

Fine-tuning the performance of abundance estimation based on environmental DNA (eDNA) focusing on eDNA particle size and marker length

Toshiaki Jo¹ and Hiroki Yamanaka²

¹Ryukoku Daigaku

²Ryukoku University

March 10, 2022

Abstract

Reliable abundance estimation is a primary challenge in environmental DNA (eDNA) analysis, which has been addressed by considering the effects of eDNA transport and degradation. However, these eDNA spatial dynamics depend on the cellular and molecular structure of eDNA, with its persistence state (particle size and DNA fragment length) being essential for improved abundance estimation. This existing knowledge gap is bridged by utilizing datasets obtained from two types of aquarium experiments (targeting zebrafish [*Danio rerio*] and Japanese jack mackerel [*Trachurus japonicus*]) and comparing the relationships between eDNA concentration and species abundance among different eDNA size fractions and target marker lengths. We reared the fish in experimental tanks with different individual numbers or biomass densities, filtered rearing water using different pore size filters, and quantified eDNA concentrations targeting different fragment lengths or genetic regions. Consequently, both experiments showed that the accuracy and sensitivity in abundance estimation were improved (i.e., R^2 values and slopes of linear regressions increased) when targeting eDNA at the 3–10- μm size fraction. On the other hand, targeting eDNA at the $>10\ \mu\text{m}$ size fraction yielded a lower R^2 value. This result indicates that an “appropriately” larger eDNA particle is vital for improving abundance estimation accuracy and sensitivity. Conversely, the target marker length negatively affected the R^2 value. This study proposes that the relationship between eDNA concentration and species abundance relies on the complex interactions between the particle size, persistence, and spatial heterogeneity of eDNA in water.

Title:

Fine-tuning the performance of abundance estimation based on environmental DNA (eDNA) focusing on eDNA particle size and marker length

Authors:

Toshiaki Jo^{1,2}, Hiroki Yamanaka^{1,2}

Affiliations:

¹Faculty of Advanced Science and Technology, Ryukoku University: 1-5, Yokotani, Oe-cho, Seta, Otsu City, Shiga 520-2194, Japan

²Ryukoku Center for Biodiversity Science: 1-5, Yokotani, Oe-cho, Seta, Otsu City, Shiga 520-2194, Japan

Corresponding author:

Toshiaki Jo

Faculty of Advanced Science and Technology, Ryukoku University: 1-5, Yokotani, Oe-cho, Seto, Otsu City, Shiga 520-2194, Japan

Tel: +81-77-544-7113

Email: toshiakijo@gmail.com

Running head:

Abundance estimation and eDNA state

Keywords:

abundance estimation; DNA fragment length; environmental DNA (eDNA); filter pore size; linear regression; R^2 value

Abstract

Reliable abundance estimation is a primary challenge in environmental DNA (eDNA) analysis, which has been addressed by considering the effects of eDNA transport and degradation. However, these eDNA spatial dynamics depend on the cellular and molecular structure of eDNA, with its persistence state (particle size and DNA fragment length) being essential for improved abundance estimation. This existing knowledge gap is bridged by utilizing datasets obtained from two types of aquarium experiments (targeting zebrafish [*Danio rerio*] and Japanese jack mackerel [*Trachurus japonicus*]) and comparing the relationships between eDNA concentration and species abundance among different eDNA size fractions and target marker lengths. We reared the fish in experimental tanks with different individual numbers or biomass densities, filtered rearing water using different pore size filters, and quantified eDNA concentrations targeting different fragment lengths or genetic regions. Consequently, both experiments showed that the accuracy and sensitivity in abundance estimation were improved (i.e., R^2 values and slopes of linear regressions increased) when targeting eDNA at the 3–10- μm size fraction. On the other hand, targeting eDNA at the >10 μm size fraction yielded a lower R^2 value. This result indicates that an “appropriately” larger eDNA particle is vital for improving abundance estimation accuracy and sensitivity. Conversely, the target marker length negatively affected the R^2 value. This study proposes that the relationship between eDNA concentration and species abundance relies on the complex interactions between the particle size, persistence, and spatial heterogeneity of eDNA in water.

Introduction

Environmental DNA (eDNA) is the total pool of DNA isolated from environmental samples (Pawlowski et al., 2020). Macroorganisms such as fish are believed to produce eDNA in the form of epidermis, mucus, and feces into their outer environments (Barnes & Turner, 2016; Rodriguez-Ezpeleta et al., 2021). Recently, the eDNA-based monitoring of species distribution, abundance, and composition has been developed to target a variety of taxa in aquatic and terrestrial environments (e.g., Ficetola et al., 2008; Yamanaka & Minamoto, 2016; Ushio et al., 2018; Valentin et al., 2020). The target eDNA, which is detected by a polymerase chain reaction (PCR), allows to evaluate species presence and relative abundance, making eDNA analysis a non-disruptive, cost-effective, and high-sensitivity monitoring tool compared with traditional capture-based surveys (e.g., Thomsen et al., 2012; Miya et al., 2015; Jo et al., 2020a; Lopes et al., 2021). Although there is potential for eDNA-based biomonitoring to become an essential approach for biodiversity and ecosystem conservation, knowledge on the production source and persistence state of eDNA, as well as its transport and degradation, is lacking (Barnes & Turner, 2016; Harrison et al., 2019; Jo et al., 2021a). The lack of or scarcity in understanding the characteristics and dynamics of eDNA causes uncertainty in eDNA-based species inferences, often resulting in ecological interpretation difficulty (Hansen et al., 2018).

Specifically, the estimation of species abundance is a significant challenge in eDNA analysis (Roussel et al., 2015; Harper et al., 2018). Although the eDNA concentration shows a positive correlation with target species abundance for multiple taxa (Takahara et al., 2012; Pilliod et al., 2013; Uthicke et al., 2018; Wu et al., 2018; Ponce et al., 2021), in natural environments, these relationships are weakened compared with

those in controlled laboratory conditions (Yates et al., 2019). Numerous environmental factors complicate the diffusion, retention, and degradation of eDNA in natural environments (e.g., temperature fluctuation, water chemistry, and hydrogeographic conditions), which renders eDNA quantification and its relationship with species abundance in the field unclear. A few studies previously addressed this challenge by correcting the eDNA concentration with a river flow (e.g., eDNA flux), mathematically modeling the processes of eDNA transport and degradation (Carraro et al., 2018; Levi et al., 2019; Fukaya et al., 2021). Nevertheless, the practical application of eDNA-based abundance estimation with high reliability in natural environments is still in its infancy.

Given that eDNA degradation significantly depends on the complex interactions between the eDNA state and multiple abiotic factors (Jo & Minamoto, 2021), the cellular and molecular structures of eDNA function as important factors to better understand the relationship between eDNA concentration and species abundance. Furthermore, a recent review discussed the benefit of targeting various types of eDNA beyond short mitochondrial DNA fragments (nuclear eDNA, longer eDNA fragments, and larger eDNA particles) for reliable species detection and accurate abundance estimation (Jo et al., 2021). Larger eDNA particles (eDNA detected in more significant size fractions) can more frequently possess longer eDNA fragments (Jo et al., 2020b) that persist for a shorter duration in water because of the inflow of degraded eDNA from larger to smaller size fractions (Jo et al., 2019). This finding suggests that the larger eDNA particles, likely derived from intra-cellular DNA (e.g., cell and tissue fragments), are released and assumed to be fresher and more precise biological signals than their smaller counterparts (e.g., organelle and extra-cellular DNA) (Jo et al., 2019; 2021). Therefore, the larger eDNA particles collected using a larger pore filter may accurately imitate the neighboring species abundance.

Conversely, we suspected that the larger eDNA particles do not continuously improve the species abundance estimation accuracy. Although much of macrobial eDNA is concentrated in a 1–10- μm size fraction (Turner et al., 2014; Jo et al., 2019; Zhao et al., 2021), it exists in water at various size fractions (<0.2 to >180 μm in diameter). Among them, eDNA particles collected by a filter with a considerable pore size (tens or hundreds of micrometers, e.g., tissue aggregation and scale) are clumped and distributed heterogeneously in water (Furlan et al., 2016). The stochasticity of retrieving such “huge” eDNA particles in a water sample could worsen the relationship between eDNA concentration and species abundance. Hence, the eDNA particle size and size fraction may associate nonlinearly with the relationship between eDNA concentration and species abundance.

Additionally, it is believed that longer eDNA fragments are degraded rapidly (e.g., Jo et al., 2017; Shogren et al., 2018), whereas agreement is not usually reached in previous studies (e.g., Bylemans et al., 2018). The longer eDNA fragments can persist for a shorter duration in water owing to its rapid degradation (Jo et al., 2017). Given its characteristics, the longer eDNA fragments also reflect more recent biological information than the shorter fragments. However, compared with shorter fragments, longer eDNA fragments are expected to be less frequently detected, specifically in natural environments (Jo et al., 2021b). The stochasticity and heterogeneity of longer eDNA fragments lead to higher variances of eDNA yields across replicated samplings, showing their poor applicability for accurate abundance estimation. Altogether, it should be evaluated how target marker length and particle size of target eDNA could influence the relationship between eDNA concentration and species abundance.

This study compared the relationships between eDNA concentration and species abundance among different eDNA size fractions and target marker lengths and proposed the characteristics of eDNA suitable for accurate abundance estimation. To address the issue, an aquarium experiment was performed using zebrafish (*Danio rerio*). We reared the fish in experimental tanks with different individual densities, filtered the rearing water samples using other pore size filters, and quantified eDNA concentrations targeting different fragment lengths. In this study, we focused on the R^2 values and slopes of the linear regressions as the parameters showing the relationship between eDNA concentration and species abundance; the former parameter represented the accuracy of the relationship, and the latter the sensitivity of eDNA concentration in response to changes in species abundance (Eichmiller et al., 2016). Furthermore, we reanalyzed the dataset from a previous study

that measured the particle size distribution of Japanese jack mackerel (*Trachurus japonicus*) eDNA (Jo et al., 2019), and assessed the importance of the eDNA particle size and size fraction for the abundance estimation performance.

Materials and methods

Aquarium experiment

The aquarium experiment was performed to examine how the relationships between fish abundance and target eDNA concentration varied depending on filter pore size and DNA fragment length. Five replicates of 25-L tanks, filled with 20 L of aged tap water aerated by a pump, were prepared. The water temperature was regulated (around 25°C) using a thermostat (DS150, GEX, Japan), and a 12-h/12-h light/dark cycle was adopted throughout the experiment. After a week, 500 mL of tank water was collected from each tank in advance, using a 1-L plastic beaker as a blank sample, and subsequently, five individuals of the juvenile zebrafish (total length: around 30 mm; no sex identification) were introduced into each tank. Then, after 5 days, 3 L of water samples were collected from each tank using a 10-L plastic container, introducing an additional five juvenile zebrafish (i.e., ten individuals per tank in total). After another 5 days, we collected 3 L of water samples similarly as above and introduced an additional ten individuals (i.e., 20 individuals per tank). Similarly, we collected water samples after 6 days and introduced an additional ten individuals (i.e., 30 individuals per tank). After 10 days, 3 L of water samples were collected, and ten additional individuals were introduced (i.e., 40 individuals per tank). Finally, after 5 days, we collected 3 L of water samples.

After water collection, the water samples were thoroughly mixed and 250 mL of them was filtered using a 47-mm-diameter polycarbonate membrane filter (0.2 µm pore size; Merck Millipore, Germany) using magnetic filter funnels (500 mL capacity; Pall Corporation, US) and a filtration manifold (three-place; Sanplatec Corporation, Japan). A similar procedure was performed for 500 mL of the samples with the same material filters with different pore sizes (0.8, 3, and 10 µm pore size). This alternative was performed as a result of clogging of smaller pore size filters. Additionally, 250 mL of distilled water was filtered with the same material filter at each sampling time, with 0.2-µm pore size filter used as a negative filtration control. All filtered samples were covered with a commercial aluminum foil and kept at -20°C until eDNA extraction. The fish in experimental tanks were fed a small amount of commercial bait every 2 or 3 days. The bottom of each tank was cleaned immediately after feeding to eliminate the effect of the feces. Tank water was filled up to the original water volume after water sampling and bottom cleaning was performed. We supplied replacements for the dead or dying individuals as soon as possible. However, we replaced them after water sampling when the dead or dying individuals were observed on the sampling day. Disposable gloves were worn to collect and filter water samples throughout the experiments. All the equipment for fish rearing (pump, air stones, tube, and acrylic tanks), water collection (beakers, containers, polyethylene tanks), and water filtration (filter funnels and tweezers) were bleached before every use in 0.1% sodium hypochlorite solution for at least 5 min (Yamanaka et al., 2017).

eDNA extraction and quantitative real-time PCR (qPCR)

We extracted total eDNA on the filter using a DNeasy Blood and Tissue Kit (Qiagen, Germany), according to the method described in Minamoto et al. (2019). The zebrafish eDNA concentration in the water sample was estimated by quantifying the copy number of mitochondrial genes using the StepOnePlus Real-Time PCR system (Thermo Fisher Scientific, U.S.). We used the primer/probe sets to amplify the different lengths of mitochondrial genes, including cytochrome b, tRNA-Glu, and ND6 regions of target species (132, 430, 715, and 1021 bp) developed by Hirohara et al. (2021). Each 15 µL of TaqMan reaction mixture contained a 2 µL DNA template, a final 900 nM concentration of each forward and reverse primer, and 125 nM of TaqMan probe in 1 × TaqPath qPCR Master Mix, CG (Thermo Fisher Scientific). Simultaneously, 2 µL of pure water samples were analyzed as the PCR negative controls. Target eDNA concentrations were quantified based on a dilution series of standards containing 3×10^1 to 3×10^4 copies of synthesized artificial DNA fragments from zebrafish mitochondrial genes (Hirohara et al., 2021). All qPCRs were performed in triplicate, including eDNA samples, standards, and negative controls. Thermal conditions of the qPCRs were 2 min at 50°C, 10

min at 95°C, 55 cycles of 15 s at 95°C, and 1 min (132 bp) or 1.5 min (other fragment lengths) at 60°C (Hirohara et al., 2021). We calculated the eDNA concentrations by averaging the triplicates, and each PCR negative replicate (indicating non-amplification) was denoted as containing zero copies (Ellison et al., 2006).

Statistical analyses

The statistical analyses were performed using the R version 4.0.4 software program (R Core Team, 2021). Target eDNA concentrations per PCR reaction (2 μ L template DNA) were converted into per water sample (1 mL water sample) before the analyses. Linear regressions of the target eDNA concentration (log10-transformed) against the number of zebrafish individuals per tank were performed for each filter pore size and DNA marker length using *lm* and *confint* functions, with their R^2 values and slopes with 95% confidence intervals (CIs) estimated. We also conducted a linear model to assess the effects of the eDNA size fraction and target marker length on the accuracy of the relationship between eDNA concentration and species abundance (i.e., the R^2 value of linear regression). The R^2 values were Fisher’s z-transformed by the package “MAc” in advance (Del Re & Hoyt, 2018) to meet the normality (Yates et al., 2019) and included as the dependent variable. Filter pore size (μ m; categorical), marker length (bp; log10-transformed), and interactions were also explanatory variables. Additionally, a linear mixed model using the *lmer* function in the package “lmerTest” (Kuznetsova et al., 2017) was performed to assess the effects of the eDNA size fraction and marker length on the sensitivity of the eDNA concentration in response to changes in species abundance (i.e., the slope of linear regression [Eichmiller et al., 2016]). The zebrafish eDNA concentration (log10-transformed) was included as the dependent variable. The number of fish individuals per tank, filter pore size, marker length (log10-transformed), and primary interactions of fish individuals with filter pore size and log10-transformed marker length were included as the fixed effects. The tank replicates were included as the random effect.

Re-analyses of relationships between size-fractionated eDNA and fish abundance

We reanalyzed the dataset from a previous study to further examine the effects of particle size and size fraction on eDNA-based abundance estimation (Jo et al., 2019). The aforementioned study measured the size distribution of *T. japonicus* eDNA derived from mitochondrial and nuclear DNA. Briefly, rearing water was sequentially filtered using different pore size filters; the particle size distributions of target eDNA were compared among rearing temperature, fish biomass in a tank, and target genes (mitochondrial or nuclear). Detailed information on the experimental design, water sampling, and molecular analyses can be accessed in Jo et al. (2019). Here, we performed linear regressions using the dataset and assessed the variation in the relationship between eDNA concentration and fish biomass according to eDNA size fraction. The eDNA samples that were collected a day before the fish removal and passed through sequential filters with 10, 3, 0.8, and 0.4 μ m pore sizes (i.e., >10, 3–10, 0.8–3, and 0.4–0.8 μ m size fractions, respectively) were compiled. The fish biomass in 200-L tanks ranged from 5.8 to 460.1 g. We then calculated two kinds of cumulative eDNA concentrations from each size-fractionated eDNA concentration as follows:

(i) Upside-cumulative eDNA (UC) meaning the cumulative eDNA concentration at all size fractions that can be collected by a given pore size filter:

$$UC_{>10 \mu\text{m}} = C_{>10 \mu\text{m}}$$

$$UC_{>3 \mu\text{m}} = C_{>10 \mu\text{m}} + C_{3-10 \mu\text{m}}$$

$$UC_{>0.8 \mu\text{m}} = C_{>10 \mu\text{m}} + C_{3-10 \mu\text{m}} + C_{0.8-3 \mu\text{m}}$$

$$UC_{>0.4 \mu\text{m}} = C_{>10 \mu\text{m}} + C_{3-10 \mu\text{m}} + C_{0.8-3 \mu\text{m}} + C_{0.4-0.8 \mu\text{m}}$$

(ii) Downside-cumulative eDNA (DC) meaning the cumulative eDNA concentration at all size fractions that can be collected by the smallest pore size filter (0.4 μ m) following pre-filtration with a given pore size filter:

$$DC_{>0.4 \mu\text{m}} = C_{0.4-0.8 \mu\text{m}} + C_{0.8-3 \mu\text{m}} + C_{3-10 \mu\text{m}} + C_{>10 \mu\text{m}}$$

$$DC_{0.4-10 \mu\text{m}} = C_{0.4-0.8 \mu\text{m}} + C_{0.8-3 \mu\text{m}} + C_{3-10 \mu\text{m}}$$

$$DC_{0.4-3 \mu\text{m}} = C_{0.4-0.8 \mu\text{m}} + C_{0.8-3 \mu\text{m}}$$

$$DC_{0.4-0.8 \mu\text{m}} = C_{0.4-0.8 \mu\text{m}}$$

where C_X means the concentration of target eDNA at $X \mu\text{m}$ size fraction (i.e., size-fractionated eDNA). As described above, we performed linear regressions of size-fractionated (C_X) and both cumulative eDNA (UC_X or DC_X) concentrations (\log_{10} -transformed + 1) against fish biomass (\log_{10} -transformed) and estimated their R^2 values and regression slopes with 95% CIs. Additionally, the relative R^2 values were calculated to simplify the effect of size fraction on the relationships, where the R^2 values at a given size fraction were divided by the R^2 value at $>10 \mu\text{m}$ size fraction for the size-fractionated (C_X) and upside-cumulative (UC_X) eDNA and at $>0.4 \mu\text{m}$ size fraction for the downside-cumulative eDNA (DC_X) corresponding to a similar temperature level and marker type.

Moreover, we assessed the effect of eDNA size fraction on the R^2 values and slopes of linear regressions, using the three types of eDNA concentrations and linear mixed models. In the former models, the R^2 values (Fisher's z-transformed) were included as the dependent variable, each size fraction (μm) was included as the fixed effect, and temperature levels (13, 18, 23, and 28°C) and marker types were included as the random effects. On the other hand, in the latter models, every kind of eDNA concentration (C_X , UC_X , or DC_X ; \log_{10} -transformed + 1) was included as the dependent variable, fish biomass (\log_{10} -transformed), each size fraction, and their interaction were included as the fixed effects, and tank replicates, temperature levels, and marker types as the random effects. In these analyses, water temperature levels and marker types were not included as the fixed effects because these factors were not of interest in this study.

Results

Aquarium experiment

The linear relationships between zebrafish eDNA concentrations and fish abundance were observed in all filter pore sizes and DNA fragment lengths (Figure 1). However, there were substantial variations in the R^2 values and slopes in the linear regressions. The R^2 values were significantly lower for a $10 \mu\text{m}$ pore size filter than the other pore size filters and the values were higher by targeting shorter eDNA fragments (Figure 2a). A linear model showed a significantly positive effect of a $3 \mu\text{m}$ pore size filter relative to a $10 \mu\text{m}$ pore size filter ($P < 0.05$) and a negative effect of DNA marker length ($P < 0.01$) on Fisher's z-transformed R^2 values (Table 1). Furthermore, the slopes tended to be larger for 3 and $0.2 \mu\text{m}$ pore size filters than for 10 and $0.8 \mu\text{m}$ pore size filters, regardless of marker length (Figure 2b). A linear mixed model showed significant positive effects in the interactions of 3 and $0.2 \mu\text{m}$ pore size filters with the number of fish individuals (both $P < 0.01$) on the slopes. Concurrently, no significant interaction between zebrafish abundance and DNA fragment length was observed (Table 2). The overall R^2 values and PCR efficiencies of the standard curves in qPCR were 0.996 ± 0.002 and $93.814\% \pm 6.174\%$, respectively (detailed information on each marker can be seen in Table S1). Although one of blank samples showed PCR amplification, its calculated concentration was 0.4 copies per PCR reaction (Table S1); the potential contamination during water filtration was considered negligible. Amplification was not observed in any of the filtration or PCR negative controls throughout the experiment.

Re-analyses of relationships between size-fractionated eDNA and fish abundance

With regard to the size-fractionated eDNA (C_X), the 95% CIs of the relative R^2 values at the $0.8-3 \mu\text{m}$ size fraction were markedly lower than those at the $>10 \mu\text{m}$ size fraction and did not include $R^2 = 1$ (relative $R^2 = 0.389$ [-0.026, 0.804]; Figure 3a). Linear mixed models displayed that the R^2 values and slopes (i.e., the interaction between fish biomass and size fraction) were significantly lower at $0.8-3 \mu\text{m}$ than at $>10 \mu\text{m}$ size fractions (both $P < 0.01$) (Tables 3a & 4a; Figures S1 & S2). On the one hand, regarding the upside-cumulative eDNA (UC_X), the 95% CIs of the relative R^2 values at the $>3 \mu\text{m}$ size fraction were substantially higher than those at the $>10 \mu\text{m}$ size fraction and without including $R^2 = 1$ (relative $R^2 = 1.257$ [1.020, 1.495]; Figure 3b). Additionally, linear mixed models showed that the R^2 values were marginally higher at $>3 \mu\text{m}$ than at $>10 \mu\text{m}$ size fractions ($P = 0.065$) and the slope was slightly lower at $>0.8 \mu\text{m}$ than at $>10 \mu\text{m}$ size fractions ($P = 0.090$) (Tables 3b & 4b; Figures S3 & S4). On the other hand, regarding the downside-cumulative eDNA (DC_X), the 95% CIs of the relative R^2 values at the $0.4-3 \mu\text{m}$ size fraction were

considerably lower than those at the $>0.4 \mu\text{m}$ size fraction and did not include $R^2 = 1$ (relative $R^2 = 0.587$ [0.376, 0.797]; Figure 3c). The linear mixed models showed that the R^2 -values were significantly lower at $0.4\text{--}3 \mu\text{m}$ than at $>0.4 \mu\text{m}$ size fractions ($P < 0.01$), with no significant interactions between fish biomass and size fractions (Tables 3c & 4c; Figures S5 & S6).

Discussion

Abundance estimation depending on eDNA particle size and size fraction

The aquarium experiment using zebrafish revealed that regardless of DNA marker length, eDNA particles collected by a $10 \mu\text{m}$ pore size filter could not explain the fish abundance more accurately than those managed by other smaller pore size filters. Notably, we observed higher R^2 values between the shortest eDNA fragments (132 bp) and fish abundance by increasing the filter pore sizes. However, the value radically decreased when using the $10 \mu\text{m}$ pore size filter, suggesting that the relationship between eDNA concentration and species abundance can be worsened when such “huge” eDNA particles and extremely larger size fractions are targeted. As expected in the Introduction, massive eDNA particles, such as tissue clumps, reflect the species abundance less precisely because of their potential heterogeneous dispersion and distribution in water (Turner et al., 2014; Furlan et al., 2016; Jo et al., 2020b). Moreover, mainly in natural environments, eDNA particle sizes can be increased by adsorbing suspended organic matter, making DNA molecules physiochemically stable and able to survive longer in water (Levy-Booth et al., 2007; Torti et al., 2015). Therefore, detecting such eDNA is likely to reflect past biological information and obscures current species distribution and abundance (Barnes & Turner, 2016; Jo et al., 2021).

On the contrary, this study indicates that eDNA particles at the $3\text{--}10 \mu\text{m}$ size fractions could be important for better performance of abundance estimation. First, the zebrafish eDNA particles collected by a $3 \mu\text{m}$ pore size filter reflected the fish abundance significantly better than those managed by a $10 \mu\text{m}$ pore size filter, corresponding to the result from the experiment targeting the upside-cumulative eDNA from *T. japonicus*. Second, for the downside-cumulative eDNA from Japanese jack mackerel, the R^2 values between the target eDNA concentration and fish biomass were significantly lower at the $0.4\text{--}3 \mu\text{m}$ size fraction (i.e., excluding the $>3 \mu\text{m}$ size fraction) than at the $>0.4 \mu\text{m}$ size fraction, whereas little difference in the R^2 values between the >0.4 and $0.4\text{--}10 \mu\text{m}$ size fractions existed (i.e., even excluding the $>10 \mu\text{m}$ size fraction). Third, in the zebrafish experiment, the slopes of the linear regressions (the sensitivity of the eDNA concentration in response to changes in species abundance) were significantly higher for the eDNA particles collected by a $3 \mu\text{m}$ pore size filter than those managed by a $10 \mu\text{m}$ pore size filter. Summarily, these results suggest that an “appropriately” larger eDNA particle is suitable for accurate and sensitive estimation of the species abundance via eDNA.

The $3\text{--}10 \mu\text{m}$ size fraction is the fraction in which microbial eDNA in water is most abundant (Turner et al., 2014; Jo et al., 2019; Zhao et al., 2021). It is possible that the size fraction mainly includes a single cell and intact nuclei and mitochondria of known sizes. Therefore, eDNA sampling focused on the specific size fraction where target eDNA is abundant may lessen false-negative eDNA detection and variance of eDNA quantification, consequently estimating species abundance more accurately and sensitively. Alternatively, different physiological origins and physiochemical structures of eDNA might result in its different persistence and spatial dispersion, hence influencing the relationship between eDNA concentration and species abundance. Additionally, the $>3 \mu\text{m}$ pore size filter generally reduces filter clogging, increasing the filtration volume of the water samples considerably compared with smaller pore size filters (e.g., Kumar et al., 2021). Moreover, Takasaki et al. (2021) recently compared the pre-filtration performance on eDNA detection using various pore size filters ($10\text{--}840 \mu\text{m}$ pore sizes). The result showed that $10 \mu\text{m}$ pre-filtration effectively reduced PCR inhibitors (e.g., humic substances) in river water samples without decreasing the detectability of target eDNA. From the above, it would be promising to collect and analyze target eDNA focusing on the $3\text{--}10 \mu\text{m}$ size fraction to better estimate species abundance.

Conversely, the performance of abundance estimation was poor for eDNA particles at the $<3 \mu\text{m}$ size fraction. According to the experiment targeting size-fractionated eDNA from *T. japonicus*, the selective collection

of eDNA particles at the 0.8–3 μm size fraction tremendously worsened its relationship with fish biomass and sensitivity in response to changes in fish biomass (i.e., the R^2 value and slope). This trend could partly contribute to the result that the slopes of linear regression tended to be lower for the Japanese jack mackerel’s eDNA particles at the $>0.8 \mu\text{m}$ size fraction than that at the $>10 \mu\text{m}$ size fraction. Furthermore, variations of the abundance estimation accuracy via eDNA across filter pore sizes in some previous studies are likely attributed to eDNA particles at the 0.8–3 μm size fraction. Takahara et al. (2012) assessed the relationships between common carp (*Cyprinus carpio*) eDNA concentration and fish biomass using different filtration methods, reporting a better R^2 value for the 3 μm pore size filter (i.e., $>3 \mu\text{m}$ size fraction; $R^2 = 0.93$) than the 0.8 μm pore size filter following a 12- μm pre-filtering (i.e., 0.8–12 μm size fraction; $R^2 = 0.85$). Eichmiller et al. (2016) also found that the R^2 values between common carp eDNA concentration and fish biomass were higher for 5 μm pore size filters ($R^2 = 0.86$) than for 1 μm pore size filters ($R^2 = 0.71$). Moreover, with the frequent non-detection (zero copies) of *T. japonicus* eDNA at the 0.4–0.8 μm size fraction, the R^2 values and slopes in the linear regressions at the 0.4–0.8 μm size fraction were biased and overestimated by a zero-inflation of the target eDNA concentration. Sequel to this, it agrees to a lower proportion of fish eDNA at the $<0.8 \mu\text{m}$ size fraction (20% at most; Turner et al., 2014).

Abundance estimation depending on target marker length

The zebrafish aquarium experiment showed that target marker length negatively influenced the abundance estimation accuracy (R^2 value), regardless of filter pore size. Although eDNA detectability declined with longer DNA fragment length, the detection of longer eDNA fragments may help identify the exact location of target species because of its shorter persistence in water (Hänfling et al., 2016). However, this advantage of longer eDNA fragments was not fully utilized in a closed and small-scale environment like our experiment. According to the studies by Jo et al. (2017; 2021b), who used longer eDNA fragments in the field, the more satisfactory relationship between longer eDNA fragments and fish abundance could be accounted for by (i) the narrower spatiotemporal range of longer eDNA fragments and (ii) the removal of highly degraded and “non-fresh” eDNA signals probably derived from fish carcasses, resuspension from bottom sediment, and contamination from boat and fishing gear. Further empirical studies are required in natural environments to clarify the situation where abundance estimation via longer eDNA fragments is adequate and to investigate the detailed mechanism of eDNA degradation depending on DNA fragment length.

Conclusions and perspectives

On the basis of the findings in the present study, we summarized the performance of eDNA-based abundance estimation depending on eDNA particle size and size fraction (Figure 4). As shown in this figure, the relationship between eDNA concentration and species abundance could complicatedly rely on the cellular structure of eDNA; the cellular structure of eDNA (e.g., intra-/extra-cellular, particle size, dissolved/adsorbed) might multifacetedly impact the processes of its degradation and persistence in water and the heterogeneity of spatial dispersion and distribution, eventually influencing the correlation between the steady-state eDNA concentration and species abundance. What is especially important in this study is that better understanding and utilizing the characteristics and dynamics of eDNA can promote the refinement of its application for biomonitoring. We here proposed a hypothesis that the accuracy and sensitivity of eDNA-based abundance estimation could be improved by targeting “appropriately” larger eDNA particles (indicated as 3–10 μm in diameter).

Nonetheless, given the substantial lack of knowledge on its production source and persistence state in water, it remains unknown in what physiological origin and physiochemical structure eDNA exists in each size fraction. To explain the mechanism underlying the findings in this study, future empirical research on the production source and persistence state of eDNA is required. For instance, detecting tissue-specific messenger RNA in water could identify the production source of eDNA (Tsuru et al., 2021). Alternatively, visualization of eDNA particles via fluorescence *in situ* hybridization (FISH) techniques and immunostaining (e.g., Kogure et al., 2006; Amann & Fuchs, 2008) may directly offer information on its cellular and the molecular structure. Although different particle sizes of eDNA can result in other transport and diffusion dynamics, the effect of the eDNA state (e.g., particle size, weight) on its transport and diffusion dynamics

has not been elucidated (Shogren et al., 2016; Jo et al., 2021a). The additional information would explain the heterogeneous distribution of eDNA in water, associating its multiple states, and help support our hypothesis above. Clumped eDNA particles, such as aggregates of cells and tissues, might show more limited diffusion but more rapid settling.

Furthermore, toward a practical application of our findings, evaluating whether the use of appropriately larger eDNA particles can be adequate for improved abundance estimation in the field is required. Various environmental parameters can complicatedly influence the production, transport, and degradation processes of different eDNA particles (Harrison et al., 2019; Jo & Minamoto, 2021; Yates et al., 2021), consequently driving the relationship between eDNA concentration and species abundance. It may also cause the potential gap between eDNA yields and its particle size distribution between controlled and natural environments, possibly complicating the improvement of abundance estimation accuracy via eDNA sampling focused on appropriately larger eDNA particles in the field. Therefore, even if our findings were proper in principle, the significance of eDNA sampling considering its state for improved species abundance estimation should be further assessed in mesocosm experiments and natural environments. Additionally, the eDNA particle size may recount a spatiotemporal range of biological information inferred by eDNA. For instance, if a larger eDNA particle shows the presence of target species near a given eDNA sampling site, it may be suitable for the accurate estimation of instantaneous and local species abundance. Conversely, if a research interest is a comprehensive abundance estimation at the ecosystem level (i.e., possibly on timescales of weeks to months), the use of a smaller pore size filter and multiple particle sizes of eDNA may be acceptable (Jo et al., 2021; Yates et al., 2021). Addressing these issues and optimizing the methodology of eDNA analysis will ensure the reliable and sensitive monitoring of species distribution and abundance, contributing to the efficient management of biodiversity conservation and fisheries resources.

Acknowledgments

This work was supported by the Environment Research and Technology Development Fund from the Ministry of the Environment, Japan (Grant Number JPMEERF20204004) and the Grant-in-Aid for Scientific Research of JSPS KAKENHI (Grant Number JP20H03326), and partly by the Fund for the Promotion of Joint International Research of JSPS KAKENHI (Grant Number JP19KK0189). Additionally, we thank Crimson Interactive Japan Co., Ltd. (www.enago.jp) for English proofreading. There is no conflict of interest to declare.

Data accessibility

All the eDNA quantification in the aquarium experiment conducted in this study is available in the Supplemental Information.

Author contributions

Both authors conceived the study. T.J. performed aquarium experiments, compiled the datasets in previous studies, analyzed the data, and wrote the first draft of the manuscript. Both authors edited and provided feedback on the manuscript.

Supporting Information

Table S1. Raw values of zebrafish eDNA concentrations quantified in the aquarium experiment.

Figure S1. The relationships between the size-fractionated Japanese jack mackerel eDNA concentration and fish biomass among temperature levels and marker types.

Figure S2. The R^2 values and slopes in the linear regressions for the size-fractionated Japanese jack mackerel eDNA among temperature levels and marker types.

Figure S3. The relationships between the upside-cumulative Japanese jack mackerel eDNA concentration and fish biomass among temperature levels and marker types.

Figure S4. The R^2 values and slopes in the linear regressions for the upside-cumulative Japanese jack mackerel eDNA among temperature levels and marker types.

Figure S5. The relationships between the downside-cumulative Japanese jack mackerel eDNA concentration and fish biomass among temperature levels and marker types.

Figure S6. The R^2 values and slopes in the linear regressions for the downside-cumulative Japanese jack mackerel eDNA among temperature levels and marker types.

References

- Amann, R., & Fuchs, B. M. (2008). Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. *Nature Reviews Microbiology*, 6(5), 339-348.
- Barnes, M. A. & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, 17(1), 1-17.
- Bylemans, J., Furlan, E. M., Gleeson, D. M., Hardy, C. M., & Duncan, R. P. (2018). Does size matter? An experimental evaluation of the relative abundance and decay rates of aquatic environmental DNA. *Environmental Science & Technology*, 52(11), 6408-6416.
- Carraro, L., Hartikainen, H., Jokela, J., Bertuzzo, E., & Rinaldo, A. (2018). Estimating species distribution and abundance in river networks using environmental DNA. *Proceedings of the National Academy of Sciences*, 115(46), 11724-11729.
- Del Re, A. C. & Hoyt, W. T. (2018). MAc: Meta-Analysis with Correlations. R package version 1.1.1. <https://CRAN.R-project.org/package=Mac>
- Eichmiller, J. J., Miller, L. M., & Sorensen, P. W. (2016). Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. *Molecular Ecology Resources*, 16(1), 56-68.
- Ellison, S. L., English, C. A., Burns, M. J., & Keer, J. T. (2006). Routes to improving the reliability of low level DNA analysis using real-time PCR. *BMC Biotechnology*, 6(1), 1-11.
- Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, 4(4), 423-425.
- Fukaya, K., Murakami, H., Yoon, S., Minami, K., Osada, Y., Yamamoto, S., ... & Kondoh, M. (2021). Estimating fish population abundance by integrating quantitative data on environmental DNA and hydrodynamic modelling. *Molecular Ecology*, 30(13), 3057-3067.
- Furlan, E. M., Gleeson, D., Hardy, C. M., & Duncan, R. P. (2016). A framework for estimating the sensitivity of eDNA surveys. *Molecular Ecology Resources*, 16(3), 641-654.
- Hänfling, B., Lawson Handley, L., Read, D. S., Hahn, C., Li, J., Nichols, P., ... & Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, 25(13), 3101-3119.
- Hansen, B. K., Bekkevold, D., Clausen, L. W., & Nielsen, E. E. (2018). The sceptical optimist: challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries*, 19(5), 751-768.
- Harper, L. R., Buxton, A. S., Rees, H. C., Bruce, K., Brys, R., Halfmaerten, D., ... & Hänfling, B. (2019). Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. *Hydrobiologia*, 826(1), 25-41.
- Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. *Proceedings of the Royal Society B*, 286(1915), 20191409.

- Hirohara, T., Tsuru, K., Miyagawa, K., Paine, R. T., & Yamanaka, H. (2021). The Application of PMA (Propidium Monoazide) to Different Target Sequence Lengths of Zebrafish eDNA: A New Approach Aimed Toward Improving Environmental DNA Ecology and Biological Surveillance. *Frontiers in Ecology and Evolution*, 9, 277.
- Jo, T., Arimoto, M., Murakami, H., Masuda, R., & Minamoto, T. (2019). Particle size distribution of environmental DNA from the nuclei of marine fish. *Environmental Science & Technology*, 53(16), 9947-9956.
- Jo, T., Fukuoka, A., Uchida, K., Ushimaru, A., & Minamoto, T. (2020a). Multiplex real-time PCR enables the simultaneous detection of environmental DNA from freshwater fishes: a case study of three exotic and three threatened native fishes in Japan. *Biological Invasions*, 22(2), 455-471.
- Jo, T., & Minamoto, T. (2021). Complex interactions between environmental DNA (eDNA) state and water chemistries on eDNA persistence suggested by meta-analyses. *Molecular Ecology Resources*, 21(5), 1490-1503.
- Jo, T., Murakami, H., Masuda, R., & Minamoto, T. (2020b). Selective collection of long fragments of environmental DNA using larger pore size filter. *Science of the Total Environment*, 735, 139462.
- Jo, T., Murakami, H., Masuda, R., Sakata, M. K., Yamamoto, S., & Minamoto, T. (2017). Rapid degradation of longer DNA fragments enables the improved estimation of distribution and biomass using environmental DNA. *Molecular Ecology Resources*, 17(6), e25-e33.
- Jo, T., Sakata, M. K., Murakami, H., Masuda, R., & Minamoto, T. (2021b). Universal performance of benzalkonium chloride for the preservation of environmental DNA in seawater samples. *Limnology and Oceanography: Methods*, 19(11), 758-768.
- Jo, T., Takao, K., & Minamoto, T. (2021a). Linking the state of environmental DNA to its application for biomonitoring and stock assessment: Targeting mitochondrial/nuclear genes, and different DNA fragment lengths and particle sizes. *Environmental DNA*, in press. <https://doi.org/10.1002/edn3.253>
- Kogure, T., Karasawa, S., Araki, T., Saito, K., Kinjo, M., & Miyawaki, A. (2006). A fluorescent variant of a protein from the stony coral *Montipora* facilitates dual-color single-laser fluorescence cross-correlation spectroscopy. *Nature Biotechnology*, 24(5), 577-581.
- Kumar, G., Farrell, E., Reaume, A. M., Eble, J. A., & Gaither, M. R. (2021). One size does not fit all: Tuning eDNA protocols for high-and low-turbidity water sampling. *Environmental DNA*, in press. <https://doi.org/10.1002/edn3.235>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1-26.
- Levi, T., Allen, J. M., Bell, D., Joyce, J., Russell, J. R., Tallmon, D. A., ... & Yu, D. W. (2019). Environmental DNA for the enumeration and management of Pacific salmon. *Molecular Ecology Resources*, 19(3), 597-608.
- Levy-Booth, D. J., Campbell, R. G., Gulden, R. H., Hart, M. M., Powell, J. R., Klironomos, J. N., ... & Dunfield, K. E. (2007). Cycling of extracellular DNA in the soil environment. *Soil Biology and Biochemistry*, 39(12), 2977-2991.
- Lopes, C. M., Baeta, D., Valentini, A., Lyra, M. L., Sabbag, A. F., Gasparini, J. L., ... & Zamudio, K. R. (2021). Lost and found: Frogs in a biodiversity hotspot rediscovered with environmental DNA. *Molecular Ecology*, 30(13), 3289-3298.
- Minamoto, T., Hayami, K., Sakata, M. K., & Imamura, A. (2019). Real-time polymerase chain reaction assays for environmental DNA detection of three salmonid fish in Hokkaido, Japan: Application to winter surveys. *Ecological Research*, 34(1), 237-242.

- Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., ... & Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, 2(7), 150088.
- Pawlowski, J., Apotheloz-Perret-Gentil, L., & Altermatt, F. (2020). Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. *Molecular Ecology*, 29(22), 4258-4264.
- Pilliod, D. S., Goldberg, C. S., Arkle, R. S., & Waits, L. P. (2013). Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Canadian Journal of Fisheries and Aquatic Sciences*, 70(8), 1123-1130.
- Ponce, J. J., Arismendi, I., & Thomas, A. (2021). Using in-situ environmental DNA sampling to detect the invasive New Zealand Mud Snail (*Potamopyrgus antipodarum*) in freshwaters. *PeerJ*, 9, e11835.
- R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rodriguez-Ezpeleta, N., Morissette, O., Bean, C. W., Manu, S., Banerjee, P., Lacoursiere-Roussel, A., ... & Deiner, K. (2021). Trade-offs between reducing complex terminology and producing accurate interpretations from environmental DNA: Comment on "Environmental DNA: What's behind the term?" by Pawlowski et al., (2020). *Molecular Ecology*, 30(19), 4601-4605.
- Roussel, J. M., Paillisson, J. M., Treguier, A., & Petit, E. (2015). The downside of eDNA as a survey tool in water bodies. *Journal of Applied Ecology*, 52(4), 823-826.
- Shogren, A. J., Tank, J. L., Andruszkiewicz, E. A., Olds, B., Jerde, C., & Bolster, D. (2016). Modelling the transport of environmental DNA through a porous substrate using continuous flow-through column experiments. *Journal of the Royal Society Interface*, 13(119), 20160290.
- Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., & Kawabata, Z. (2012). Estimation of fish biomass using environmental DNA. *PLoS ONE*, 7(4), e35868.
- Takasaki, K., Aihara, H., Imanaka, T., Matsudaira, T., Tsukahara, K., Usui, A., ... & Doi, H. (2021). Water pre-filtration methods to improve environmental DNA detection by real-time PCR and metabarcoding. *PLoS ONE*, 16(5), e0250162.
- Thomsen, P. F., Kielgast, J., Iversen, L. L., Moller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS ONE*, 7(8), e41732.
- Torti, A., Lever, M. A., & Jorgensen, B. B. (2015). Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Marine Genomics*, 24, 185-196.
- Tsuri, K., Ikeda, S., Hirohara, T., Shimada, Y., Minamoto, T., & Yamanaka, H. (2021). Messenger RNA typing of environmental RNA (eRNA): A case study on zebrafish tank water with perspectives for the future development of eRNA analysis on aquatic vertebrates. *Environmental DNA*, 3(1), 14-21.
- Turner, C. R., Barnes, M. A., Xu, C. C., Jones, S. E., Jerde, C. L., & Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous microbial eDNA. *Methods in Ecology and Evolution*, 5(7), 676-684.
- Ushio, M., Murata, K., Sado, T., Nishiumi, I., Takeshita, M., Iwasaki, W., & Miya, M. (2018). Demonstration of the potential of environmental DNA as a tool for the detection of avian species. *Scientific reports*, 8, 4493.
- Uthicke, S., Lamare, M., & Doyle, J. R. (2018). eDNA detection of corallivorous seastar (*Acanthaster cf. solaris*) outbreaks on the Great Barrier Reef using digital droplet PCR. *Coral Reefs*, 37(4), 1229-1239.
- Valentin, R. E., Fonseca, D. M., Gable, S., Kyle, K. E., Hamilton, G. C., Nielsen, A. L., & Lockwood, J. L. (2020). Moving eDNA surveys onto land: Strategies for active eDNA aggregation to detect invasive forest

insects. *Molecular Ecology Resources*, 20(3), 746-755.

Wu, Q., Kawano, K., Uehara, Y., Okuda, N., Hongo, M., Tsuji, S., ... & Minamoto, T. (2018). Environmental DNA reveals nonmigratory individuals of *Palaemon paucidens* overwintering in Lake Biwa shallow waters. *Freshwater Science*, 37(2), 307-314.

Yamanaka, H. & Minamoto, T. (2016). The use of environmental DNA of fishes as an efficient method of determining habitat connectivity. *Ecological Indicators*, 62, 147-153.

Yamanaka, H., Minamoto, T., Matsuura, J., Sakurai, S., Tsuji, S., Motozawa, H., ... & Kondo, A. (2017). A simple method for preserving environmental DNA in water samples at ambient temperature by addition of cationic surfactant. *Limnology*, 18(2), 233-241.

Yates, M. C., Cristescu, M. E., & Derry, A. M. (2021). Integrating physiology and environmental dynamics to operationalize environmental DNA (eDNA) as a means to monitor freshwater macro-organism abundance. *Molecular Ecology*, 30(24), 6531-6550.

Yates, M. C., Fraser, D. J., & Derry, A. M. (2019). Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. *Environmental DNA*, 1(1), 5-13.

Zhao, B., van Bodegom, P. M., & Trimbos, K. (2021). The particle size distribution of environmental DNA varies with species and degradation. *Science of the Total Environment*, 797, 149175.

Tables

Table 1. Summary of a linear model to assess the effects of filter pore size and target marker length on the relationship (z-transformed R^2 value) between zebrafish eDNA concentration and fish abundance.

Variable	Estimate	SE	P value
Intercept	1.397	0.307	0.002**
Filter pore size (3 μm)	1.220	0.434	0.023*
Filter pore size (0.8 μm)	0.498	0.434	0.285
Filter pore size (0.2 μm)	0.039	0.434	0.930
log10(Marker length)	-0.424	0.115	0.006**
Filter pore size (3 μm): log10(Marker length)	-0.219	0.162	0.215
Filter pore size (0.8 μm): log10(Marker length)	0.062	0.162	0.714
Filter pore size (0.2 μm): log10(Marker length)	0.186	0.162	0.284

Note: The variable “Filter pore size” estimates were calculated against the 10 μm pore size filter. Asterisks represent the statistical significances of the variables (** $P < 0.01$; * $P < 0.05$).

Table 2. Summary of a linear mixed model to assess the effects of filter pore size and target marker length on the sensitivity (slope; interactions with individuals) of the zebrafish eDNA concentration in response to changes in the number of fish individuals.

Random effect	Groups	Name	Variance	SD
	Tank	Intercept	0.013	0.116
	Residual		0.101	0.318
Fixed effects	Variable	Estimate	SE	P value
	Intercept	3.825	0.254	0.000***
	Individuals	0.025	0.010	0.016*
	Filter pore size (3 μm)	0.481	0.086	0.000***
	Filter pore size (0.8 μm)	0.587	0.086	0.000***
	Filter pore size (0.2 μm)	0.660	0.086	0.000***

Random effect	Groups	Name	Variance	SD
	log10(Marker length)	-0.530	0.091	0.000***
	Individuals: Filter pore size (3 μm)	0.010	0.004	0.004**
	Individuals: Filter pore size (0.8 μm)	0.004	0.004	0.222
	Individuals: Filter pore size (0.2 μm)	0.009	0.004	0.008**
	Individuals: log10(Marker length)	-0.004	0.004	0.306

Note: The estimates in the fixed effects “Filter pore size” were calculated against the 10 μm pore size filter. Asterisks represent the statistical significance of the variables (*** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$).

Table 3. Summary of linear mixed models to assess the effects of size fraction on the relationship (z-transformed R^2 -value) between each type of *T. japonicus* eDNA concentration and fish biomass.

(a) Size-fractionated eDNA concentration (C_X)

Random effect	Groups	Name	Variance	SD
	Temperature	Intercept	0.037	0.192
	Marker	Intercept	0.001	0.038
	Residual		0.031	0.177
Fixed effects	Variable	Estimate	SE	P value
	Intercept	0.511	0.118	0.007**
	Size fraction (3–10 μm)	0.078	0.088	0.385
	Size fraction (0.8–3 μm)	-0.319	0.088	0.001**
	Size fraction (0.4–0.8 μm)	0.104	0.088	0.249

(b) Upside-cumulative eDNA concentration (UC_X)

Random effect	Groups	Name	Variance	SD
	Temperature	Intercept	0.120	0.346
	Marker	Intercept	0.008	0.088
	Residual		0.024	0.155
Fixed effects	Variable	Estimate	SE	P value
	Intercept	0.511	0.192	0.056
	Size fraction (>3 μm)	0.150	0.078	0.065
	Size fraction (>0.8 μm)	0.110	0.078	0.168
	Size fraction (>0.4 μm)	0.121	0.078	0.131

(c) Downside-cumulative eDNA concentration (DC_X)

Random effect	Groups	Name	Variance	SD
	Temperature	Intercept	0.069	0.262
	Marker	Intercept	0.002	0.041
	Residual		0.035	0.187
Fixed effects	Variable	Estimate	SE	P value
	Intercept	0.632	0.150	0.011*
	Size fraction (0.4–10 μm)	-0.037	0.094	0.693
	Size fraction (0.4–3 μm)	-0.305	0.094	0.003**
	Size fraction (0.4–0.8 μm)	-0.017	0.094	0.858

Note: The estimates in the fixed effects “Size fraction” were calculated against the $>10 \mu\text{m}$ size fraction (a & b) and the $>0.4 \mu\text{m}$ size fraction (c). Asterisks represent the statistical significances of the variables (** $P < 0.01$; * $P < 0.05$).

Table 4. Summary of linear mixed models to assess the effects of size fraction on the sensitivity (slope; interactions between fish biomass and size fraction) of each type of *T. japonicus* eDNA concentration in response to changes in fish biomass.

(a) Size-fractionated eDNA concentration (C_X)

Random effect	Groups	Name	Variance	SD
	Tank	Intercept	0.508	0.713
	Temperature	Intercept	0.000	0.000
	Marker	Intercept	0.019	0.137
	Residual		0.405	0.636
Fixed effects	Variable	Estimate	SE	P value
	Intercept	-0.147	0.333	0.662
	log10(Biomass)	1.313	0.181	0.000***
	Size fraction (3–10 μm)	0.713	0.240	0.003**
	Size fraction (0.8–3 μm)	0.174	0.240	0.469
	Size fraction (0.4–0.8 μm)	-0.856	0.240	0.000***
	log10(Biomass): Size fraction (3–10 μm)	-0.185	0.137	0.176
	log10(Biomass): Size fraction (0.8–3 μm)	-0.578	0.137	0.000***
	log10(Biomass): Size fraction (0.4–0.8 μm)	-0.124	0.137	0.364

(b) Upside-cumulative eDNA concentration (UC_X)

Random effect	Groups	Name	Variance	SD
	Tank	Intercept	0.575	0.758
	Temperature	Intercept	0.000	0.000
	Marker	Intercept	0.059	0.244
	Residual		0.244	0.494
Fixed effects	Variable	Estimate	SE	P value
	Intercept	-0.147	0.359	0.689
	log10(Biomass)	1.313	0.179	0.000***
	Size fraction ($>3 \mu\text{m}$)	0.921	0.187	0.000***
	Size fraction ($>0.8 \mu\text{m}$)	1.001	0.187	0.000***
	Size fraction ($>0.4 \mu\text{m}$)	0.993	0.187	0.000***
	log10(Biomass): Size fraction ($>3 \mu\text{m}$)	-0.152	0.106	0.152
	log10(Biomass): Size fraction ($>0.8 \mu\text{m}$)	-0.180	0.106	0.090
	log10(Biomass): Size fraction ($>0.4 \mu\text{m}$)	-0.169	0.106	0.112

(c) Downside-cumulative eDNA concentration (DC_X)

Random effect	Groups	Name	Variance	SD
	Tank	Intercept	0.513	0.716
	Temperature	Intercept	0.000	0.000
	Marker	Intercept	0.035	0.188
	Residual		0.261	0.511
Fixed effects	Variable	Estimate	SE	P value
	Intercept	0.847	0.331	0.017*

log10(Biomass)	1.144	0.172	0.000***
Size fraction (0.4–10 μm)	-0.159	0.193	0.410
Size fraction (0.4–3 μm)	-1.056	0.193	0.000***
Size fraction (0.4–0.8 μm)	-1.849	0.193	0.000***
log10(Biomass): Size fraction (0.4–10 μm)	-0.028	0.110	0.800
log10(Biomass): Size fraction (0.4–3 μm)	-0.141	0.110	0.201
log10(Biomass): Size fraction (0.4–0.8 μm)	0.045	0.110	0.682

Note: The estimates in the fixed effects “Size fraction” were calculated against the $>10 \mu\text{m}$ size fraction (a & b) and the $>0.4 \mu\text{m}$ size fraction (c). Asterisks represent the statistical significances of the variables (** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$).

Figure legends

Figure 1. Relationships between the zebrafish eDNA concentration and the number of fish individuals among filter pore sizes and fragment lengths. From left to right, each four-panel shows the relationships using larger (10 μm) to smaller (0.2 μm) pore size filters. Additionally, each four-panel reveals the relationships targeting shorter (132 bp) to longer (1021 bp) eDNA fragments from the top to bottom. Solid and dashed lines represent the linear regressions of target eDNA concentrations (log-transformed) against fish individuals and their 95% confidence intervals (CIs). The abbreviation “Adj.R2” means R^2 values adjusted by the degree of freedom.

Figure 2. R^2 values (a) and slopes (b) in the linear regressions for zebrafish eDNA among filter pore sizes and fragment lengths. Each color of the bar plots (a) and plots (b) indicate the filter pore size. Plots and error bars in (b) represent the mean values and their 95% CIs.

Figure 3. Relative R^2 values among eDNA size fractions targeting the size-fractionated (C_X , a), upside-cumulative (UC_X , b), and downside-cumulative eDNA concentration (DC_X , c) from Japanese jack mackerel. Each relative R^2 value is calculated based on the eDNA concentration at $>10 \mu\text{m}$ (a & b) and $>0.4 \mu\text{m}$ size fractions (c). Plots and error bars represent the mean relative R^2 values and 95% CIs. Gray dashed lines represent the standard R^2 values (relative R^2 value = 1). All the datasets originated from Jo et al. (2019).

Figure 4. Schematic depiction of the relationship between eDNA concentration and species abundance depending on the eDNA particle size inferred by this study. As the particle size of the target eDNA increases, its apparent persistence in water could be shorter, whereas its spatial dispersion and distribution could be more heterogeneous. Thus, larger eDNA particles would not necessarily explain the abundance of target species more accurately and sensitively. However, considering the multiple factors regarding the eDNA characteristics and dynamics, “appropriately” larger eDNA particles (indicated as 3–10 μm in our study) can be suitable for accurate and sensitive abundance estimation.

Hosted file

Figure 1.docx available at <https://authorea.com/users/343681/articles/559397-fine-tuning-the-performance-of-abundance-estimation-based-on-environmental-dna-edna-focusing-on-edna-particle-size-and-marker-length>

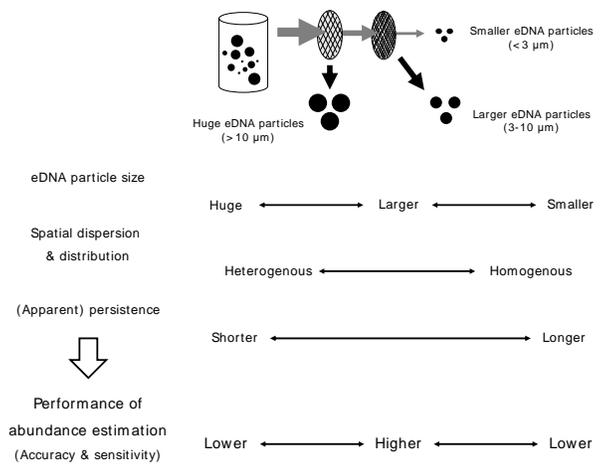
Hosted file

Figure 2.docx available at <https://authorea.com/users/343681/articles/559397-fine-tuning-the-performance-of-abundance-estimation-based-on-environmental-dna-edna-focusing-on-edna-particle-size-and-marker-length>

Hosted file

Figure_3 (1).docx available at <https://authorea.com/users/343681/articles/559397-fine-tuning-the-performance-of-abundance-estimation-based-on-environmental-dna-edna-focusing-on-edna-particle-size-and-marker-length>

on-edna-particle-size-and-marker-length



1
2 Figure 4.
3