Estimation of biodiversity metrics by environmental DNA metabarcoding compared with visual and capture surveys of river fish communities

Hideyuki Doi[1,*,†], Ryutei Inui[2,3*,†], Shunsuke Matsuoka[1, †], Yoshihisa Akamatsu[2], Masuji Goto[2,4], and Takanori Kono[2,5]

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

[2] Graduate School of Science and Engineering, Yamaguchi University, 2-16-1 Tokiwadai, Ube, Yamaguchi, 755-8611, Japan

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

[4]Nippon Koei Research and Development Center, 2304 Inarihara, Tsukuba City, Ibaraki 300-1259, JAPAN

[5]Aqua Restoration Research Center, Public Works Research Institute National Research and Development Agency, Kawashima, Kasada-mathi, Kakamigahara-City, Gifu, 501-6021, Japan

[†]These authors equally contributed to this study.

[*]Corresponding authors:

Hideyuki Doi (hideyuki.doi@icloud.com)

Ryutei Inui (inuiryutei@gmail.com)

Running head: eDNA metabarcoding for fish diversity

Abstract

1. Information on alpha (local), beta (between habitats), and gamma (regional) diversity is fundamental to understanding biodiversity as well as the function and stability of community dynamics. Methods like environmental DNA (eDNA) metabarcoding are currently considered useful to investigate biodiversity.

2. We compared the performance of eDNA metabarcoding with visual and capture surveys for estimating alpha and gamma diversity of river fish communities, and nestedness and turnover in particular.

3. In five rivers across west Japan, by comparison to visual/capture surveys, eDNA metabarcoding detected more species in the study sites (i.e., alpha diversity). Consequently the overall number of species in the region (i.e., gamma diversity) was higher. In particular, the species found by visual/capture surveys were encompassed by those detected by eDNA metabarcoding.

4. Estimates of community diversity within rivers differed between survey methods. Although we found that the methods show similar levels of community nestedness and turnover within the rivers, visual/capture surveys showed more distinct community differences from upstream to downstream. Our results suggest that eDNA metabarcoding may be a suitable method for community assemblage analysis, especially for understanding regional community patterns, for fish monitoring in rivers.

Key words: eDNA, community, alpha and gamma diversity, nestedness, turnover

Introduction

The maintenance of biodiversity underpins the stability of ecosystem processes in constantly changing environments (Primack, 1993; Margules & Pressey, 2000; Pecl et al., 2017). Moreover, biodiversity loss affects ecosystem functions and services and, consequently, human society (Primack 1993; Margules & Pressey, 2000, Pecl et al. 2017). Ecologists have made efforts to conserve biodiversity based on essential biodiversity survey methods, especially for species richness and distribution (Primack, 1993; Margules & Pressey, 2000, Doi & Takahara, 2016, Pecl et al., 2017). Biodiversity can be evaluated in different levels: e.g., by estimating alpha (local), beta (between habitats), and gamma (regional) diversity and the variation of the community assemblages. To conserve local communities, ecologists incorporate these diversity metrics into management decisions (Primack 1993; Margules & Pressey, 2000, Socolar et al., 2016). For example, measuring variation among community assemblages can identify biodiversity loss and inform the placement of protected areas and the management of biological invasions and landscapes (Socolar et al., 2016). Thus, robust methods for monitoring biodiversity are fundamental for biodiversity and environmental management.

Environmental DNA (eDNA) analysis is considered a useful tool to investigate the distribution and richness of aquatic and terrestrial organisms (Takahara et al., 2012, 2013; Rees et al., 2014; Goldberg et al., 2015; Miya et al., 2015; Thomsen & Willerslev, 2015; Doi et al., 2017; Doi et al., 2019; Fujii et al., 2019). High-throughput sequencing derived from eDNA, called “eDNA metabarcoding”, is an

75 exceptionally useful and powerful tool for community biodiversity surveys (Taberlet et
76 al., 2012; Deiner et al., 2016, 2017; Sato et al., 2017; Bylemans et al., 2018; Fujii et al.,
77 2019). eDNA metabarcoding has recently been applied in fish community surveys, e.g.
78 Miya et al. (2015) designed and applied universal PCR primers (the MiFish primers) to
79 survey marine fish communities. To confirm the usefulness of eDNA metabarcoding for
80 community assessment, many studies have compared eDNA metabarcoding to a species
81 list generated by traditional surveys including visual and capture methods (Deiner et al.,
82 2016; Bylemans et al., 2018; Nakagawa et al., 2018; Yamamoto et al., 2018; Fujii et al.,
83 2019). However, comparisons of eDNA and traditional survey methods for estimating
84 regional (gamma) diversity or differences in community assemblage composition
85 among locations are more limited (but see Drummond et al., 2015; Deiner et al., 2016;
86 Staehr et al. 2016, Maechler et al. 2019).
87 Variation in community assemblage composition reflects two components,
88 nestedness and species turnover for considering regional structure of community
89 assemblage (Harrison et al., 1992; Baselga et al., 2007; Baselga, 2010). Nestedness
90 occurs when the community at the sites with fewer species are subsets of the
91 community at the sites with higher species richness (Wright & Reeves, 1992; Ulrich &
92 Gotelli, 2007) and generally reflects a non-random process of species loss (Gaston &
93 Blackburn, 2000). Contrastingly, species turnover implies the replacement of some
94 species by others because of environmental sorting or spatial/historical constraints
95 (Baselga, 2010). Statistical separation methods for nestedness and species turnover were
96 applied for evaluating variation of the community assemblages in various systems
97 (Baselga, 2010; Baselga et al., 2012). Moreover, Baselga's (2010) framework can be
98 applied to compare the performance among methods when evaluating beta diversity via

99 nestedness and species turnover. Nestedness and species turnover can also occur due to
100 differences in sampling methods, even when the surveys are conducted on the same
101 community. For example, differing biases in species detection between two methods
102 can result in apparent species turnover. Evaluating nestedness and species turnover
103 would allow us to reveal the potential bias of community analysis of eDNA
104 metabarcoding compared to traditional surveys. Therefore, we can quantitatively
105 compare the performance of eDNA metabarcoding and traditional surveys for
106 alpha/gamma diversity evaluation of biological communities and the variation of the
107 community assemblages among the study sites.
108 Using statistical methods, we tested the performance of eDNA metabarcoding
109 in five river systems in different regions with various fish species. We conducted eDNA
110 metabarcoding using universal MiFish primers that target fish and identified the fish by
111 visual snorkeling and hand-net capture surveys. We evaluated the performance of
112 eDNA metabarcoding by comparing the obtained fish community structure to that
113 evaluated by visual/capture survey with special regard to nestedness and species
114 turnover.
115
116 **Methods**
117
118 Site description
119 In 2016, we conducted field surveys in five river systems across Japan (river map in
120 Fig. S1): the Kyuragi River on October 10, the Koishiwara River on October 21, the
121 Yato River on October 25, the Hazuki River on November 2, and the Oze River on
122 November 6. The survey sites were set at a site at each of three river segments

123 (upstream, mid-stream, and downstream, the internal distances ranged from 4.5 to 25.8
124 km, Fig. S1) for each river. Each site was set so that the length in the up-down direction
125 was approximately 100 m with a riffle at the downstream end (e.g., Fig. S2).
126
127 Water collection for eDNA survey
128 In each site, we collected 1 L of surface water in bleached bottles at two points, in
129 flowing water near the downstream end of the site, and in stagnant or nearly stagnant
130 water near the shore (Fig. S2) immediately before visual and capture surveys. Prior to
131 sample collection, eDNA was removed from the bottles and filtering equipment using
132 10% commercial bleach (ca. 0.6% hypochlorus acid) and washing with DNA-free
133 distilled water. One milliliter of benzalkonium chloride (BAC, 10% w/v) was added per
134 liter of water sample to avoid a decrease in eDNA concentration in the samples
135 (Yamanaka et al., 2016). During transport, samples were stored in a cooler with ice
136 packs. The 'cooler blank' contained 1 L DNA-free water, which we brought to the field
137 and treated identically to the other water samples, except that it was not opened at the
138 field sites.
139
140 Visual observation and capture methods
141 After water sampling, the fish fauna survey was conducted by visual observation with
142 snorkel and collection with hand net. For visual observation, we observed and recorded
143 fish species by snorkeling in a 100-m transect (snorkeling by 1 person for 1 h, Fig. S2).
144 We observed at various micro habitats, including the riffle, pool, and shore bank from
145 the downstream end to upstream end. We also conducted a hand-net capture survey (1
146 person × 1 h) using a D-frame net (2 mm mesh, net opening: 0.16 m^2) in the various

147 habitats in the river, including the riffle, pool, and shore bank. Fishes were identified
148 according to Nakabo et al. (2013) at the survey site. We used the combined taxa list
149 from both traditional surveys to compare to that of eDNA metabarcoding. In order to
150 prevent contamination of eDNA samples, the investigator who collected and identified
151 the fish (i.e., handled the fish) and the investigator who sampled the water were
152 different.
153
154 eDNA collection, extraction and measurements
155 Collected water samples were vacuum-filtered into GF/F glass filters (47 mm diameter,
156 pore size: 0.7 μm, GE Healthcare, Little Chalfont, UK) in the laboratory within 24 h of
157 sampling. After filtration, all filters were stored at -20 °C before eDNA extraction. The
158 cooler blank was also processed in the same manner. A liter of Milli-Q water was used
159 as the filtering blank to monitor contamination during filtering in each site and during
160 subsequent DNA extraction.
161 To extract the DNA from the filters, we followed the methods described in
162 Uchii, Doi, & Minamoto (2016). We incubated the filter by submerging the mixed
163 buffer of 400 μL of Buffer AL in DNeasy Blood & Tissue Kit (Qiagen, Hilden,
164 Germany) and 40 μL of Proteinase K (Qiagen, Hilden, Germany), using a Salivette tube
165 (Sarstedt, Nümbrecht, Germany) at 56 °C for 30 min. The Salivette tube with filters was
166 centrifuged at 5000 × *g* for 5 min. Then, we added 220 μL of TE buffer (pH: 8.0) onto
167 the filter and again centrifuged at 5000 × *g* for 5 min. The DNA was purified using a
168 DNeasy Blood & Tissue Kit with extracted the DNA in 200 μL in Buffer AE. Samples
169 were stored at -20 °C until the 1st-PCR assay.
170

171 *Library preparation and MiSeq sequencing*

172 The detailed molecular methods are described in Fujii et al. (2019) with a two-step
173 PCR-procedure for Illumina MiSeq sequencing. Briefly, we performed 1st-PCR
174 withMiFish-U-F and MiFish-U-R primers (Miya et al., 2015), which were designed to
175 contain Illumina sequencing primer regions and 6-mer Ns;

176 Forward: 5'-*ACACTCTTTCCCTACACGACGCTCTTCCGATCT* NNNNNN
177 GTCGGTAAAACTCGTGCCAGC-3′ ,

178 Reverse: 5′ -*GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT* NNNNNN
179 CATAGTGGGGTATCTAATCCCAGTTTG-3′

180 The italicized and non-italicized letters represent MiSeq sequencing primers
181 and MiFish primers, respectively, and the six random bases (N) were used to enhance
182 cluster separation on the flow cells during initial base call calibrations on the MiSeq
183 (Miya et al. 2015, Doi et al. 2019).

184 We performed the 1st-PCR with a 12 μL reaction volume containing 1× PCR
185 Buffer for KOD FX Neo (Toyobo, Osaka, Japan), 0.4 mM dNTP mix, 0.24 U KOD FX
186 Neo polymerase, 0.3 μM of each primer, and 2 μL template. The thermocycling
187 conditions for this step were as follows: initial denaturation at 94 °C for 2 min, followed
188 by 35 cycles of denaturation at 98 °C for 10 s, annealing at 65 °C for 30 s, and
189 elongation at 68 °C for 30 s, followed by final elongation at 68 °C for 5 min. The first
190 PCRs were performed using eight replicates (Doi et al. 2019) and individual first PCR
191 replicates were pooled and purified using AMPure XP (Beckman Coulter, Brea CA,
192 USA) as templates for the 2nd-PCR. The Illumina sequencing adaptors and the eight bp
193 identifier indices (XXXXXXXX) were added to the subsequent PCR process using a
194 forward and reverse fusion primer:

195 Forward: 5′ -*AATGATACGGCGACCACCGAGATCTACA* XXXXXXXX

196 ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′

197 Reverse: 5′ -*CAAGCAGAAGACGGCATACGAGAT* XXXXXXXX

198 GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3′

199 The italicized and non-italicized letters represent MiSeq P5/P7 adapter and

200 sequencing primers, respectively. The 8X bases represent dual-index sequences inserted

201 to identify different samples (Hamady et al. 2008). We performed the 2nd-PCR with 12

202 cycles of a 12 μL reaction volume containing 1× KAPA HiFi HotStart ReadyMix, 0.3

203 μM of each primer, and 1.0 μL of the first PCR production. The thermocycling

204 conditions profile after an initial 3 min denaturation at 95 °C was as follows:

205 denaturation at 98 °C for 20 s, annealing, and extension combined at 72 °C (shuttle

206 PCR) for 15 s, with the final extension at the same temperature for 5 min. We

207 confirmed the positive bands of the targeted 1st-PCR amplicons by electrophoresis. The

208 2nd-PCR products were pooled in equal volumes and purified using AMPure XP.

209 The purified PCR products were loaded on a 2% E-Gel SizeSelect (Thermo

210 Fisher Scientific, Waltham, MA, USA) and the target size of the amplified DNA

211 (approximately 370 bp) was collected. The samples concentration and quality were

212 estimated by a Qubit dsDNA HS assay kit and a Qubit 2.0 (Thermo Fisher Scientific).

213 The amplicons were sequenced by 2 × 250 bp paired-end sequencing on the MiSeq

214 platform using the MiSeq v2 Reagent Kit. Note that the sequencing run contained a total

215 of 339 samples including 40 of our samples (30 samples plus five cooler and five filter

216 negative controls) and 299 samples from other research projects. The MiSeq sequencing

217 was conducted in the Department of Environmental Solution Technology, Faculty of

218 Science and Technology, Ryukoku University. All sequence data were deposited in
219 DNA Data Bank of Japan (DRA, Accession number: DRA008090).
220
221 Bioinformatic analysis for MiSeq sequencing
222 The detailed procedures used for bioinformatics analysis are described in Fujii et al.
223 (2019). First, low-quality tails were trimmed from each read and paired-end reads were
224 then merged. For the obtained 1,823,446 reads, primer sequences were removed and
225 identical sequences (i.e., 100% sequence similarity) were merged using UCLUST
226 (usearch 7.0.1001, Edgar, 2010). The sequences with 10 or more identical reads were
227 subjected to the downstream processes. To annotate the taxonomy, local BLASTN
228 search using BLAST 2.2.29 was conducted with the reference database of fish species
229 for processed reads (Miya et al., 2015). The top BLAST hit with a sequence identity ≥
230 97% was applied to species detection of each sequence. Note that the species were
231 mostly identified with a ≥ 99% match. From the BLAST results, we identified the
232 species using methods previously described (Sato et al., 2017). Also, we detect most of
233 the fish species known to occur in the rivers of the region, according to Kawanabe et al.
234 (2001).
235
236 Statistical analyses
237 All statistical analyses and graphics were conducted in R ver. 3.4.4 (R Core Team,
238 2018). All statistics were set at the significance level $\alpha = 0.05$. To compare between
239 eDNA metabarcoding and visual survey data, the taxonomic levels in the species list
240 from visual survey were compared to the lists from eDNA metabarcoding (Table S1,
241 S2) in reference to previous studies using the MiFish primer (Sato et al., 2017; Fujii et

242 al., 2019). Before statistical analysis, we confirmed that the sequencing depth was
243 sufficient to detect alpha diversity in the samples by “iNEXT” and “ggiNEXT”
244 functions in the “iNEXT” ver. 2.0.19 package (Chao et al. 2014, Fig. S3). We merged
245 the community data from two points, the stream near the downstream end and the shore,
246 to compare with visual/capture surveys.

247 We tested the differences in fish richness of sites, segments, and rivers
248 between both methods using generalized linear mixed models (GLMMs) with the
249 “lmer” function in the “lme4” ver. 1.1-21 package (Bates et al. 2015). In the GLMM
250 models, the method was treated as a fixed effect with Poisson distribution, and the
251 rivers and segments were treated as random effects.

252 The differences in community compositions were visualized using nonmetric
253 multidimensional scaling (NMDS) with 500 separate runs of real data. For NMDS, the
254 two data sets (eDNA metabarcoding and visual/capture survey) were combined into one
255 ordination (but showed divided into two panels by "facet_grid" function in ggplot2).
256 The community dissimilarity was calculated based on incidence-based Jaccard indices.
257 We evaluated the differences in community structures between methods and sites using
258 permutational multivariate analysis of variance (PERMANOVA). For PERMANOVA,
259 we used Jaccard and Raup-Crick dissimilarity indices and calculated the statistical
260 values with 999 permutations. The Raup-Crick index estimates the probability that two
261 samples have greater differences in species composition than would be expected by
262 chance draws from the regional species pool, and accounts for differences in alpha
263 diversity between samples. We used “metaMDS” and “adonis” functions in the “vegan”
264 ver. 2.5-6 package (https://github.com/vegandevs/vegan) for NMDS ordination and
265 PERMANOVA, respectively.

266 Indicator taxa analysis (Cáceres & Legendre, 2009) was performed to
267 determine which taxa had significantly different frequencies between eDNA and
268 visual/capture methods. The analysis was performed using the "signassoc" function in
269 the "indicspecies" ver. 1.7.8 package on the presence/absence data for the testing with
270 regarding the package description and Cáceres & Legendre (2009). The "signassoc"
271 function can calculate the index with both presence/absence and abundance data. We
272 used the presence/absence data with mode = 1 (group-based) and calculated the *P*-
273 values with 999 permutations after Sidak's correction of the multiple testing.
274 To compare community nestedness and turnover (Baselga, 2010; Baselga,
275 2012) as estimated by eDNA and visual/capture surveys, we used the "beta.pair"
276 function in the "betapart" ver. 1.5.1 package (Baselga and Orme 2012). Standardized
277 effect sizes (SESs) were calculated to show the degree of nestedness and turnover
278 structure. The significance was defined by deviation from zero, and the expectation of
279 random assembly (a null model) was estimated with 999 random sampling replicates.
280 The SES was defined as follows: $(\beta_{obs} - \beta_{null})/\beta_{sd}$, where β_{obs} is the observed beta
281 diversity (here, the variation of the community assemblages among the sites), β_{null} is the
282 mean of the null distribution of beta diversity, and β_{sd} is the standard deviation of the
283 null distribution. SES values greater than zero indicate statistically stronger nestedness
284 or turnover structure than expected under a random model of community assembly,
285 while negative values indicate weaker nestedness or turnover than expected. The
286 randomized community data were generated with independent swap algorithm (Gotelli,
287 2000) using "randomizeMatrix" function in the "picante" ver. 1.8 package (Kembel et
288 al. 2010). First, to evaluate the differences in nestedness and species turnover between
289 the survey methods (eDNA metabarcoding vs. visual/capture survey) at the same

290 segments, the SES of pairwise nestedness and turnover were calculated for each sample
291 pair within the river. Then, the fish community longitudinal nestedness and turnover
292 structure along with river flow (i.e., upstream to downstream) by each method was
293 evaluated with NODF and pairwise indices of nestedness. First, a nestedness metric
294 (NODF, Almeida-Neto et al., 2008) and their SES value were calculated with 999
295 randomizations using the "nestednodf" and the “oecosimu” function in the “vegan”
296 package. Then, we tested the differences in SES of pairwise indices between the survey
297 methods (eDNA metabarcoding vs. visual/capture survey) by generalized linear mixed
298 model (GLMM, with Gaussian distribution) with the “lmer” function in the “lme4”
299 package. In the GLMM models for the SES of pairwise indices, the survey method was
300 treated as a fixed effect, and the rivers-pairs and segment-pairs (i.e., three pairs in each
301 river) were treated as random effects.
302
303 **Results**
304
305 Overview
306 We detected 53 fish taxa, almost all identified to the species or genus level, by eDNA
307 metabarcoding in five rivers (Table S1, and S2) and visually observed 38 fish taxa in
308 total. MiSeq paired-end sequencing for the library (30 samples plus five cooler and five
309 filter negative controls) yielded a total of 1,601,816 reads (53,351 ± 17,639; mean ± S.
310 D. for each sample, Table S2). We confirmed very low reads from negative controls
311 (Table S2) with only one detection of a fish species, *Tridentiger* sp. in the blank of the
312 Yato River, probably because of the cross-contamination among the samples. The reads

313 of *Tridentiger* sp. in water samples (4783-9818 reads) were higher than in the blank
314 sample (1286 reads), and so we retained this species in the analysis.
315
316 Diversity indices between methods
317 We consistently detected more taxa at a site (i.e., higher alpha diversity) with eDNA
318 metabarcoding than with visual/capture surveys (Fig. 1; GLMM, t = –5.45, P =
319 0.000018). Estimated taxa richness generally increased from upstream to downstream
320 (Fig. 1, t = –5.85, P = 0.000004, random effect for variation among segments: 26.879 ±
321 5.185, SD). The random effect for "river" was not significantly different from zero (t =
322 1.737, P = 0.0942, random effect for variation among river: 3.896 ± 1.974).
323 We ordinated the differences in community structure between the two
324 methods by NMDS ordination by Jaccard index (Fig. 2) as well as Raup-Crick (Fig.
325 S4). The patterns of differences in the both Jaccard and Raup-Crick indices were
326 similar. The ordination showed the differences in community structure between the two
327 methods, probably due to the differences in species detection in the communities. The
328 PERMANOVA results with Jaccard and Raup-Crick indices for the ordination
329 suggested there were differences in community composition evaluated by each method,
330 eDNA metabarcoding and visual/capture survey (P < 0.012, Table S3). Moreover,
331 communities from the combined results of eDNA metabarcoding and visual/capture
332 survey were significantly different among rivers and segments (P < 0.029, Table S3).
333 We found different patterns in ordinated river sites for each method (Fig. 2 for Jaccard
334 index, Fig. S4 for Raup-Crick index).
335 The PERMANOVA results with Jaccard index determined that communities
336 were significantly different among the five rivers by eDNA metabarcoding (P = 0.001)

337 but not by visual/capture survey ($P = 0.12$, Table S4). Conversely, the communities
338 were significantly different among river segments (from downstream to upstream) by
339 visual/capture survey ($P = 0.011$) but not significantly different (albeit marginally) by
340 eDNA metabarcoding ($P = 0.061$, Table S4). The differences in PERMANOVA results
341 with Jaccard index suggested that differences in the variation of the community
342 assemblages among rivers across regions can be detected by eDNA metabarcoding but
343 not by visual/capture survey. While, the PERMANOVA results with Raup-Crick index
344 determined that communities were significantly different among the rivers by both
345 eDNA metabarcoding ($P = 0.001$) and visual/capture survey ($P = 0.018$, Table S4).
346 Indicator taxa analysis comparing the communities estimated by both methods
347 showed that eDNA metabarcoding more frequently detected several species than
348 visual/capture survey, including Japanese eel (*Anguilla japonica*), salmon (e.g.,
349 *Oncorhynchus masou)*, and Amur catfish (*Silurus asotus*) ($P < 0.05$, Table 1 for
350 statistically significant taxa, Table S5 for all taxa).
351
352 Nestedness and species turnover
353 We evaluated nestedness and turnover between communities as assessed using eDNA
354 compared with visual/capture surveys using pairwise standardized effect size (pSES)
355 calculated for all sites. Nestedness pSES was significantly positive without overlapping
356 the zero-pSES 95% confidence interval (Fig. 3). The significantly negative pSES in
357 species turnover indicated that taxa were not replaced in communities as assessed by
358 eDNA compared to visual/capture surveys any more than would be expected to occur
359 with random sampling.

360 We compared the longitudinal nested structure and turnover strength
361 (upstream to downstream) of the fish communities as assessed by eDNA metabarcoding
362 and visual/capture surveys (Fig. 4). The longitudinal nested structure was not
363 significantly different between both methods. The pSES values were significantly
364 positive for both eDNA metabarcoding and visual/capture survey, indicating significant
365 nested structures between segments, regardless of sampling method. Moreover, the
366 pSES values were not significantly different between the eDNA metabarcoding and
367 visual/capture survey, so the strength of the nested structures did not differ between the
368 two methods (GLMM, $P = 0.302$). The species per segment were nested as downstream
369 > mid-stream > upstream according to both methods (Fig. S5 for all species and S6 for
370 each river, NOFD, $P < 0.001$). Longitudinal species turnover pSES values were not
371 different between sampling methods (GLMM, p=0.28) and were significantly negative
372 (Fig. 4b) indicating a lack of turnover whether communities were sampled by eDNA or
373 visual/capture survey.
374
375 **Discussion**
376
377 We here examined the nestedness and species turnover patterns of fish communities
378 evaluated by eDNA metabarcoding in five rivers to evaluate the performance of eDNA
379 metabarcoding. Our results suggested that the nestedness and species turnover from
380 upstream to downstream did not significantly differ within a river. Therefore, we found
381 that the both methods show similar levels of community nestedness and turnover within
382 the rivers. In general, nestedness of species assemblages occurs when the communities
383 obtained by the method estimating lower number of species are subsets of the

384 communities estimated by other methods with higher species richness (Baselga 2010).
385 Furthermore, the species turnover in the communities was very weak between the both
386 methods.

387 eDNA metabarcoding has been reported to perform better than traditional
388 methods in evaluating species richness (Deiner et al., 2016; Sato et al., 2017; Bylemans
389 et al., 2018; Nakagawa et al., 2018; Yamamoto et al., 2018; Fujii et al., 2019).
390 Nakagawa et al. (2018) investigated freshwater fish communities in 100 rivers and
391 confirmed that the communities detected by eDNA metabarcoding were similar to the
392 species lists observed in government-authorized monitoring. Furthermore, several
393 eDNA metabarcoding studies on fish communities have been performed in other river
394 systems (Bylemans et al., 2018), marine habitats (Yamamoto et al., 2018), and
395 freshwater lakes (Sato et al., 2017; Fujii et al., 2019). Deiner et al. (2016) showed that
396 river eDNA metabarcoding can reflect the regional community in a watershed,
397 indicating that eDNA metabarcoding has high performance for gamma diversity
398 evaluation. These studies indicated the great potential of eDNA metabarcoding as a
399 useful tool for alpha and gamma diversity assessment by simply comparing the
400 communities detected by eDNA metabarcoding and traditional surveys. Here, we
401 support the previous literature (e.g., Nakagawa et al. 2018, Maechler et al. 2019) by
402 showing patterns similar to those observed in alpha and gamma diversity in river
403 macroinvertebrates. However, previous studies, excluding Maechler et al. (2019), did
404 not evaluate performance in terms of nestedness and species turnover between eDNA
405 metabarcoding and other community data. We also estimated nestedness, species
406 turnover, and the capture bias of communities detected by eDNA metabarcoding and
407 traditional methods in segment scale. We found that communities detected by

408 traditional methods were nested within those detected by eDNA metabarcoding in a
409 given river locale, with scarce species turnover between methods.
410 We especially focused on the variation of the community assemblages
411 evaluated by eDNA metabarcoding compared to visual/capture survey. The two survey
412 methods showed different sources of community variation. That is, eDNA
413 metabarcoding detected significant differences between the rivers, while visual/capture
414 survey detected significant differences between the segments rather than between the
415 rivers. These differences may lead us to interpret the variation in community
416 assemblages using the results from both survey methods. We recommend using eDNA
417 for screening surveys because the method performs similarly to visual/capture surveys
418 for estimating alpha and gamma diversity, although eDNA may fail to detect some
419 species that can be captured or observed using traditional methods.
420 Using Jaccard index, a higher variation of the river fish communities was
421 statistically detected by eDNA metabarcoding than by visual/capture survey and the
422 variation of the community assemblages between segments could be significantly
423 detected by visual/capture surveys but not by eDNA metabarcoding. While the results
424 of Raup-Crick index, considering the alpha diversity of sites, showed the same results of
425 the both methods. We further compared the indicator taxa for the communities obtained
426 from both eDNA metabarcoding and visual/capture survey and concluded that several
427 taxa, including eel, salmon, and catfish, were significantly better detected by eDNA
428 metabarcoding, whereas non-indicator taxa were detected by visual/capture surveys.
429 These results indicated that eDNA metabarcoding had higher detection frequency than
430 visual/capture surveys in fish taxa detection. The community structures estimated by
431 eDNA metabarcoding and visual/capture survey were slightly different, as reported in

432 previous studies (e.g., Sato et al., 2017; Fujii et al., 2019), probably because of the
433 differences in taxa-detection performances.
434 Discriminated taxa in this analysis included eel, salmon, and catfish, which
435 mostly had larger body size and lower abundances in these rivers (Kawanabe, 2001;
436 Nakabo, 2013), although we here could not separately examine the detected capture bias
437 and actual pattern differences in species detection. In fact, the Japanese eel *Anguilla*
438 *japonica*, was difficult to find by visual observation, probably due to its hiding behavior
439 (Itakura et al., 2019). Such endangered species would be important as top predators
440 (Nakabo, 2013). eDNA metabarcoding can evaluate the distribution of such rare and
441 important taxa in fish communities better than traditional surveys. While we did not
442 detect any indicator taxa by visual/capture surveys, two species *Lepomis macrochirus*
443 and *Biwia zezera* were only detected by this method. *Lepomis macrochirus* often inhabit
444 lentic systems and are rarely observed in rivers (Nakabo, 2013) and *Biwia zezera* are not
445 widely distributed in this region (Hosokawa et al., 2007). In fact, we only observed *L.*
446 *macrochirus* and *B. zezera* in downstream segments of the Kyuragi and Oze river,
447 respectively.
448 We sampled eDNA from only two habitats at each study locale and examined
449 the performances of eDNA metabarcoding. Thus, our understanding of some aspects of
450 the fish community spatial structure in the rivers and the performance of community
451 evaluation in local habitats, such as backwater, was still limited. In fact, Bylemans et al.
452 (2018) found that river morphology in these habitats influenced the optimal sampling
453 strategy for eDNA metabarcoding. Moreover, in backwater lakes, the performance of
454 eDNA metabarcoding varied with different lake morphologies (Fujii et al. 2019).
455 However, testing the usefulness of the eDNA metabarcoding for assessing river fish

456 community biodiversity has been limited. Further research is needed to evaluate fish
457 community spatial structure in rivers. In addition, we should consider whether eDNA
458 recovered from a water sample came from an individual in the survey area. Previous
459 studies have suggested the eDNA came from upstream (Deiner, K., & Altermatt 2014;
460 Deiner et al. 2016). Therefore, the comparisons between eDNA metabarcoding and
461 visual/capture methods could be using community data with different spatial scales. For
462 example, visual/capture methods obtained community data from a 100-m reach but
463 eDNA potentially detected the community in larger area than surveyed. However, the
464 visual/capture survey also does not cover all species due to limited sampling time and
465 efforts.

466 Community assessment with eDNA metabarcoding required much less effort
467 in the field and detected the community in broader area than visual/capture survey.
468 Therefore, eDNA metabarcoding may be a suitable method, especially for regional
469 community patterns. Biodiversity testing using statistical frameworks, especially
470 community nestedness and turnover, provided the quantitative evidence to compare the
471 performance of eDNA metabarcoding and traditional surveys. eDNA methods for
472 biodiversity assessment may provide more information to us, as shown here, but we
473 should also pay attention to the unknown characteristics of eDNA, such as the origins,
474 degradation, and transport of eDNA in water, which are still unknown in various
475 habitats (Barnes & Turner, 2016; Seymour et al., 2018).

476

477 **Acknowledgments**

478 MiSeq sequencing was conducted in the Department of Environmental Solution
479 Technology, Faculty of Science and Technology, Ryukoku University, and we thank to

480 H. Yamanaka and H. Sato for supporting the experiments involving MiSeq sequencing.
481 Fish visual/capture surveys were permitted by the prefectures (No. 38 for Fukuoka, No.
482 3024 for Yamaguchi prefecture) for the survey period; 23 Sep. to 30 Nov. 2016. We
483 conducted the fish survey under the guidance of Fisheries Adjustment Rules of the
484 prefectures and Yamaguchi University. This study was supported by the Environment
485 Research and Technology Development Fund (4-1602) of Environmental Restoration
486 and Conservation Agency, Japan and JST-CREST (JPMJCR13A2).
487

488 **Data availability**

489 All data of the MiSeq sequencing was shared in DRA (Accession number:
490 DRA008090). Data was available in the supplementary material Table S1 and S2.
491

492 **Author contributions**

493 HD, RI, and YA designed the study, RI, MG, TK, and YA contributed to field survey
494 and sampling. SM, HD, RI, and MG contributed to molecular experiments. SM and HD
495 analyzed the data and interpreted the results. HD, SM, and RI wrote the initial draft of
496 the manuscript. All other authors critically reviewed the manuscript.
497

498 **Conflict of Interest**

499 The authors declare no competing financial interests.
500
501

502 **References**

Almeida-Neto, M., Guimaraes, P., Guimaraes, P. R., Loyola, R. D., & Ulrich, W. (2008). A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. Oikos, 117, 1227–1239.

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics, 17(1), 1–17.

Baselga, A., Jimenez-Valverde, A., & Niccolini, G. (2007). A multiple-site similarity measure independent of richness. Biology Letters, 3, 642–645.

Baselga, A. (2010). Multiplicative partition of true diversity yields independent alpha and beta components; additive partition does not. Ecology, 91(7), 1974–1981.

Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. Global Ecology and Biogeography, 21, 1223–1232.

Baselga, A., Gómez-Rodríguez, C., & Lobo, J. M. (2012). Historical legacies in world amphibian diversity revealed by the turnover and nestedness components of beta diversity. PloS ONE, 7(2), e32341.

Baselga, A., & Orme, C. D. L. (2012). betapart: an R package for the study of beta diversity. Methods in Ecology and Evolution, 3, 808-812.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67, 1-48.

Bylemans, J., Gleeson, D. M., Lintermans, M., Hardy, C. M., Beitzel, M., … Furlan, E. M. (2018). Monitoring riverine fish communities through eDNA metabarcoding: determining optimal sampling strategies along an altitudinal and biodiversity gradient. Metabarcoding and Metagenomics, 2, e30457.

Cáceres, M. D., & Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. Ecology, 90(12), 3566–3574.

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecological Monographs, 84, 45-67.

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, Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., ... Pfrender, M. E. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology, 26, 5872–5895. doi: 10.1111/mec.14350.

Doi, H., Fukaya, K., Oka, S. I., Sato, K., Kondoh, M., & Miya, M. (2019). Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. Scientific Reports, 9, 3581.

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, 6, 30–39.

Doi, H., & Takahara, T. (2016). Global patterns of conservation research importance in different countries of the world. PeerJ, 4, e2173.

Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C., Heled, J., Ross, H. A., Tooman, L., Grosser, S. & Park, D. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. GigaScience, 4:s13742-015.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26, 2460–2461. doi: 10.1093/bioinformatics/btq461.

Fujii, K., Doi, H., Matsuoka, S., Nagano, S., Sato, H., & Yamanaka, H. (2019). Environmental DNA metabarcoding for fish community analysis in backwater lakes: A comparison of capture methods. PLoS ONE, 14, e0210357.

Gaston, K. J., & T. M. Blackburn. (2000). Pattern and process in macroecology. Blackwell Scientific, Oxford.

Goldberg, C. S., Strickler, K. M., & Pilliod, D. S. (2015). Moving environmental DNA methods from concept to practice for monitoring aquatic macroorganisms. Biological Conservation, 183, 1–3. doi: 10.1016/j.biocon.2014.11.040.

Gotelli, N. J. (2000). Null model analysis of species co-occurrence patterns. Ecology, 81, 2606–2621.

Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J., & Knight, R. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods, 5, 235–237.

Harrison, S., Ross, S. J., & Lawton, J. H. (1992). Beta-diversity on geographic gradients in Britain. Journal of Animal Ecology, 61, 151–158.

Horikawa, M., Nakajima, J. and Mukai, T. (2007). Distribution of indigenous and non-indigenous mtDNA haplotypes of *Biwia zezera* (Cyprinidae) in northern Kyushu. Japanese Journal of Ichthyology, 54, 149-159.

Itakura, H., Wakiya, R., Yamamoto, S., Kaifu, K., Sato, T., & Minamoto, T. (2019). Environmental DNA analysis reveals the spatial distribution, abundance, and biomass of Japanese eels at the river-basin scale. Aquatic Conservation: Marine and Freshwater Ecosystems, 29(3), 361–373.

Kawanabe, H., Mizuno, N. & K. Hosoya (eds.) (2001) Freshwater Fishes of Japan Yama-Kei, Tokyo (in Japanese).

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics, 26, 1463-1464.

Margules, C. R., & Pressey, R. L. (2000). Systematic conservation planning. Nature, 405, 243.

Maechler, E., Little, C.J., Wuethrich, R., Alther, R., Fronhofer, E.A., Gounand, I., Harvey, E., Huerlemann, S., Walser, J.-C. & Altermatt, F. (2019). Assessing different components of diversity across a river network using eDNA. Environmental DNA, 1, 290-301.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., ... Kondoh, M. (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.

Nakabo, T. (2013). Fishes of Japan with pictorial keys to the species, 3rd ed. Tokai University Press (in Japanese).

Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T., & Miya, M. (2018). Comparing local- and regional-scale estimations of the diversity of stream fish

using eDNA metabarcoding and conventional observation methods. Freshwater Biology, 63, 569–580.

Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I. C., ... Falconi, L. (2017). Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science, 355, 9214.

Primack, R. B. (1993). Essentials of conservation biology (Vol. 23). Sunderland, Massachusetts: Sinauer Associates.

R Development Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R., & Gough, K. C. (2014). Review: 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. doi: 10.1111/1365-2664.12306.

Sato, H., Sogo, Y., Doi, H., & Yamanaka, H. (2017). Usefulness and limitations of sample pooling for environmental DNA metabarcoding of freshwater fish communities. Scientific Reports, 7, 14860. doi: 10.1038/s41598-017-14978-6.

Seymour, M., Durance, I., Cosby, B. J., Ransom-Jones, E., Deiner, K., Ormerod, S. J., ... Edwards, F. (2018). Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. Communications Biology, 1(1), 4.

Socolar, J. B., Gilroy, J. J., Kunin, W. E., & Edwards, D. P. (2016). How should beta-diversity inform biodiversity conservation? Trends in Ecology & Evolution, 31(1), 67–80.

Stæhr, P. A., Santos, S., Hansen, L. H., Lundsteen, S., Haraguchi, L., Dahl, K., Ávila, M. P., Winding, A. & Boutrup, S. (2016). Comparison of eDNA and conventional techniques for monitoring species diversity of boulder reefs in Danish waters. DCE-Danish Centre for Environment and Energy. Aarhus University, p.22.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology, 21, 2045–2050.

Takahara T., Minamoto T., Yamanaka H., Doi H. & Kawabata Z. (2012). Estimation of fish biomass using environmental DNA. PLOS ONE, 7, e35868

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

Thomsen, P. F., Willerslev, E. (2015). Environmental DNA–an emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation, 183, 4–18. doi: 10.1016/j.biocon.2014.11.019.

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. doi: 10.1111/1755-0998.12460.

Ulrich, W., & Gotelli, N. J. (2007). Null model analysis of species nestedness patterns. Ecology, 88(7), 1824–1831.

Wright, D. H., & Reeves, J. H. (1992). On the meaning and measurement of nestedness of species assemblages. Oecologia, 92(3), 416–428.

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

Yamanaka, H., Motozawa, H., Tsuji, S., Miyazawa, R. C., Takahara, T., & Minamoto, T. (2016). On-site filtration of water samples for environmental DNA analysis to avoid DNA degradation during transportation. Ecological Research, 31(6), 963–967.

649 Table 1 Indicator taxa analysis for the taxa which had significantly different detection
650 frequencies between eDNA metabarcoding and visual/capture methods ($P < 0.05$). Best
651 means preferred methods. P-values was calculated with 999 permutations after Sidak's
652 correction of the multiple testing.
653

Taxa	P-value for eDNA metabarcoding	P-value for visual and capturing survey	best	P-value for multiple testing
Anguilla japonica	0.001	1.000	eDNA metabarcoding	0.002
Oncorhynchus masou	0.001	1.000	eDNA metabarcoding	0.002
Oncorhynchus mykiss	0.016	1.000	eDNA metabarcoding	0.032
Cyprinus carpio	0.001	1.000	eDNA metabarcoding	0.002
Silurus asotus	0.003	1.000	eDNA metabarcoding	0.006
Misgurnus anguillicaudatus	0.009	1.000	eDNA metabarcoding	0.018

654

655

656

657 **Figure legends**

658

659 Figure 1. Numbers of taxa detected by eDNA metabarcoding (red), visual/capture
660 surveys (blue), or both methods (white) at individual sites and for sites combined (ALL)
661 within each of five river systems in Japan. Total taxa detected by eDNA metabarcoding
662 is the sum of values included within the larger circle. Total taxa detected by
663 visual/capture surveys is the sum of values in blue and white areas.

664

665 Figure 2. Nonmetric multidimensional scaling (NMDS) ordination of fish communities
666 evaluated by study river (shapes) and segment (colors), using Jaccard index (ordination
667 with Raup-Crick index shown in Fig. S4). A single ordination is divided into two panels
668 to display differences in communities as measured by sampling methods. MDS stress
669 was 0.158.

670

671 Figure 3. Species nestedness and turnover between the communities detected by eDNA
672 metabarcoding and visual/capture methods, expressed as pairwise standardized effect
673 sizes (SES). Error bars indicate 95% confidence intervals. The horizontal line represents
674 SES=0, indicating a non-significant effect.

675

676 Figure 4. Effect sizes for (a) community nestedness and (b) species turnover among
677 river segments as assessed by eDNA metabarcoding and visual/capture methods. The
678 error bars indicate 95% confidence intervals

679

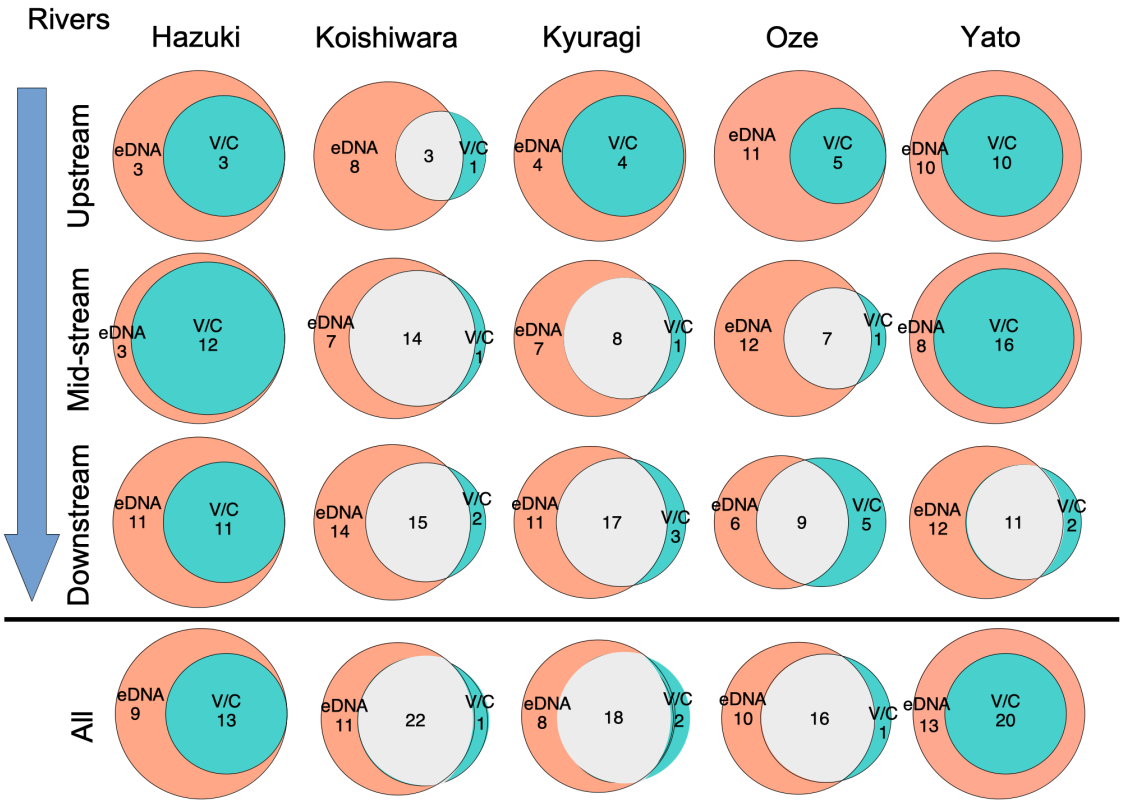
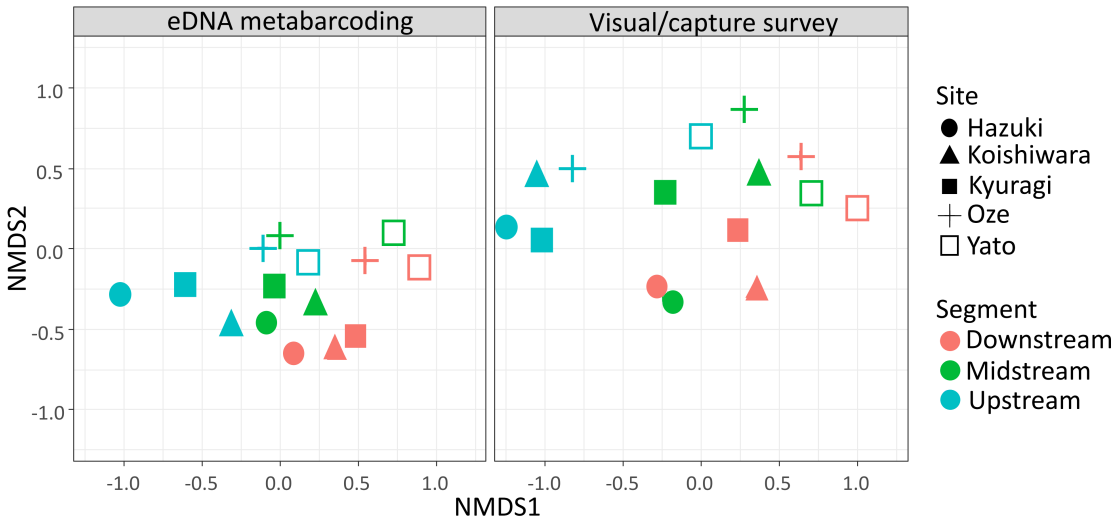

Rivers
Hazuki
Koishiwara
Kyuragi
Oze
Yato
Upstream
Mid-stream
Downstream
All
eDNA 3
V/C 3
eDNA 8
3
V/C 1
eDNA 4
V/C 4
eDNA 11
V/C 5
eDNA 10
V/C 10
eDNA 3
V/C 12
eDNA 7
14
V/C 1
eDNA 7
8
V/C 1
eDNA 12
7
V/C 1
eDNA 8
V/C 16
eDNA 11
V/C 11
eDNA 14
15
V/C 2
eDNA 11
17
V/C 3
eDNA 6
9
V/C 5
eDNA 12
11
V/C 2
eDNA 9
V/C 13
eDNA 11
22
V/C 1
eDNA 8
18
V/C 2
eDNA 10
16
V/C 1
eDNA 13
V/C 20

Figure 1

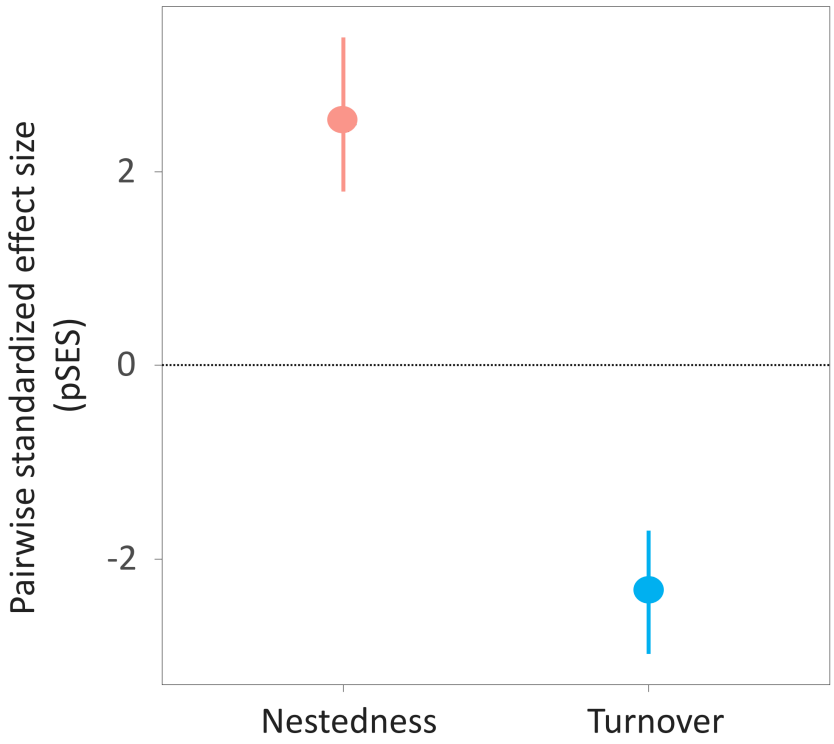
683



684 Figure 2

685

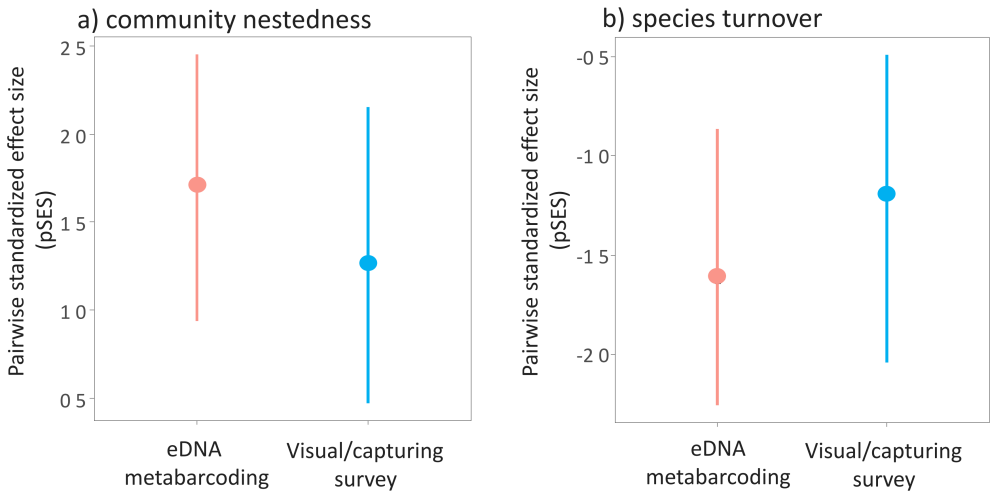
686



687

688 Figure 3

689


a) community nestedness
Pairwise standardized effect size (pSES)
2 5
2 0
1 5
1 0
0 5
eDNA metabarcoding
Visual/capturing survey
b) species turnover
Pairwise standardized effect size (pSES)
-0 5
-1 0
-1 5
-2 0
eDNA metabarcoding
Visual/capturing survey

690 Figure 4