons. In this eDNA metabarcoding results, we found that fish species under the order of Mugiliformes (e.g. *Mugil cephalus*, *Planiliza haematocheila*) can adapt the changes in a wide range of water temperature and salinity. The *Mugil cephalus* is a commonly consumed fish in the Busan area (Jang et al., 2002; Kwak et al., 2015), and this fish species identified at the downstream sampling sites (station C and D) of the Suyeong River throughout the sampling period.

In this study, the 54 identified species were clustered into two large clades, namely the summer clade (April, June, and August) while the remaining clades were winter (October, December, and February). During the summer season, In June, the highest number of species identified in each station, followed by August and April, while the lost number of fish species were identified in February.

From the present study by the eDNA metabarcoding analysis indicated that there is a changing pattern of fish diversity in summer and winter season in the Suyeong River. The availability of feed for fish is one of the essential variables in the food chain, especially the primary productivity (phytoplankton). The presence of phytoplankton in summer is very significantly higher than in winter (Ryu, 2011). The nutrients delivered from the Suyeong River carry enough Nitrate and the Phosphate components to

support the growth of phytoplankton, besides the availability of sunlight in the photosynthesis process also completes the availability of feed for zooplankton and other tropics above it in the summer season, which triggered the higher abundance of freshwater fish in a riverine environment.

Cryptic species is a major problem in morphological based identification (Piggott et al., 2011), species identification errors often occur in a study. The use of molecular approaches has managed to overcome this problem even if the fish is at the larvae stage (Hubert et al., 2012; Rezagholinejad et al., 2016). In addition, rare and invasive species in an aquatic system were described successfully in molecular base surveillance and monitoring through Environmental DNA (eDNA) detection (Mahon and Jerde, 2016; Nevers et al., 2018).

In this research, we detected 65 fish haplotypes by using the eDNA metabarcoding approach from the water samples collected at the four sampling sites of the Suyeong River. Some or few of these detected fish species may be at their larval stage or juvenile stage, and these stages are difficult to detect by visual observation. Another important benefit of the eDNA metabarcoding is allowing us to survey or monitor more sampling sites in a cheaper and faster way; feasibility study of suitable environments for potential species also possible (Bista et al., 2017; Deiner et al., 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.

However, there are still several drawbacks to eDNA metabarcoding in its use to overcome. First, low MiFish primer region (12S rRNA) sequence data, researchers should upload the sequence of the region, numbers of indigenous fish species in the MiFish database. Secondly, the 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, sometimes it is difficult to differentiate some close related fish species e.g. *Sebastes* spp. and *Takifugu* spp., so the high-quality database is most important for the successful species assignment (Sato et al., 2018; Yamamoto et al., 2017).

Thirdly, the taxonomic assignment of the MiFish pipeline is a bit different from the taxonomy of FishBase (https://www.fishbase.org), for example, from the MiFish pipeline the *Coreoperca herzi* were assigned to the family Sinipercidae and order Centrarchiformes; but FishBase assigned this fish belongs to family Percichthyidae and order Perciformes. Moreover,

in the Mifish database the family Cichlidae, Gobiidae, Sparidae were assigned under the order Cichliformes, Gobiiformes, Spariformes respectively, but these all three family are belongs to the order Perciformes in the FishBase database (Table 3.3.7).

Fourthly, the transported genetic materials (eDNA) affects the accurate community structure. Environmental DNA transported from the upstream may lead to false positive in the downstream areas and it becomes difficult to infer species location relative to the sampling location. The above points are influential obstacles to biological or ecological research, for designing or implementing the conservation strategies based on eDNA metabarcoding results. Additional mechanistic tests of such environmental variables are needed to understand when and how the environmental conditions in aquatic ecosystems influence the detection rate and transport of eDNA. However, these drawbacks might be solved by designing a research plan carefully.

Table 3.3.6 Comparison of the taxonmic assignment between MiFish pipeline and FishBase website

No.	Species name	Family level taxonomic assigned		Order level taxonomic assigned	
		MiFish pipeline	FishBase	MiFish pipeline	FishBase
1	*Trachurus japonicus*	Carangidae	Carangidae	Carangiformes	Perciformes
2	*Lepomis macrochirus*	Centrarchidae	Centrarchidae	Centrarchiformes	Perciformes
3	*Micropterus salmoides*	Centrarchidae	Centrarchidae	Centrarchiformes	Perciformes
4	*Channa argus*	Channidae	Channidae	Anabantiformes	Perciformes
5	*Oreochromis niloticus*	Cichlidae	Cichlidae	Cichliformes	Perciformes
6	*Acanthogobius flavimanus*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
7	*Chaenogobius annularis*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
8	*Chaenogobius gulosus*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
9	*Gymnogobius breunigii*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
10	*Mugilogobius abei*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
11	*Rhinogobius brunneus*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
12	*Rhinogobius giurinus*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
13	*Tridentiger obscurus*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
14	*Tridentiger trigonocephalus*	Gobiidae	Gobiidae	Gobiiformes	Perciformes
15	*Odontobutis platycephala*	Odontobutidae	Odontobutidae	Gobiiformes	Perciformes
16	*Coreoperca herzi*	Sinipercidae	Percichthyidae	Centrarchiformes	Perciformes
17	*Acanthopagrus latus*	Sparidae	Sparidae	Spariformes	Perciformes
18	*Acanthopagrus schlegelii*	Sparidae	Sparidae	Spariformes	Perciformes
19	*Rhynchopelates oxyrhynchus*	Terapontidae	Terapontidae	Centrarchiformes	Perciformes

3.5 Future directions for the eDNA metabarcoding research

Although we identified that eDNA metabarcoding analysis by using the MiFish pipeline is a useful tool to monitor the biodiversity of freshwater fish species, we also found several points to consider the future research plan to applying the platform for fish biodiversity analysis. First, MiFish sequence data for endemic fish species in Korean water should be supplemented and we can study few endemic fish species very intensively and repeated monitoring of their abundance, migration and biomass can be studied by the application real-time PCR technique.

Secondly, the quantitative analytic methods should be introduced for the biodiversity study. In fact, the presence or absence of species by using real-time PCR (qPCR) have been undertaken successfully. Moreover, based on the concentration of environmental DNA in water, we can also try to estimate species abundance and biomass in different ecosystem. Thirdly, we can use the eDNA metabarcoding technique to study about the distribution and abundance of the exotic or invasive species in aquatic ecosystem. Normally, using the 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 to 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.

3.6 Conclusion

The combination with the Next Generation Sequencing (NGS) method, the environmental DNA metabarcoding approach is a potential technique to identify multiple species and/or to monitor species diversity in aquatic habitat. It will offer a more precise estimation of fish biodiversity rather than single or a handful of fish species surveillance in a waterbody. Historically freshwater fish living separately in their habitat (river, lake, reservoir, stream), so each country should have its database of genetic information e.g. DNA sequences for this kind of environmental DNA metabarcoding study. The present study revealed that the eDNA metabarcoding by using the MiFish universal primer sets can uncover the fish biodiversity in a Korean river. 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 the Korean peninsula. These results may be able to utilize for various purposes, e.g. take some effective steps to enhance efforts from the government and improve public conversance for the better management of the freshwater resources. Our findings suggest that environmental DNA metabarcoding required less time and taxonomic expertise, and it is better to understand the fish distribution and biodiversity in rivers/streams than the traditional survey systems.

IV. 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.

Baird, D.J., Sweeney, B.W., 2011. Applying DNA barcoding in benthology: the state of the science. Journal of the North American Benthological Society 30, 122-124.

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.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Douglas, W.Y., De Bruyn, M., 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in ecology & evolution 29, 358-367.

Bonar, S.A., Hubert, W.A., Willis, D.W., 2009. Standard methods for sampling North American freshwater fishes.

Bruton, M.N., 1995. Have fishes had their chips? The dilemma of threatened fishes. Environmental Biology of Fishes 43, 1-27.

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.

Colautti, R.I., Niimi, A.J., van Overdijk, C.D., Mills, E.L., Holeck, K., MacIsaac, H.J., 2003. Spatial and temporal analysis of transoceanic shipping vectors to the Great Lakes. Island Press: Washington DC.

Communities, C.o.t.E., 2009. Communication From the Commission to the Council, the European Parliament, the European Economic and Social Committee, and the Committee of the Regions: A Mid-Term Assessment of Implementing the EC Biodiversity Action Plan. Journal of International Wildlife Law & Policy 12, 108-120.

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'neill, R.V., Paruelo, J., 1997. The value of the world's ecosystem services and natural capital. nature 387, 253.

Díaz, S., Fargione, J., Chapin III, F.S., Tilman, D., 2006. Biodiversity loss threatens human well-being. PLoS biology 4, e277.

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.

Deiner, K., Knapp, R.A., Boiano, D.M., May, B., 2013. Increased accuracy of species lists developed for alpine lakes using morphology and cytochrome oxidase I for identification of specimens. Molecular ecology resources 13, 820-831.

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.

DeWalt, R.E., 2011. DNA barcoding: a taxonomic point of view. Journal of the North American Benthological Society 30, 174-181.

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, 30-39.

Doi, H., Takahara, T., Minamoto, T., Matsuhashi, S., Uchii, K., Yamanaka, H., 2015. Droplet digital polymerase chain reaction (PCR) outperforms real-time PCR in the detection of environmental DNA from an invasive fish species. Environmental science & technology 49, 5601-5608.

Dong–Myung, K., Sunyoung, K., Jiwon, S., Young–Mog, K., Kyong Joon, R., Seok Jin, O., Oshima, Y., KANG, I.J., 2013. A Study on the Distribution of Nonylphenol in Estuaries in Busan, Korea. J. Fac. Agr., Kyushu Univ 58, 313-318.

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.

Dulvy, N.K., Metcalfe, J.D., Glanville, J., Pawson, M.G., Reynolds, J.D., 2000. Fishery stability, local extinctions, and shifts in community structure in skates. Conservation Biology 14, 283-293.

Edwards, J., Schindler, E., 2017. A Video Monitoring System to Evaluate Ocean Recreational Fishing Effort in Astoria, Oregon.

Egerton, J., Johnson, A., Turner, J., LeVay, L., Mascareñas-Osorio, I., Aburto-Oropeza, O., 2018. Hydroacoustics as a tool to examine the effects of Marine Protected Areas and habitat type on marine fish communities. Scientific reports 8, 47.

Elliott, M., Whitfield, A.K., Potter, I.C., Blaber, S.J.M., Cyrus, D.P., Nordlie, F.G., Harrison, T.D., 2007. The guild approach to categorizing estuarine fish assemblages: a global review. Fish and Fisheries 8, 241-268.

Evans, N.T., Lamberti, G.A., 2017. Freshwater fisheries assessment using environmental DNA: A primer on the method, its potential, and shortcomings as a conservation tool. Fisheries Research.

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.

Ficetola, G.F., Miaud, C., Pompanon, F., Taberlet, P., 2008. Species detection using environmental DNA from water samples. Biology letters 4, 423-425.

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.

Fukumoto, S., Ushimaru, A., Minamoto, T., 2015. A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: a case study of giant salamanders in Japan. Journal of Applied Ecology 52, 358-365.

Gabriel, O., Lange, K., Dahm, E., Wendt, T., 2008. Von Brandt's fish catching methods of the world. John Wiley & Sons.

Goldberg, C.S., Pilliod, D.S., Arkle, R.S., Waits, L.P., 2011. Molecular detection of vertebrates in stream water: a demonstration using Rocky Mountain tailed frogs and Idaho giant salamanders. PloS one 6, e22746.

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.

Hiddink, J.G., MacKenzie, B.R., Rijnsdorp, A., Dulvy, N., Nielsen, E.E., Bekkevold, D., Heino, M., Lorance, P., Ojaveer, H., 2008. Importance of fish biodiversity for the management of fisheries and ecosystems. Fisheries Research 90, 6-8.

Hilborn, R., Quinn, T.P., Schindler, D.E., Rogers, D.E., 2003. Biocomplexity and fisheries sustainability. Proceedings of the National Academy of Sciences 100, 6564-6568.

Hong, Y.-K., Sung, H.-C., Ko, M.-H., Kim, K.-S., Bang, I.-C., 2017. Distribution status and habitat characteristics of the endangered freshwater fish, Microphysogobio rapidus (Cyprinidae). Animal Cells and Systems 21, 286-293.

Hooper, D.U., Chapin, F., Ewel, J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J., Lodge, D., Loreau, M., Naeem, S., 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecological monographs 75, 3-35.

Hopkins, G., Freckleton, R.P., 2002. Declines in the numbers of amateur and professional taxonomists: implications for conservation. Animal Conservation 5, 245-249.

Hubert, N., Meyer, C.P., Bruggemann, H.J., Guerin, F., Komeno, R.J., Espiau, B., Causse, R., Williams, J.T., Planes, S., 2012. Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA-barcoding provides new support to the centre-of-overlap hypothesis. PloS one 7, e28987.

Hutchings, J.A., 2000. Collapse and recovery of marine fishes. Nature 406, 882.

Ikeda, K., Doi, H., Tanaka, K., Kawai, T., Negishi, J.N., 2016. Using environmental DNA to detect an endangered crayfish Cambaroides japonicus in streams. Conservation Genetics Resources 8, 231-234.

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., Chadderton, W.L., Mahon, A.R., Renshaw, M.A., Corush, J., Budny, M.L., Mysorekar, S., Lodge, D.M., 2013. Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. Canadian Journal of Fisheries and Aquatic Sciences 70, 522-526.

Jerde, C.L., Mahon, A.R., Chadderton, W.L., Lodge, D.M., 2011. "Sight-unseen" detection of rare aquatic species using environmental DNA. Conservation Letters 4, 150-157.

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.

Jones, F.C., 2008. Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environmental Reviews 16, 45-69.

Justić, D., Rabalais, N.N., Turner, R.E., Dortch, Q., 1995. Changes in nutrient structure of river-dominated coastal waters: stoichiometric nutrient balance and its consequences. Estuarine, Coastal and Shelf Science 40, 339-356.

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., 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, U.-J., Kim, H.Y., Alvarez, D., Lee, I.-S., Oh, J.-E., 2014. Using SPMDs for monitoring hydrophobic organic compounds in urban river water in Korea compared with using conventional water grab samples. Science of the Total Environment 470, 1537-1544.

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, 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, 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.

Kwak, S.-N., Kim, D.-M., Chung, Y.-H., 2015. Management for Improvement in Water Quality and Change of Fish Assemblage in Urban Dong Stream with Input of Seawater. Journal of Fisheries and Marine Sciences Education 27, 253-261.

Lahoz-Monfort, J.J., Guillera-Arroita, G., Tingley, R., 2016. Statistical approaches to account for false-positive errors in environmental DNA samples. Molecular Ecology Resources 16, 673-685.

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.

Lenat, D.R., Resh, V.H., 2001. Taxonomy and stream ecology—the benefits of genus-and species-level identifications. Journal of the North American Benthological Society 20, 287-298.

Li, Y., Evans, N.T., Renshaw, M.A., Jerde, C.L., Olds, B.P., Shogren, A.J., Deiner, K., Lodge, D.M., Lamberti, G.A., Pfrender, M.E., 2018. Estimating fish alpha-and beta-diversity along a small stream with environmental DNA metabarcoding. Metabarcoding and Metagenomics

2, e24262.

Mace, G., Masundire, H., Baillie, J., Ricketts, T., Brooks, T., Hoffmann, M., Stuart, S., Balmford, A., Purvis, A., Reyers, B., 2005. Biodiversity: millennium ecosystem assessment: current state and trends: findings of the condition and trends working group. Ecosystems and human well-being. Island Press, Washington, DC.

Maclaurin, J., Sterelny, K., 2008. What is biodiversity? University of Chicago Press.

Magurran, A.E., 1988. Ecological diversity and its measurement. Princeton university press.

Mahapatra, B., Sarkar, U., Lakra, W., 2014. A review on status, potentials, threats and challenges of the fish biodiversity of West Bengal. Biodiversity, Bioprospecting and Development, 2376-0214.1000140.

Mahon, A.R., Jerde, C.L., 2016. Using Environmental DNA for Invasive Species Surveillance and Monitoring, Marine Genomics. Springer, 131-142.

Mahon, A.R., Jerde, C.L., Galaska, M., Bergner, J.L., Chadderton, W.L., Lodge, D.M., Hunter, M.E., Nico, L.G., 2013. Validation of eDNA surveillance sensitivity for detection of Asian carps in controlled and field experiments. PloS one 8, e58316.

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.

Merkes, C.M., McCalla, S.G., Jensen, N.R., Gaikowski, M.P., Amberg, J.J., 2014. Persistence of DNA in carcasses, slime and avian feces may affect interpretation of environmental DNA data. PLoS One 9, e113346.

Minamoto, T., Yamanaka, H., Takahara, T., Honjo, M.N., Kawabata, Z.i.,

2012. Surveillance of fish species composition using environmental DNA. Limnology 13, 193-197.

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.

Moyle, P.B., Leidy, R.A., 1992. Loss of biodiversity in aquatic ecosystems: evidence from fish faunas, Conservation biology. Springer, 127-169.

Murphy, B.R., Willis, D.W., 1996. Fisheries techniques. American Fisheries Society Bethesda, Maryland.

Nations, U., 2002. Plan of implementation of the World Summit on Sustainable Development. UN.

Nevers, M.B., Byappanahalli, M.N., Morris, C.C., Shively, D., Przybyla-Kelly, K., Spoljaric, A.M., Dickey, J., Roseman, E.F., 2018. Environmental DNA (eDNA): A tool for quantifying the abundant but elusive round goby (Neogobius melanostomus). PloS one 13, e0191720.

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, M., 1997. A study on the water quality characteristics of the Suyeong river. Master Thesis, Korea national university of education, Cheongju,

Korea.

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.

Pauly, D., Christensen, V., Guénette, S., Pitcher, T.J., Sumaila, U.R., Walters, C.J., Watson, R., Zeller, D., 2002. Towards sustainability in world fisheries. Nature 418, 689.

Pfrender, M., Hawkins, C., Bagley, M., Courtney, G., Creutzburg, B., Epler, J., Fend, S., Ferrington Jr, L., Hartzell, P., Jackson, S., 2010. Assessing macroinvertebrate biodiversity in freshwater ecosystems: advances and challenges in DNA-based approaches. The Quarterly Review of Biology 85, 319-340.

Piggott, M.P., Chao, N.L., Beheregaray, L.B., 2011. Three fishes in one: cryptic species in an Amazonian floodplain forest specialist. Biological Journal of the Linnean Society 102, 391-403.

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

Pilliod, D.S., Goldberg, C.S., Laramie, M.B., Waits, L.P., 2013b. 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.

Redfield, A.C., 1963. The influence of organisms on the composition of

seawater. The sea 2, 26-77.

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, 1450-1459.

Revenga, C., Brunner, J., Henninger, N., Kassem, K., Payne, R., 2000. Freshwater Systems. World Resources Institute Washington, DC.

Revenga, C., Mock, G., 2000. Freshwater biodiversity in crisis. Earth Trends World Resources Institute, 1-4.

Rezagholinejad, S., Ehteshami, F., Arshad, A., Amin, S., Ara, R., 2016. Identification and morphometric descriptions of four dominant families of larval fishes in the estuary of Marudu Bay, Sabah, East Malaysia. Iranian Journal of Fisheries Sciences 15, 1262-1268.

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.

Ryu, J.-C., 2011. Abstracts of the 4 th International Conference on Environmental Health Science, 27–28 October 2011, Incheon Korea. Toxicology and Environmental Health Sciences 3, 1-88.

Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., 2000.

Global biodiversity scenarios for the year 2100. science 287, 1770-1774.

Sato, M., Kawaguchi, Y., Nakajima, J., Mukai, T., Shimatani, Y., Onikura, N., 2010. A review of the research on introduced freshwater fishes: new perspectives, the need for research, and management implications. Landscape and Ecological Engineering 6, 99-108.

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.

Sinclair, M., Arnason, R., Csirke, J., Karnicki, Z., Sigurjonsson, J., Skjoldal, H.R., Valdimarsson, G., 2002. Responsible fisheries in the marine ecosystem. Fisheries Research 58, 255-265.

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.

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

Thompson, P.A., Bonham, P.I., Swadling, K.M., 2008. Phytoplankton blooms in the Huon Estuary, Tasmania: top-down or bottom-up control? Journal of Plankton Research 30, 735-753.

Thomsen, P.F., Kielgast, J., Iversen, L.L., Møller, P.R., Rasmussen, M., Willerslev, E., 2012a. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PLoS one 7, e41732.

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

Thomsen, P.F., Møller, P.R., Sigsgaard, E.E., Knudsen, S.W., Jørgensen, O.A., Willerslev, E., 2016. Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater fishes. PloS one 11, e0165252.

Townsend, D.W., Keller, M.D., Sieracki, M.E., Ackleson, S.G., 1992. Spring phytoplankton blooms in the absence of vertical water column stratification. Nature 360, 59.

Uchii, K., Doi, H., Minamoto, T., 2016. A novel environmental DNA approach to quantify the cryptic invasion of non-native genotypes. Molecular Ecology Resources 16, 415-422.

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.

Valentini, A., Pompanon, F., Taberlet, P., 2009. DNA barcoding for ecologists. Trends in Ecology & Evolution 24, 110-117.

Wheeler, Q.D., Raven, P.H., Wilson, E.O., 2004. Taxonomy: impediment or expedient? American Association for the Advancement of Science.

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.

WON, J.-H., LEE, B.-J., SIM, M.-G., PARK, H.-S., 1979. EFFECT OF SUYEONG RIVER WATER ON THE WATER POLLUTION OF SUYEONG BAY. Korean Journal of Fisheries and Aquatic Sciences 12, 267-276.

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.

Yoon, J., Kim, J., Park, S., Jang, M., 2018. The distribution and diversity of freshwater fishes in Korean Peninsula. Kor J Ecol Environ 51, 71-85.

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.

V. Summary in Korean

환경 DNA (eDNA) 메타 바 코딩은 생태 및 보존 생물학 연구에서 상대적으로 비용 효율적인 신규 바이오 평가 접근법으로, 높은 효율을 가진 생태계의 생물 다양성을 추정합니다. 이 방법은 종의 존재 / 분포를 감지하는데있어 놀랍고 효과적인 도구가 될 것입니다. 여기에서, 우리는 MiFish Pipeline을 채택하여 강물 샘플로부터 물고기의 생물 다양성을 추정 할 수있는 시스템인지를 확인했습니다. 우리는 한국의 4 개 강 (형산, 태화, 섬진, 낙동)의 16 개 물 샘플을 분석 한 결과, 한 번의 조사만으로 미 피쉬 파이프 라인이 생태계에 지장을주지 않으면 서 생물 다양성을 추정하는 유용한 도구라는 단일 조사만으로 73 종의 어류를 식별했습니다. 4 개의 강 중에서 섬진강 (52 종)에서 가장 높은 생물 다양성이 확인되었으며, 그 뒤에 태화강 (42 종), 형산강 (40 종), 낙동강 (26 종)이 낙동 생태계를 시사 강은 대도시에 비해 건강에 좋지 않습니다. 그러나, 우리는 또한 고유종 어류의 대표적인 일배 체형 정보가 더 나은 종 식별을 위해 보충되어야한다는 것을 알 수 있었다. 5 종의 침습 종 (*Carassius cuvieri, Cyprinus carpio, Cyprinus megalophthalmus, Lepomis macrochirus, Micropterus salmoides*) 도 조사 된 모든 하천에 널리 분포되어있어 한국 강 생태계에 문제가 될 수있다.

이 논문의 두 번째 부분에서, 우리는 수영 강에서 1 년 동안 종의 풍부 도와 계절 변화를 비교했다. 이 연구를 위해 우리는 2017 년 8 월부터 2018 년 6 월까지 수강의 4 개 샘플링 스테이션에서 물 샘플을 수집했습니다. MiFish 범용 프라이머 세트는이 강에서 eDNA 메타 바 코딩 분석에 사용되었습니다. 여기에서 우리는 수강의 4 개 시료 채집 스테이션에서 50 종의 어류를 식별했으며, 그 중 어류 중 21 개는 Cyprinidae과에서 식별되었으며 그 뒤에 Gobiidae (9), Cobitidae (2), 나머지 18 종은 다른 15 가구에서 4 개의 샘플링 스테이션 중에서 가장 높은 평균 Shannon-Wiener (H ') 지수는 스테이션 D (2.599)에서 발견 된 반면 가장 낮은 것은 스테이션 B (2.088)에서 발견되었습니다. 모든 샘플링 스테이션 중에서 스테이션 D는 가장 높은 평균 종 탐지율 (17.33)을 보인 반면 가장 낮은 (9.33)은 스테이션 B에있었습니다. 연구 기간 동안 6 월

(24.25) 에 가장 높은 평균 종 수가 발견되었습니다. 가장 낮은 (8.5)은 2 월에있었습니다. 우리는 또한 수강에서 5 종의 침습성 종 (*Carassius cuvieri, Cyprinus carpio, Micropterus salmoides, Lepomis macrochirus, Oreochromis niloticus*) 을 확인하여 현재 한국 해역에 널리 분포하고있다. 우리의 연구 결과는 환경 DNA 메타 바 코딩이 더 적은 시간과 분류 학적 전문 지식을 필요로했으며, 기존의 조사 시스템보다 하천 / 하천의 어류 분포와 생물 다양성을 이해하는 것이 좋습니다. MiFish 메타 바 코딩이 한국 강에서 서식하는 어종을 성공적으로 제시했지만 더 나은 결과를 얻으려면 추가 서열 데이터를 보충해야합니다. 환경 DNA 분석에 의한 고유종, 멸종 위기종, 침입 종 및 어류 분포의 탐지 정확도는 다른 생태계에서 전통적인 모니터링 접근법을 보완하는 데 매우 효과적입니다. 결론적으로, 이번 발견은 한반도 내륙의 해역에서 어류 자원의 효과적인 관리 또는 보존에 유용한 정보를 제공 할 것으로 기대된다.

VI. Acknowledgment

My deepest appreciation to the government of the Republic of Korea, for providing scholarship and research funding during my doctoral program at Pukyong National University, Busan, Republic of Korea.

I express my deepest gratitude to Professor Hyun-Woo Kim, my supervisor, who has been following the progression of this thesis with dedication. Thank you for your feedback on this dissertation. I am so honored to be able to practice and learn about the molecular ecology approaches and techniques during my doctoral study here. I am also thankful to Prof. Won-Gyu Park, Prof. Young-Mog Kim, Prof. Do-Hyung Kim, and Dr. Kyeong-Dong Park, my examiner, for their valuable input and encouraging comments during the examination.

Many thanks and gratefulness to my brother like friend Dr. Sapto Andriyono, for his indebted cooperation during my study period. I am also thankful to all of my colleagues of Molecular Physiology Laboratory, (Dr. Hye-Eun Kang, Dr. Ji-Hyun Lee, Dr. Sinar Pagi Sektiana, Ah-Ran Kim, Soo Rin Lee, Nack-Keun Kim, Eun-Bi Kim, Nazia, Hilman, Sun-Ok Lee, Tae-Ho Yoon, A-Young Jeon, Hyun Sa Gong, Gungryul Kim, SunYoung Yoon) for their warm support and assistance, where many idea and interesting discussions come to pass. I am eternally grateful to my respected supervisor and all of my lab colleagues; I dedicated this dissertation to all of them.