

Article type : Special Issue

How quantitative is metabarcoding: a meta-analytical approach

Authors: Philip D. Lamb¹, Ewan Hunter^{2,3}, John K. Pinnegar^{2,3}, Simon Creer⁴, Richard G. Davies¹, Martin I. Taylor¹

¹School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR4 7TJ, UK

²School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR4 7TJ, UK

³Cefas, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK

⁴School of Biological Sciences, Bangor University, Bangor, Gwynedd LL57 2UW, UK

Corresponding author: Philip D. Lamb (p.lamb@uea.ac.uk)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14920

This article is protected by copyright. All rights reserved.

Abstract

Metabarcoding has been used in a range of ecological applications such as taxonomic assignment, dietary analysis, and the analysis of environmental DNA.

However, after a decade of use in these applications there is little consensus on the extent to which proportions of reads generated corresponds to the original proportions of species in a community. To quantify our current understanding we conducted a structured review and meta-analysis. The analysis suggests that a weak quantitative relationship may exist between the biomass and sequences produced (slope = 0.52 ± 0.34 , $p < 0.01$), albeit it with a large degree of uncertainty. None of the tested moderators: sequencing platform type, the number of species used in a trial, or the source of DNA were able to explain the variance. Our current understanding of the factors affecting the quantitative performance of metabarcoding is still limited: additional research is required before metabarcoding can be confidently utilised for quantitative applications. Until then, we advocate the inclusion of mock communities when metabarcoding as this facilitates direct assessment of the quantitative ability of any given study.

Keywords

Metabarcoding biomass high-throughput next-generation sequencing meta-analysis

Introduction

Metabarcoding, the use of a polymerase chain reaction (PCR) and high throughput sequencing (HTS) to characterise organisms present in a sample, has been used to address an array of ecological questions (Creer et al., 2016) (PCR-free sequencing is an emerging technology (Paula et al., 2015; Srivathsan, Ang, Vogler, & Meier, 2016) but is not the focus of this analysis). For example, metabarcoding has allowed the taxonomic identification of many specimens simultaneously using a standardised DNA region (Valentini, Pompanon, & Taberlet, 2009) without the need for on-the-ground taxonomic expertise. Similarly, environmental DNA (eDNA) studies, which sequence DNA in soil and water (Yu et al., 2012) without first isolating any organisms, facilitate rapid biodiversity monitoring with only small sediment or water samples. Metabarcoding has also played an important role in uncovering diets and resolving food webs (Pompanon et al., 2012), as well as reconstructing community dynamics temporally using ancient DNA preserved in sedimentary layers (Thomsen & Willerslev, 2015).

Early adopters of metabarcoding were hopeful that outputs would be quantitative, *i.e.* that reads obtained from a sequencing run would correlate with biomass in the original sample (Symondson & Harwood, 2014) in a similar manner to other applications such as RNA sequence analysis (Mohorianu et al., 2017) and the characterisation of microbial communities (where it is referred to as meta-genomics). However, several factors, detailed in figure 1, can introduce bias into the results and yield inaccurate biomass estimates. Yet, despite these factors being well documented, after more than a decade of use there is no clear consensus as to what extent metabarcoding is quantitative. Many studies report their findings in a quantitative manner where the relative read abundance (RRA) (Deagle et al., 2018) is interpreted as the relative abundance of biomass (Kowalczyk et al., 2011; Soininen et al., 2015; Sousa et al., 2016; Vaz et al., 2017). Others use a frequency of occurrence (FOO) approach, also referred to as weighted occurrence (Deagle et al., 2018), where the proportion of samples in which a given sequence was detected is used to infer a different sort of quantitative measure (Bohmann et al., 2011; De Barba et al., 2014). It is also common to incorporate a qualitative approach (detected / not detected), sometimes simply referred to as occurrence (Deagle et al., 2018) or a 'species list', alongside these quantitative approaches (Vesterinen et al., 2016).

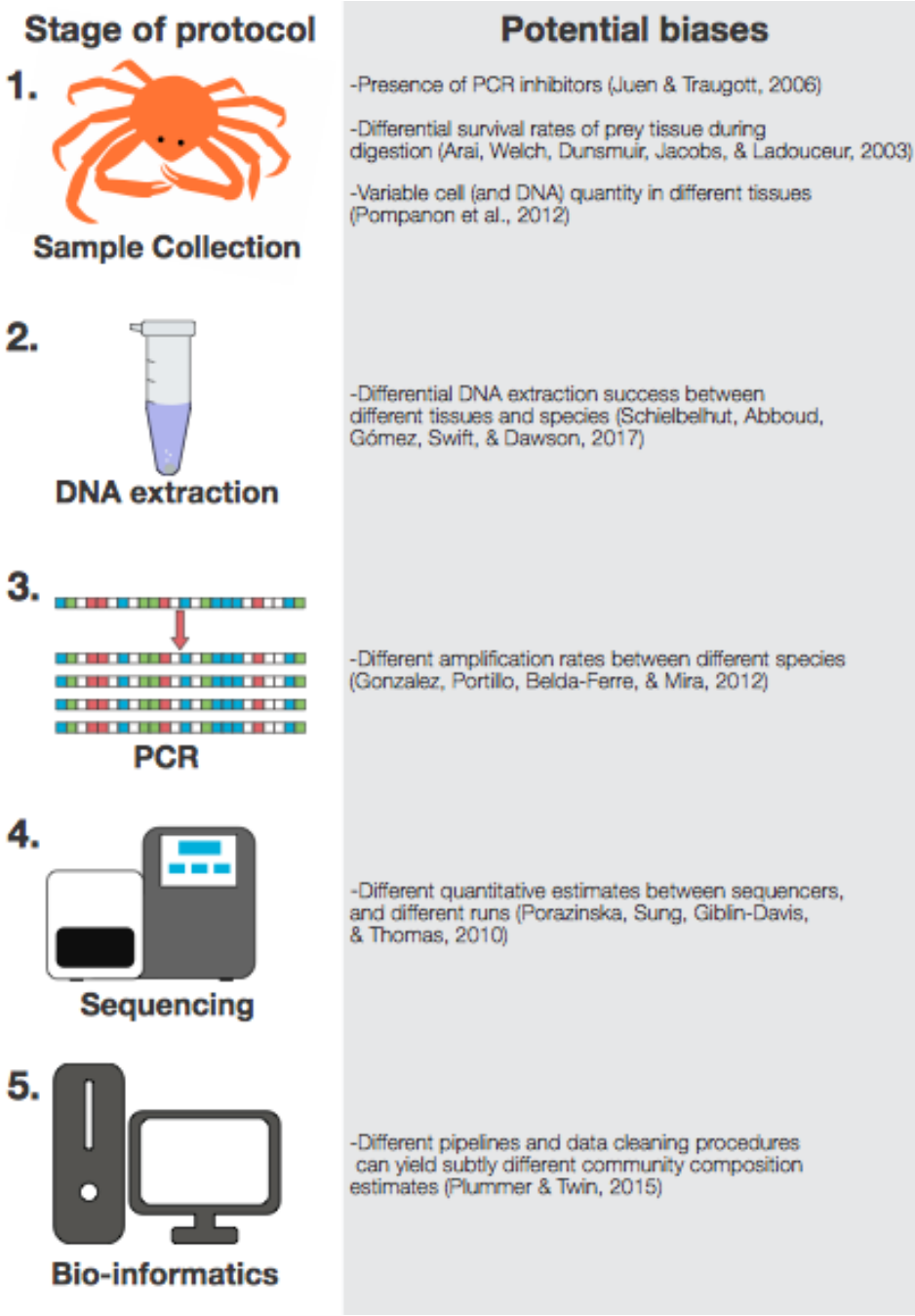


Figure 1) Overview of HTS procedure, and factors that can influence the quantitative output.

Empirically determining the extent to which metabarcoding is quantitative should be relatively simple: take a mixture of organisms with known biomass, PCR and sequence, then compare the results of the HTS run to the original biomass of each community member. Indeed, many studies have used this approach (Leray & Knowlton, 2017; Piñol, Mir, Gomez-Polo, & Agustí, 2015). However, often only one primer set is used and the output may be a result of primer bias (the differential amplification of target DNA due to different numbers of nucleotide mismatches between the primer and target DNA between samples) rather than a reflection of the ability of metabarcoding techniques. Even if multiple primers are used, they are normally used on the same sequencing run, in which case results cannot be considered independent. An experiment featuring enough sequencing runs to gather sufficient statistical power to disentangle the various factors that may affect quantitative performance would be prohibitively expensive for most research groups. Consequently, there is an *ad hoc* collection of methodologies that provide different levels of quantitative performance, but little certainty as to whether the variance is due to unique parameters in the experimental set up or a result of more general drivers.

In this study, we aim to address this knowledge gap. A structured review was conducted to collate our knowledge about the extent to which metabarcoding for taxonomic assignment is quantitative. Subsequently a meta-analysis was conducted to investigate the degree to which metabarcoding is quantitative across multiple independent studies. Factors affecting the quantitative performance such as platform choice, the experimental set up (does using biomass, individuals, or DNA as the input unit affect quantitative estimates?) and the number of species incorporated in a study were also investigated. Factors that could not be addressed are also discussed to direct future research.

Methods

Search strategy

Articles that used quantified multi-species assemblages, PCR, and HTS platforms for taxonomic assignment with metabarcoding were targeted using specific search terms. Identifying optimised search terms was important since metabarcoding is now widely used across evolutionary, ecological, and medical research. After assessing a variety of search terms an appropriate combination was finalised: the Web of Science

was searched on 31/10/2017 for English language articles for all available years using the following search terms: ((quant* OR diet OR biomass) AND (barcod* OR metabarcod*)). In total, 1262 articles were retrieved.

Article screening

Initial filtering of the articles was based on their titles: any articles that obviously had no relevance to quantification of biomass using metabarcoding were discarded. After initial filtering 262 articles remained. These articles were manually inspected and any that included a quantified community of biomass, individuals, or DNA as starting material and reported the proportion of reads obtained from a HTS platform were used for data extraction. Since the slope of a fitted linear model was to be used as an effect-size (see below), variation in the amount of input material was also required (equal amount of starting biomass could not be used). In total, 22 articles (table 1) were used in the meta-analysis.

Data extraction

The composition of the community assessed (either biomass, number of individuals, or concentration of DNA) and the proportions of reads corresponding to the relevant species in the test community obtained from the sequencing platform were recorded for each trial within an experiment. The sequencing platform, number of species used, and the source of input material for each trial within any given study was recorded. The main manuscript and supplementary information were inspected: if possible data were taken from a table, if tables were unavailable the data were manually extracted from figures using *Web Plot Digitizer* (Rohatgi, 2017). If data were not presented in the main article, the corresponding author was emailed to obtain the data.

The composition of the mock community, and corresponding sequence data were converted in percentage values (see figure 2 (a)). For the Elbrecht et al. (2017) study using individuals of varying sizes (Elbrecht, Peinert, & Leese, 2017), the composition of individuals in the mock community, and the output of reads, was presented grouped by size (large, medium, and small individuals) and unsorted. In this instance,

we calculated input and output percentages by the sorted size groupings as this was most similar to the approaches used in other included studies.

Slope is a commonly used effect size when the relationship between two continuous variables is being investigated (Rosenberg, Rothstein, & Gurevitch, 2013). In this instance, it was chosen as it is easy to interpret and meets the statistical assumptions of the meta-analysis model without transformation (in this instance because slopes did not approach vertical asymptotes and little skewness was present in the data).

Meta-analysis model fitting

Slope (the effect size) was calculated by fitting a linear model for each trial detected in the review using *R* (R Core Team, 2017), such that the proportion of reads produced from the sequencing run would be a function of the proportion of starting material in the experiment. The variance of the slope was calculated in *R*, and used as the sampling variance as described by Rosenberg, Rothstein, & Gurevitch (2013).

Figure 2 illustrates how the results of a mock community experiment are incorporated into this analysis.

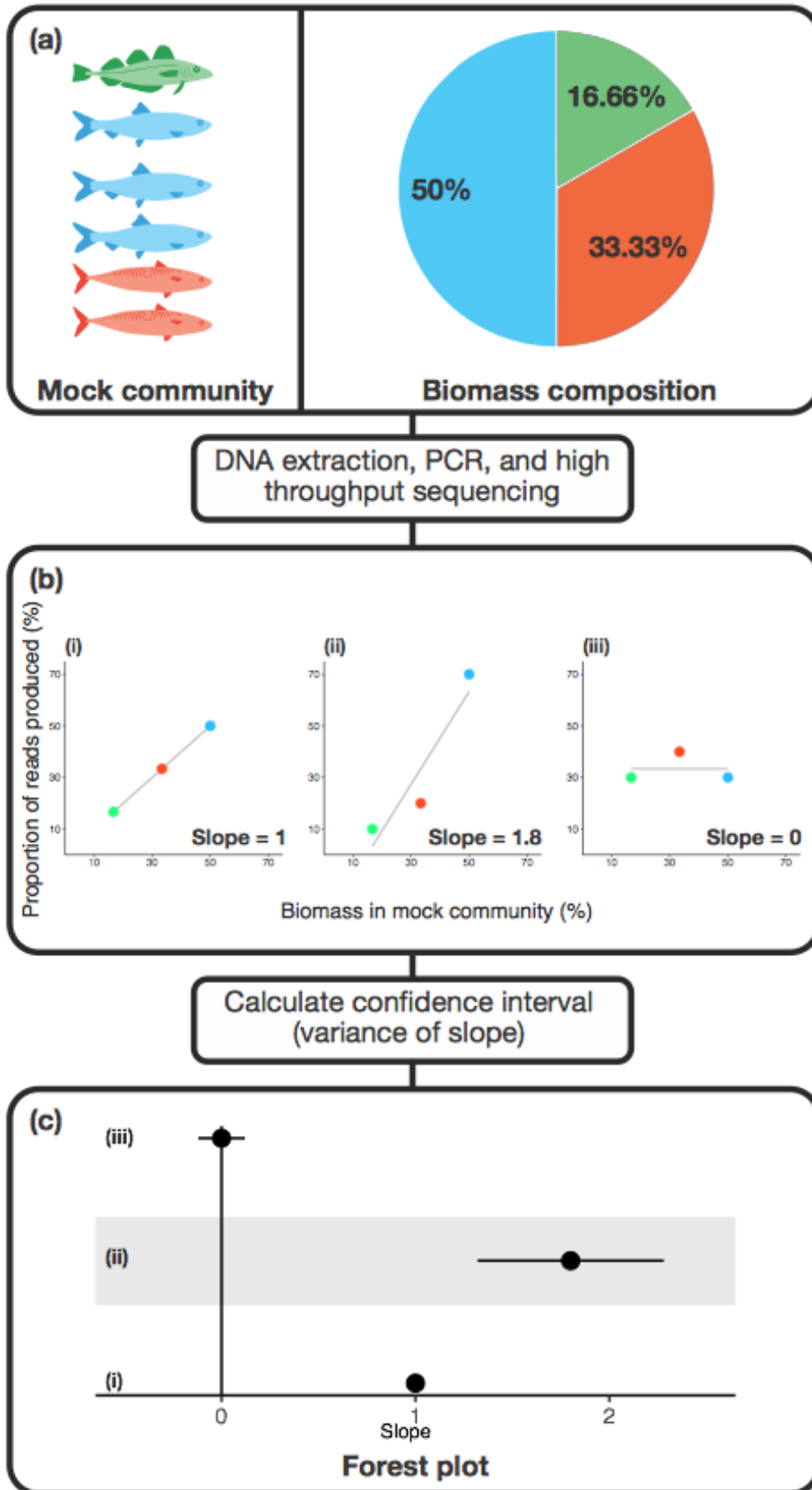


Figure 2) A schematic illustrating how data are utilised in the meta-analysis. (a) The mock community with quantified biomass. (b) three hypothetical outcomes of the metabarcoding step: (i) a perfect quantitative relationship between biomass and sequencing yield *i.e.* a 1% increase in biomass yields a 1% increase in reads, generating a slope = 1 (ii) a quantitative signal in which rank abundance is same in the mock community, but with over-representation of common sequences and under-representation of rare sequences resulting in a slope greater than 1. A slope of between 0 and 1 would be produced when common sequences are under-represented and rare sequences over-represented (not shown). (iii) no quantitative information, with a slope close to 0. Negative slopes would also be indicative of non-quantitative signals. (c) shows how (i),(ii), and (iii) would be visualised in a forest plot with corresponding variance of slope denoted by error bars.

All meta-analysis was conducted in the customisable, open-source, meta-analysis package 'metafor' (Viechtbauer, 2010) in R. Many studies used multiple trials within a single study, however, these trials cannot be treated as statistically independent from one another. To account for this non-independence, a cross-study slope-estimate was determined using a two-level nested random effects model using a restricted likelihood function. Trials within an experiment were nested at the study level. The influence of sequencing platform, and DNA source material, were tested by including them as moderating factors in the model. Terms were iteratively omitted from the model, and AIC was used to select the final model.

Weighting of each study in the meta-analysis model was determined solely by the number of sequencing runs used in each study (e.g. 1 for 1 run, 2 for 2 runs etc).

However, when multiple trials were conducted within a single study, the weight of each trial was calculated by dividing the number of reads produced for the *trial* by the total number of sequences produced by the sequencing run within the study.

This allows different sequencing depths within a single study to be accounted for (using a nested model) whilst maintaining sequencing runs as independent data

points. For example in Saitoh *et al.* (2016) a single sequencing run was used and a

meta-analysis model study weight of one was assigned. Within this study, there were two trials: the 16s *trial* produced 45% of the reads, therefore it accounted for 45% of model weight within the nested model (at the study level).

Sensitivity testing

Assessing publication biases (the increased probability of positive results being accepted for publication) in meta-analytical models is challenging for nested-models:

Funnel plots are difficult to interpret: studies cluster together due to statistical

dependencies rather than genuine biases (Lau, Ioannidis, Terrin, Schmid, & Olkin,

2006). Egger' s regression test (Egger, Smith, Schneider, & Minder, 1997), another

commonly used metric, is not supported for nested-models in the current version of metafor. Consequently, it was not possible to assess if publication bias may be present in the data set. However, influential trials in the meta-analysis were visually identified using hat values, which show the importance of any given trial in relation to the model as a whole (Krahn, Binder, & König, 2013), plotted against the standardised residuals of the meta-analysis model.

Results

Across all studies a significant ($p < 0.01$) relationship existed between the proportion of input material for each species present and the proportions of sequences obtained from metabarcoding. A large amount of observed variation was due to actual differences in the inter-study slope estimate ($I^2 = 88.5\%$). Across all studies, an effect size estimate (slope) of 0.52 (± 0.34 variance of slope) was identified.

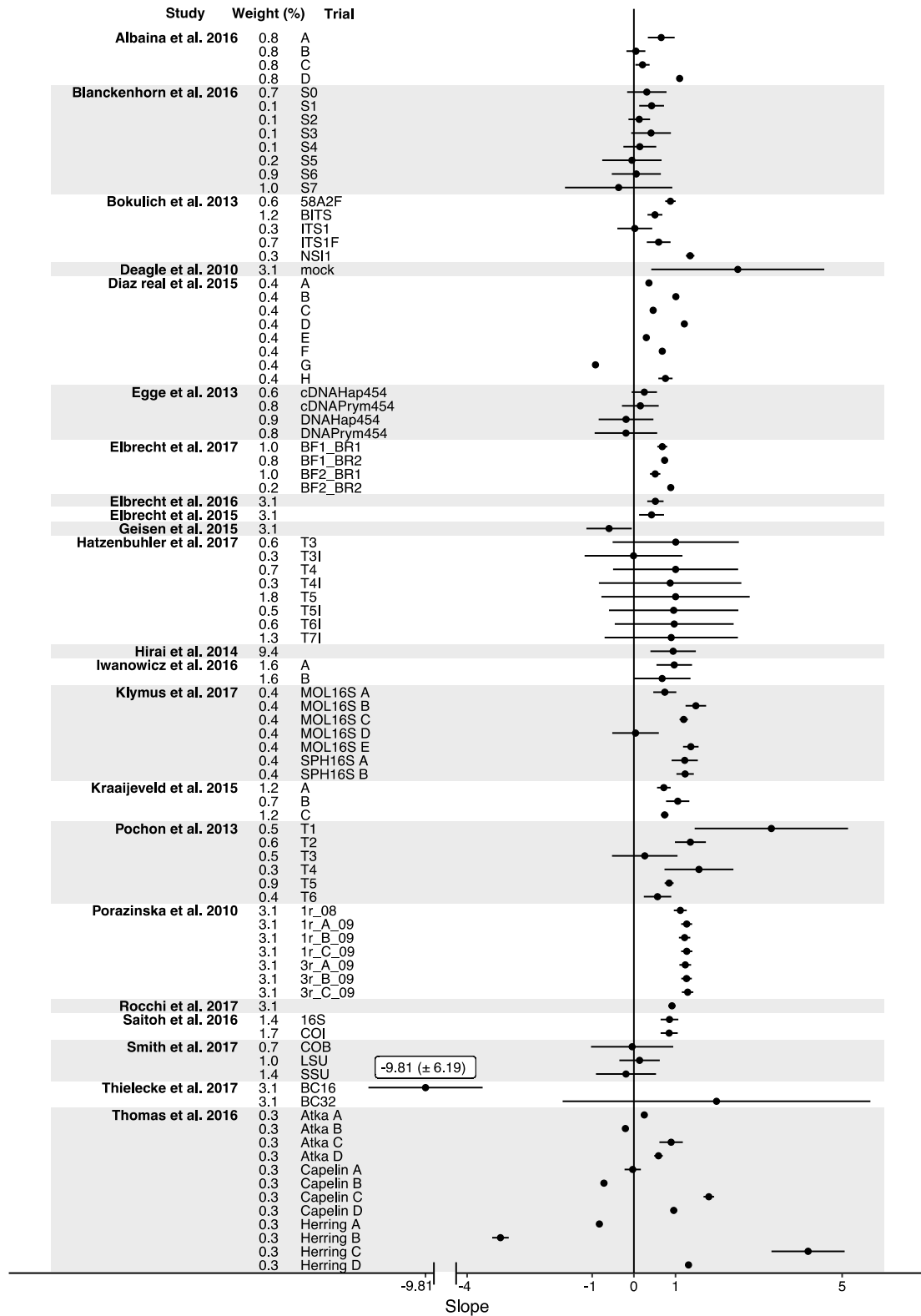


Figure 3) Forest plot showing the slope estimates for all trials in the meta-analysis (\pm variance of slope). Trials are clustered at the paper level denoted by the grey and white shading.

None of the tested moderators: type of sequencer, number of species used in a trial, or type of starting material had a significant effect ($p > 0.05$ in all instances) on the estimate provided by the meta-analysis model. Figure 4 illustrates the lack of difference in quantitative ability (a) between the materials used for meta-barcoding, (b) among the sequencers, and (c) the number of species used in a trial.

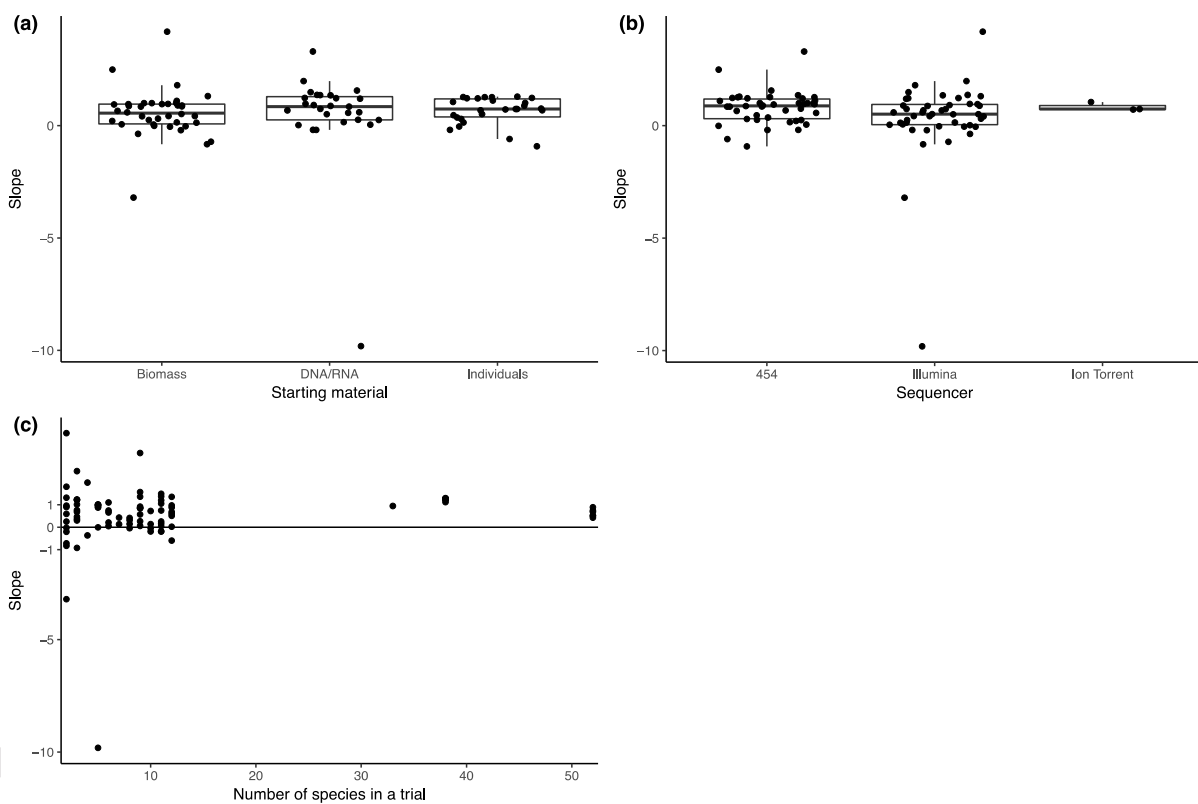


Figure 4) The quantitative ability of metabarcoding using (a) various starting materials, (b) different sequencing platforms, and (c) different number of species within in a trial. Note that each point represents a trial, which may not be fully independent from one another. However, this non-independence is accounted for in the meta-analysis model.

Sensitivity testing, using hat values and residuals (figure 5) appear to show a single trial (Hirai, Kuriyama, Ichikawa, Hidaka, & Tsuda, 2015) was having a large influence on the final output of the model. However three sequencing runs were used for a single trial in this study, and as such it has a relatively greater weight in the meta-analysis compared to most other trials that only used a single sequencing run.

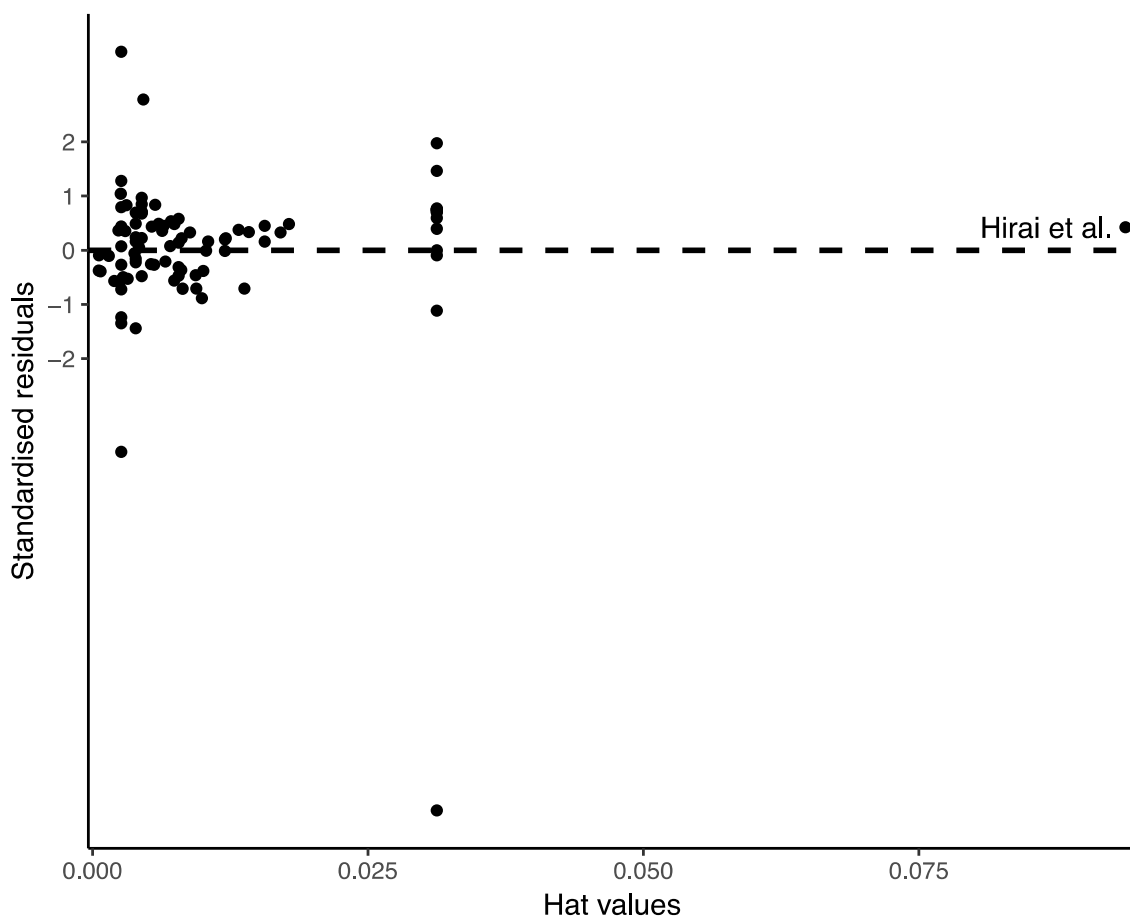


Figure 5) Hat values of trials included within the meta-analysis (a measure of influence on the meta-analysis model plotted against standardised residuals. Outlying trials are labelled. Note the points correspond to trial-level influence, not study level influence.

Discussion

How quantitative is DNA metabarcoding?

Across all studies, a slope estimate of 0.52 was identified as the relationship between biomass and sequence read number. This shows that the RRA produced from a metabarcoding loosely corresponds to the relative occurrence of species in the starting material. If no data about composition of a sampled community exists, metabarcoding data interpreted quantitatively could therefore be more informative than treating it in a strictly detected / not-detected manner even if the accuracy is low. This supports evidence from simulations presented in Deagle *et al.* (2018), which suggest that a more accurate interpretation of communities can be achieved by treating metabarcoding data quantitatively rather than relying solely on qualitative measures. However, this estimate has a large degree of uncertainty: ± 0.34 variance of slope suggests that in real world applications metabarcoding can be either somewhat quantitative, or produce a very weak signal. This uncertainty is reflected within some of the experiments themselves: figure 3 shows that while many of the included trials appeared to produce quantitative results, their variance of slopes were sufficiently large, overlapping with 0, that a non-trivial probability exists that non-quantitative

Accepted Article

data will be produced on any given sequencing run. Furthermore, there are several trials included in this meta-analysis in which metabarcoding produced extremely poor quantitative performance. With such variation between studies, and no easy way to diagnose whether any given metabarcoding study has produced quantitative results, it is easy to see how different opinions on the quantitative ability of metabarcoding has arisen. Focusing on the factors influencing the quantitative performance is essential to further clarify this situation.

The influence of sequencing technology, and initial experimental design were included as moderating factors in the initial model. The sequencing platforms did not significantly differ in quantitative ability. This was unexpected, as the different platforms have different technical approaches towards sequencing (Mardis, 2008), and different levels of bias were expected. Additionally, Illumina platforms produce many more reads than other platforms, so a greater level of precision might have been expected. This is not to say platform choice is not important when undertaking a metabarcoding study: read length, sequencing accuracy, and cost will all play a role in determining the best choice for a given study. However, these results suggest that in terms of attaining quantitative data, any difference between sequencing

technologies is too subtle to be detected in this meta-analysis, and the factors driving quantitative performance perhaps lie elsewhere in the experimental set up.

It has been hypothesised that including a higher number of species in a metabarcoding study will improve the quantitative performance as different amplification efficiencies will have diminishing effects on the overall quantitative performance as the number of species used increases (Deagle et al., 2018; Piñol, Senar, & Symondson, 2018). However, this relationship was not detected here. This may be due to most of the included studies using relatively few species: only three studies had more than 30 species. Thus, the lack of relationship may be driven by lack of variation in the data. Additionally, it is expected that different primer sets, or other factors tested here, would explain much of the variation. Our ability to detect subtle trends in a noisy dataset is limited with relatively few studies. This relationship could be better characterised with empirical studies, or if the amount of data available for meta-analysis were to increase substantially.

Different input materials had no explanatory ability in the final model: sequences were able to replicate the original biomass, quantity of DNA, or individuals in a study equally well. We believe this may be because counts of individuals were frequently used for species of similar size: if there is little variation in size of individuals between different species, count data can be regarded as a proxy for biomass. A notable exception, in using counts of individuals from species of similar sizes was the Elbrecht *et al.* 2017 study: here species were sorted by size prior to sequencing. The authors demonstrated that sorting individuals by size affected the quantitative ability of metabarcoding by comparing a mock community sorted by size, and a mock community where individual size was not considered. We used the sorted data treatment as this was most similar to other studies in the meta-analysis. However, given that counts of individuals and biomass were proxies in many studies, and empirical evidence suggests that the RRA does not correspond with the number of individuals if significant size differences are present (Elbrecht *et al.*, 2017) we would advocate caution when inferring count data from metabarcoding data without *a priori* knowledge of minimal size variation between individuals.

No difference in quantitative performance existed between studies using quantified DNA as a starting material, and those that used biomass. Given DNA extraction is the only step (figure 1) separating these points in the protocol, this suggests it is not a source of significant bias in the studies included in the meta-analysis. However, it must be noted that this is not always the case: Pornon *et al.* (2016) reported a 300-fold difference in DNA concentrations after extraction. It is possible that structural differences in the exine (the tough protective coating of pollen) may have driven the variable DNA yield. Although not a significant factor in this study, best practice would dictate that quantifying the relationship between biomass and DNA yield in the target organisms is advised prior to metabarcoding.

Future directions

This analysis has shed light onto some, but not all, of the factors that influence the quantitative performance of metabarcoding. Although not considered here, primer bias is likely a large source of variation in the quantitative performance of metabarcoding studies: Piñol *et al.* (2015) empirically tested the relationship between primer mismatch and amplification efficiency and found mismatches accounted for 75% of variation. We had hoped to explore the effect of primer bias on the

quantitative performance of metabarcoding by using the nucleotide pairwise

diversity at the primer binding site of the mock community as a moderating factor in

the final model. Unfortunately, this was not possible: the sequence in the target DNA

at the primer site could not be inferred from the studies included in this meta-

analysis as, at most, only the primer sequence can be obtained. For a number of

studies, sequences covering the primer binding-region were not present in DNA

databases. Additionally, even for those species which had relevant sequences, inter-

individual variation was a concern: amplification-efficiency is very sensitive to both

the type of nucleotide mismatch between the target DNA and the primer, and the

location of the mismatches in the primer sequence (Stadhouders et al., 2010).

Without knowing the actual sequence present in the individuals used in the studies,

we opted to omit primer site mismatches from this analysis. However, the effect of

nucleotide mismatches in primer sequences on quantitative performance of

metabarcoding is explored in detail through the use of simulations in this issue (Piñol

et al., 2018). This topic will be an ongoing research area, and until we accurately

determine the quantitative performance of any given primer set we would advocate

reporting all *in-silico* testing to assess the quantitative ability of primers, and the

inclusion of a mock community control on each sequencing run to gauge how accurately RRA corresponds with the starting material.

Reflection on meta-analysis

It is important to remember what is entailed in a meta-analysis: a consensus of studies included in the analysis, weighted by sample size. Studies were included based on their detection in a structured review; although this presents a transparent, repeatable, way of including literature, our approach may have missed some relevant studies. Indeed, not all of the high-quality literature detected in the structured review was included (Leray & Knowlton, 2017; Piñol et al., 2015), due to their experimental-design being incompatible with our analytical framework, rather than any shortcomings of the work or relevance to contribute further understanding on the topic.

It should be noted, that incorporating results into meta-analysis necessitates some loss of nuance in the results. Most notably, in this study, we used the slope derived from a linear model as an effect size to facilitate synthesis. However, the quantitative

nature of the relationships reported in this analysis may well be more complex than reported by a linear model. As such, we would encourage readers to use this manuscript only as reference material, and assess the cited literature themselves, as a perfect distillation of included literature is inherently not possible.

Furthermore, publication bias remains an unknown factor. Using a nested-model to account for non-independence makes using most common tests for publication bias problematic as they detect the structure implemented in the model, not genuine publication bias. Not accounting for the non-independence of trials run on the same sequencing run was, we felt, a more immediate flaw than accounting for publication bias. That unfortunately leaves us in a position where the extent of any publication bias is unknown, and we are unable to say how important, or trivial, the issue may be: as such we reiterate that any synthesis drawn from this model may have been influenced by the omission of unpublished data, as much as the studies included.

Another issue worth considering is the relative weighting given to each study. Meta-analyses differ from a simple vote-count by assigning increased weighting to studies with a larger sample size. Here, weighting was assigned based on the number of sequencing runs used in a study. We feel this weighting is more appropriate than a

simple vote count but it is worth highlighting the results presented here are influenced more heavily by some studies than others, e.g. Porazinska et al. (2010) had the greatest influence on the model (21.7%) due to the study's use of seven sequencing runs.

Finally, it should be noted this analysis quantifies the understanding of the field at a point in time rather than attempting to be a final point of authority. As highlighted above, much more research is still to be done in this area, and we hope the shortcomings and gaps highlighted will be filled as exciting new research reveals a more mechanistic understanding of this topic.

Conclusion

Our meta-analysis suggests that metabarcoding possesses some quantitative ability: a cross-study slope-estimate of 0.52 was found, suggesting a weak quantitative signal is present, albeit with a large degree of uncertainty (± 0.34 variance of slope). Quantitative ability did not appear to differ among sequencing platforms, the

amount of species included in a trial, or with different starting materials: biomass, individuals, or DNA. We remain sceptical that individual count data can be reliably inferred from metabarcoding if there are large size differences between the individuals being assessed and would advise against count-based inferences without *a priori* knowledge of the community being assessed. All presented results have probably been influenced by the relatively small sample sizes: additional research is warranted to reveal the mechanistic factors driving quantitative performance. While metabarcoding may eventually become a quantitative tool, many uncertainties remain. Moving forward, we suggest explicitly testing the relationship between read abundance and input biomass using mock communities included as quantitative controls during metabarcoding. Not only will this allow researchers to assess their own study, but it will also assist future meta-analyses. We also recommend presenting all trials and simulations used in primer selection to make the rationale behind primer choice transparent. Finally, we would encourage additional empirical research into the mechanistic factors behind primer bias in metabarcoding since this is difficult to study using meta-analytical techniques, yet potentially holds the key to truly quantitative metabarcoding.

References

- Albaina, A., Aguirre, M., Abad, D., Santos, M., & Estonba, A. (2016). 18S rRNA V9 metabarcoding for diet characterization: A critical evaluation with two sympatric zooplanktivorous fish species. *Ecology and Evolution*, *6*(6), 1809–1824.
- Arai, M. N., Welch, D. W., Dunsmuir, A. L., Jacobs, M. C., & Ladouceur, A. R. (2003). Digestion of pelagic Ctenophora and Cnidaria by fish. *Canadian Journal of Fisheries and Aquatic Sciences*, *60*, 825–829.
- Blanckenhorn, W. U., Rohner, P. T., Bernasconi, M. V., Haugstetter, J., & Buser, A. (2016). Is qualitative and quantitative metabarcoding of dung fauna biodiversity feasible? *Environmental Toxicology and Chemistry*, *35*(8), 1970–1977.
- Bohmann, K., Monadjem, A., Noer, C. L., Rasmussen, M., Zeale, M. R. K., Clare, E., Jones, G., Willerslev, E., & Gilbert, M. T. P. (2011). Molecular diet analysis of two African free-tailed bats (Molossidae) using high throughput sequencing. *PLoS ONE*, *6*(6), e21441.
- Bokulich, N. A., & Mills, D. A. (2013). Improved selection of internal transcribed spacer-specific primers enables quantitative, ultra-high-throughput profiling of fungal communities. *Applied and Environmental Microbiology*, *79*(8), 2519–2526.
- Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Kelley Thomas, W., Potter, C., & Bik, H.M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, *5*(6), 68–74.
- De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, *14*(2), 306–323.
- Deagle, B. E., Chiaradia, A., McInnes, J., & Jarman, S. N. (2010). Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? *Conservation Genetics*, *11*, 2039–2048.

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., Kartzinel, T.R. & Eveson, J. P. (2018). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*. (This issue).

Diaz-Real, J., Serrano, D., Piriz, A., & Jovani, R. (2015). NGS metabarcoding proves successful for quantitative assessment of symbiont abundance: the case of feather mites on birds. *Experimental and Applied Acarology*, 67(2), 209–218.

Egge, E., Bittner, L., Andersen, T., Audic, S., de Vargas, C., & Edvardsen, B. (2013). 454 Pyrosequencing to Describe Microbial Eukaryotic Community Composition, Diversity and Relative Abundance: A Test for Marine Haptophytes. *PLoS ONE*, 8(9), e74371.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634.

Elbrecht, V., & Leese, F. (2015). Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. *Plos ONE*, 10(7), e0130324.

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, 7(17), 6918–6926.

Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-polatera, P., Beisel, J.-N., Coissac, E., Boyer, F. & Leese, F. (2016). Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ*, 4, e1966.

Geisen, S., Laros, I., Vizcaíno, A., Bonkowski, M., & De Groot, G. A. (2015). Not all are free-living: High-throughput DNA metabarcoding reveals a diverse community of protists parasitizing soil metazoa. *Molecular Ecology*, 24(17), 4556–4569.

Gonzalez, J. M., Portillo, M. C., Belda-Ferre, P., & Mira, A. (2012). Amplification by PCR artificially reduces the proportion of the rare biosphere in microbial communities. *PLoS ONE*, 7(1), e29973.

Hatzenbuehler, C., Kelly, J. R., Martinson, J., Okum, S., & Pilgrim, E. (2017). Sensitivity and accuracy of high-throughput metabarcoding methods for early detection of invasive fish species. *Scientific Reports*, *7*, 46393.

Hirai, J., Kuriyama, M., Ichikawa, T., Hidaka, K., & Tsuda, A. (2015). A metagenetic approach for revealing community structure of marine planktonic copepods. *Molecular Ecology Resources*, *15*(1), 68–80.

Iwanowicz, D. D., Vandergast, A. G., Cornman, R. S., Adams, C. R., Kohn, J. R., Fisher, R. N., & Brehme, C. S. (2016). Metabarcoding of fecal samples to determine herbivore diets: A case study of the endangered Pacific pocket mouse. *PLoS ONE*, *11*(11), e0165366.

Juen, A., & Traugott, M. (2006). Amplification facilitators and multiplex PCR: Tools to overcome PCR-inhibition in DNA-gut-content analysis of soil-living invertebrates. *Soil Biology and Biochemistry*, *38*(7), 1872–1879.

Klymus, K. E., Marshall, N. T., & Stepien, C. A. (2017). Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLoS ONE*, *12*(5), e0177643.

Kowalczyk, R., Taberlet, P., Coissac, E., Valentini, A., Miquel, C., Kamiński, T., & Wójcik, J. M. (2011). Influence of management practices on large herbivore diet-Case of European bison in Białowieża Primeval Forest (Poland). *Forest Ecology and Management*, *261*(4), 821–828.

Kraaijeveld, K., de Weger, L. A., Ventayol García, M., Buermans, H., Frank, J., Hiemstra, P. S., & den Dunnen, J. T. (2015). Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Molecular Ecology Resources*, *15*(1), 8–16.

Krahn, U., Binder, H., & König, J. (2013). A graphical tool for locating inconsistency in network meta-analyses. *BMC Medical Research Methodology*, *13*(35).

Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *British Medical Journal*, *333*, 597–600.

Leray, M., & Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ*, *5*, e3006.

Mardis, E. R. (2008). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*, *9*, 387–402.

Mohorianu, I., Bretman, A., Smith, D. T., Fowler, E. K., Dalmay, T., & Chapman, T. (2017). Comparison of alternative approaches for analysing multi-level RNA-seq data. *PLoS ONE*, *12*(8), e0182694.

Paula, D. P., Linard, B., Andow, D. A., Sujii, E. R., Pires, C. S. S., & Vogler, A. P. (2015). Detection and decay rates of prey and prey symbionts in the gut of a predator through metagenomics. *Molecular Ecology Resources*, *15*(4), 880–892.

Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, *15*(4), 819–830.

Piñol, J., Senar, M. A., & Symondson, W. O. C. (2018). The choice of universal primers and the characteristics of the species mixture determines when DNA metabarcoding can be quantitative. *Molecular Ecology*, (*This issue*).

Plummer, E., & Twin, J. (2015). A Comparison of Three Bioinformatics Pipelines for the Analysis of Preterm Gut Microbiota using 16S rRNA Gene Sequencing Data. *Journal of Proteomics & Bioinformatics*, *8*(12), 283–291.

Pochon, X., Bott, N. J., Smith, K. F., & Wood, S. A. (2013). Evaluating Detection Limits of Next-Generation Sequencing for the Surveillance and Monitoring of International Marine Pests. *PLoS ONE*, *8*(9), e73935.

Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology*, *21*(8), 1931–1950.

Porazinska, D. L., Sung, W., Giblin-Davis, R. M., & Thomas, W. K. (2010). Reproducibility of read numbers in high-throughput sequencing analysis of

nematode community composition and structure. *Molecular Ecology Resources*, 10(4), 666–676.

Pornon, A., Escaravage, N., Burrus, M., Holota, H., Khimoun, A., Mariette, J., Pellizzari, C., Iribar, A., Etienne, R., Taberlet, P., Vidal, M., Winterton, P., Zinger, L. & Andalo, C. (2016). Using metabarcoding to reveal and quantify plant-pollinator interactions. *Scientific Reports*, 6, 27282.

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.r-project.org/>.

Rohatgi, A. (2017). WebPlotDigitizer. Available at: <https://automeris.io/webplotdigitizer>.

Rocchi, S., Valot, B., Reboux, G., & Millon, L. (2017). DNA metabarcoding to assess indoor fungal communities: Electrostatic dust collectors and Illumina sequencing. *Journal of Microbiological Methods*, 139, 107–112.

Rohatgi, A. (2017). WebPlotDigitizer. Available at: <https://automeris.io/WebPlotDigitizer>.

Rosenberg, M. S., Rothstein, H. R., & Gurevitch, J. (2013). Effect sizes: conventional choices and calculations. In *Handbook of Meta-analysis in Ecology and Evolution* (pp. 61–71). Oxford: Princeton University Press.

Saitoh, S., Aoyama, H., Fujii, S., Sunagawa, H., Nagahama, H., Akutsu, M., Shinzato, N., Kaneko, N., & Nakamori, T. (2016). A quantitative protocol for DNA metabarcoding of springtails (Collembola). *Genome*, 59, 705–723.

Schiebelhut, L. M., Abboud, S. S., Gómez Daglio, L. E., Swift, H. F., & Dawson, M. N. (2017). A comparison of DNA extraction methods for high-throughput DNA analyses. *Molecular Ecology Resources*, 17(4), 721–729.

Smith, K. F., Kohli, G. S., Murray, S. A., & Rhodes, L. L. (2017). Assessment of the metabarcoding approach for community analysis of benthic-epiphytic dinoflagellates using mock communities. *New Zealand Journal of Marine and*

Freshwater Research, 51(4), 555–576.

- Soininen, E. M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., Gussarova, G., Bellemain, E., Hassel, K., Stenøien, H.K., Epp, L., Schrøder-Nielsen, A., Brochmann, C., & Yoccoz, N. G. (2015). Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. *PLoS ONE*, 10(1), e0115335.
- Sousa, L. L., Xavier, R., Costa, V., Humphries, N. E., Trueman, C., Rosa, R., Sims, D.W. & Queiroz, N. (2016). DNA barcoding identifies a cosmopolitan diet in the ocean sunfish. *Scientific Reports*, 6, 28762.
- Srivathsan, A., Ang, A., Vogler, A. P., & Meier, R. (2016). Fecal metagenomics for the simultaneous assessment of diet, parasites, and population genetics of an understudied primate. *Frontiers in Zoology*, 13(17).
- Stadhouders, R., Pas, S. D., Anber, J., Voermans, J., Mes, T. H. M., & Schutten, M. (2010). The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5' nuclease assay. *Journal of Molecular Diagnostics*, 12(1), 109–117.
- Symondson, W. O. C., & Harwood, J. D. (2014). Special issue on molecular detection of trophic interactions: unpicking the tangled bank. *Molecular Ecology*, 23, 3601–3604.
- Thielecke, L., Aranyosy, T., Dahl, A., Tiwari, R., Roeder, I., Geiger, H., Fehse, B., Glauche, I. & Cornils, K. (2017). Limitations and challenges of genetic barcode quantification. *Scientific Reports*, 7, 43249.
- Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., & Trites, A. W. (2016). Quantitative DNA metabarcoding: Improved estimates of species proportional biomass using correction factors derived from control material. *Molecular Ecology Resources*, 16(3), 714–726.
- Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. *Biological*

Conservation, 183, 4–18.

Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists.

Trends in Ecology & Evolution, 24(2), 110–117.

Vaz, A. B. M., Fonseca, P. L., Leite, L. R., Badotti, F., Salim, A. C. M., Araujo, F. M. G., Cuadros-Orellana, S., Duarte, A.A., Rosa, C.A., Oliveira, G. & Goes-Neto, A. (2017). Using next-generation sequencing (NGS) to uncover diversity of wood-decaying fungi in neotropical atlantic forests. *Phytotaxa*, 295(1), 1–21.

Vesterinen, E. J., Ruokolainen, L., Wahlberg, N., Peña, C., Roslin, T., Laine, V. N., Vasko, V., Sääksjärvi, I.E., Norrdahl, K. & Lilley, T. M. (2016). What you need is what you eat? Prey selection by the bat *Myotis daubentonii*. *Molecular Ecology*, 25(7), 1581–1594.

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package.

Journal of Statistical Software, 36(3), 1–48.

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012).

Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3, 613–623.

Acknowledgements

We would like to thank the NERC EnvEast DTP (NE/L002582/1) and Cefas for funding P.D.L studentship.

Data Accessibility

The data used in this meta-analysis is uploaded to the Dryad data repository:

doi:10.5061/dryad.085jj60

Author Contributions

P.D.L conceived the study, conducted review and meta-analysis, and wrote the first draft of the manuscript. The supervisory team (S.C., E.H., J.K.P., R.G.D. and M.I.T.)

critiqued and commented on the analyses and edited the manuscript.

Tables & Figures

Table 1. Articles included in the meta-analysis.

Author	Species per trial	Sequencer	Starting material	Organisms	Marker
Albaina, Aguirre, Abad, Santos, & Estonba (2016)	6	454	Biomass	Marine invertebrates (Crustaceans, Annelids)	18s
Blanckenhorn, Rohner, Bernasconi, Haugstetter, & Buser (2016)	4 to 9	Illumina	Biomass	Macroinvertebrates (Coleoptera, Diptera, Hymenoptera)	COI
(Bokulich & Mills, 2013)	12	Illumina	DNA/RNA	Yeast	ITS
(Deagle, Chiaradia, McInnes, & Jarman, 2010)	3	454	Biomass & Faecal	Fish	16s
(Diaz-Real, Serrano, Piriz, & Jovani, 2015)	3	454	Individuals	Feather mites	COI
(Egge et al., 2013)	11	454	DNA/RNA	Haptophytes	18s
(Elbrecht & Leese, 2015)	52	Illumina	Biomass	Macroinvertebrates (freshwater)	COI

(Elbrecht et al., 2016)	52	Illumina	Individuals	Macroinvertebrates (freshwater)	COI
(Elbrecht et al., 2017)	52	Illumina	Biomass	Macroinvertebrates (freshwater)	16s
(Geisen, Laros, Vizcaíno, Bonkowski, & De Groot, 2015)	8	454	Individuals	Protist culture	18s
(Hatzenbuehler, Kelly, Martinson, Okum, & Pilgrim, 2017)	5	454	Biomass	Fish	COI
(Hirai et al., 2015)	33	454	Biomass	Copepods	LSU
(Iwanowicz et al., 2016)	12	Illumina	DNA/RNA	Plants	ITS
(Klymus, Marshall, & Stepien, 2017)	11	Illumina	DNA/RNA	Bivalves, Gastropods	16s
(Kraaijeveld et al., 2015)	6 to 11	Ion Torrent	Individuals	Plants (pollen)	TrnL
(Pochon, Bott, Smith, & Wood, 2013)	9	454	DNA/RNA	Marine invertebrates (Echinoderms, crustaceans, ascidians, molluscs, annelids)	18s
(Porazinska, Sung, Giblin-Davis, & Thomas, 2010)	38	454	Individuals	Nematodes	18s

(Rocchi, Valot, Reboux, & Millon, 2017)	9	Illumina	DNA/RNA		Fungus	ITS2
(Saitoh et al., 2016)	9	454	Biomass	Macroinvertebrates (springtails)		16s, COI
(Smith, Kohli, Murray, & Rhodes, 2017)	10	Illumina	Individuals		Dinoflagellates	Cyt b, LSU, 18s
(Thielecke et al., 2017)	5	Illumina	DNA/RNA		Plasmid constructs	n/a
(Thomas, Deagle, Eveson, Harsch, & Trites, 2016)	3	Illumina	Biomass		Fish	16s