

Efficiency of eDNA and iDNA in assessing vertebrate diversity and its abundance

Carolina Carvalho¹, Marina Oliveira¹, Karen Rodriguez-Castro¹, Bruno Saranholi¹, and Pedro Galetti Jr¹

¹Universidade Federal de São Carlos

March 07, 2024

Abstract

Environmental DNA (eDNA) and invertebrate-derived DNA (iDNA) have been increasingly recognized as a powerful tool for biodiversity assessment and conservation management. However, because of uncertainties on the efficiency of eDNA/iDNA approach in comparison to conventional methods, its use to assess the vertebrate diversity is still rare. Here we assessed the efficiency of eDNA/iDNA in comparison to conventional methods to survey vertebrate diversity across several type of samplers, vertebrate groups, and location (tropical vs temperate zones), as well as its efficiency to be used as a proxy for relative abundance or biomass across different molecular methods (qPCR and metabarcoding) and types of experiments (in the lab or in the field). The metaanalysis showed that, in general, there is no difference in the number of species detected or number of sites that a target species was detected when using eDNA/iDNA or conventional methods, suggesting that eDNA/iDNA and conventional methods were equally efficient in characterizing the biodiversity. However, for water sampler and fish, separately, the risk of not finding a species was greater using conventional method than eDNA, suggesting that eDNA/iDNA was more efficient in finding the target species. Abundance and biomass showed similar correlation patterns, and there was a positive correlation between eDNA/iDNA and abundance/biomass data, suggesting that eDNA/iDNA can be used as a proxy for abundance and biomass. Therefore, eDNA/iDNA has proved to be an efficient tool to assess vertebrate diversity in terms of both diversity of species and abundance or biomass.

Introduction

Earth is facing a biodiversity crisis as a consequence of habitat loss and climatic change (Bellard, Bertelsmeier, Leadley, Thuiller, & Courchamp, 2012; Brooks et al., 2002; Haddad et al., 2015) and the inadequate knowledge of biodiversity has hampered species conservation worldwide (Hortal, Jiménez-Valverde, Gómez, Lobo, & Baselga, 2008). Traditional methods for surveying species are generally limited to sampling on a local scale, require high sampling effort, and extensive taxonomic expertise; new technologies have been needed to support faster, more accurate and timely survey of biodiversity in natural ecosystems (Baird & Hajibabaei, 2012). Environmental DNA (eDNA) has proven to be a powerful tool for biodiversity surveys mainly because of its ability to assess species diversity in a large numbers of samples, reducing the amount of time currently required by traditional methods and enabling researchers to conduct long-term, large-scale biodiversity surveys more easily and with a great reduction in labor costs (Bohmann et al., 2014; Rees, Maddison, Middleditch, Patmore, & Gough, 2014; Valentini et al., 2016). Indeed, eDNA has provided more accurate data than traditional methods mainly due to its higher ability to detect rare, secretive, and elusive species (Bohmann et al., 2014; Valentini et al., 2016). Although many studies already showed the benefits of eDNA over traditional methods for sampling biodiversity, especially in marine and freshwater ecosystems, eDNA-based methods are still rarely used as a recognized tool for biodiversity assessment and conservation management (Cristescu & Hebert, 2018; Darling, 2019). This happens because of uncertainties and errors associated with those methods, mainly in relation to variation in time of DNA persistence in the environ-

ment that can increase rates of false positives (species is absent, but its DNA is detected) and false negative (non-detection of a species when it is present) (Cristescu & Hebert, 2018). However, no method is flawless and, therefore, it is necessary to evaluate the efficiency of eDNA approach in comparison to conventional methods that are more widely accepted to increase eDNA methods acceptance by environmental managers.

eDNA has been used to survey vertebrate diversity in aquatic and terrestrial ecosystems (Rees et al., 2014; Sales et al., 2020). For that, several samplers of eDNA have been tested, including water, soil, snow, and sediments (Buxton, Groombridge, & Griffiths, 2018; Kinoshita, Yonezawa, Murakami, & Isagi, 2019; Leempoel, Hebert, & Hadly, 2020). In addition, to environmental samplers, vertebrate diversity can also be assessed by extracting the DNA of the content ingested by invertebrates such as mosquitos, leeches, ticks, beetles, and flies, in this case named as ingested DNA or invertebrate-derived DNA (iDNA) (Calvignac-Spencer, Merkel, et al., 2013; Schnell et al., 2015). Each sampler presents its advantages and disadvantages, for example, sediment or leeches can preserve DNA better than water or mosquitos, but this can result in greater difficulty in determining whether the DNA sampled is contemporary or not (Pedersen et al., 2015; Schnell et al., 2015). Although there is a metanalysis that evaluated the efficiency of eDNA for sampling fishes using water as sampler (McElroy et al., 2020), to our knowledge, there is no study that compares biodiversity data obtained from eDNA/iDNA and traditional methods across different samplers and different vertebrate groups. Moreover, no study has yet evaluated the efficiency of eDNA over conventional methods by comparing tropical and temperate zones. It is expected that the efficiency of eDNA/iDNA will be lower in tropical zones, as the high biodiversity may impair the use of universal primers with sufficient taxonomic resolution to identify all the species diversity (Doble et al., 2020; Zhan, Bailey, Heath, & Macisaac, 2014). Moreover, high temperatures in the tropics may increase DNA degradation, a factor that can lead to higher rates of false negatives (Barnes et al., 2014).

Biodiversity surveys based on eDNA/iDNA can be carried out using several molecular methods, such as those based on real time quantitative PCR (qPCR) and those based on DNA metabarcoding. While the former is generally used to detect a target species using species-specific primers, the latter uses universal primers followed by high-throughput next-generation sequencing to determine the species composition of a sample (Freeland, 2017). Along with determining the presence/absence of species, the data generated by both molecular methods have been used as proxy for relative abundance (Lacoursière-Roussel, Côté, Leclerc, & Bernatchez, 2016; Yates, Fraser, & Derry, 2019). However, the use of the data generated from DNA metabarcoding appears to be less promising due to the amplification bias among species during PCR (Cristescu & Hebert, 2018). Many studies have evaluated the efficiency of eDNA in estimating relative abundance through the comparison between eDNA (eDNA concentration or number of reads) and abundance or biomass data, however, few broad compilations were done regarding this thematic. Lamb et al. (2019) carried out a metanalysis encompassing only studies that used eDNA and metabarcoding to determine the factors affecting the quantitative performance of such as sequencing platform choice and the number of species incorporated in a study. Yates et al. (2019), in turn, carried out a metanalysis using only studies that used qPCR method to evaluate the efficiency of eDNA as a proxy for relative abundance in laboratory and field experiments.

Here we aimed to assess the comprehensiveness and efficiency of the use of eDNA and iDNA to survey vertebrate diversity. By compiling studies of eDNA and iDNA, we aimed to detect the tendency of studies in terms of type of samplers, vertebrate groups most assessed, location, molecular methods, and genes used for species identification. Also, by carrying out a metanalysis, we aimed to assess the efficiency of eDNA/iDNA in comparison to conventional methods to survey vertebrate diversity across several type of samplers, vertebrate groups, and location (tropical vs temperate zones). Finally, we assessed the efficiency of eDNA/iDNA as a proxy for relative abundance or biomass across different molecular methods (qPCR and metabarcoding) and the type of experiment (in the lab or in the field).

Materials and Methods

Database Search and Screening

We surveyed publications that reported the use of eDNA and iDNA to characterize vertebrate biodiversity in the following database: Web of Science (<http://www.webofscience.bom>), Scopus (<http://www.scopus.com>), Pubmed (<https://pubmed.ncbi.nlm.nih.gov>), and Science direct (<https://www.sciencedirect.com>). The survey covered all English language publications available up to March 2020 using the following search terms: Metabarcod* OR “Environmental DNA” OR eDNA OR “ingested DNA” OR iDNA in combination with vertebrate OR mammal OR bird OR fish OR reptile OR amphibian OR monitoring. In total, 2629 articles were retrieved and peer-reviewed in the screening process to assess each paper eligibility. We removed all papers that did not aim to characterize vertebrate biodiversity (i.e., invertebrates, plants, and microorganisms), that did not use eDNA and iDNA samplers or that used them but were applied in diet studies. After that, 552 articles remained (Fig. 1).

Data extraction

We manually inspected each article and extracted the following information: location of the study (country, continent, and if temperate or tropical zone), type of eDNA (soil, water, sediment, or snow) or iDNA sampler (carrion fly, mosquito, beetles, or leeches), target vertebrate group surveyed (mammal, fish, bird, amphibian, or reptile), molecular method (metabarcoding, barcoding, qPCR, or ddPCR), gene (12S, 16S, 18S, COI, or CytB), type of experiment (in the laboratory or in the field) (Fig. 1). Barcoding molecular method comprises studies that used Sanger instead of high-throughput next-generation sequencing as is in metabarcoding. Moreover, we inspected if the articles compared the number and identity of species obtained with eDNA/iDNA with other conventional methods for vertebrate biodiversity assessment (i.e., camera trap, electrofishing, visual surveys, etc.); or if the articles correlated DNA concentration or number of reads obtained from eDNA/iDNA approach with known species biomass or abundance (Fig. 1).

Metanalysis

We carried out a metanalysis using only the data extracted from studies that compared eDNA/iDNA with conventional methods for assessing vertebrate biodiversity (127 studies) and from studies that correlated eDNA/iDNA data with species biomass or abundance (53 studies) (Fig. 1).

Estimating effect size

Studies comparing eDNA/iDNA and conventional methods provided data of i) the number of species detected with eDNA/iDNA and conventional methods in a target area (55 studies), ii) the number of sites where a target species was detected with eDNA and conventional methods (81 studies), and iii) the mean number of species (and standard deviation) detected across several sites with eDNA/iDNA and conventional methods (20 studies) (Fig. 1). We analyzed these three types of data separately using the log ratio risk for the first and second datasets and the standardized mean difference for the third one as the effect sizes (Fig. 1). A negative log ratio risk value indicates that the risk of not detecting a species or a site with a target species is higher in conventional methods whereas a positive value indicates that the risk is higher with the eDNA/iDNA approach. For studies that correlated eDNA/iDNA data with species biomass or abundance, we used the fisher z-score as the effect size by extracting the correlation coefficient of Pearson (r) or coefficient of determination (R^2) and the sample size. Biomass and abundance data were analyzed separately and, after removing studies with a sample size smaller than four, 35 and 23 studies remained, respectively. Some studies were included in both datasets and those that did not provide the type of data described above were removed from the analysis. Studies that presented more than one effect size were aggregated into a single combined effect size using the function `aggregate` of R package *metafor* (Viechtbauer, 2010).

Model fitting

All model fitting was conducted in R package *metafor* (Viechtbauer, 2010) using random effect models, which assumes that true effect sizes vary among studies, and restricted maximum-likelihood estimator. We used random effect models because studies included are not identical in their methods and characteristics, and because we aimed to “provide an unconditional inference about a larger set of studies from which the studies included in the metanalysis are assumed to be a random sample” (Viechtbauer, 2010). To test for residual

heterogeneity among the true effects (τ^2) we carried out a Cochran's Q -test and, when significant, we included moderators in the model to account for the heterogeneity among the true effects. For the studies that compared data obtained from eDNA/iDNA and conventional methods, we used as moderators the type of samplers, target vertebrate group and location (tropical vs temperate zones). For the studies that correlated eDNA/iDNA data with species abundance or biomass, the moderators were the molecular method and the type of experiment (in the lab or in the field). The significance of all coefficients of models that included a moderator was tested using Omnibus test.

Publication bias

First, we assessed publication bias using funnel plot of effect sizes against standard errors. The funnel plot visually indicates whether there is a publication bias (plot asymmetrical) or not (plot symmetrical). Then, we used trim-and-fill method to estimate the number of studies missing from the meta-analysis through the suppression of the most extreme results on one side of the funnel. A new weighted mean effect size and variance was estimated by reinserting the suppressed extreme results along with their corresponding “mirror” effect sizes. The analyses were done using the function *funnel* and *trimfill* in metafor R package (Viechtbauer, 2010).

Results

We inspected 552 articles which aims were: the assessment of vertebrate biodiversity (347), literature review (102), evaluation of eDNA degradation, release, and detection probability (99), evaluation of laboratory procedure (47), primer development and validation (159), sampling test (including sampling effort and design, 44), building of reference library (14) and evaluation of bioinformatic protocol (12). Most of the studies focused on fishes, followed by amphibians, mammals, reptiles, and birds (Fig. 2A). Water was expressively the most used sampler for all vertebrate groups (Fig. 2B). For water and sediment samplers, fishes were the most studied group while mammals were the most studied group for the remaining samplers (Fig. 2B). qPCR was the most used molecular method, followed by metabarcoding, barcoding and ddPCR (Fig. 2C), and COI and Cytb were the most used genes (Fig. 2D). It is important to highlight that the gene Cytb and COI were mainly used for qPCR molecular methods while 12S and 16S were the most used genes for metabarcoding approach (Fig. 2E). While most studies were carried out in North America, Asia, and Europe; Antarctic, Central America, Africa, and South America were the least studied areas (Fig. 2F).

Comparing eDNA/iDNA and conventional methods

Number of species in a target area

On average 50.5% of the species found in a target area were detected by both eDNA/iDNA and conventional methods (ranging from 0 to 100% shared detection). Most studies (54%) detected more exclusive species using eDNA/iDNA, while 29% of studies detected more exclusive species using conventional methods. Log risk ratio ranged from -3.71 to 3.89 (eight studies with positive values, 17 with negative values and 30 studies which log risk ratio values were not significant) and the weighted mean effect size was equal to -0.32 (95% CI ± 0.34 , $k=55$, $z=-1.81$, $p=0.07$, Fig. 3). This result means that the risk of not finding a species using conventional methods was not greater than using eDNA/iDNA approach, suggesting that eDNA/iDNA and conventional methods were equally efficient in characterizing the biodiversity, in terms of number of species in a target area. The heterogeneity between studies (τ^2) was equal to 1.08 and statistically significant ($Q=377.22$, $df=54$, $p<0.001$). The risk of not finding a species differed among samplers ($Q=6.31$, $df=2$, $p=0.04$), among vertebrate groups ($Q=7.89$, $df=3$, $p=0.04$), and among zones ($Q=9.31$, $df=1$, $p=0.002$). For water sampler and fish, separately, the risk of not finding a species was greater using conventional method than eDNA, while for the other samplers and vertebrate groups the risk was not statically significant (Fig. 3). Temperate zones presented higher risk than tropical zones of not finding a species using conventional method than eDNA/iDNA. The heterogeneity between studies was statistically significant even with the inclusion of moderators, showing that other factors not tested here may influence such heterogeneity.

Number of sites that a target species was detected

Log risk ratio ranged from -4.74 to 4.34 (seven studies with positive values, 18 with negative values and 56 studies which log risk ratio values were not significant) and the weighted mean effect size was equal to -0.54 (95% CI ± 0.35 , $k=81$, $z= -2.99$, $p<0.0001$, Fig. 3). This result means that the risk of not detecting a target species in a site where it is present using conventional methods was greater than when using eDNA/iDNA approach, suggesting that eDNA/iDNA was more efficient in finding the target species. The heterogeneity between studies (τ^2) was equal to 1.56 and statistically significant ($Q=309.56$, $df=80$, $p<0.001$). The risk of not detecting the target species did not differ among samplers ($Q=0.64$, $df=2$, $p=0.72$), vertebrate groups ($Q=6.97$, $df=3$, $p=0.07$), or world climate zones ($Q=0.71$, $df=1$, $p=0.39$). Although the risk of not detecting the target species did not differ between the moderators, for water sampler, amphibian, and temperate zone, separately, the risk of not finding the target species was greater using conventional method than eDNA/iDNA approach. The heterogeneity between studies was statistically significant even with the inclusion of moderators.

Mean number of species detected across several sampling sites

Standardized mean difference ranged from -6.97 to 3.89 (six studies with positive values, three with negative values and eleven studies which standardized mean difference values were not significant) and the weighted mean effect size was equal to 0.29 (95% CI ± 0.97 , $k=20$, $z= 0.59$, $p=0.55$, Fig. 3). This result means that the mean number of species found using conventional methods was not different from that found using eDNA/iDNA approach, suggesting that eDNA/iDNA and conventional methods were equally efficient in characterizing the biodiversity, in terms of mean number of species. The heterogeneity between studies (τ^2) was equal to 4.26 and statistically significant ($Q=394.61$, $df=19$, $p<0.001$). The mean number of species did not differ among samplers ($Q=0.49$, $df=2$, $p=0.77$) and among world climate zones ($Q=0.08$, $df=1$, $p=0.77$), but differed among vertebrate groups ($Q=6.49$, $df=2$, $p=0.03$). The mean number of fish species sampled using eDNA was higher than using conventional methods. The heterogeneity between studies was statistically significant even with the inclusion of moderators.

Abundance and biomass correlations

Abundance and biomass showed similar correlation patterns. Fisher z-score ranged from 0.00 to 2.73 for abundance and from 0.55 to 2.31 for biomass and the weighted mean effect size was equal to 1.21 (95% CI ± 0.20 , $k=35$, $z= 12.31$, $p<0.0001$, Fig. 4) and 1.27 (95% CI ± 0.26 , $k=23$, $z=9.92$, $p<0.0001$, Fig. 4), respectively. This result means that there was a positive correlation between eDNA and abundance/biomass data, suggesting that eDNA/iDNA can be used as a proxy for abundance and biomass. The heterogeneity between studies (τ^2) was equal to 0.25 and 0.28 for abundance and biomass, respectively, and statistically significant in both cases (abundance $Q=253.32$, $df=34$, $p<0.001$ and biomass $Q=152.11$, $df=22$, $p<0.001$). The Fisher z-score differed between experiment in the field and in the laboratory (abundance $Q=18.44$, $df=1$, $p<0.0001$ and biomass $Q=12.95$, $df=1$, $p=0.003$), but did not differ among molecular methods (abundance $Q=0.65$, $df=1$, $p=0.41$ and biomass $Q=0.001$, $df=1$, $p=0.97$). The correlation between abundance/biomass and eDNA was higher for experimental studies in the laboratory than in the field. The heterogeneity between studies was statistically significant even with the inclusion of moderators.

Publication bias

Funnel plots visually approached symmetry in all datasets (Fig. S1). The results of the trim-and-fill method indicated that no study was missed using the dataset *number of sites that a target species was detected*, *mean number of species detected across several sites* and *biomass correlations*. For the datasets *number of species in a target area* and *abundance correlations*, one and two studies were potentially missed, respectively. The impact of the missing studies on the weighted mean effect size was low because confidence intervals overlap with those we observed (Table S1). These results suggested there is no publication bias in all datasets.

Discussion

By compiling studies of eDNA and iDNA that aimed to assess the vertebrate diversity, we detected tendencies in terms of type of samplers, target vertebrate groups, location, molecular methods, and genes used. Moreover,

we assessed the efficiency of eDNA/iDNA in comparison to conventional methods to survey vertebrate diversity, as well as its efficiency to be used as a proxy for relative abundance or biomass. By comparing results generated by different laboratories under multiple conditions, eDNA/iDNA has proved to be an efficient tool to assess vertebrate diversity in terms of both species diversity and abundance or biomass. Because the sampling efforts for eDNA are generally smaller when compared to conventional methods, the use of eDNA can reduce the time currently required by conventional methods, allowing a wide and holistic view of the species being monitored in more areas and generating more data to be used in management and conservation programs.

We found that water was the most used sampler for eDNA studies and fish the most studied group of vertebrates, likely due to their high biodiversity (number of species and biomass) and economic importance (Hiddink et al., 2008; Lynch et al., 2016). Because a great sampling effort and high costs are required to survey fishes by conventional methods (Rees et al., 2014; Valentini et al., 2016), the ease of sampling eDNA may also be contributing to the high use of this method in fish. Fish survey by eDNA is already a well-established method and many countries, mainly in North America and Europe, have been using this approach to assess fish communities and to monitor invasive species (Bohmann et al., 2014; Rees et al., 2014). Water sampling processing usually requires collecting samples in free-DNA bottles, filtering them in the field or in the laboratory and storing the filters in a preservation buffer until the DNA extraction process (Rees et al., 2014; Valentini et al., 2016). Samplers like soil and invertebrates can be more challenging than water to use because they require larger volume of samples and different transport and storage strategies (Leempoel et al., 2020; Ruppert, Kline, & Rahman, 2019) or specific sampling traps (Kocher, de Thoisy, Catzeflis, Valière, et al., 2017), for example, for leeches the collector is offered as bait (Schnell et al., 2015). Moreover, unlike water in which the eDNA has a relatively homogeneous distribution (Cristescu & Hebert, 2018), in soils and sediments the DNA is not distributed uniformly throughout the area, which requires greater sampling effort (number of sampling points) to better describe the biodiversity of the area (Ruppert et al., 2019). In addition, it was demonstrated that the eDNA present in water better reflects the current species composition than sediment and soil (Cristescu & Hebert, 2018; Ruppert et al., 2019). Using invertebrates as samplers of vertebrate diversity can also be challenging because, for example, mosquitos and leeches have feeding preferences which could bias the results of biodiversity surveys (Calvignac-Spencer, Leendertz, Gilbert, & Schubert, 2013; Schnell et al., 2015). The challenges of using soil, sediment and insects may have favored the use of water in eDNA studies.

Most of eDNA/iDNA studies were conducted in temperate zones, such as in North America, Europe, and Asia and a much smaller number of studies in tropical zones such as countries in Central and South America and Africa. This results probably reflects the limited funding for science in many tropical countries, the lack of qPCR and sequencing equipment, and the high cost of molecular reagents since most of them are imported (personal observation based on experience working in tropical and developing countries). Due to these limitations, the vertebrate species from most high biodiversity areas (tropical regions) have been less assessed by eDNA/iDNA approach, reinforcing the need of more eDNA/iDNA studies in these regions. Most studies were carried out using qPCR as molecular method. Although qPCR can be much cheaper than other molecular methods based on sequencing, it usually focuses on one species and requires the development of species-specific primers.

Fragments of the genes COI and Cytb were the most used, but they were mainly used with qPCR while 12S and 16S were the most used genes in studies with metabarcoding. eDNA/iDNA studies require the use of primers that amplify fragments of small size because of the high degradation degree of DNA present in the environment (Rees et al., 2014). Primers amplifying small fragments of 12S and 16S genes were developed for vertebrates and showed high-resolution power for the target taxonomic group (Collins et al., 2019; Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014). Unfortunately, there is no reference library for 12S and 16S, like the BOLD system (<https://www.boldsystems.org/>) for COI sequences, applying to vertebrate groups. Thus, most studies using 12S and 16S built their own reference library (Kocher, de Thoisy, Catzeflis, Huguin, et al., 2017; Sato, Miya, Fukunaga, Sado, & Iwasaki, 2018) or used data deposited in the NCBI database (<https://www.ncbi.nlm.nih.gov/genbank/>). Leempoel et al. (2020), however, found

that 59% of all known mammals do not have the 12S sequence deposited in any public database. Although COI presents the most extensive reference database for barcode studies (Hebert, Ratnasingham, & de Waard, 2003), there is a challenge in the development of primers that amplify a small fragment capable of identifying a high proportion of taxa (Deagle et al., 2014), avoiding COI primer mismatch (Kocher, de Thoisy, Catzeflis, Huguin, et al., 2017). Mini-COI sequences are usually developed to identify species within vertebrate groups, as recently developed to leporids (Rodrigues et al., 2020) and elasmobranchs (Zahn, Silva, & Hellberg, 2020). Nevertheless, a substantial amplification bias towards non-metazoan taxa has been identified with COI, which can hinder the recovery of all vertebrate taxa in an environmental sample (Harper et al., 2019). As any set of primers may have some bias in amplifying species DNA, multi-markers have been suggested to increase species detection in eDNA samples (Leempoel et al., 2020; McElroy et al., 2020). It is important to highlight that the cost for sequencing has only recently decreased, simplifying the process of building reference databases that improve the species molecular identification and eDNA/iDNA approach.

Our metanalysis showed that there is no difference in the number of species detected or number of sites where a target species is detected when using eDNA/iDNA or conventional methods, corroborating with a metanalysis carried out by McElroy et al. (2020) using a smaller dataset of studies ($n=37$). Moreover, we found that 50.5% of the species detected in a target area were shared between eDNA/iDNA and conventional methods and most studies (54%) detected more exclusive species when using eDNA/iDNA than conventional methods. These results show that eDNA/iDNA can be used as a tool to survey vertebrate diversity despite some flaws that the method may have. For example, DNA release or degradation in the water is affected by temperature (Saito & Doi, 2020; Tsuji, Ushio, Sakurai, Minamoto, & Yamanaka, 2017) while iDNA degradation (from mosquitoes) increases over time becoming undetectable after 30-72 hours post-feeding (Reeves, Gillett-Kaufman, Kawahara, & Kaufman, 2018). Conventional methods, however, also have flaws such as the difficulty in identifying rare and elusive species and sampling bias due to logistic constraints (Darling, 2019; Rees et al., 2014; Valentini et al., 2016). Although we have found that there is no difference between eDNA/iDNA and conventional methods, the latter requires greater field effort and staff expertise to identify vertebrate species (Rees et al., 2014; Valentini et al., 2016). By combining these factors with the lower cost of sequencing, many studies claim that the cost of sampling biodiversity is lower with eDNA/iDNA than with conventional methods (Bohmann et al., 2014; Sigsgaard, Carl, Møller, & Thomsen, 2015; Valentini et al., 2016). It also represents an important advantage of eDNA/iDNA approach over conventional methods because favor long-term biodiversity monitoring. Nevertheless, the use of combined methods (eDNA/iDNA and conventional methods) could be useful in surveys that seek to make a comprehensive list of species.

Our results also showed that for water sampler, fish as a target species, and studies done in temperate areas eDNA is more efficient than conventional methods to survey vertebrate diversity (the risk of not detecting a species or a site with a target species were higher using conventional methods than eDNA). These results may arise because (i) eDNA is a robust method despite methodological variation (DNA extraction methods, primers, sequencing platform), (ii) conventional methods underestimate species diversity, mainly rare species, due to sampling limitations, or (iii) eDNA overestimates species diversity by detecting species on a larger geographic scale or due to false positives emerging from contamination or errors in reference databases (Cristescu & Hebert, 2018; McElroy et al., 2020). For the other samplers, vertebrate groups, and in tropical zones there were no differences between eDNA/iDNA and conventional methods, however, more studies are needed due to small sample size. For example, due to the high species diversity and temperature in tropical areas, further comparative studies are needed to assess whether the genes chosen can identify all species diversity and whether DNA degradation due to high temperature (and high humidity) impairs eDNA detection. We have not found that total number of species in an area influence the effect size variation ($Q=0.04$, $df=1$, $p=0.83$), but further studies in tropical areas are needed to confirm that eDNA is efficient in areas with high diversity.

Many studies have shown that DNA release and degradation are impacted by temperature, pH, and light intensity, affecting the amount of DNA found in the environment (Harper et al., 2019; Saito & Doi, 2020; Tsuji et al., 2017). Besides that, our results showed that eDNA can be used as proxy for relative abundance or biomass of species because most studies found a positive correlation between eDNA and abundance or biomass

data. We also found that this correlation was higher for laboratory than for field studies, suggesting a loss of efficiency of eDNA in evaluating relative abundance or biomass when applied in field conditions, corroborating with a meta-analysis carried out by Yates et al. (2019) using a smaller dataset (n=19). Nevertheless, the effect sizes were still positive and significant for field studies, indicating that eDNA can be used as a proxy for relative abundance or biomass even in uncontrolled environmental conditions. Here we analyzed abundance and biomass separately to avoid data dependency. Yates et al. (2019), by combining abundance and biomass data, found no evidence that the latter was more strongly correlated with eDNA than abundance. Finally, we found no difference in the effect size between studies that used eDNA concentration from qPCR and number of reads from metabarcoding, suggesting that both molecular methods are efficient and can be used as proxy for relative abundance and biomass. As the number of studies that used metabarcoding was three times lower than the number that used qPCR, more studies are needed to assess its efficiency in relation to different samplers and vertebrate groups. Finally, although eDNA showed to be a powerful proxy for relative abundance, its efficiency in estimating absolute abundance remain unknown.

Conclusions

The number of studies using eDNA/iDNA to evaluate vertebrate diversity has increased worldwide, but few studies so far have conducted meta-analysis to assess its efficiency in comparison to conventional methods and its use as a proxy for relative abundance and biomass. Here we found that eDNA/iDNA is as efficient as conventional methods in assessing the number of species and the number of sites containing a target species. Due to much less field effort, eDNA/iDNA is more advantageous than conventional methods, especially in studies of fish in temperate zones. Thus, if the objective is to obtain a list of the species present in an area or to detect a specific species of interested across its range, eDNA/iDNA is the most efficient tool. But if more data are required, such as sex and age, conventional methods are needed. For the other vertebrate groups and tropical zones, more studies must be carried out to evaluate the efficiency of eDNA/iDNA over conventional methods. We conclude that eDNA can be used as a proxy for relative abundance and biomass even in uncontrolled environmental conditions. Finally, our study was important to evaluate the efficient of eDNA/iDNA methods in assessing vertebrate diversity and to identify potential gaps. For example, there is an urgency in studies comparing eDNA/iDNA and conventional methods using soil, mosquitoes, carryon flies, among others, as well as studies that aim to complete the public reference databases. Moreover, there is an urgency in assessing the efficiency of eDNA/iDNA in tropical zones that harbor much of the world biodiversity but are still poorly studied.

Acknowledgments

PMGJ and MEO thank to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 303524/2019-7 and 141085/2019-3, respectively). CSC, KGRC and BHS received Fundação de Amparo à Pesquisa do estado de São Paulo (FAPESP 2019/26436-6), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES, Finance Code 001) and Natural Environment Research Council (NERC, NE/S011811/1) fellowships, respectively.

References

- Baird, D. J., & Hajibabaei, M. (2012). Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology* , 21 (8), 2039–2044. doi: 10.1111/j.1365-294X.2012.05519.x
- Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D. M. (2014). Environmental Conditions Influence eDNA Persistence in Aquatic Systems. *Environmental Science & Technology* , 48 (3), 1819–1827. doi: 10.1021/es404734p
- Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters* , 15 (4), 365–377. doi: 10.1111/j.1461-0248.2011.01736.x
- Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., . . . de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution* , 29

(6), 358–367. doi: 10.1016/j.tree.2014.04.003

Brooks, T. M., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., Rylands, A. B., Konstant, W. R., ... Hilton-Taylor, C. (2002). Habitat Loss and Extinction in the Hotspots of Biodiversity. *Conservation Biology* , 16 (4), 909–923. doi: 10.1046/j.1523-1739.2002.00530.x

Buxton, A. S., Groombridge, J. J., & Griffiths, R. A. (2018). Seasonal variation in environmental DNA detection in sediment and water samples. *PLoS ONE* , 13 (1), 1–14. doi: 10.1371/journal.pone.0191737

Calvignac-Spencer, S., Leendertz, F. H., Gilbert, M. T. P., & Schubert, G. (2013). An invertebrate stomach's view on vertebrate ecology: Certain invertebrates could be used as “vertebrate samplers” and deliver DNA-based information on many aspects of vertebrate ecology. *BioEssays* , 35 (11), 1004–1013. doi: 10.1002/bies.201300060

Calvignac-Spencer, S., Merkel, K., Kutzner, N., Kühl, H., Boesch, C., Kappeler, P. M., ... Leendertz, F. H. (2013). Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Molecular Ecology* , 22 (4), 915–924. doi: 10.1111/mec.12183

Collins, R. A., Bakker, J., Wangensteen, O. S., Soto, A. Z., Corrigan, L., Sims, D. W., ... Mariani, S. (2019). Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution* , 10 (11), 1985–2001. doi: 10.1111/2041-210X.13276

Cristescu, M. E., & Hebert, P. D. N. (2018). Uses and misuses of environmental DNA in biodiversity science and conservation. *Annual Review of Ecology, Evolution, and Systematics* , 49 (July), 209–230. doi: 10.1146/annurev-ecolsys-110617-062306

Darling, J. A. (2019). How to learn to stop worrying and love environmental DNA monitoring. *Aquatic Ecosystem Health and Management* , 22 (4), 440–451. doi: 10.1080/14634988.2019.1682912

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters* , 10 (9), 2–5. doi: 10.1098/rsbl.2014.0562

Doble, C. J., Hipperson, H., Salzburger, W., Horsburgh, G. J., Mwita, C., Murrell, D. J., & Day, J. J. (2020). Testing the performance of environmental DNA metabarcoding for surveying highly diverse tropical fish communities: A case study from Lake Tanganyika. *Environmental DNA* , 2 (1), 24–41. doi: 10.1002/edn3.43

Freeland, J. R. (2017). The importance of molecular markers and primer design when characterizing biodiversity from environmental DNA. *Genome* , 60 (4), 358–374. doi: 10.1139/gen-2016-0100

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., ... Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances* , 1 (2), e1500052. doi: 10.1126/sciadv.1500052

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. doi: 10.1007/s10750-018-3750-5

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences* , 270 (suppl.1). doi: 10.1098/rsbl.2003.0025

Hiddink, J. G., MacKenzie, B. R., Rijnsdorp, A., Dulvy, N. K., Nielsen, E. E., Bekkevold, D., ... Ojaveer, H. (2008). Importance of fish biodiversity for the management of fisheries and ecosystems. *Fisheries Research* , 90 (1–3), 6–8. doi: 10.1016/j.fishres.2007.11.025

Hortal, J., Jiménez-Valverde, A., Gómez, J. F., Lobo, J. M., & Baselga, A. (2008). Historical bias in biodiversity inventories affects the observed environmental niche of the species. *Oikos* , 117 (6), 847–858. doi: 10.1111/j.0030-1299.2008.16434.x

- Kinoshita, G., Yonezawa, S., Murakami, S., & Isagi, Y. (2019). Environmental DNA Collected from Snow Tracks is Useful for Identification of Mammalian Species. *Zoological Science* ,36 (3), 198. doi: 10.2108/zs180172
- Kocher, A., de Thoisy, B., Catzeflis, F., Huguin, M., Valière, S., Zinger, L., ... Murienne, J. (2017). Evaluation of short mitochondrial metabarcodes for the identification of Amazonian mammals. *Methods in Ecology and Evolution* , 8 (10), 1276–1283. doi: 10.1111/2041-210X.12729
- Kocher, A., de Thoisy, B., Catzeflis, F., Valière, S., Bañuls, A.-L., & Murienne, J. (2017). iDNA screening: Disease vectors as vertebrate samplers. *Molecular Ecology* , 26 (22), 6478–6486. doi: 10.1111/mec.14362
- Lacoursière-Roussel, A., Côté, G., Leclerc, V., & Bernatchez, L. (2016). Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. *Journal of Applied Ecology* ,53 (4), 1148–1157. doi: 10.1111/1365-2664.12598
- Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology* , 28 (2), 420–430. doi: 10.1111/mec.14920
- Leempoel, K., Hebert, T., & Hadly, E. A. (2020). A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. *Proceedings of the Royal Society B: Biological Sciences* ,287 (1918), 20192353. doi: 10.1098/rspb.2019.2353
- Lynch, A. J., Cooke, S. J., Deines, A. M., Bower, S. D., Bunnell, D. B., Cowx, I. G., ... Beard, T. D. (2016). The social, economic, and environmental importance of inland fish and fisheries. *Environmental Reviews* , 24 (2), 115–121. doi: 10.1139/er-2015-0064
- McElroy, M. E., Dressler, T. L., Titcomb, G. C., Wilson, E. A., Deiner, K., Dudley, T. L., ... Jerde, C. L. (2020). Calibrating Environmental DNA Metabarcoding to Conventional Surveys for Measuring Fish Species Richness. *Frontiers in Ecology and Evolution* ,8 , 0–12. doi: 10.3389/fevo.2020.00276
- Pedersen, M. W., Overballe-Petersen, S., Ermini, L., Der Sarkissian, C., Haile, J., Hellstrom, M., ... Willerslev, E. (2015). Ancient and modern environmental DNA. *Philosophical Transactions of the Royal Society B: Biological Sciences* , 370 (1660). doi: 10.1098/rstb.2013.0383
- Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology* , 51 (5), 1450–1459. doi: 10.1111/1365-2664.12306
- Reeves, L. E., Gillett-Kaufman, J. L., Kawahara, A. Y., & Kaufman, P. E. (2018). Barcoding blood meals: New vertebrate-specific primer sets for assigning taxonomic identities to host DNA from mosquito blood meals. *PLOS Neglected Tropical Diseases* , 12 (8), e0006767. doi: 10.1371/journal.pntd.0006767
- Rodrigues, N. T., Saranholi, B. H., Angeloni, T. A., Pasqualotto, N., Chiarello, A. G., & Galetti Jr, P. M. (2020). DNA mini-barcoding of leporids using noninvasive fecal DNA samples and its significance for monitoring an invasive species. *Ecology and Evolution* ,10 (12), 5219–5225. doi: 10.1002/ece3.5863
- Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation* , 17 , e00547. doi: 10.1016/j.gecco.2019.e00547
- Saito, T., & Doi, H. (2020). *Environmental DNA degradation simulation from water temperature and DNA fragment length : A meta-analysis approach* .
- Sales, N. G., McKenzie, M. B., Drake, J., Harper, L. R., Browett, S. S., Coscia, I., ... McDevitt, A. D. (2020). Fishing for mammals: Landscape-level monitoring of terrestrial and semi-aquatic communities using eDNA from riverine systems. *Journal of Applied Ecology* ,57 (4), 707–716. doi: 10.1111/1365-2664.13592
- Sato, Y., Miya, M., Fukunaga, T., Sado, T., & Iwasaki, W. (2018). MitoFish and mifish pipeline: A mitochondrial genome database of fish with an analysis pipeline for environmental DNA metabarcoding. *Molecular*

Biology and Evolution , 35 (6), 1553–1555. doi: 10.1093/molbev/msy074

Schnell, I. B., Sollmann, R., Calvignac-Spencer, S., Siddall, M. E., Yu, D. W., Wilting, A., & Gilbert, M. T. P. (2015). iDNA from terrestrial haematophagous leeches as a wildlife surveying and monitoring tool - prospects, pitfalls and avenues to be developed. *Frontiers in Zoology* , 12 (1), 1–14. doi: 10.1186/s12983-015-0115-z

Sigsgaard, E. E., Carl, H., Moller, P. R., & Thomsen, P. F. (2015). Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. *Biological Conservation* , 183 , 46–52. doi: 10.1016/j.biocon.2014.11.023

Tsuji, S., Ushio, M., Sakurai, S., Minamoto, T., & Yamanaka, H. (2017). Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. *PLoS ONE* , 12 (4), 1–13. doi: 10.1371/journal.pone.0176608

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., ... Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology* , 25 (4), 929–942. doi: 10.1111/mec.13428

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* , 36 (3), 1–48. Retrieved from <https://www.jstatsoft.org/v36/i03/>

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. doi: 10.1002/edn3.7

Zahn, R. J., Silva, A. J., & Hellberg, R. S. (2020). Development of a DNA mini-barcoding protocol targeting COI for the identification of elasmobranch species in shark cartilage pills. *Food Control* , 109 , 106918. doi: 10.1016/j.foodcont.2019.106918

Zhan, A., Bailey, S. A., Heath, D. D., & Macisaac, H. J. (2014). Performance comparison of genetic markers for high-throughput sequencing-based biodiversity assessment in complex communities. *Molecular Ecology Resources* , n/a-n/a. doi: 10.1111/1755-0998.12254

Data Accessibility

All data will be deposited in Figshare and a url address provided upon the acceptance of this manuscript.

Authors Contributions

CSC, MEO, KGRC, BHS and PMGJ conceptualized and performed the study design. CSC, MEO, KGRC, BHS collected the data. CSC performed the data analysis. The first draft of the paper was written by CSC and all authors contributed to discussing the results and editing the paper.

Figures

Figure 1: Flowchart showing the steps involved in this study.

Figure 2: Summary of data extracted from studies that used eDNA/iDNA to characterize vertebrate biodiversity. Frequency of studies for each vertebrate group (A), samplers (B), molecular method (C), genes used for species identification (D), genes used in different molecular methods (E), and continent where the studies were carried out (F). Colors in panels B, C and D correspond to vertebrate groups (see colors in panel A).

Figure 3: Variation in the mean effect size within different groups of moderators: sampler types, vertebrate groups, and zones. Negative log ratio risk or standardized mean difference values indicate that the risk of not detecting a species (or a site with a target species) or the mean number of species are higher in conventional methods while positive values indicate that the risk and the mean number of species are higher with eDNA/iDNA approach. Circles and horizontal lines represent the weighted mean effect size and 95%

confidence interval. The dashed lines indicate an effect equal to zero. We omitted studies that presented more than one subgroup level to ensure data independence. The number of studies is above the circles.

Figure 4: Variation in the mean effect size within different groups of moderators: experiment types and molecular methods. Circles and horizontal lines represent the weighted mean effect size and 95% confidence interval. Positive values indicate positive correlation between eDNA and abundance/biomass data. The dashed lines indicate an effect equal to zero. The number of studies is above the circles.



