Illumina iSeq 100 and MiSeq exhibit similar performance in freshwater fish environmental DNA metabarcoding

Running title: Fish eDNA metabarcoding using iSeq and MiSeq

Ryohei Nakao[1], Ryutei Inui[2], Yoshihisa Akamatsu[1], Masuji Goto[3], Hideyuki Doi[4], Shunsuke Matsuoka[4]

[1]Graduate School of Science and Technology for Innovation, Yamaguchi University, 2-16-1, Tokiwadai, Ube, Yamaguchi, 755-8611, Japan.

[2]Faculty of Socio-Environmental Studies, Fukuoka Institute of Technology, 3-30-1, Wajiro-higashi, Higashi-Ku, Fukuoka, 811-0295, Japan.

[3]Research and Development Center of Nippon Koei Co. Ltd., 2304, Inarihara, Tsukuba, Ibaraki, 300-1259, Japan.

[4]Graduate School of Simulation Studies, University of Hyogo, 7-1-28, Minatojima-minamimachi, Chuo-ku, Kobe, 650-0047, Japan.

Correspondence

Ryohei Nakao, Graduate School of Science and Technology for Innovation, Yamaguchi University, 2-16-1, Tokiwadai, Ube, Yamaguchi, 755-8611, Japan.

E-mail: lineck92@yahoo.co.jp

Abstract

Environmental DNA (eDNA) analysis is a method of detecting DNA from environmental samples, and it is used as a biomonitoring tool. In recent studies, Illumina MiSeq has been the most extensively used tool for eDNA metabarcoding, one of the eDNA analysis approaches. The Illumina iSeq 100 (hereafter, iSeq) is one of the numerous high-throughput sequencers (HTS). It has a relatively simple workflow and is potentially more affordable than other sequencers for deployment in HTS environments. However, to date, only a few studies have adopted iSeq, and its utility in eDNA metabarcoding has still not been investigated comprehensively. In the present study, we applied fish eDNA metabarcoding to river and lake environmental samples using iSeq and MiSeq approaches. We also assessed differences in fish species detectability among iSeq, MiSeq, and conventional approaches. Twenty-seven river and 13 lake samples were amplified using MiFish primers and sequenced with iSeq and MiSeq, respectively. The iSeq and MiSeq metabarcoding achieved high detectability for fish taxa in the ecosystems. Species numbers and compositions in each river detected using iSeq were almost consistent with those of MiSeq, indicating detectability of both techniques was comparable. The comparison of the species compositions of the two HTSs with those of conventional methods showed that the common species between each HTS and the conventional methods were exactly similar. According to the results, if the same amplicon

37 library were used for sequencing, there would be negligible detectability differences between

38 iSeq and MiSeq based on eDNA metabarcoding.

39 **Keywords:** environmental DNA, metabarcoding, MiFish, iSeq 100, MiSeq

40

41

Introduction

Environmental DNA (eDNA) analysis methods can detect the DNA fragments shed from macro-organisms in environmental samples (water, sediment, or air). The eDNA methods can provide information on the distribution, abundance, seasonal change, and migration of species (Takahara et al., 2012; Doi et al., 2017; Stockle et al., 2017; Wu et al., 2019) and can facilitate biodiversity monitoring activities (Tabarlet et al., 2012; Barns & Turner, 2016). Furthermore, eDNA methods use environmental samples for DNA detection, permitting non-invasive and non-destructive surveys in target species, habitats, and ecosystems (Jerde et al., 2011). One of the eDNA methods, eDNA metabarcoding, can detect multiple species from an environmental sample simultaneously using high-throughput sequencing (HTS) (Thomsen et al., 2012; Miya et al., 2015; Yamamoto et al., 2017; Deiner et al., 2017). eDNA metabarcoding has been applied to detect both vertebrate (Port et al., 2016; Closek et al., 2019; Ushio et al., 2017; 2018, Miya et al., 2015) and invertebrate (Komai et al., 2019; Thomsen & Sigsgaard 2019; Mychek-Londer et al., 2019) compositions in communities. In addition, eDNA metabarcoding can detect higher levels of species diversity (Olds et al., 2016; Shaw et al., 2016) or complementary species diversity compared to conventional monitoring methods (Yamamoto et al., 2017; Hänfling et al., 2016).

Illumina MiSeq is the mainstream HTS platform for the detection species composition

using eDNA metabarcoding (Miya et al., 2015; Port et al., 2016; Deiner et al., 2016; Stoeckle et al., 2017; Komai et al., 2019). In early 2019, the Illumina iSeq 100 (iSeq), which is a simpler and more affordable HTS system, was released (Illumina, 2019a). The differences between iSeq and MiSeq are as follows. First, iSeq and MiSeq have different base-calling systems, with iSeq sequencing in a single-color signal, while MiSeq has four-color signals corresponding to each sequence (Illumina, 2019a). Secondly, the structure of the flow cell for loading the sequencing library varies between iSeq and MiSeq, with iSeq being a patterned flow cell, while MiSeq is a random flow cell. Finally, the sequencing workflow between the iSeq and MiSeq is different. The iSeq is simpler and requires less preparation using the cartridge, while the MiSeq requires relatively more preparation steps, including pre- and post-run wash of the flow channel. Such differences between the two sequencing approaches could influence species detectability and the sequencing quality during eDNA metabarcoding. However, to the best of our knowledge, no eDNA studies have been conducted using iSeq, and no comparative studies between iSeq and MiSeq have been performed.

In the present study, we compared the sequence quality and taxonomic assignment of eDNA metabarcoding between iSeq and MiSeq using the fish-specific primer set (MiFish primers; Miya et al. 2015). In addition, to evaluate the capacity of iSeq and MiSeq to detect species based on eDNA metabarcoding, we compared fish species compositions between

78 eDNA metabarcoding (iSeq and MiSeq) and conventional methods.

79

80 **Materials and Methods**

81 *Sample collection and filtration*

82 We used 40 water samples for eDNA metabarcoding from 27 sites in 9 rivers and 13 lakes
83 in Japan from 2016 to 2018 (Fig. 1). Sampling ID and detailed information for each site are
84 listed in Supplementary Table S1. In the river water sampling, 1-L water samples were
85 collected from the surface of at the shore of each river using bleached plastic bottles. In the
86 field, a 1-ml Benzalkonium chloride solution (BAC, Osuban S, Nihon Pharmaceutical, Tokyo,
87 Japan; Yamanaka et al., 2017) was added to each water sample to suppress eDNA
88 degeneration before filtering the water samples. We did not include field negative control
89 samples in the HTS library, considering the aim of the presents study. The lake samples were
90 provided by Doi et al. (2020) as DNA extracted samples. In the lake samples, 1-L water
91 samples were collected from the surface at shore sites at each lake. The samples were then
92 transported to the laboratory in a cooler at 4°C. Each of the 1-L water samples was filtered
93 through GF/F glass fiber filter (normal pore size = 0.7 μm; diameter = 47 mm; GE Healthcare
94 Japan Corporation, Tokyo, Japan) and divided into two parts (maximum 500-ml water per 1
95 GF/F filter). To prevent cross-contamination among the water samples, the filter funnels, and

96 the measuring cups were bleached after filtration. All filtered samples were stored at –20□ in
97 the freezer until the DNA extraction step.
98
99 *DNA extraction and library preparation*
100 The total eDNA was extracted from each filtered sample using the DNeasy Blood and
101 Tissue Kit (QIAGEN, Hilden, Germany). Extraction methods were according to Uchii et al.
102 (2016), with a few modifications. A filtered sample was placed in the upper part of a Salivette
103 tube and 440 μL of a solution containing 400 μL Buffer AL and 40 μL Proteinase K added.
104 The tube with the filtered sample was incubated at 56°C for 30 min. Afterward, the tube was
105 centrifuged at 5000 ×*g* for 3 min, and the solution at the bottom part of the tube was collected.
106 To increase eDNA yield, 220-μL Tris-EDTA (TE) buffer was added to the filtered sample and
107 the sample re-centrifuged at 5000 ×*g* for 1 min. Subsequently, 400 μL of ethanol was added to
108 the collected solution, and the mixture was transferred to a spin column. Afterward, the total
109 eDNA was eluted in 100-μL buffer AE according to the manufacturer's instructions. All
110 eDNA samples were stored at -20°C until the library preparation step.
111 In the present study, we used a universal primer set "MiFish" (Miya et al. 2015) for eDNA
112 metabarcoding. The amplicon library was prepared according to the following protocols. In
113 the first PCR, the total reaction volume was 12 μL, containing 6.0μL 2× KOD buffer, 2.4 μL

dNTPs, 0.2 μL KOD FX Neo (TOYOBO, Osaka, Japan), 0.35 μL MiFish-U-F (5'-*ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNN*GTCGGTAAAACTCGTGCCAGC -3'), MiFish-U-R (5'-*GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNN*CATAGTGGGGTATCTAATCCCAGTTTG -3'), MiFish-E-F (5'-*ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNN*RGTTGGTAAATCTCGTGCCAGC -3') and MiFish-E-R (5'-*GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNN*GCATAGTGGGGTATCTAATCCTAGTTTG -3') primers with Illumina sequencing primer region and 6-mer Ns, and 2 μL template DNA. The thermocycling conditions were 94□ for 2 min, 35 cycles of 98□ for 10 s, 65□ for 30 s, 68□ for 30 s, and 68□ for 5 min. The first PCR was repeated four times for each sample, and the replicated samples were pooled as a single first PCR product for use in the subsequent step. The pooled first PCR products were purified using the Solid Phase Reversible Immobilization select Kit (AMPure XP; BECKMAN COULTER Life Sciences, Indianapolis, IN, USA) according to the manufacturer's instructions. The DNA concentrations of purified first PCR products were measured using a Qubit dsDNA HS assay kit and a Qubit 3.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). All purified first PCR products were diluted to 0.1 ng/μL with H_2O, and the diluted samples were used as templates

for the second PCR. In the first PCR step, the PCR negative controls (four replicates) were included in each experiment. A total of three PCR negative controls were included in the library (PCR Blank 1–3 samples in Supplementary Table S1, S2, S4, and S5).

The second PCR was performed to add HTS adapter sequences with 8-bp dual indices. The total reaction volume was 12 μL, containing 6.0 μL 2× KAPA HiFi HotStart ReadyMix, 1.4 μL forward and reverse primer (2.5 μM), 1 μL purified first PCR product, and 2.2 μL H_2O. The thermocycling conditions were 95□ for 3 min, 12 cycles of 98□ for 20 s, 72□ for 15 s, and 72□ for 5 min.

Each Indexed second PCR product was pooled in the equivalent volume, and 25 μL of the pooled libraries were loaded on a 2% E-Gel SizeSelect agarose gels (Thermo Fisher Scientific), and a target library size (ca. 370 bp) was collected. The quality of the amplicon library was checked using an Agilent 2100 Bioanalyzer and Agilent 2100 Expert (Agilent Technologies Inc., Santa Clara, CA, USA), and the DNA concentrations of the amplicon library were measured using Qubit dsDNA HS assay Kit using a Qubit 3.0 fluorometer.

High-throughput sequencing

Amplicon library was sequenced using iSeq and MiSeq platforms (Illumina, San Diego, CA, USA). To normalize the percentage of pass-filtered read numbers, the sequencing runs

using the same libraries were performed using iSeq i1 Reagent and MiSeq Reagent Kit v2 Micro. Both sequencing was performed with 8 million pair-end reads and 2×150 bp read lengths. Each library was spiked with approximately 20% PhiX control (PhiX Control Kit v3, Illumina, San Diego, CA, USA) before sequencing runs according to the recommendation of Illumina. The wells of cartridges in the iSeq run were loaded with 20 μL of 50 pM library pool, and sequencing performed at Yamaguchi University, Yamaguchi, Japan. The wells of cartridges for MiSeq runs were loaded with 600 μL of 16 pM library pool, and sequencing performed at Illumina laboratories (Minato-ku, Tokyo, Japan). Subsequently, the sequencing dataset outputs from iSeq and MiSeq were subjected to pre-processing and taxonomic assignments. All sequence data are registered in the DNA Data Bank of Japan (DDBJ) Sequence Read Archive (DRA, Accession number: DRA10593).

Pre-processing and taxonomic assignments

We used the USEARCH v11.0667 (Edger, 2010) for all data pre-processing activities and taxonomic assignment of the HTS datasets obtained from the iSeq and MiSeq platforms (Komai et al., 2019; Takeuchi et al., 2019). First, pair-end reads (R1 and R2 reads) generated from iSeq and MiSeq platforms were assembled using the “fastq_mergepairs” command (overlapped reads are not written). In the process, the low-quality tail reads with a cut-off

threshold at a Phred score of 2, and the paired reads with too many mismatches (> 5 positions) in the aligned regions were discarded. Secondly, the primer sequences were removed from the merged reads using the "fastx_truncate" command. Afterward, read quality filtering was performed using the "fastq_filter" command with thresholds of max expected error > 1.0 and > 50 bp read length. The pre-processed reads were dereplicated using the "fastx_uniques" command, and the chimeric reads and less than 10 reads were removed from all samples as the potential sequence errors. Finally, an error-correction of amplicon reads, which checks and discards the PCR errors and chimeric reads, was performed using the "unoise3" command in the unoise3 algorithm (Edger & Flyvbjerg, 2016). Before the taxonomic assignment, the processed reads from the above steps were subjected to sequence similarity search using the "usearch_global" command against reference databases of fish species that had been established previously (MiFish local database v34). The sequence similarity and cut off E-value were 99% and 10^{-5}, respectively. If there was only one species with ≧ 99 % similarity, the sequence was assigned to the top-hit species. Conversely, sequences assigned to two or more species in the ≧ 99 % similarity were merged as species complex and listed in the synonym group. Generally, the species complexes were assigned to the genus level (e.g. Asian crucian carp *Carassius* spp.). Species that were unlikely to inhabit Japan were excluded from the candidate list of species complexes. For example, the sequence of one of bitterling

186 *Acheilignathus macropterus* included other different two species, *Acheilognathus barbatus*
187 and *Acheilignathus chankaensis*, as the species of the 2nd hit candidate; however, the two
188 species are not currently found in Japan. Therefore, the sequence was assigned to *A.*
189 *macropterus* in the present study. Because we used only freshwater fish species, we removed
190 the OTUs assigned to marine and brackish fishes from each sample. Finally, sequence reads
191 of each fish species were arranged into the matrix, with the rows and columns representing
192 the number of sites and fish species (or genus), respectively.
193
194 *Comparing sequence quality and fish fauna between iSeq and MiSeq*
195 We evaluated sequence quality based on 1) the percentage of clustering passing filter (%
196 PF) and 2) sequencing quality score ≧ % Q30 (Read1 and Read2) between iSeq and MiSeq
197 platforms. The % PF value is an indicator of signal purity for each cluster. When an HTS
198 library with higher DNA concentrations is loaded, flow cells typically have high numbers of
199 overlapped clusters and are over clustered. The condition leads to poor template generation,
200 which decreases the % PF value (Illumina, 2019b). In the present study, a >80 % PF value
201 was set as the threshold of sequence quality in iSeq and MiSeq runs. Sequence quality scores
202 (Q score) measure the probability that a base is called incorrectly. With sequencing by
203 synthesis (SBS) technology, each base in a read is assigned a quality score by a Phred-like

204 algorithm, similar to that originally developed for Sanger sequencing experiments (Ewing et
205 al. 1998). Higher Q scores indicate lower probability of sequencing error, and lower Q scores
206 indicate probability of false-positive variant calls resulting in inaccurate conclusions. In the
207 present study, the % Q30 values (error rate = 0.001 %) were used for the comparison of
208 sequence quality between iSeq and MiSeq. The parameters were collected directly using
209 Illumina BaseSpace Sequence Hub.

210 We also evaluated changes in sequence reads in pre-processing steps between iSeq and
211 MiSeq platforms. Sequence reads were assessed based 1) merge pairs, 2) quality filtering, and
212 3) denoising. Sequence R1 and R2 reads generated from HTS were first merged into single
213 sequence reads using the “fastq_mergepairs” command; however, non-merged sequence reads
214 were discarded in this step. In the marge pairs step, the changes in sequence reads from raw
215 sequence reads to merged sequence reads were calculated. In the quality filtering step, the
216 changes in sequence reads from merged sequence reads to sequence reads through the quality
217 filtering step were calculated. Finally, in denoising steps, the changes in sequence reads from
218 sequence reads through the quality filtering step to sequence reads through the denoising step
219 were calculated. The calculated numbers of sequence reads are listed in Supplementary Table
220 S2 and S3 in series.

221 Before the comparison of fish fauna, rarefaction curves were illustrated for each sample in

both iSeq and MiSeq to confirm that the sequencing depth adequately covered the species composition using the "rrarefy" function of the "vegan" package ver. 2.5-6 (https://github.com/vegandevs/vegan) in R ver. 3.6.2 (R Core Team 2019). We evaluated the species detection capacities of iSeq and MiSeq based on environmental DNA metabarcoding by comparing the differences in species composition and the number of species detected between the two sequencers. To test the differences in the number of species between iSeq and MiSeq, we performed paired t-tests based on the number of species at each sampling site.

Comparison of fish species detectability between eDNA metabarcoding and conventional methods

We evaluated species detectability between HTS platforms by comparing the fish species lists of HTS platforms with lists from conventional methods. Five sampling sites were selected from Kyushu and Chugoku districts (R23–27 in Fig.1). The fish fauna data obtained by conventional methods were based on the results of a previous study (Doi et al., 2020). The conventional surveys were conducted through hand-net sampling and visual observation by snorkeling (see Doi et al., 2020 for the detailed methods). The count data of each species were replaced with the incidence-based datasets (presence or absence) for comparing with the eDNA metabarcoding datasets. Fish sequence reads of each sampling site obtained by eDNA

metabarcoding were also replaced with the incidence-based data.

To test the detectability of species observed with detectability under conventional methods, the fish species compositions in five rivers were compared between the eDNA metabarcoding (iSeq and MiSeq) and the conventional methods. Using fish faunal data obtained from iSeq, MiSeq, and conventional methods, non-metric multidimensional scaling (NMDS) was performed in 1000 separate runs. For NMDS, the dissimilarity of the fish fauna was calculated based on the incidence-based Jaccard indices. Permutational multivariate analysis of variance (PERMANOVA) was performed with 1000 permutations to assess the differences in fish fauna among the methods and sites. Furthermore, to evaluate variance across sites among methods, the permutational analyses of multivariate dispersions (PERMDISP) was also performed with 1000 permutations. The “metaMDS”, “adonis”, and “betadisper” functions in the “vegan” package ver. 2.5-6 were used for NMDS ordination, PERMNOVA, and PERMDISP, respectively. To visualize the number of species in each method and the number of common species between methods, Venn diagrams were illustrated for each river using the “VennDiagram” package ver. 1.6.2 (Chen & Boutros, 2011) in R. To assess differences in the number of species among methods at each river, the repeated measures analysis of variance (ANOVA) was performed among iSeq, MiSeq, and conventional methods. If a significant difference was found in repeated measures ANOVA, the Tukey-Kramer

multiple comparison test was performed to analyze differences among methods.

Results

HTS using iSeq and MiSeq

In iSeq, the passing Filter (% PF), ≧ % Q30 (Read 1), and ≧ % Q30 (Read 2) were 80.80, 96.80, and 95.30%, respectively (Fig S1). In MiSeq, Passing Filter (% PF), ≧ % Q30 (Read 1), and ≧ % Q30 (Read 2) were 95.05, 97.30, and 96.48%, respectively (Fig. S1). The % PF value of iSeq was slightly lower than that of MiSeq; however, this was due to differences in the %PF calculation methods (Illumina, 2017). The Q30 values of Read 1 and 2 were not remarkably different between iSeq and MiSeq.

In total, 3,325,177 and 2,154,367 read sequences were determined using iSeq and MiSeq, respectively. The sequencing depth and sequence per sample were consistent between iSeq and MiSeq in each processing step (Supplementary Table 2 and 3, Fig. S2–S5). In the present study, the differences in the numbers of sequence reads among samples were confirmed in the two sequencers. The rarefaction curves were saturated in all iSeq and MiSeq samples (Supplementary Fig. S6 and S7). Therefore, the rarefying of read numbers among the samples was not performed.

Taxonomic assignment of sequence read

Sequence reads obtained from iSeq and MiSeq in the river and lake samples after pretreatment are listed in Table 1. After the denoising step, most sequence reads in iSeq and MiSeq could be taxonomically assigned to fish species. In total, 154 and 168 representative iSeq and MiSeq sequences, respectively, were assigned to fish species (Assigned Reads in Table 1). After the taxonomic assignment, 102 and 101 freshwater fish sequences with ≧ 99 % identity were retained in iSeq and MiSeq, respectively. Low numbers of sequence reads of two species (*Rhinogobius* sp. and *Tridentiger* sp.) were retained from negative control samples (NC41–43) of iSeq and MiSeq (Supplementary Table S4 and S5), which could be due to cross-contamination among samples. The species were commonly observed throughout the samples, and the source of contamination could not be identified. Therefore, we did not assess cross-contamination across the samples using negative control samples.

Based on the sequencing results, retained iSeq and MiSeq sequences were assigned to 69 and 68 freshwater fish species, subspecies, or genera, respectively (Supplementary Table S4 and S5). The genera that were assigned to multiple candidate species are listed in Supplementary Table S6. Species compositions in all samples were almost similar between iSeq and MiSeq. However, Japanese striped loach *Cobitis biwae* typeB was only detected in L11 by iSeq sequencing. The number of species per sample ranged from 4 to 27, in iSeq, and

294 4 to 26, in MiSeq. The paired *t*-test revealed significant differences in the number of species
295 between iSeq and MiSeq at all sites (t = 5.6488, $p < 0.001$). The differences in the number of
296 species between iSeq and MiSeq ranged from 0 to 4. However, species that were detected by
297 iSeq only often had low read counts (11 to 32 reads per species, Supplementary Table S). For
298 example, 15 and 12 species were detected in the iSeq and MiSeq at site 5; however, the
299 numbers of reads of the three species detected only in the iSeq were 11 (*Lepomis*
300 *macrochirus*), 11 (*Oncorhynchus masou* subsp.), and 28 (*Tachysurus nudiceps*), respectively
301 (Supplementary Table S7).
302
303 *Comparison of fish species composition between eDNA methods and conventional methods*
304 R23–27 river samples were used for the comparisons between the two sequencers and the
305 conventional methods (Fig. 2). A total of 30, 30, and 29 species were detected in R23–27
306 using the iSeq, MiSeq, and conventional methods, respectively (Supplementary Table S8).
307 The number of species detected by iSeq, MiSeq, and conventional methods were 14–19, 14-
308 –19, and 8–16 in each site in R23–27, respectively (Supplementary Table S8). The number of
309 species detected by iSeq and MiSeq was higher than that detected the conventional methods
310 in all survey sites (deep green in Fig. 2). High numbers of species were commonly detected
311 by iSeq, MiSeq, and conventional methods. In addition, no species that was common between

312 iSeq or MiSeq and conventional methods was identified. Based on the results of repeated
313 measured ANOVA, there were significant differences among methods ($F = 9.061$, $p = 0.0088$),
314 among sites ($F = 4.827$, $p = 0.0282$ in Supplementary Table S9). However, based on the
315 results of the subsequent Tukey-Kramer test, there were no significant differences among the
316 three methods ($p > 0.05$, Supplementary Table S10).
317 To assess differences in species composition among the three methods, we used NMDS
318 ordination (Fig. 3). In NMDS ordination, the fish communities of iSeq and MiSeq at each site
319 were plotted at almost similar coordinates. Based on the results of PERMANOVA and
320 PERMDISP analyses, there were no significant differences in species composition evaluated
321 by iSeq, MiSeq, and conventional methods (PERMANOVA, $p = 0.48450$ and PERMDISP, p
322 $= 0.18510$ in Table 2). Therefore, there were no significant differences in assemblage structure
323 among the three methods.
324

325 **Discussion**

326 Here, we observed that iSeq and MiSeq could obtain similar sequence qualities and fish
327 fauna in eDNA metabarcoding. As mentioned previously, the SBS chemistry of iSeq is
328 distinct from that of MiSeq in sequencing workflow and base-calling, as well as distinct
329 cartridge structures and flow cell mechanisms. However, such differences would not influence

sequence quality in eDNA metabarcoding activities. In addition, we obtained high-quality sequence data and a few sequencing errors were observed in both iSeq and MiSeq. Furthermore, the rate of change in the number of sequence reads in each process of the analytical pipeline was almost similar between iSeq and MiSeq.

The freshwater fish fauna in the rivers and lakes identified by MiFish metabarcoding exhibited minimal differences in the number of species or species composition between iSeq and MiSeq. Generally, the number of species per sample in iSeq was higher than that in MiSeq due to the differences in the obtained sequence reads between the two sequencers. Species detected only in the iSeq had relatively low numbers of reads (11–32 reads, Supplementary Table S10). Since the same sequence library was used in the iSeq and MiSeq, the reads of such species could have been detected in the MiSeq by increasing the sequencing depth. Such results indicate that the number of species that can be detected by iSeq and MiSeq are approximately similar if the same sequence library is used.

Based on the results of the comparisons of the numbers of species among the three methods, the number of species detected by eDNA metabarcoding was higher than those of conventional methods in each river, including many species that were only detected by HTS (Fig. 2). Conversely, some species were only observed in each river using the conventional methods. In previous studies, eDNA metabarcoding has exhibited "higher diversity" or

348 "complementary" results when compared to the results of conventional methods (Olds et al.
349 2016, Yamamoto et al. 2017, Deiner et al. 2017). Such complementary results in iSeq and
350 MiSeq support the findings of previous studies. Nevertheless, there were no statistically
351 significant differences in the numbers of observed species among the three methods.
352 Studies of eDNA metabarcoding have been increasing annually, and it is attracting the
353 attention of researchers and stakeholders as a time- and cost-efficient method for detecting
354 species composition and diversity (Tsuji et al. 2019). Traditionally, MiSeq has been used
355 extensively for eDNA metabarcoding (Miya et al. 2015, Port et al. 2016, Deiner et al. 2016,
356 Stoeckle et al. 2017, Komai et al. 2019). Our results indicate that iSeq can be used for eDNA
357 metabarcoding and has similar levels of species detectability to MiSeq. Despite the
358 differences in sequencing depths, iSeq and MiSeq revealed similar fish fauna at each site. The
359 results suggest that fish fauna from iSeq and MiSeq can be compared directly if library
360 preparation is performed using similar processes. MiSeq has an advantage over iSeq because
361 it can be used for a wide range of taxa and purposes due to the availability of different kits
362 with different lead lengths (max. 2 × 300 bp) and the number of leads (max. pair-end 25
363 million leads). However, MiSeq has to be adhere to a few complex procedures during
364 sequencing, such as the need to clean the channels for decontamination between sequencing
365 runs. Such procedural complexity may lead to cross-contamination between library samples

366 or between sequencing runs. In contrast, the iSeq has less working procedures than MiSeq
367 because of the use of cartridges and no need to clean the flow path. Therefore, iSeq may have
368 lower cross-contamination risk between sequencing runs than MiSeq. Future research should
369 evaluate cross-contamination risk between the two technologies.
370

371 **Acknowledgments**

372 We thank T. Kono and K. Yamaguchi for assisting in the field sampling. We also thank H.
373 Fujii for assistance in the laboratory work and S. Miyazono for providing constructive
374 comments on an earlier version of the manuscript. We thank K. Naka for assisting in iSeq
375 running, and T. Kobayashi for running the MiSeq platform at the Illumina laboratory and for
376 manuscript corrections. We also thank Illumina K. K. (Japan) for providing us with a
377 sequencing cartridge kit for the iSeq run.
378

379 **References**

380 Barnes, M.A. & Turner, C. R. (2016). The ecology of environmental DNA and implications
381 for conservation genetics. *Conservation Genetics*, *17*, 1–17.
382 https://doi.org/10.1007/s10592-015-0775-4
383
384 Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D.
385 M. (2014). Environmental conditions influence eDNA persistence in aquatic systems.
386 *Environmental Science & Technology*, *48*, 1819–1827.
387 https://doi.org/10.1021/es404734p
388
389 Bylemans, J., Furlan, E. M., Hardy, C. M., McGuffie, P., Lintermans, M., & Gleeson, D. M.

(2017). An environmental DNA-based method for monitoring spawning activity: a case study, using the endangered Macquarie perch (*Macquaria australasica*). *Methods in Ecology and Evolution*, *8*, 646–655. https://doi.org/10.1111/2041-210X.12709

Chen, H., & Boutros, P.C. (2011). Venn Diagram: a package for the generation of highly-customizable Venn & Euler diagrams in R. *BMC Bioinformatics*, *12*, 35. https://doi.org/10.1186/1471-2105-12-35

Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M., Bograd, S. J., Hazen, E. L., Field, J. C., & Boehm, A. B. (2019). Marine vertebrate biodiversity and distribution within the Central California Current using environmental DNA (eDNA) Metabarcoding and Ecosystem Surveys. *Frontiers in Marine Science*, *16*. https://doi.org/10.3389/fmars.2019.00732

Deiner, K., Bik, H.M., Mächler, E., Seymour, M, Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M, Vere, N., Pfrender, M.E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Doi, H., Inui R., Akamatsu, Y., Kannno, 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. https://doi.org/10.1111/fwb.12846

Doi, H., Inui R., Matsuoka S., Akamatsu Y., Goto M., & Kono T. (2020). Evaluation of biodiversity metrics through environmental DNA metabarcoding compared with visual and capture surveys in river fish community. *BioRxiv* preprint, https://doi.org/10.1101/617670.

Doi, H., Matsuoka S., Matsuzaki S. S., Nagano M., Sato H., Yamanaka H., Matsuhashi S., Yamamoto S., Minamoto T., Araki H., Ikeda K., Kato A., Kumei K., Maki M., Mitsuzuka T., Takahara T., Toki K., Ueda N., Watanabe T., Yamazoe K., & Miya M. The effects of ecosystem characteristics and species traits on species detection by eDNA metabarcoding compared to species records in Japanese lake fish communities. *Molecular Ecology*, submitted.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST.

Bioinformatics, *26*, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Edgar, R. C. and Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, 31, 3476–3482. https://doi.org/10.1093/bioinformatics/btv401

Ewing, B., Hillier L., Wendl M. C., & Green P. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Research*, 8: 175–185. https://doi.org/10.1101/gr.8.3.175

Hänfling, B., Handley, L. L., Read, D. S., Hahn, C., Li, J., Nichols, P., Blackman, R. C., Oliver, A., & Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, 25, 3101–3119. https://doi.org/10.1111/mec.13660

Hinlo, R., Furlan, E., Suitor, L., & Gleeson D. (2017). Environmental DNA monitoring and management of invasive fish: comparison of eDNA and fyke netting. *Management of Biological Invasions*, *8*: 89–100. https://doi.org/10.3391/mbi.2017.8.1.09

Illumina. (2017). *Calculating percent passing filter for patterned and nonpatterned flow cells-A comparison of methods for calculating percent passing filter on Illumina flow cells* Illumina. https://www.illumina.com/content/dam/illumina-marketing/documents/products/technotes/hiseq-x-percent-pf-technical-note-770-2014-043.pdf

Illumina. (2019a). *iSeq 100 sequencing system guide*. Illumina. https://support.illumina.com/downloads/iseq-100-system-guide-1000000036024.html

Illumina. (2019b). *Cluster optimization overview guide*. Illumina. https://support.illumina.com/downloads/cluster-optimization-overview-guide-1000000071511.html

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. https://doi.org/10.1111/j.1755-263X.2010.00158.x

Komai, T., Goto, O. R., Sado, T., & Miya M. (2019). Development a new set of PCR primers for eDNA metabarcoding decapod crustaceans, *Metabarcoding & Metagenomics*, *3*,

1–19. https://doi.org/10.3897/mbmg.3.33835

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. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, *2*, 150088. http://dx.doi.org/10.1098/rsos.150088

Mychek-Londer, J. G., Balasingham, K. D., & Heath, D. D. (2019). Using environmental DNA metabarcoding to map invasive and native invertebrates in two Great Lakes tributaries. *Environmental DNA*. https://doi.org/10.1002/edn3.56

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., Deiner, K., Mahon, A. R., Bruesekel, M. A., Shirey, P. D., Pfrender, M. E., Lodge, D. M., & Lamberti, G. A. (2016). Estimating species richness using environmental DNA. *Ecology and Evolution*, *6*(12), 4214–4226. https://doi.org/10.1002/ece3.2186

Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary P. R., Litvin, S. Y., Nickols, K. J. Yamahara, K. M. and Kelly R. P. (2016). Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, *25*, 527–541. https://doi.org/10.1111/mec.13481

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. https://doi.org/10.1016/j.biocon.2016.03.010

Stoeckle, M. Y., Soboleva L., & Charlop-Powers Z. (2017). Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PLoS ONE 12*(4): e0175186. https://doi.org/10.1371/journal. pone.0175186

Tabarlet, P., Coissac, E., Hajibabei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*, 1789–1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., & Kawabata Z. (2012). Estimation of fish biomass using Environmental DNA. *PLoS ONE*, *7*(4), e35868. https://doi.org/10.1371/journal.pone.0035868

Takeuchi, A., Sado, T., Gotoh, R.O., Watanabe, S., Tsukamoto, K., & Miya M. (2019). New PCR primers for metabarcoding environmental DNA from freshwater eels, genus *Anguilla*. *Scientific Reports*, *9*, 7977. https://doi.org/10.1038/s41598-019-44402-0

Thomsen P. F., Kielgast, K., Iversen, L. L., Møller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using Environmental DNA from seawater samples. *PLoS ONE*, *7*(8), e41732. https://doi.org/10.1371/journal.pone.0041732

Thomsen, P. F., & Sigsgaard, E. E. (2019). Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. *Ecological Evolution*, *9*, 1665–1679. https://doi.org/10.1002/ece3.4809

Tsuji, S., Takahara, T., Doi, H., Shibata, N., & Yamanaka, H. (2019). The detection of aquatic macroorganisms using environmental DNA analysis—A review of methods for collection, extraction, and detection. *Environmental DNA*, *1*, 99–108. https://doi.org/10.1002/edn3.21

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. https://doi.org/10.1111/1755-0998.12460

Ushio, M., Fukuda, H., Inoue, T., Kobayashi, M., Kishida, O., Sato, K., Murata, K., Nikaido, M., Sado, T., Sato, Y., Takeshita, M., Iwasaki, W., Yamanaka, H., Kondoh, M., & Miya, M. (2017). Environmental DNA enables detection of terrestrial mammals from forest pond water. *Molecular Ecology Resources*, *17*, e63–e75. https://doi.org/10.1111/1755-0998.12690

Ushio, M., Murata, K., Sado, T., Nishiumi, I., Takeshita, M., Iwasaki, W., & Miya M. (2018). Demonstration of the potential of environmental DNA as a tool for the detection of avian species. *Scientific Reports*, *8*, 4493. https://doi.org/10.1038/s41598-018-22817-5

Yamamoto, S., Masuda, M., Sato, Y., Sado, T., Araki, H., Kondoh6, M., Minamoto, T. and Miya, M. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports*, *7*, 40368. https://doi.org/10.1038/srep40368

Yamanaka, H., Minamoto, T., Matsuura, J., Sakurai, S., Tsuji, S., Motozawa, H., Hongo4, M., Sogo, Y., Kakimi, N., Teramura, I., Sugita, M., Baba, M., & Kondo, A. (2017). A simple

method for preserving environmental DNA in water samples at ambient temperature by addition of cationic surfactant. *Limnology*, *18*, 233–241. https://doi.org/10.1007/s10201-016-0508-5

Wu, Q., Kawano, K., Ishikawa, T., Sakata, M. K., Nakao, R., Hiraiwa, M. K. Tsuji, S., Yamanaka, H., & Minamoto T. (2019). Habitat selection and migration of the common shrimp, *Palaemon paucidens* in Lake Biwa, Japan—An eDNA-based study. *Environmental DNA*, *1*, 54–63. https://doi. org/10.1002/edn3.6

Data Accessibility

All data of the iSeq and MiSeq sequencing was shared in DRA (Accession number: DRA10593), and all used data, including all detected species by iSeq, MiSeq, and conventional methods, were shared in Supplemental Tables.

Author Contributions

Study Concept: Yoshihisa Akamatsu, Hideyuki Doi

Field Sampling: Ryutei Inui, Masuji Goto

Data Cleaning: Masuji Goto, Shunsuke Matsuoka, Hideyuki Doi

Data Analysis: Ryohei Nakao

Manuscript writing and Visualization: Ryohei Nakao

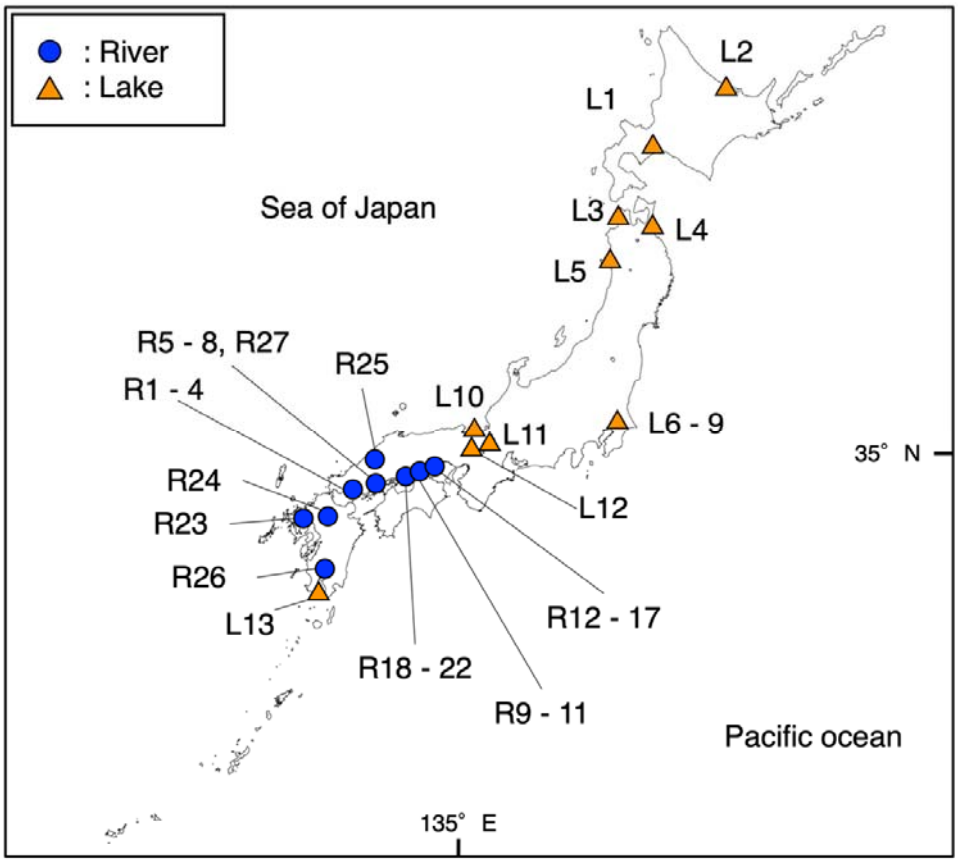

Figure 1. Sampling sites used in the present study. Blue circles and orange triangles show the locations of the river and lake samples, respectively. Detailed information on each site is listed in Supplementary Table S1.

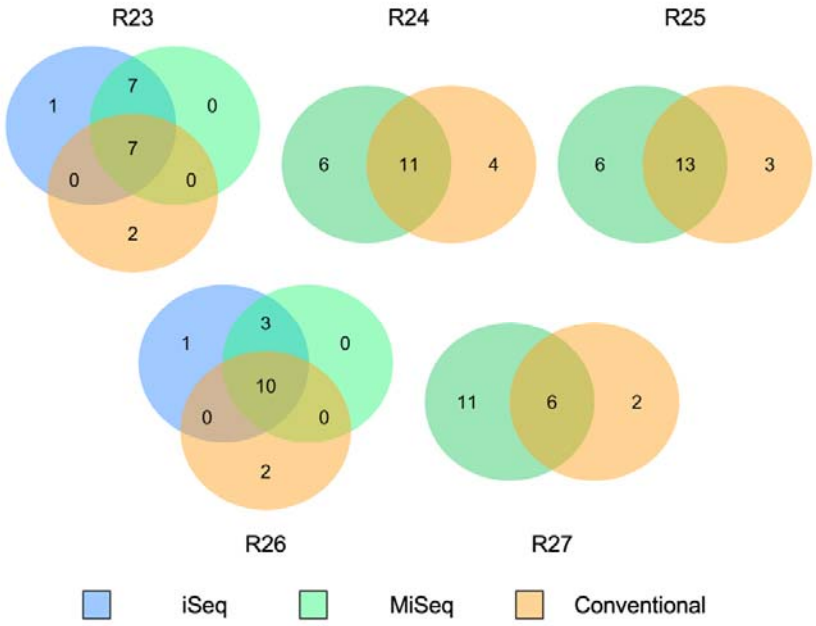

Figure 2. Venn diagrams for comparison of the number of species between high throughput sequencers and conventional methods in 5 river samples (R23-27 in Fig. 1). Blue, green, and orange circles indicate the species of eDNA of iSeq, eDNA of MiSeq, and conventional methods, respectively. Deep green circles in R24, 25, and 27 indicate the species of two sequencers because of fully-overlapping species composition between iSeq and MiSeq. Note that the size of each circle does not represent a difference in the number of species.

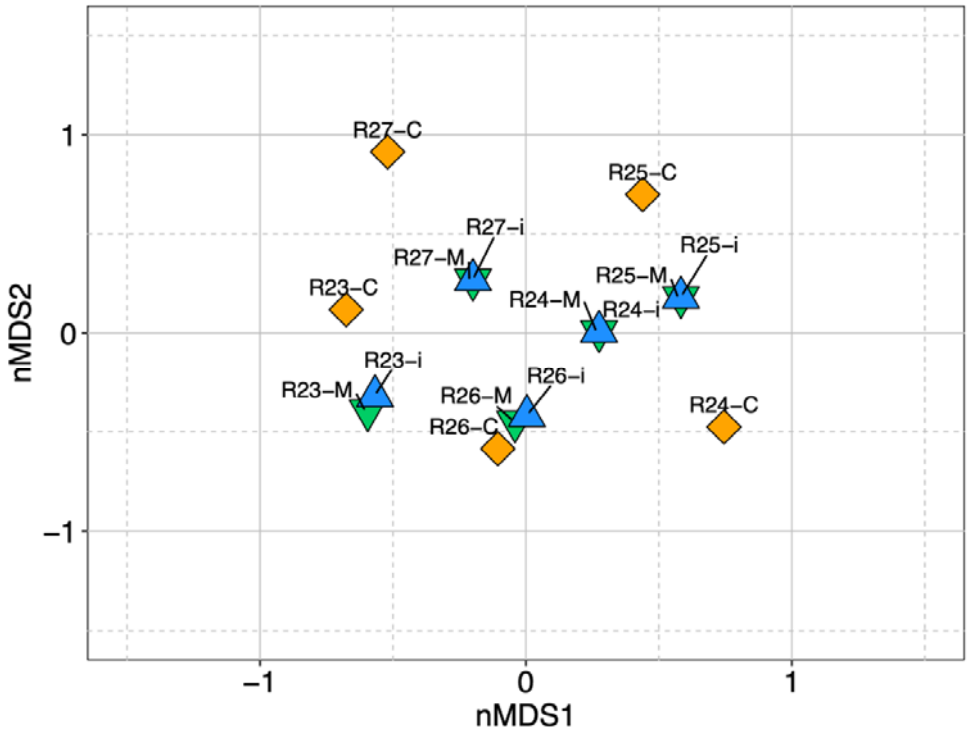

Figure 3. Non-metric multidimensional scaling (NMDS) plot using 5 river samples. Blue triangles, green inverted triangles, and diamond shapes indicate the points of iSeq, MiSeq, and conventional methods, respectively. Labels with these shapes also indicate iSeq (-i), MiSeq (-M), and conventional methods (-C).

Table 1 Ratio of sequence reads and Operational Taxonomic Unit (OTU) count after denoising

	iSeq		MiSeq	
Read Ratio				
Denoised Read	3010341	100.0%	1914412	100.0%
Assigned Read	2732599	90.8%	1687629	88.2%
OTU Ratio				
Total OTUs	154	100.0%	168	100.0%
OTU(Freshwater)	102	66.2%	101	60.1%
OTU (Marine and Brackish)	24	15.6%	22	13.1%
Under 99 % Identity	28	18.2%	45	26.8%

Table 2 Statistical results of Permutational Multivariate Analysis of Variance (PERMANOVA) and Permutational Multivariate Analysis of Dispersion (PERMDISP) for comparisons of species composition between iSeq and MiSeq

	Df†	Sum of Square	Mean Square	F value	R^2	Pr (>F)
PERMANOVA						
Methods‡	2	0.35707	0.17853	0.96117	0.13808	0.48450
Residuals	12	2.22894	0.18574		0.86192	
Total	14	2.58600			1.00000	
PERMDISP						
Methods‡	2	0.01696	0.00848	1.94810		0.18510
Residuals	12	0.05224	0.00435			

† Degree of Freedom (Df)

‡ Species detection methods (iSeq, MiSeq, and Conventional)

600