Evaluation of biodiversity metrics through environmental DNA metabarcoding compared with visual and capture surveys in river fish community

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

[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

[†]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. The 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 in estimating alpha/gamma diversity and the variation of the community assemblages of river fish communities, particularly considering community nestedness and turnover.

3. In five rivers across west Japan, with comparing to visual/capture surveys, eDNA metabarcoding detected more species in the study sites, 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 by eDNA metabarcoding.

4. With analyzing the community assemblages between the rivers, we showed the different results between the both methods. While, in the same river, the nestedness and species turnover changing from upstream to downstream did not significantly differ between the both methods. Our results suggest that eDNA metabarcoding may be

suitable method, especially for understanding regional community patterns, for fish monitoring in rivers.

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

53

54 **Introduction**

55

56 The maintenance of biodiversity underpins the stability of ecosystem processes in
57 constantly changing environments (Primack, 1993; Margules & Pressey, 2000; Pecl et
58 al., 2017). Moreover, biodiversity loss affects ecosystem functions and services and,
59 consequently, human society (Primack 1993; Margules & Pressey, 2000, Pecl et al.
60 2017). Ecologists have made efforts to conserve biodiversity based on essential
61 biodiversity survey methods, e.g., species richness and distribution (Primack, 1993;
62 Margules & Pressey, 2000, Doi & Takahara, 2016, Pecl et al., 2017). Biodiversity can
63 be evaluated in different levels: e.g., by estimating alpha (local), beta (between
64 habitats), and gamma (regional) diversity and the variation of the community
65 assemblages. To conserve local communities, ecologists incorporated these diversity
66 measurements into management decision-making (Primack 1993; Margules & Pressey,
67 2000, Socolar et al., 2016). For example, the variation of the community assemblages
68 can quantify biodiversity loss and inform the placement of protected areas and the
69 management of biological invasions and landscapes (Socolar et al., 2016). Thus, robust
70 methods for monitoring biodiversity are fundamental for biodiversity and environmental
71 management.

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

77 exceptionally useful and powerful tool for community biodiversity surveys (Taberlet et
78 al., 2012; Deiner et al., 2016, 2017; Sato et al., 2017; Bylemans et al., 2018; Fujii et al.,
79 2019). eDNA metabarcoding has recently been applied in fish community surveys, e.g.
80 Miya et al. (2015) designed and applied universal PCR primers (the MiFish primers) to
81 survey marine fish communities. To confirm the usefulness of eDNA metabarcoding for
82 community assessment, many studies compared eDNA metabarcoding to a species list
83 generated by traditional surveys including visual and capture methods (Deiner et al.,
84 2016; Bylemans et al., 2018; Nakagawa et al., 2018; Yamamoto et al., 2018; Fujii et al.,
85 2019). However, evaluating the performance of eDNA metabarcoding estimating alpha
86 diversity was still limited quantitively and statistically (but Drummond et al., 2015;
87 Deiner et al., 2016; Staehr et al. 2016, Maechler et al. 2019), especially when gamma
88 diversity and the variation of the community assemblages evaluating and comparing to
89 traditional surveys.
90 Variation of the community assemblages is a fundamental aspect for
91 communities, and it is important for evaluating community responses to environmental
92 gradients and the spatial locations of community across different habitats and
93 ecosystems (Chase et al., 2004; Chase, 2010; McCune & Vellend, 2013; Socolar et al.,
94 2016). The variation itself and trends along environmental gradients, such as
95 productivity and disturbance, can be influenced by alpha diversity (Chase, 2010).
96 Furthermore, variation of the community assemblages can be influenced by alpha and
97 gamma diversity (Olden & Poff, 2003; Van Calster et al., 2007; Keith et al., 2009).
98 Therefore, the variation evaluation of community assemblages can vary considerably
99 depending on differences in alpha and gamma diversity evaluation by different survey
100 methods, i.e. eDNA metabarcoding vs. traditional surveys.

101 Variation of the community assemblages is considered to reflect two different
102 components: nestedness and species turnover (Harrison et al., 1992; Baselga et al.,
103 2007; Baselga, 2010). Nestedness occurs when the community at the sites with less
104 species are subsets of the community at the sites with higher species richness (Wright &
105 Reeves, 1992; Ulrich & Gotelli, 2007) and generally reflects a non-random process of
106 species loss (Gaston & Blackburn, 2000). Contrastingly, species turnover implies the
107 replacement of some species by others because of environmental sorting or
108 spatial/historical constraints (Baselga, 2010). Statistical separation methods for
109 nestedness and species turnover were applied for evaluating variation of the community
110 assemblages in various systems (Baselga, 2010; Baselga et al., 2012). However, the
111 method has never been applied to evaluate the performance of eDNA metabarcoding
112 estimating community assemblages of fish species. Moreover, Baselga's (2010)
113 framework can be applied to compare the performance among methods when evaluating
114 alpha diversity via nestedness and species turnover.
115 Using statistical methods, we can quantitatively compare the performance of
116 eDNA metabarcoding and traditional surveys for alpha/gamma diversity evaluation of
117 biological communities and the variation of the community assemblages among the
118 study sites. Here, we tested the performance of eDNA metabarcoding in five river
119 systems in different regions with various fish species. We conducted eDNA
120 metabarcoding using universal MiFish primers that target fish and identified the fish by
121 visual snorkeling and hand-net capture surveys. We evaluated the performance of
122 eDNA metabarcoding by comparing the obtained fish community structure to that
123 evaluated by visual/capture survey with special regard to nestedness and species
124 turnover.

Methods

Site description

In 2016, we conducted field surveys in five river systems across Japan (river map in Fig. S1): the Kyuragi River on October 10, the Koishiwara River on October 21, the Yato River on October 25, the Hazuki River on November 2, and the Oze River on November 6. The survey sites were set at a site at each of three river segments (upstream, mid-stream, and downstream, the internal distances ranged from 4.5 to 25.8 km, Fig. S1) for each river. Each site was set so that the length in the up-down direction was approximately 100 m with a riffle at the downstream end (e.g., Fig. S2).

Water collection for eDNA survey

In each site, we collected 1 L of surface water in bleached bottles at two points, the stream near the downstream end and the shore with static or semi-static water (Fig. S2) immediately before visual and capture surveys. eDNA was removed from the bottles and filtering equipment using 10% commercial bleach (ca. 0.6% hypochlorous acid) and washing with DNA-free distilled water. One milliliter of benzalkonium chloride (BAC, 10% w/v) was added per liter of water sample to avoid a decrease in eDNA concentration in the samples (Yamanaka et al., 2016). During transport, samples were stored in a cooler with ice packs. The ‘cooler blank’ contained 1 L DNA-free water, which we brought to the field and treated identically to the other water samples, except that it was not opened at the field sites.

149 Visual observation and capture methods

150 After water sampling, the fish fauna survey was conducted by visual observation with
151 snorkel and collection with hand net. For visual observation, we observed and recorded
152 fish species by snorkeling in a 100-m transect (snorkeling by 1 person for 1 h, Fig. S2).
153 We observed at various micro habitats, including the riffle, pool, and shore bank from
154 the downstream end to upstream end. We also conducted a hand-net capture survey (1
155 person × 1 h) using a D-frame net (2 mm mesh, net opening: 0.16 m^2) in the various
156 habitats in the river, including the riffle, pool, and shore bank. Fishes were identified
157 according to Nakabo et al. (2013) at the survey site. We used the combined taxa list
158 from both traditional surveys to compare to that of eDNA metabarcoding. In order to
159 prevent contamination of eDNA samples, the investigator who collected and identified
160 the fish and the investigator who sampled the water were different.

161

162 eDNA collection, extraction and measurements

163 Collected water samples were vacuum-filtered into GF/F glass filters (47 mm diameter,
164 pore size: 0.7 µm, GE Healthcare, Little Chalfont, UK) in the laboratory within 24 h of
165 sampling. After filtration, all filters were stored at -20 °C before eDNA extraction. The
166 cooler blank was also processed in the same manner. A liter of Milli-Q water was used
167 as the filtering blank to monitor contamination during filtering in each site and during
168 subsequent DNA extraction.

169 To extract the DNA from the filters, we followed the methods described in
170 Uchii, Doi, & Minamoto (2016). We incubated the filter by submerging the mixed
171 buffer of 400 µL of Buffer AL in DNeasy Blood & Tissue Kit (Qiagen, Hilden,
172 Germany) and 40 µL of Proteinase K (Qiagen, Hilden, Germany), using a Salivette tube

173 (Sarstedt, Nümbrecht, Germany) at 56 °C for 30 min. The Salivette tube with filters was
174 centrifuged at 5000 × *g* for 5 min. Then, we added 220 µL of TE buffer (pH: 8.0) onto
175 the filter and again centrifuged at 5000 × *g* for 5 min. The DNA was purified using a
176 DNeasy Blood & Tissue Kit with extracted the DNA in 200 µL in Buffer AE. Samples
177 were stored at -20 °C until the 1st-PCR assay.
178
179 *Library preparation and MiSeq sequencing*
180 The detailed molecular methods are described in Fujii et al. (2019) with a two-step
181 PCR-procedure for Illumina MiSeq sequencing. Briefly, we performed 1st-PCR
182 withMiFish-U-F and MiFish-U-R primers (Miya et al., 2015), which were designed to
183 contain Illumina sequencing primer regions and 6-mer Ns;
184 Forward: 5'-*ACACTCTTTCCCTACACGACGCTCTTCCGATCT* NNNNNN
185 GTCGGTAAAACTCGTGCCAGC-3′,
186 Reverse: 5′-*GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT* NNNNNN
187 CATAGTGGGGTATCTAATCCCAGTTTG-3′
188 The italicized and non-italicized letters represent MiSeq sequencing primers
189 and MiFish primers, respectively, and the six random bases (N) were used to enhance
190 cluster separation on the flow cells during initial base call calibrations on the MiSeq
191 (Miya et al. 2015, Doi et al. 2019).
192 We performed the 1st-PCR with a 12 µL reaction volume containing 1× PCR
193 Buffer for KOD FX Neo (Toyobo, Osaka, Japan), 0.4 mM dNTP mix, 0.24 U KOD FX
194 Neo polymerase, 0.3 µM of each primer, and 2 µL template. The thermocycling
195 conditions for this step were as follows: initial denaturation at 94 °C for 2 min, followed
196 by 35 cycles of denaturation at 98 °C for 10 s, annealing at 65 °C for 30 s, and

197 elongation at 68 °C for 30 s, followed by final elongation at 68 °C for 5 min. The first
198 PCRs were performed using eight replicates (Doi et al. 2019) and individual first PCR
199 replicates were pooled and purified using AMPure XP (Beckman Coulter, Brea CA,
200 USA) as templates for the 2nd-PCR. The Illumina sequencing adaptors and the eight bp
201 identifier indices (XXXXXXXX) were added to the subsequent PCR process using a
202 forward and reverse fusion primer:
203 Forward: 5′-*AATGATACGGCGACCACCGAGATCTACA* XXXXXXXX
204 ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′
205 Reverse: 5′-*CAAGCAGAAGACGGCATACGAGAT* XXXXXXXX
206 GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3′
207 The italicized and non-italicized letters represent MiSeq P5/P7 adapter and
208 sequencing primers, respectively. The 8X bases represent dual-index sequences inserted
209 to identify different samples (Hamady 2008). We performed the 2nd-PCR with 12
210 cycles of a 12 μL reaction volume containing 1× KAPA HiFi HotStart ReadyMix, 0.3
211 μM of each primer, and 1.0 μL of the first PCR production. The thermocycling
212 conditions profile after an initial 3 min denaturation at 95 °C was as follows:
213 denaturation at 98 °C for 20 s, annealing, and extension combined at 72 °C (shuttle
214 PCR) for 15 s, with the final extension at the same temperature for 5 min. We
215 confirmed the positive bands of the targeted 1st-PCR amplicons by electrophoresis. The
216 2nd-PCR products were pooled in equal volumes and purified using AMPure XP.
217 The purified PCR products were loaded on a 2% E-Gel SizeSelect (Thermo
218 Fisher Scientific, Waltham, MA, USA) and the target size of the libraries
219 (approximately 370 bp) was collected. The samples concentration and quality were
220 estimated by a Qubit dsDNA HS assay kit and a Qubit 2.0 (Thermo Fisher Scientific).

221 The amplicon libraries were sequenced by 2 × 250 bp paired-end sequencing on the
222 MiSeq platform using the MiSeq v2 Reagent Kit. Note that the sequencing run
223 contained a total of 339 libraries including 40 of our library (30 samples plus five cooler
224 and five filter negative controls) and 299 libraries from other research projects. The
225 MiSeq sequencing was conducted in the Department of Environmental Solution
226 Technology, Faculty of Science and Technology, Ryukoku University. All sequence
227 data were deposited in DNA Data Bank of Japan (DRA, Accession number:
228 DRA008090).
229
230 Bioinformatic analysis for MiSeq sequencing
231 The detailed procedures used for bioinformatics analysis are described in Fujii et al.
232 (2019). First, low-quality tails were trimmed from each read and paired-end reads were
233 then merged. For the obtained 1,823,446 reads, primer sequences were removed and
234 identical sequences (i.e., 100% sequence similarity) were merged using UCLUST
235 (usearch 7.0.1001, Edgar, 2010). The sequences with 10 or more identical reads were
236 subjected to the downstream processes. To annotate the taxonomy, local BLASTN
237 search using BLAST 2.2.29 was conducted with the reference database of fish species
238 for processed reads (Miya et al., 2015). The top BLAST hit with a sequence identity ≥
239 97% was applied to species detection of each sequence. Note that the species were
240 mostly identified with a ≥ 99% match. From the BLAST results, we identified the
241 species using methods previously described (Sato et al., 2017). Also, we detected the
242 most of fish species inhabiting the rives in the regions with regarding to the known
243 distributions (Kawanabe et al. 2001).
244

245 Statistical analyses

246 All statistical analyses and graphics were conducted in R ver. 3.4.4 (R Core Team,
247 2018). All statistics were set at the significance level $\alpha = 0.05$. To compare between
248 eDNA metabarcoding and visual survey data, the taxonomic levels in the species list
249 from visual survey were compared to the lists from eDNA metabarcoding (Table S1,
250 S2) in reference to previous studies using the MiFish primer (Sato et al., 2017; Fujii et
251 al., 2019). Before statistical analysis, we confirmed that the sequencing depth was
252 sufficient to detect alpha diversity in the samples by "iNEXT" and "ggiNEXT"
253 functions in the "iNEXT" ver. 2.0.19 package (Chao et al. 2014, Fig. S3). We merged
254 the community data from two points, the stream near the downstream end and the shore,
255 to compare with visual/capture surveys.

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

261 The differences in community compositions were visualized using nonmetric
262 multidimensional scaling (NMDS) with 500 separate runs of real data. For NMDS, the
263 community dissimilarity was calculated based on incidence-based Jaccard indices. We
264 evaluated the differences in community structures between methods and sites using
265 permutational multivariate analysis of variance (PERMANOVA). For PERMANOVA,
266 we used Jaccard and Raup-Crick similarity matrix and calculated the statistical values
267 with 999 permutations. Raup-Crick index is the probability that compared the sampling
268 sites have non-identical species composition (i.e., considering the alpha diversity of

269 sites). We used “metaMDS” and “adonis” functions in the “vegan” ver. 2.5-6 package
270 (https://github.com/vegandevs/vegan) for NMDS ordination and PERMANOVA,
271 respectively.

272 For the communities evaluated by both methods, indicator taxa analysis
273 (Cáceres & Legendre, 2009) was performed to determine which taxa had significantly
274 different frequencies between both methods. The analysis was performed using the
275 “signassoc” function in the “indicspecies” ver. 1.7.8 package on the present/absence
276 data for the testing with regarding the package description and Cáceres & Legendre
277 (2009). The “signassoc” function can calculate the index with both present/absence and
278 abundance data. We used mode = 1 (group-based) and calculated the *P*-values with 999
279 permutations after Sidak’s correction of the multiple testing.

280 To compare community composition (i.e., the variation of the community
281 assemblages) focusing on nested and turnover structures between eDNA metabarcoding
282 and visual/capture survey, we calculated the pairwise indices for nestedness and species
283 turnover (Baselga, 2010; Baselga, 2012) using the “beta.pair” function in the “betapart”
284 ver. 1.5.1 package (Baselga and Orme 2012). Standardized effect sizes (SESs) of
285 nestedness and turnover indices were calculated to show the degree of nestedness and
286 turnover structure. The significance was defined by deviation from zero, and the
287 expectation of random assembly (a null model) was estimated with 999 random
288 sampling replicates. The SES was defined as follows: $(\beta_{obs} - \beta_{null})/\beta_{sd}$, where β_{obs} is the
289 observed beta diversity (here, the variation of the community assemblages among the
290 sites), β_{null} is the mean of the null distribution of beta diversity, and β_{sd} is the standard
291 deviation of the null distribution. SES values greater than zero indicate statistically
292 stronger nestedness or turnover structure than expected under a random model of

community assembly, while negative values indicate weaker nestedness or turnover than expected. The randomized community data were generated with independent swap algorithm (Gotelli, 2000) using “randomizeMatrix" function in the “picante” ver. 1.8 package (Kembel et al. 2010). First, to evaluate the differences in nestedness and species turnover between the survey methods (eDNA metabarcoding vs. visual/capture survey) at the same segments, the SES of pairwise nestedness and turnover were calculated for each sample pair within the river. Then, the fish community longitudinal nestedness and turnover structure along with river flow (i.e., upstream to downstream) by each method was evaluated with NODF and pairwise indices of nestedness. First, a nestedness metric (NODF, Almeida-Neto et al., 2008) and their SES value were calculated with 999 randomizations using the "nestednodf" and the “oecosimu” function in the “vegan” package. Then, we tested the differences in SES of pairwise indices between the survey methods (eDNA metabarcoding vs. visual/capture survey) by generalized linear mixed model (GLMM, with Gaussian distribution) with the “lmer” function in the “lme4” package. In the GLMM models, the SES of pairwise indices was treated as a fixed effect, and the rivers-pairs and segment-pairs (i.e., three pairs in each river) were treated as random effects.

For the map graphics in Supplemental Materials, we used the stream and coast line data from the National Land Information Division, Ministry of Land, Infrastructure, Transport and Tourism of Japan (http://nlftp.mlit.go.jp/ksj/).

Results

Overview

317 We detected 53 fish taxa, almost all identified to the species or genus level, by eDNA
318 metabarcoding in five rivers (Table S1, and S2) and visually observed 38 fish taxa in
319 total. MiSeq paired-end sequencing for the library (30 samples plus five cooler and five
320 filter negative controls) yielded a total of 1,601,816 reads (53,351 ± 17,639; mean ± S.
321 D. for each sample, Table S2). We confirmed very low reads from negative controls
322 (Table S2) with only detecting a fish species, *Tridentiger* sp. in the blank of the Yato
323 River, probably because of the cross-contamination among the samples. The read of
324 *Tridentiger* sp.in thee samples (4783-9818 reads) were high than the detected sites
325 (1286 reads), therefore we used the species for the analysis.
326
327 Diversity indices between methods
328 We found significant differences in fish local richness (alpha diversity) between eDNA
329 metabarcoding and visual/capture surveys (Fig. 1, GLMM with river as a random effect,
330 t = –5.45, P = 0.000018). Richness was significantly different among river segments (t
331 = –5.85, P = 0.000004), indicating higher alpha diversity estimated by eDNA
332 metabarcoding than by visual/capture surveys. While, richness was not significantly
333 different among rivers (t = 1.737, P = 0.0942).
334 We found differences in community structure between the two methods by
335 NMDS ordination by Jaccard index (Fig. 2) as well as Raup-Crick (Fig. S4), especially
336 among the study rivers. The patterns of differences in the both Jaccard and Raup-Crick
337 indices were similar.
338 The PERMANOVA results with Jaccard and Raup-Crick indices for the
339 ordination suggested there were differences in community composition evaluated by
340 each method, eDNA metabarcoding and visual/capture survey (P < 0.012, Table S3).

341 Moreover, communities from the combined results of eDNA metabarcoding and
342 visual/capture survey were significantly different among rivers and segments ($P <$
343 0.029, Table S3). We found different patterns in ordinated river sites for each method
344 (Fig. 2 for Jaccard index, Fig. S4 for Raup-Crick index).
345 To test the community differences among the segments and rivers, the
346 PERMANOVA results with Jaccard index determined that communities were
347 significantly different among the rivers by eDNA metabarcoding ($P = 0.001$) but not by
348 visual/capture survey ($P = 0.12$, Table S4). Conversely, the communities were
349 significantly different among river segments by visual/capture survey ($P = 0.011$) but
350 not significantly different (albeit marginally) by eDNA metabarcoding ($P = 0.061$,
351 Table S4). The differences in PERMANOVA results with Jaccard index suggested that
352 differences in the variation of the community assemblages among rivers across regions
353 can be detected by eDNA metabarcoding but not by visual/capture survey. While, the
354 PERMANOVA results with Raup-Crick index determined that communities were
355 significantly different among the rivers by the both eDNA metabarcoding ($P = 0.001$)
356 and visual/capture survey ($P = 0.018$, Table S4).
357 Indicator taxa analysis comparing the communities estimated by both
358 methods, eDNA metabarcoding and visual/capture survey, detected capture bias by both
359 methods for several several taxa, including Japanese eel (*Anguilla japonica*), salmon
360 (e.g., *Oncorhynchus masou)*, and Amur catfish (*Silurus asotus*) ($P < 0.05$, Table 1 for
361 statistically significant taxa, Table S5 for all taxa).
362
363 Nestedness and species turnover

364 We compared the pairwise standardized effect size (pSES) between eDNA
365 metabarcoding and visual/capture survey in nestedness and species turnover (Fig. 3).
366 Nestedness pSES was significantly positive without overlapping the zero-pSES 95%
367 confidence interval, indicating that the visual/capture survey community was nested in
368 that detected by eDNA metabarcoding. The significantly negative pSES in species
369 turnover indicated that the taxa did not significantly turnover in the community
370 evaluated by eDNA metabarcoding and visual/capture survey compared to that in
371 random communities in the segment-scale comparisons.
372 We compared the longitudinal nested structure (upstream to downstream) of the fish
373 community in the study rivers using nested pSES (Fig. 4a). The longitudinal nested
374 structure was not significantly different between both methods (GLMM, $P = 0.302$).
375 The positive pSES indicated that downstream communities were significantly nested in
376 the upstream ones. The species per segment were nested as downstream > mid-stream >
377 upstream both methods (Fig. S5 and S6, NOFD, $P < 0.001$). We also compared the
378 longitudinal species turnover of fish communities in the study rivers using turnover
379 pSES (Fig. 4b). The longitudinal species turnover was not significantly different
380 between both methods (GLMM, P = 0.280) and the negative turnover values indicated
381 no turnover observed.
382
383 **Discussion**
384
385 We found that river fish communities estimated by eDNA metabarcoding were
386 significantly nested in the communities estimated by visual/capture survey.
387 Furthermore, the species turnover in the communities was very weak between methods.

388 Nestedness of species assemblages occurs when the communities obtained by the
389 method estimating lower number of species are subsets of the communities estimated by
390 other methods with higher species richness (Baselga 2010). The eDNA metabarcoding
391 estimated higher alpha diversity than visual/capture survey for river fish communities
392 and provide the similar gamma diversity and species composition with visual/capture
393 survey. In fact, the fish local and regional richness (alpha and gamma diversity)
394 evaluated by eDNA metabarcoding was significantly higher, including almost all taxa
395 evaluated by visual/capture survey.
396 eDNA metabarcoding has been reported to perform better than traditional
397 methods in evaluating species richness (Deiner et al., 2016; Sato et al., 2017; Bylemans
398 et al., 2018; Nakagawa et al., 2018; Yamamoto et al., 2018; Fujii et al., 2019).
399 Nakagawa et al. (2018) investigated freshwater fish communities in 100 rivers and
400 confirmed that the community detected by eDNA metabarcoding were similar to the
401 species lists observed in government-authorized monitoring. Furthermore, several
402 eDNA metabarcoding studies on fish communities have been performed in other river
403 systems (Bylemans et al., 2018), marine habitats (Yamamoto et al., 2018), and
404 freshwater lakes (Sato et al., 2017; Fujii et al., 2019). Deiner et al. (2016) showed that
405 river eDNA metabarcoding can reflect the community in a watershed, indicating that
406 eDNA metabarcoding has high performance for gamma diversity evaluation. These
407 studies indicated the great potential of eDNA metabarcoding as a useful tool for alpha
408 and gamma diversity assessment by simply comparing the community data obtained
409 from eDNA metabarcoding and traditional surveys. However, previous studies did not
410 evaluate performance in terms of nestedness and species turnover between eDNA
411 metabarcoding and other community data. Here, we support the previous literature

412 (Maechler et al. 2019) by showing patterns observed in alpha and gamma diversity are
413 also observed in nestedness, species turnover, and the capture bias of communities
414 detected by eDNA metabarcoding and traditional methods in segment scale. With
415 comparing to the same segment of the same river we found that the community detected
416 by visual/capture survey was nested with that by eDNA metabarcoding, with scarce
417 species turnover in the community.

418 We especially focused on the variation of the community assemblages
419 evaluated by eDNA metabarcoding compared to visual/capture survey. For evaluating
420 the fish communities, the community compositions were different between eDNA
421 metabarcoding and visual/capture survey. That is, eDNA metabarcoding detected
422 significant differences between the rivers, while visual/capture survey detected
423 significant differences between the segments rather than between the rivers. On the
424 other hand, with focusing on a single river, the patterns of nestedness and species
425 turnover were not significantly different between eDNA metabarcoding and
426 visual/capture survey.

427 Using Jaccard index, a higher variation of the river fish communities was
428 statistically detected by eDNA metabarcoding than by visual/capturing survey and the
429 variation of the community assemblages between segments could be significantly
430 detected by visual/capture surveys but not by eDNA metabarcoding. While the results
431 of Raup-Crick index, considering the alpha diversity of sites, showed the same results of
432 the both methods. This might suggest that Jaccard dissimilarity detect significant
433 differences between the segments for visual/capturing survey due to the low number of
434 species detected in the upstream. These differences in the variation of the community
435 assemblage evaluation may lead us to interpret the variation of the community

436 assemblages using the results from both survey methods, eDNA and visual/capture
437 survey.
438 We further compared the indicator taxa for the communities obtained from
439 both eDNA metabarcoding and visual/capture survey and concluded that several taxa,
440 including eel, salmon, and catfish, were significantly better detected by eDNA
441 metabarcoding, whereas non-indicator taxa were detected by visual/capture surveys.
442 These results indicated that eDNA metabarcoding had higher detection frequency of
443 visual/capture surveys in fish taxa detection. The community structures estimated by
444 eDNA metabarcoding and visual/capture survey were slightly different, as reported in
445 previous studies (e.g., Sato et al., 2017; Fujii et al., 2019), probably because of the
446 differences in taxa-detection performances. Discriminated taxa in this analysis included
447 eel, salmon, and catfish, which mostly had larger body size and lower abundances in
448 these rivers (Kawanabe, 2001; Nakabo, 2013). In fact, the Japanese eel *Anguilla*
449 *japonica*, was difficult to find by visual observation, probably due to its hiding behavior
450 (Itakura et al., 2019). Such endangered species would be important as top predators
451 (Nakabo, 2013). eDNA metabarcoding can evaluate the distribution of such rare and
452 important taxa in fish communities better than traditional surveys. While we did not
453 were detect any indicator taxa by visual/capture surveys, we found couple species
454 detected only by visual/capture surveys, for example, *Lepomis macrochirus* and *Biwia*
455 *zezera* (see Table S1). *Lepomis macrochirus* usually inhabit in lotic systems and rarely
456 observed in rivers (Nakabo, 2013). *Biwia zezera* distributed in this region (Hosokawa et
457 al., 2007). We only observed the both species in a downstream segment of Kyuragi and
458 Oze river, respectively, thus, this species was rarely observed in this study. Also, the
459 BLAST identification for *L. macrochirus* and *B. zezera* were preliminary confirmed

460 (Biodiversity Center of Japan, Ministry of Environment, Japan, in Japanese;
461 http://www.biodic.go.jp/edna/edna_top.html). Therefore, we speculated that we could
462 not capture the eDNA of *L. macrochirus* and *B. zezera*, due to the sampling timing or
463 points. Further study needs to the species detection which we captured, especially
464 rarely-observed species.

465 To evaluate the comparison among the local sites and rivers, we examined the
466 performances of eDNA metabarcoding at three sites of five rivers with eDNA sampling
467 from only two habitats. Thus, our understanding of some aspects of the fish community
468 spatial structure in the rivers and the performance of community evaluation in local
469 habitats, such as backwater, was still limited. In fact, Bylemans et al. (2018) found that
470 river morphology in these habitats influenced the optimal sampling strategy for eDNA
471 metabarcoding. Moreover, in backwater lakes, the performance of eDNA
472 metabarcoding varied with different lake morphologies (Fujii et al. 2019). However,
473 testing the usefulness of the eDNA metabarcoding for assessing river fish community
474 biodiversity has been limited. Further research is needed to evaluate fish community
475 spatial structure in rivers. In addition, we should consider that eDNA recovered from a
476 water sample came from an individual in the survey area. The previous studies
477 suggested the eDNA came from the upstream (Deiner, K., & Altermatt 2014; Deiner et
478 al. 2016). Therefore, the comparisons between eDNA metabarcoding and visual/capture
479 methods were using the community data with different spatial scales, for example,
480 community data from a 100-m reach by visual/capture methods, but eDNA potentially
481 detected the community in larger area than surveyed. We should carefully consider the
482 phenomenon, especially for alpha diversity of community.

483 In conclusion, eDNA metabarcoding have similar species detection
484 performance with the visual/capture survey in evaluating the community among the
485 rivers. The eDNA metabarcoding was much less effort in the field and detected the
486 community in broader area than visual/capture survey, therefore, eDNA metabarcoding
487 may be suitable method, especially for regional community patterns. Biodiversity
488 testing using statistical frameworks, especially community nestedness and turnover,
489 provided the quantitative evidence to compare the performance of eDNA
490 metabarcoding and traditional surveys. eDNA methods for biodiversity assessment may
491 provide more information to us, as shown here, but we should also pay attention to the
492 unknown characteristics of eDNA, such as the origins, degradation, and transport of
493 eDNA in water, which are still unknown in various habitats (Barnes & Turner, 2016;
494 Seymour et al., 2018). To routinely use eDNA methods to assess biodiversity, we
495 encourage testing the effect of eDNA degradation and transport on the performance of
496 eDNA methods for biodiversity evaluation.
497

498 **Acknowledgments**

499 MiSeq sequencing was conducted in the Department of Environmental Solution
500 Technology, Faculty of Science and Technology, Ryukoku University, and we thank to
501 H. Yamanaka and H. Sato for supporting the experiments involving MiSeq sequencing.
502 This study was supported by the Environment Research and Technology Development
503 Fund (4-1602) of Environmental Restoration and Conservation Agency, Japan and JST-
504 CREST (JPMJCR13A2).
505

506 **Data availability**

All data of the MiSeq sequencing was shared in DRA (Accession number: DRA008090), and the observed species data was shared in Table S2 in Supplemental Materials.

Author contributions

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

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.

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

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.

Chase, J. M., & Ryberg, W. A. (2004). Connectivity, scale-dependence, and the productivity–diversity relationship. Ecology Letters, 7(8), 676–683.

553 Chase, J. M. (2010). Stochastic community assembly causes higher biodiversity in more
554 productive environments. Science, 328, 1388–1391.
555 Cox, M. P., Peterson, D. A., & Biggs, P. J. (2010). SolexaQA: at-a-glance quality
556 assessment of Illumina second-generation sequencing data. BMC Bioinform., 11,
557 485.
558 Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental
559 DNA in a natural river. PLOS ONE, 9, e88786.
560 Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., & Altermatt, F. (2016).
561 Environmental DNA reveals that rivers are conveyer belts of biodiversity
562 information. Nature Communications, 7, 12544.
563 Deiner K, Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt,
564 F., ... Pfrender, M. E. (2017). Environmental DNA metabarcoding:
565 Transforming how we survey animal and plant communities. Molecular
566 Ecology, 26, 5872–5895. doi: 10.1111/mec.14350.
567 Doi, H., Fukaya, K., Oka, S. I., Sato, K., Kondoh, M., & Miya, M. (2019). Evaluation
568 of detection probabilities at the water-filtering and initial PCR steps in
569 environmental DNA metabarcoding using a multispecies site occupancy model.
570 Scientific Reports, 9, 3581.
571 Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T., & Minamoto,
572 T. (2017). Environmental DNA analysis for estimating the abundance and
573 biomass of stream fish. Freshwater Biology, 6, 30–39.
574 Doi, H., & Takahara, T. (2016). Global patterns of conservation research importance in
575 different countries of the world. PeerJ, 4, e2173.

576 Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST.
577 Bioinformatics, 26, 2460–2461. doi: 10.1093/bioinformatics/btq461.
578 Fujii, K., Doi, H., Matsuoka, S., Nagano, S., Sato, H., & Yamanaka, H. (2019).
579 Environmental DNA metabarcoding for fish community analysis in backwater
580 lakes: A comparison of capture methods. PLoS ONE, 14, e0210357.
581 Gaston, K. J., & T. M. Blackburn. (2000). Pattern and process in macroecology.
582 Blackwell Scientific, Oxford.
583 Goldberg, C. S., Strickler, K. M., & Pilliod, D. S. (2015). Moving environmental DNA
584 methods from concept to practice for monitoring aquatic macroorganisms.
585 Biological Conservation, 183, 1–3. doi: 10.1016/j.biocon.2014.11.040.
586 Gotelli, N. J. (2000). Null model analysis of species co-occurrence patterns. Ecology,
587 81, 2606–2621.
588 Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J., & Knight, R. (2008). Error-
589 correcting barcoded primers for pyrosequencing hundreds of samples in multiplex.
590 Nature Methods, 5, 235–237.
591 Harrison, S., Ross, S. J., & Lawton, J. H. (1992). Beta-diversity on geographic gradients
592 in Britain. Journal of Animal Ecology, 61, 151–158.
593 Horikawa, M., Nakajima, J. and Mukai, T. (2007). Distribution of indigenous and non-
594 indigenous mtDNA haplotypes of *Biwia zezera* (Cyprinidae) in northern
595 Kyushu. Japanese Journal of Ichthyology, 54, 149-159.
596 Itakura, H., Wakiya, R., Yamamoto, S., Kaifu, K., Sato, T., & Minamoto, T. (2019).
597 Environmental DNA analysis reveals the spatial distribution, abundance, and
598 biomass of Japanese eels at the river-basin scale. Aquatic Conservation: Marine
599 and Freshwater Ecosystems, 29(3), 361–373.

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

Keith, S. A., Newton, A. C., Morecroft, M. D., Bealey, C. E., & Bullock, J. M. (2009) Taxonomic homogenization of woodland plant communities over 70 years. Proceedings of the Royal Society of London. Series B, Biological science, 276, 3539–3544.

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.

McCune, J. L., & Vellend, M. (2013). Gains in native species promote biotic homogenization over four decades in a human-dominated landscape. Journal of Ecology, 101, 1542–1551.

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.

Olden, J. D., & Poff, N. L. (2003). Toward a mechanistic understanding and prediction of biotic homogenization. The American Naturalist, 162, 442–460.

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.

Schmieder, R., Lim, Y. W., Rohwer, F., & Edwards, R. (2010). TagCleaner: Identification and removal of tag sequences from genomic and metagenomic datasets. BMC Bioinformatics, 11, 341, doi: 10.1186/1471-2105-11-341.

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.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology, 21(8), 2045–2050.

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 doi: 10.1371/journal.pone.0056584.

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.

Van Calster, H., Baeten, L., De Schrijver, A., De Keersmaeker, L., Rogister, J. E., Verheyen, K., & Hermy, M. (2007). Management driven changes (1967–2005)

in soil acidity and the understorey plant community following conversion of a coppic-with-standards forest. Forest Ecology and Management, 241, 258–271.

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.

684 Table 1 Indicator taxa analysis for the taxa had significantly different frequency
685 between eDNA metabarcoding and visual/capture methods ($P < 0.05$). Best means
686 preferred methods. P-values was calculated with 999 permutations after Sidak's
687 correction of the multiple testing.
688

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

689

690

691

692 **Figure legends**

693

694 Figure 1. Venn diagrams for the number of detected taxa. eDNA metabarcoding
695 (eDNA, blue) and visual/capture methods (V/C, red) in each site of the five study river
696 systems. The bottom diagrams (All) showed the number of detected taxa in the river
697 systems. The numbers represent the number of taxa.

698

699 Figure 2. Nonmetric multidimensional scaling (NMDS) ordination with Jaccard index
700 (with Raup-Crick index in Fig. S4). Fish communities evaluated by the study rivers
701 (shape) and each segment (colored) of the river. MDS stress was 0.158.

702

703 Figure 3. Pairwise effect size (SES, standard effect size). Comparison between
704 community nestedness and species turnover between the communities detected by
705 eDNA metabarcoding and visual/capture methods. Error bars indicate 95% confidence
706 interval. The horizontal dotted line represents SES = 0, indicating a non-significant
707 effect.

708

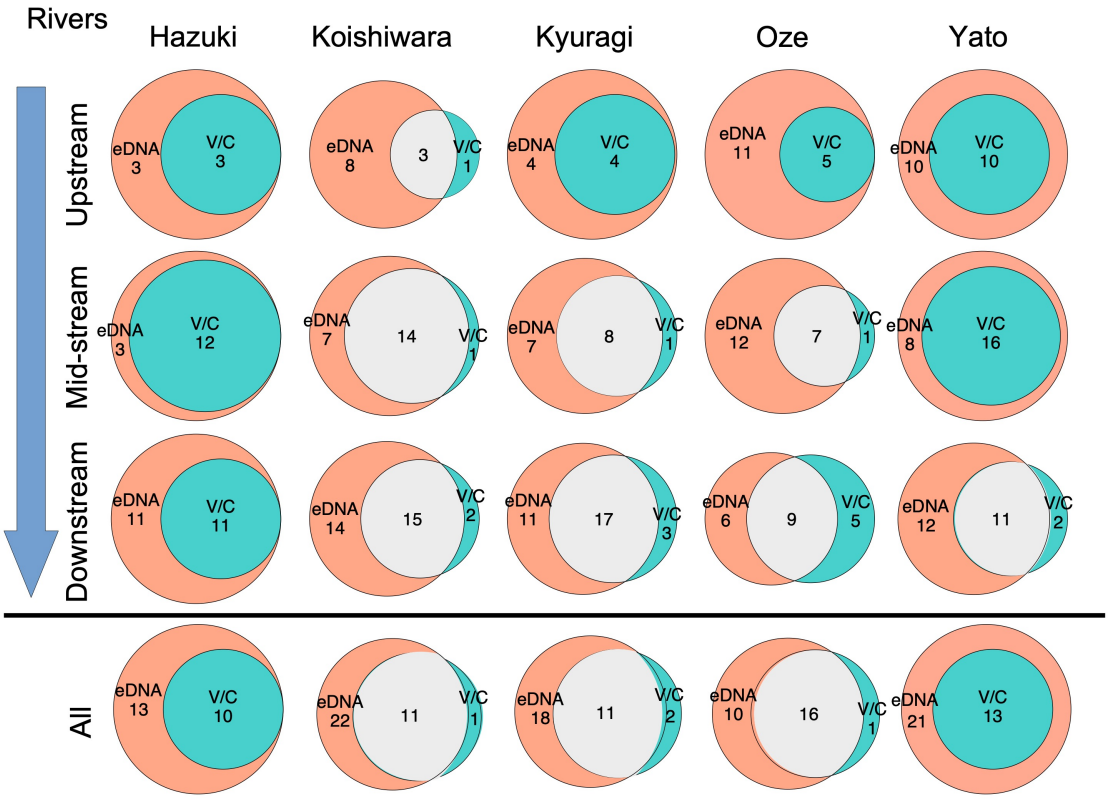
709 Figure 4. Effect size (SES, standard effect size). a) community nestedness and b)
710 species turnover SES among the segment communities evaluated by eDNA
711 metabarcoding and visual/capture methods. The error bars indicate 95% confidence
712 interval. The horizontal dotted line represents SES = 0, indicating non-significant effect.

713

714

715

716

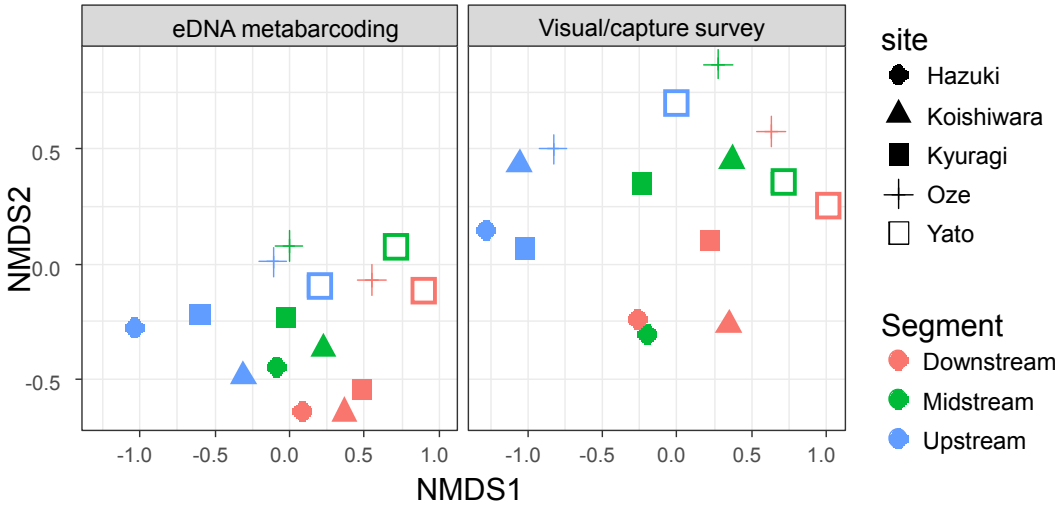


717

718 Figure 1

719

720



721

722 Figure 2

723

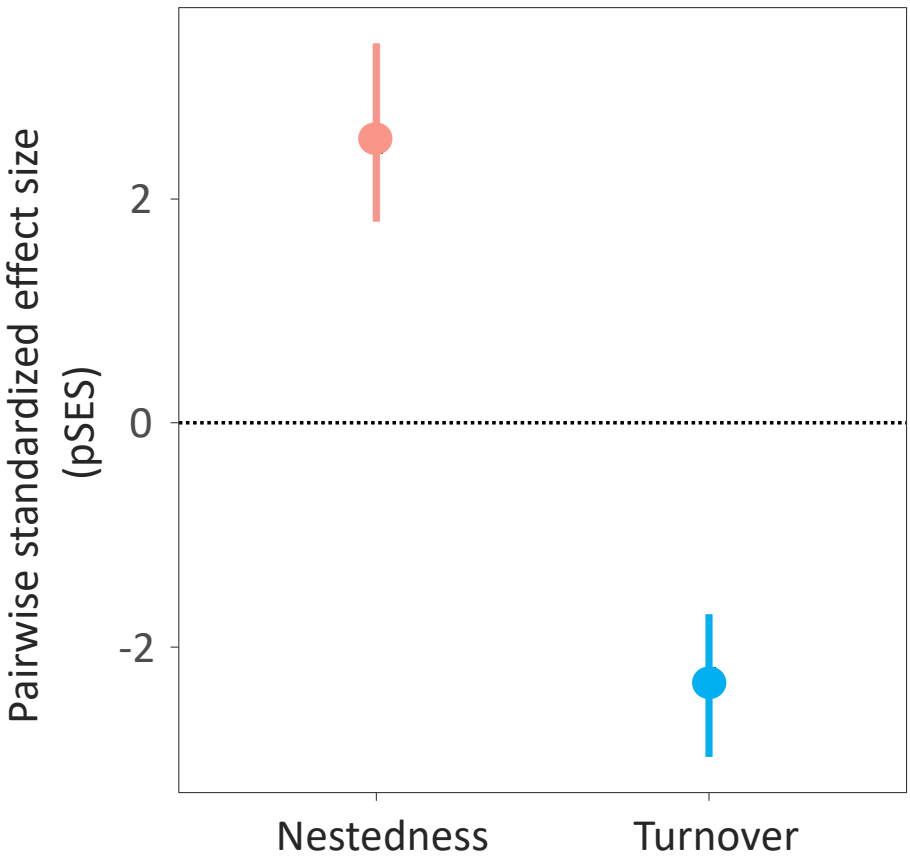

Pairwise standardized effect size
(pSES)
2
0
-2
Nestedness
Turnover

Figure 3

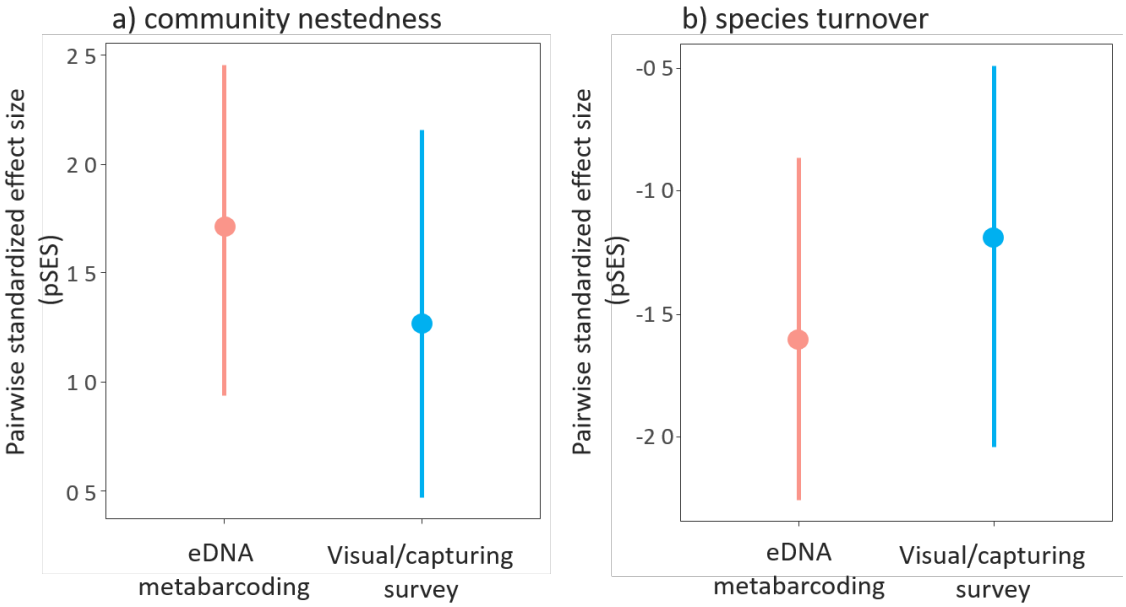

a) community nestedness
Pairwise standardized effect size (pSES)
2 5
2 0
1 5
1 0
0 5
eDNA metabarcoding
Visual/capturing survey
b) species turnover
Pairwise standardized effect size (pSES)
-0 5
-1 0
-1 5
-2 0
eDNA metabarcoding
Visual/capturing survey

Figure 4