RESEARCH ARTICLE

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

Scrutinizing key steps for reliable metabarcoding of environmental samples

Antton Alberdi[1] (iD) | Ostaizka Aizpurua[1] | M. Thomas P. Gilbert[1,2,3] | Kristine Bohmann[1,4]

[1]Section for Evolutionary Genomics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

[2]Trace and Environmental DNA Laboratory, Department of Environment and Agriculture, Curtin University, Perth, WA, Australia

[3]NTNU University Museum, Trondheim, Norway

[4]School of Biological Sciences, University of East Anglia, Norwich, Norfolk, UK

Correspondence
Antton Alberdi
Email: antton.alberdi@snm.ku.dk

Abstract

1. Metabarcoding of environmental samples has many challenges and limitations that require carefully considered laboratory and analysis workflows to ensure reliable results. We explore how decisions regarding study design, laboratory set-up, and bioinformatic processing affect the final results, and provide guidelines for reliable study of environmental samples.

2. We evaluate the performance of four primer sets targeting COI and 16S regions characterizing arthropod diversity in bat faecal samples, and investigate how metabarcoding results are affected by parameters including: (1) number of PCR replicates per sample, (2) sequencing depth, (3) PCR replicate processing strategy (i.e. either additively, by combining the sequences obtained from the PCR replicates, or restrictively, by only retaining sequences that occur in multiple PCR replicates for each sample), (4) minimum copy number for sequences to be retained, (5) chimera removal, and (6) similarity thresholds for Operational Taxonomic Unit (OTU) clustering. Lastly, we measure within- and between-taxa dissimilarities when using sequences from public databases to determine the most appropriate thresholds for OTU clustering and taxonomy assignment.

3. Our results show that the use of multiple primer sets reduces taxonomic biases and increases taxonomic coverage. Taxonomic profiles resulting from each primer set are principally affected by how many PCR replicates are carried out per sample and how sequences are filtered across them, the sequence copy number threshold and the OTU clustering threshold. We also report considerable diversity differences between PCR replicates from each sample. Sequencing depth increases the dissimilarity between PCR replicates unless the bioinformatic strategies to remove allegedly artefactual sequences are adjusted according to the number of analysed sequences. Finally, we show that the appropriate identity thresholds for OTU clustering and taxonomy assignment differ between markers.

4. Metabarcoding of complex environmental samples ideally requires (1) investigation of whether more than one primer sets targeting the same taxonomic group is needed to offset primer biases, (2) more than one PCR replicate per sample, (3) bioinformatic processing of sequences that balance diversity detection with removal of artefactual sequences, and (4) empirical selection of OTU clustering and taxonomy assignment thresholds tailored to each marker and the obtained taxa.

1 | INTRODUCTION

Metabarcoding is one of the most broadly employed molecular approaches for detection of biodiversity in environmental samples such as soil, water, gut contents and faeces (reviewed in Bohmann et al., 2014). The range of metabarcoding studies and questions addressed (e.g. Chariton, Court, Hartley, Colloff, & Hardy, 2010; Comtet, Sandionigi, Viard, & Casiraghi, 2015) highlights the importance of ensuring reliable results, but despite a seemingly straightforward process, metabarcoding has its challenges and limitations (reviewed in Pompanon et al., 2012). It is therefore essential, that researchers acknowledge these to design experimental and analytical approaches that optimize detection of true diversity while minimizing errors.

One of the initial critical decisions in a metabarcoding study is which marker region and primers to use. Marker choice is often guided by DNA reference database coverage, thus studies that target animals usually rely on markers within the standardized barcode region of the COI gene, particularly well-represented in the Barcode of Life Database—BOLD (Ratnasingham & Hebert, 2007). However, it is virtually impossible to design metabarcoding primers that do not have mismatches to at least some of the target species' sequences, and such mismatches can result in PCR amplification biases. This problem is magnified in protein-coding genes like the COI in which it is challenging to locate conserved areas suitable as primer regions (Clarke, Soubrier, Weyrich, & Cooper, 2014; Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014).

Another key decision is how to balance detection of diversity with error removal. High-throughput sequencing magnifies the effect of PCR and sequencing errors in relation to true but low abundant sequences. In that regard, an important experimental decision is whether to make PCR replicates for each DNA extract for each marker (Robasky, Lewis, & Church, 2014). PCR replicates allow researchers to optimize diversity detection by counteracting effects of PCR stochasticity (Leray & Knowlton, 2015), which might be especially high in low template complex DNA extracts (Murray, Coghlan, & Bunce, 2015). Analytically, if PCR replicates are made, they are often used additively, that is, through pooling the sequences of a single sample's PCR replicates to maximize diversity detection (e.g. Burgar et al., 2014; Leray & Knowlton, 2015). While this certainly reduces the risk of missing taxa, false positives are a potential consequence due to accumulation of artefactual sequences. To counteract this, a few studies have used PCR replicates in a restrictive context by only retaining sequences that are shared by a number of a sample's PCR replicates (De Barba et al., 2014; Giguet-Covex et al., 2014; Hope et al., 2014). However, the most widely used strategy for eliminating erroneous sequences is to set a minimum sequence copy number below which sequences are

discarded. While many authors opt for removing singletons (e.g. Burgar et al., 2014), others set higher copy number thresholds (e.g. Giguet-Covex et al., 2014). One specific group of PCR artefacts are chimeric sequences (Judo, Wedel, & Wilson, 1998), which can be predicted *in silico* using dedicated software, although these can also misidentify real sequences as chimeric (Bjørnsgaard Aas, Davey, & Kauserud, 2016). Clustering of sequences into Operational Taxonomic Units, OTUs, is another common approach to remove artefacts. Sequences are clustered into OTUs based on a similarity threshold that mirrors natural intraspecific divergence. This approach generates "species equivalents" that can be used to approximate species numbers in metabarcoding studies irrespective of DNA reference database coverage (Clare, Chain, Littlefair, & Cristescu, 2016). However, OTU clustering can cause both oversplitting and undersplitting of sequences into species equivalents, resulting in either inflated or underestimated species counts (Clare et al., 2016). Overall, whatever the approach used to remove erroneous sequences, when it is too conservative then rare species might also be removed from the dataset, whereas when it is too lenient, artefacts might be interpreted as real diversity (De Barba et al., 2014).

A last crucial decision is taxonomy assignment using reference databases such as GenBank or BOLD. Taxonomy can be assigned using either a fixed identity threshold where species- or genus-level assignments are only made if the match is above the defined threshold, usually 98% (Clare et al., 2014), or a multi-level assignment approach, where assignments at multiple taxonomic levels are done using different identity thresholds (Munch, Boomsma, Huelsenbeck, Willerslev, & Nielsen, 2008).

In summary, researchers must make multiple decisions throughout a metabarcoding study before they finally obtain an overview of which taxa are present in the samples. With an eye on providing recommendations for future metabarcoding studies, in this study we use "real" environmental DNA (eDNA) extracts obtained from bat faeces to investigate how different decisions in a metabarcoding workflow affect the final results (Figure 1). We first compare diversity detection by four primer sets targeting COI and 16S to assess biases and complementarity. Second, we measure the effects on the final results of (1) the number of PCR replicates employed, (2) how PCR replicates are used in the bioinformatic analyses, (3) sequencing depth, (4) the minimum sequence copy number below which sequences are discarded, (5) chimera removal, and (6) the OTU clustering similarity threshold. Third, we compare the sequence outputs of PCR replicates from each sample in order to evaluate the reliability of the retained sequences. Fourth, we measure the relative amount of target DNA present in the samples using shotgun metagenomics. Finally, we analyse the within- and between-taxa identities at different taxonomic levels using
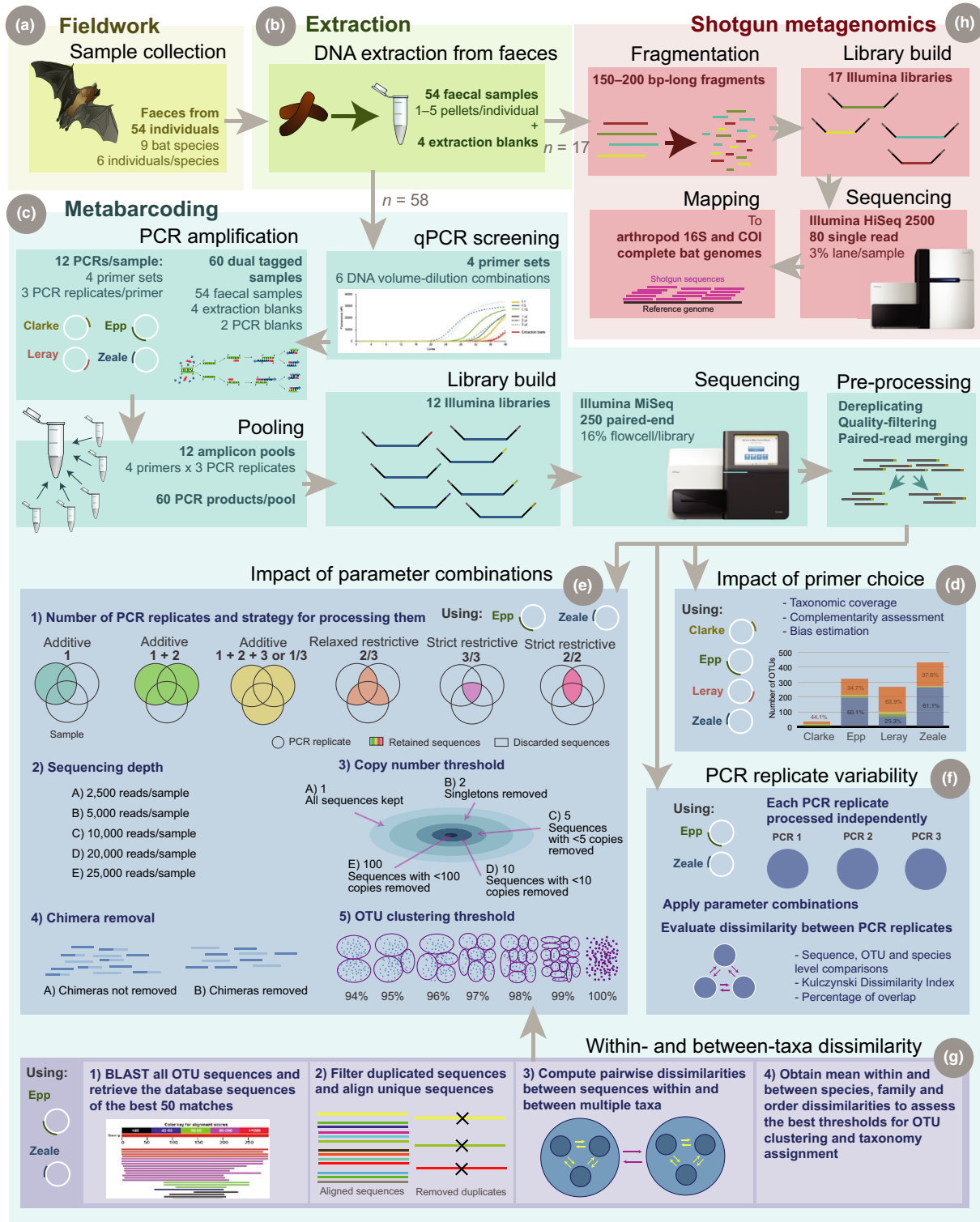
FIGURE 1 Study overview. (a) Faecal samples were obtained from wild bats captured in cave entrances. (b) DNA was extracted from the bat faecal samples in a dedicated pre-PCR laboratory. (c) The 54 samples and six extraction blanks were processed as described in Appendix S1 to obtain amplicon libraries. (d) After applying a common bioinformatic pipeline (also described in Appendix S1), we compared the diversity detected by each of the four primer sets and assessed their complementarity and taxonomic biases. (e) We evaluated the impact of different parameters on the detected diversity by comparing 2,100 parameter combinations. (f) We evaluated the effect of parameter combinations on the dissimilarity between PCR replicates. (g) Dissimilarity between and within taxonomic groups was measured using sequences from reference databases in order to assess the optimal thresholds for Operational Taxonomic Unit (OTU) clustering and taxonomy assignment. (h) Finally, we estimated the relative amount of target DNA in the faecal extracts by shotgun sequencing 17 bat faeceal extracts

sequences from reference databases to identify the most appropriate OTU clustering and taxonomy assignment thresholds.

2 | MATERIALS AND METHODS

2.1 | Laboratory work

DNA was extracted from faecal samples from 54 insectivorous bats (Figure S1). Each of the four extraction rounds included one extraction blank. DNA was PCR amplified from the extracts using four primer sets as detailed in Table 1 and Figure S2, hereafter referred to as Clarke (Clarke et al., 2014; Elbrecht et al., 2016), Epp (Epp et al., 2012), Leray (Geller, Meyer, Parker, & Hawk, 2013; Leray et al., 2013) and Zeale (Zeale, Butlin, & Barker, 2011). After optimizing PCR conditions through qPCR screening (Murray et al., 2015) (Figure S3), each of the 54 bat faecal extracts, four extraction blanks and two PCR blanks were amplified in three PCR replicates. PCR products from each primer set were subsequently combined into three amplicon pools, each containing one of the three PCR replicates of each sample. Amplicon pools were converted into Illumina sequencing libraries and the 12 indexed amplicon libraries were sequenced using 250 bp paired-end chemistry on two Illumina MiSeq flowcells, aiming for >1.5 million reads per library (25,000 reads/PCR replicate/sample). Further details in Appendix S1.

2.2 | Impact of primer choice on detected diversity

We compared arthropod detection by the Clarke, Epp, Leray and Zeale primers (Figure 1d) in the bat faecal samples by applying the bioinformatic pipeline based on DAMe (Zepeda-Mendoza, Bohmann, Carmona Baez, & Gilbert, 2016) detailed in Appendix S1 with the following parameters: (1) the number of quality-filtered sequences in the 12 datasets was randomly normalized to 0.6 million sequences (on average 10,000 sequences/PCR replicate/sample) to make them comparable to the Leray dataset, which contained fewer sequences than the other datasets, (2) chimeric sequences and sequences identical to those of the corresponding blank were removed, (3) the copy number threshold per dataset was set to two, (4) only sequences appearing in at least in two out of three PCR replicates were retained, and 5) sequences were clustered into OTUs using a 97% similarity threshold.

The effects of primer choice on biodiversity detection were measured using pairwise Kulczynski Dissimilarity Indexes—KDI (Kulczynski, 1927) of the yielded diversity profiles at order and species levels. KDI ranges from 0 (equal) to 1 (completely different). Taxonomic coverage of the primer sets was assessed as the number of sequences and OTUs assigned to arthropods (target prey taxa), bats (non-target predator taxa), and other organisms (e.g. fungi), as well as OTUs not assigned to any organism. Complementarity and biases between primer sets were assessed by comparing the order- and species-level taxonomic diversity profiles, and visualized in network diagrams using Cytoscape 3.4 (Shannon et al., 2003). Due to the Clarke and Leray primers' poor or inefficient detection of target taxa, further analyses were restricted to the Epp and Zeale datasets.

TABLE 1 Details of the four primer sets used in this study

Custom name	Region	Primer names	Forward primer 5'–3'	Reverse primer 5'–3'	Length (bp)	Target taxa	Reference(s)
Clarke	16S	F: Ins16S_1shortF R: Ins16S_1shortR	RGACGAGAAGACCCTATARA	ACGCTGTTATCCCTAARGTA	190 [188–192]	Insecta	Clarke et al. (2014) and Elbrecht et al. (2016)
Epp	16S	F: Coleop_16Sc R: Coleop_16Sd	TGCAAAGGTAGCATAATMATTAG	TCCATAGGGTCTTCTCGTC	106 [102–107]	Coleoptera	Epp et al. (2012)
Leray	COI	F: mlCOIintF R: jgHCO2198	GGWACWGGWTGAACWGTWTAYCCYCC	TAIACYTCIGGRTGICCRAARAAYCA	313 [311–314]	Metazoa	Geller et al. (2013) and Leray et al. (2013)
Zeale	COI	F: ZBJ-ArtF1c R: ZBJ-ArtR2c	AGATATTGGAACWTTATATTTTATTTTTGG	WACTAATCAATTWCCAAATCCTCC	157 [157–159]	Arthropoda	Zeale et al. (2011)

Length refers to the amplicon size excluding primers.

2.3 | Impact of parameter combinations on detected diversity

We evaluated the effect of combinations of the following parameters on the yielded diversity (Figure 1e):

1. Number of PCR replicates and strategy for processing them: We compared six approaches using one or multiple PCR replicates applying both additive and restrictive strategies. In the additive approach, the sequences of the three different PCR replicates from each sample were used in three different ways: using only the first PCR replicate (1), merging the first two PCR replicates (1 + 2) and combining the three PCR replicates (1 + 2 + 3). In the restrictive approach, only the sequences present in at least *n* out of *m* PCR replicates (*n/m*) were retained. One relaxed (2/3) and two strict (3/3 and 2/2) restrictive approaches were compared.
2. Sequencing depth: We compared sequencing depths of 0.15, 0.3, 0.6, 1.2 and 1.5 million reads per library, corresponding to an average of ca. 2,500, 5,000, 10,000, 20,000 and 25,000 reads/PCR replicate.
3. Copy number threshold: For each PCR replicate, we explored the effect of retaining sequences above a minimum number of copies threshold: 1 (=all sequences retained), 2 (=singletons removed), 5, 10 and 100.
4. Chimera removal: Whether chimera removal was performed or not.

5. OTU clustering threshold: The effect of clustering sequences into OTUs with similarity thresholds that range from 94% to 100% (=no clustering).

We thus compared 2,100 parameter combinations per primer set. The relative impact of each of the parameters on the final results was analysed using a PERMANOVA, and a subset of parameter combinations was plotted in a heatmap chart (Figure 3). Subsequently, we performed the steps described below to disentangle the variability observed between parameter combinations, thus attempt to build an understanding of the underlying causes of variation and identify the most appropriate values for each parameter. Details in Appendix S1.

2.4 | PCR replicate variability and sequence reliability

We used the PCR replicates from each sample to assess the reliability of the sequences that were retained after applying the different parameter combinations (Figure 1e,f). In doing so, we relied on the assumption that sequences that appear in multiple PCR replicates are more likely to be real rather than randomly generated artefactual sequences. To assess the effect of the different parameters on the sequence reliability, we examined whether, and how, the dissimilarity between PCR replicates from each sample (as measured by computing KDI values and overlap percentages), varied depending on sequencing depth, copy number threshold, chimera
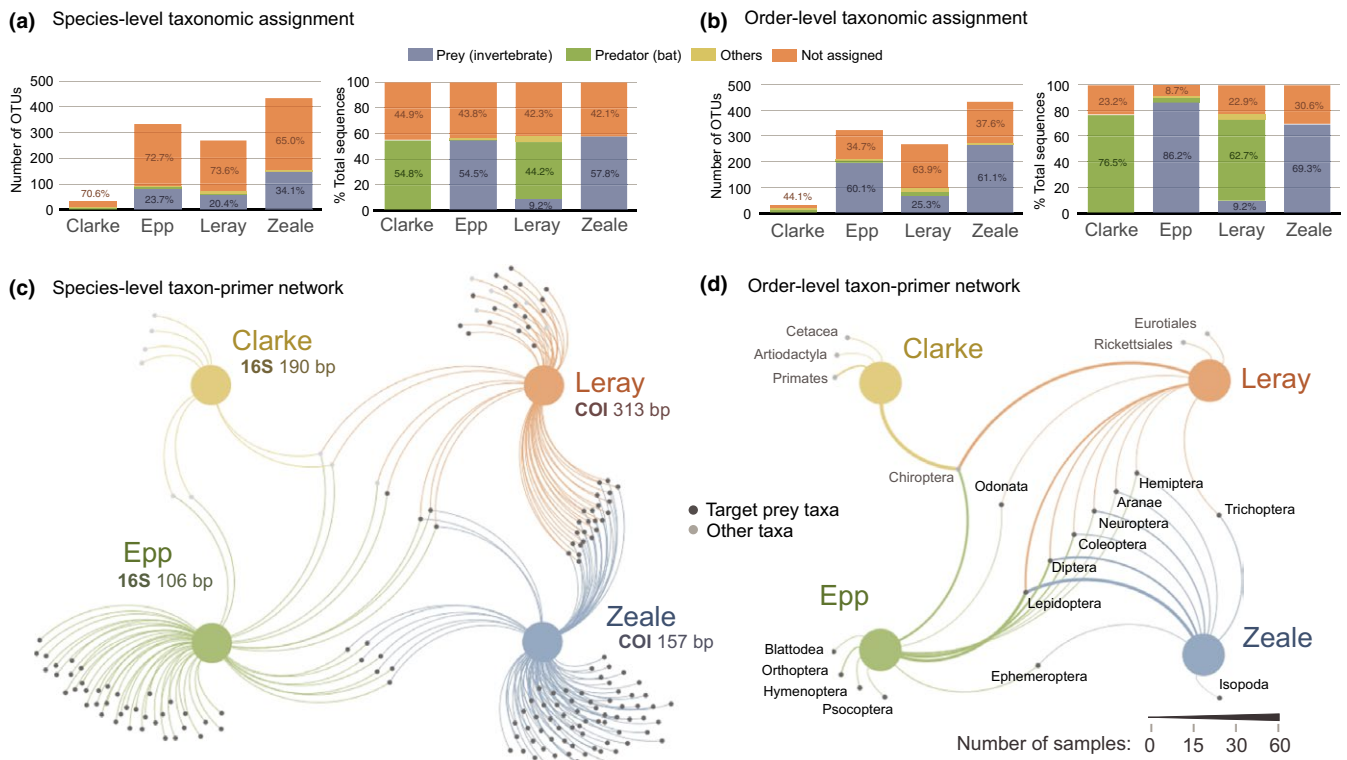
FIGURE 2 Taxonomic assignment of Operational Taxonomic Units (OTUs) generated through metabarcoding of 54 faecal samples from insect-eating bats using four primer sets. (a, b) Number of OTUs and the relative number of total sequences (a) assigned to species-level using a 98% BLAST identity threshold, and (b) assigned to order-level using a 95% identity threshold. (c, d) Network diagrams connecting primers with detected (c) species and (d) orders
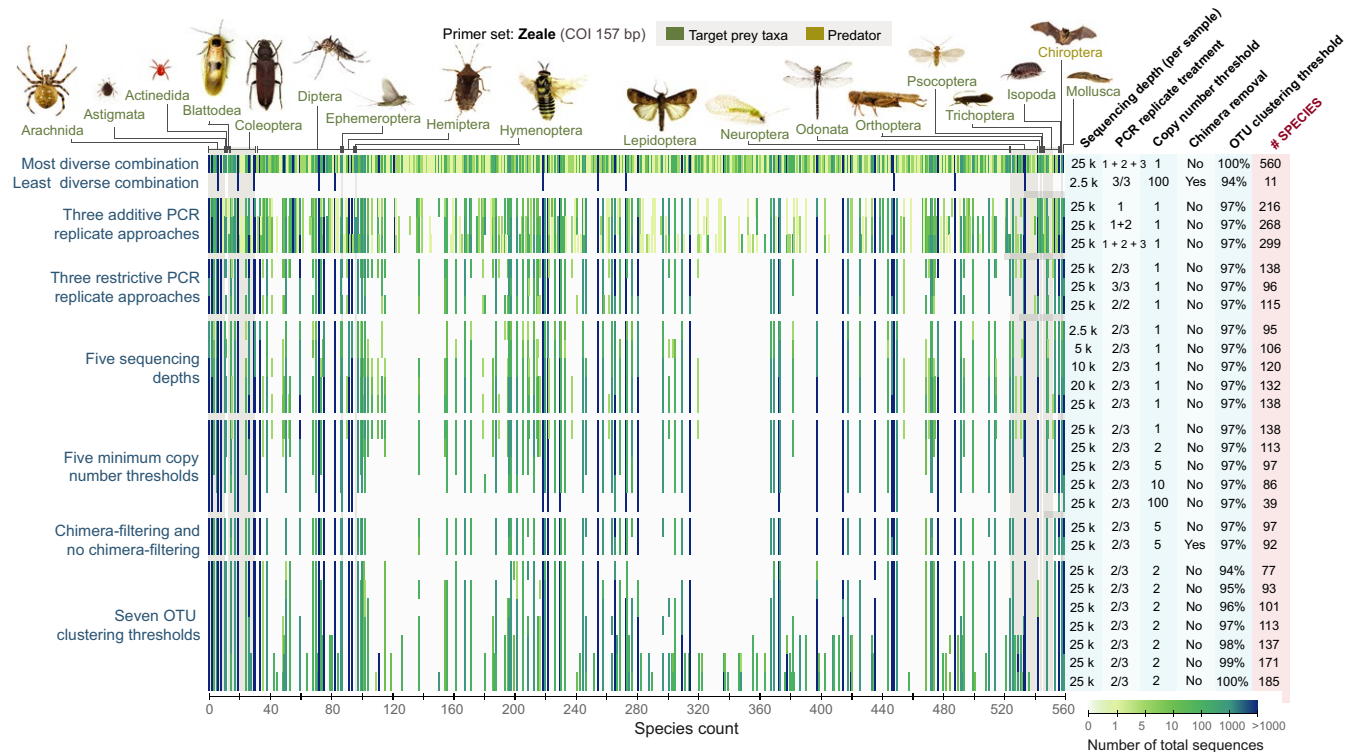
FIGURE 3 Heatmap chart showing the number of sequences assigned to detected species using multiple parameter combinations applied to the Zeale dataset. The figure shows 27 parameter combinations selected out of the analysed 2,100 options explained in Figure 1e. Vertical bars refer to species while their colour indicates the number of sequences assigned to each species considering all 54 samples

removal and OTU clustering thresholds. For each of these parameters, we carried out repeated two-way ANOVAs (rmANOVA).

2.5 | Within- and between-taxa dissimilarity

To identify the most appropriate similarity levels for OTU clustering and taxonomic assignment using the Zeale and Epp primers, we measured the within- and between-taxa dissimilarity levels for the marker regions and studied target taxa using sequences retrieved from GenBank (Figure 1g). We calculated the pairwise % identities between randomly picked sequences within different taxonomic groups and levels; namely, within species, between species, between families, and between orders. Details in Appendix S1.

2.6 | Target DNA estimation

We estimated the relative amount of target DNA in the bat faecal extracts in order to shed light on the causes of the observed PCR stochasticity (Figure 1h). We did this through shotgun sequencing a subset ($n = 17$) of the bat faecal DNA extracts. The DNA was fragmented, built into Illumina shotgun libraries, and sequenced using 80 bp single-read chemistry on 50% of an Illumina HiSeq 2500 lane. The relative amount of target template DNA was estimated by mapping the quality-filtered reads to (1) whole bat genomes and (2) arthropod COI and 16S sequences retrieved from the NCBI genome and nt databases respectively (details in Appendix S1).

3 | RESULTS

Sequencing of the 12 metabarcoding libraries yielded 29.8 million paired-end reads, while the shotgun sequencing yielded 132.3 million single reads (details in Table S1).

3.1 | Primer choice

The Zeale and Epp primers had the highest arthropod prey assignment success at species level (Figure 2a). For both markers, taxonomy assignments were almost doubled when using a 95% BLAST identity threshold (order-level assignment) instead of 98% (species-level assignment). All Zeale OTUs with ≥95% BLAST matches were assigned to invertebrates, but some Epp and Leray OTUs were assigned to non-invertebrate taxa including chiroptera (i.e. predator) and for Leray also fungi (Figure 2b). While in the Epp dataset non-invertebrate OTUs accounted for a very small fraction of the total number of sequences, in the Leray dataset over half of the sequences were assigned to non-invertebrate taxa. All taxonomic assignments in the Clarke dataset belonged to bats (Figure 2b).

The taxa identified with the four primers sets were rather different at the order level ($KDI_{order} = 0.58 \pm 0.26$) and very different at the species level ($KDI_{species} = 0.82 \pm 0.14$). Thirty-four invertebrate species were shared by at least two primer sets, while 130 were unique to a single primer set (Figure 2c). We identified 14 invertebrate orders in total, and no single primer set detected all orders. The highest target

phylogenetic diversity was recovered by the Epp primer set (12 invertebrate orders), followed by the Zeale (9) and Leray (8) primer sets (Figure 2d). The Epp primers also showed the highest uniqueness, recovering four invertebrate orders that were not detected by other primers.

3.2 | Parameter combinations

The analysed parameter combinations using the Epp and Zeale datasets had large effects on the final results (Figure 3). The PERMANOVA showed that all analysed parameters significantly ($p < .001$) altered the resulting species composition (Table S2). The highest variation in both datasets was introduced by the PCR replicate processing approach, copy number threshold and OTU clustering threshold ($R^2 > .15$). The relative importance of sequencing depth was lower ($R^2 \approx .05$), while chimera removal barely altered ($R^2 < .01$) the overall species composition.

3.2.1 | Number of PCR replicates and strategy for processing them

Additive use of PCR replicates increased the number of OTUs in both the Epp and the Zeale dataset; up to 95% more OTUs and 30% more species were detected when combining sequences that originated from three PCR replicates instead of one (Figure 4a). Conversely, when PCR replicates were processed using a restrictive strategy, the number of OTUs and identified species decreased considerably, with the highest decrease when using the strict approaches (Figure 4b).

3.2.2 | Sequencing depth

For the investigated sequencing depths, the detected diversity increased with sequencing depth. However, the degree of increment showed a tendency to level off at the highest sequencing depths, more so when processing PCR replicates restrictively (Figure 4c).

3.2.3 | Copy number threshold

The largest reduction in detected diversity was introduced by singleton removal, while at higher copy number thresholds the loss of diversity was attenuated (Figure 4d). The effect of removing sequences below a minimum copy threshold was much larger when using additive strategies for processing PCR replicates than when using restrictive strategies (Figure 4d). This is because restrictive PCR processing removed most of the singletons and many of the sequences with less than five copies (Figure 4e). It is noteworthy though, that even when applying the most restrictive filtering approach, i.e. the 3/3, 10% of the singletons and 43% of the sequences with less than five copies were retained (Figure 4e).

3.2.4 | Chimera removal

The effect of removing sequences identified as chimeras depended on the copy number threshold that was applied to the PCR replicates.

When applying no copy number thresholds, around 50% of OTU representative sequences were identified as chimeric, while higher copy number thresholds produced a marked drop in the amount of OTU representative sequences identified as chimeric (Figure 4f). This occurred because the chimeric sequences tended to have low number of copies (Figure S5).
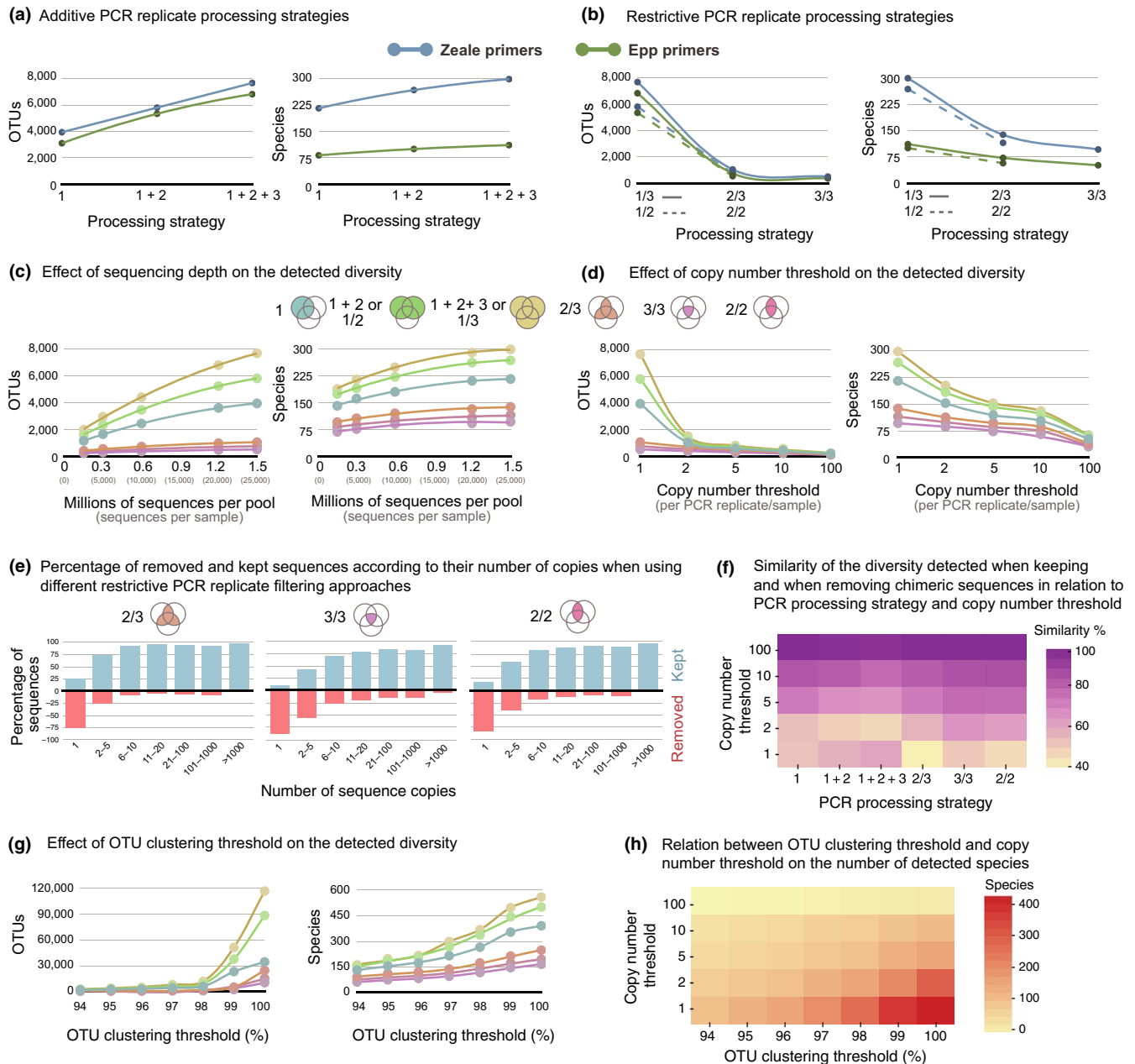
3.2.5 | OTU clustering threshold

The OTU clustering threshold had a large impact on the detected diversity. The number of species detected with the highest clustering threshold was almost tripled compared to the diversity detected with the lowest clustering threshold (Figure 4g). When no sequence copy number threshold was applied, the effect of the OTU clustering threshold was considerably larger than when high copy number thresholds were applied (Figure 4h), as many distinct sequences with low number of copies were then removed from the datasets.

3.3 | PCR replicate variability

Across all 54 ± 7.8 bat faecal samples processed using Epp and Zeale primers, the outputs of each sample's three PCR replicates were rather different: within samples, an average of 68.3 ± 10.3% of unique sequences, 69.4 ± 13.8% of OTUs and 72.3 ± 10.8% of species appeared only in a single PCR replicate. The average KDIs between PCR replicates of each sample was 0.65 ± 0.1 for sequences, 0.77 ± 0.1 for OTUs and 0.60 ± 0.2 for species. Sequencing depth increased the OTU dissimilarity between PCR replicates (rmANOVA $F = 69.82$, $p < .001$), because the number of OTUs unique to a single PCR replicate increased (Figure 5a). Raising the copy number threshold decreased the OTU dissimilarity between PCR replicates (rmANOVA $F = 119$, $p < .001$), although even at 100 copies threshold, an average of 23 ± 7.8% of OTUs were unique to a single PCR replicate (Figure 5b). Due to the antagonistic effect of the two mentioned parameters, higher sequencing depths required higher absolute copy number thresholds to obtain the same dissimilarity between PCR replicates (Figure 5c). The OTU clustering threshold also significantly altered the dissimilarity between PCR replicates (rmANOVA $F = 95.67$, $p < .001$), although its effect varied depending on the copy number threshold employed (Figure 5d). In contrast, chimera removal did not significantly alter the dissimilarity between PCR replicates (paired t-test $t = -1.55$, $p = .12$).

3.4 | Within- and between-taxa dissimilarity

The average pairwise identities between individual reference sequences, both within and between species, were different for the two markers analysed. For the Zeale marker, the overlap between intra- and interspecific dissimilarity values was relatively low, and the identity level that minimized oversplitting and undersplitting of species was 98% (Figure 6a,b). In contrast, we observed higher overlap between intra- and interspecific dissimilarity values for the Epp marker, which complicates selection of an appropriate similarity level

FIGURE 4 How PCR replicate processing strategy, sequencing depth, copy number threshold, chimera removal, and Operational Taxonomic Unit (OTU) clustering threshold affect diversity detection. Unless otherwise specified, the charts refer to the total results for the 54 faecal samples processed with Zeale primers. (a, b) The effect of processing PCR replicates additively or restrictively (see Figure 1e) on the detected diversity using the Zeale and Epp primers. (c) The effect of sequencing depth on detected diversity using different strategies for filtering of sequences across PCR replicates. (d) The effect of discarding sequences below a certain copy number threshold on the detected diversity using different strategies for processing PCR replicates. (e) The relative number of sequences according to copy number of each sequence that were kept or removed after processing PCR replicates using the 2/3, 3/3 and 2/2 restrictive strategies. (f) Similarity of the detected diversity between bioinformatic pipelines that discard or retain chimeric sequences, in relation to PCR replicate processing strategy and copy number threshold. (g) The effect of OTU clustering threshold on the detected diversity using different PCR replicate processing strategies. (h) Combined effect of OTU clustering threshold and copy number threshold below which sequences were discarded on the detected species diversity

to cluster OTUs to approximate species (Figure 6a,b). At higher taxonomic levels, between-family identity levels varied considerably, and we observed pairwise identities between families above 95% for both markers (Figure 6c). In contrast, all pairwise identity levels between insect orders were below 93% in the case of Zeale and below 87% in the case of Epp (Figure 6d).

3.5 | Target DNA

Only a small fraction of the shotgun sequences (7.7 ± 3.3 million sequences/sample) could be mapped to target 16S and COI arthropod reference sequences. Across samples, 778 ± 902 (0.009 ± 0.01%) sequences were mapped to arthropod COI sequences, while
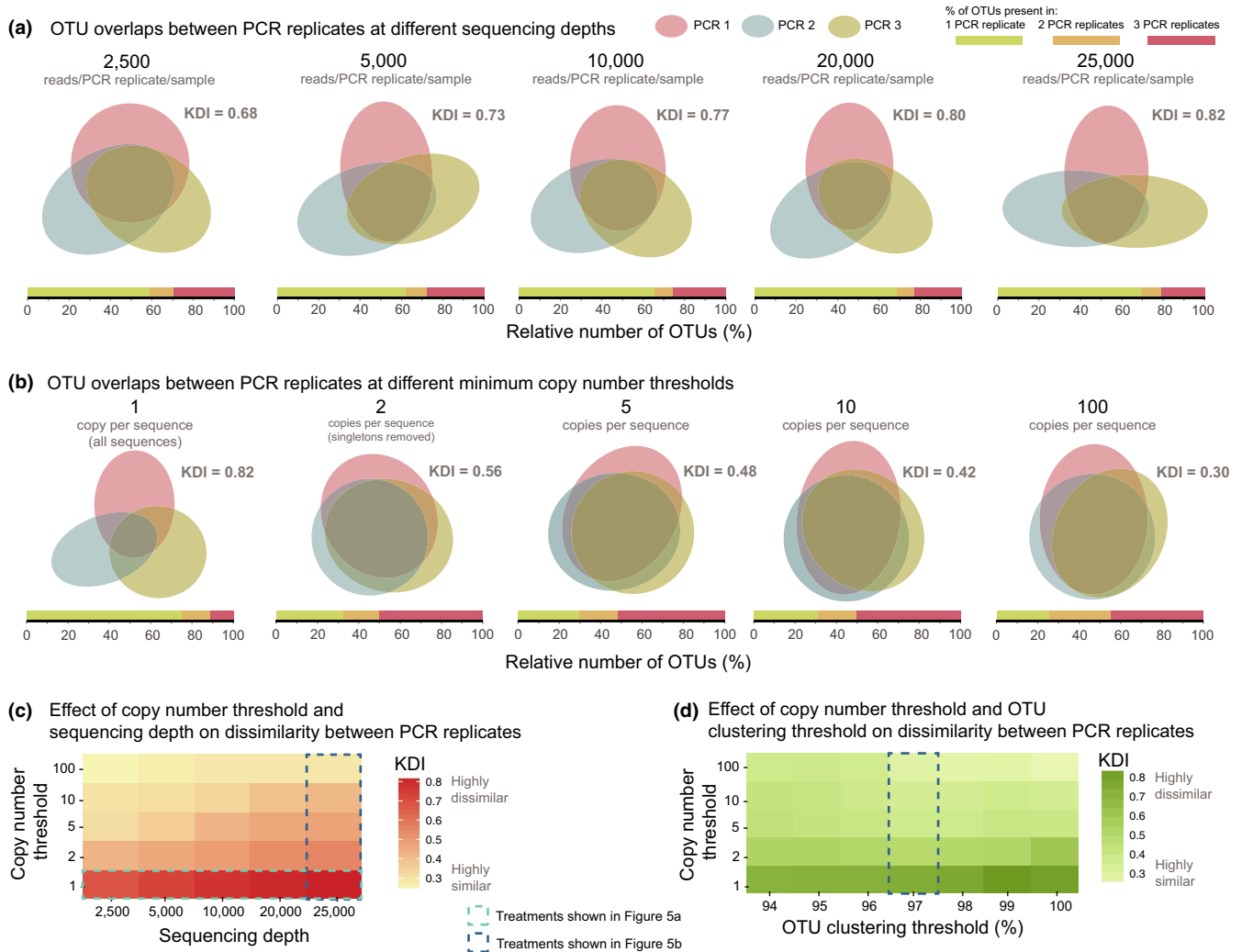
FIGURE 5 Effect of sequencing depth and copy number threshold on dissimilarity between PCR replicates. Unless otherwise stated, the charts refer to average values of the 54 bat faecal samples processed using the Zeale primers. (a, b) The average overlap of Operational Taxonomic Units (OTUs) between the three PCR replicates (a) depending on sequencing depth and (b) different minimum copy number thresholds. The horizontal bars indicate the percentage of OTUs that appeared in a single PCR replicate (green), in two PCR replicates (yellow) and the three PCR replicates (red). (c, d) Heatmaps showing (c) the combined effect of sequencing depth and copy number threshold on the average dissimilarity (KDI, Kulczynski Dissimilarity Index) values between PCR replicates, and (d) the combined effect of OTU clustering threshold and copy number threshold on the average KDI values. The higher the KDI value, the more dissimilar the OTU composition between PCR replicates.

1661 ± 1796 (0.02 ± 0.02%) were mapped to arthropod 16S sequences. In contrast, 1.6 ± 1.9 million (20.5 ± 25.0%) sequences were mapped to bat genomes (Table S3).

4 | DISCUSSION

In this study, we address a range of decisions that researchers must make when designing and carrying out metabarcoding studies (Figure 1). Unlike similar studies (Flynn, Brown, Chain, MacIsaac, & Cristescu, 2015; Leray & Knowlton, 2017; Pornon et al., 2016), our analyses are based on "real" eDNA extracts instead of mock eDNA extracts. Although the content of real eDNA extracts is unknown, the key strength is that it enables us to study the influence of

different parameters in a set-up that mirrors a true metabarcoding study in which samples consist of an unknown mix of diverse DNA (e.g. predator, prey, fungi, microbial and human DNA) in varying quantities and fragmentation stages, as well as substances that might inhibit the PCRs. As discussed later, this can lead to conclusions that differ radically from "mock eDNA" approaches. In the following, in light of our results we discuss a range of questions researchers will inevitably face during a metabarcoding study.

4.1 | Should I use multiple primer sets?

In metabarcoding studies that aim to assess general biodiversity, it is common to use multiple primer sets (e.g. Evans et al., 2016), although when a specific taxonomic group is targeted, many studies rely on a
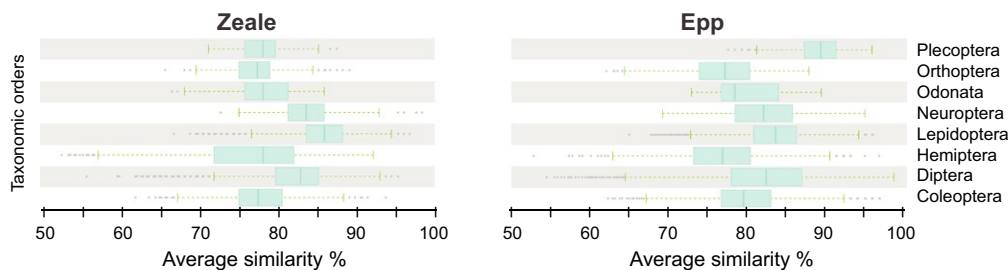
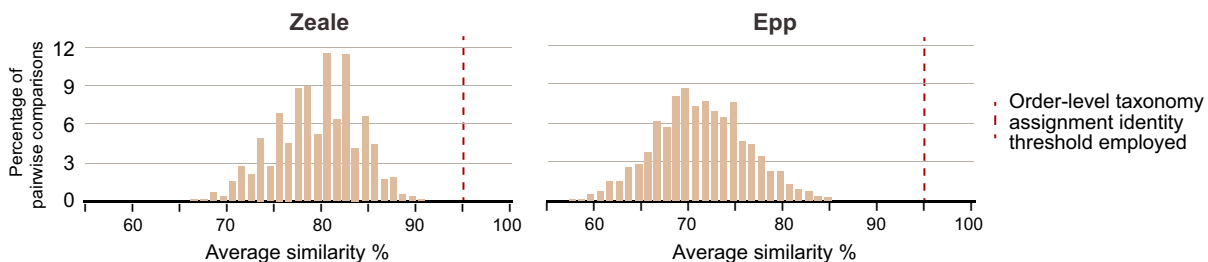**(a)** Average within- and between-species similarities
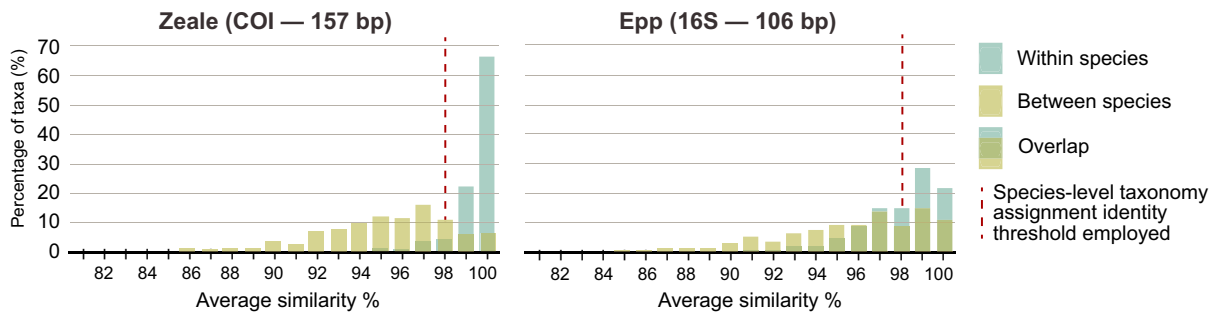
(b) Oversplitting and undersplitting of species at different similarity thresholds
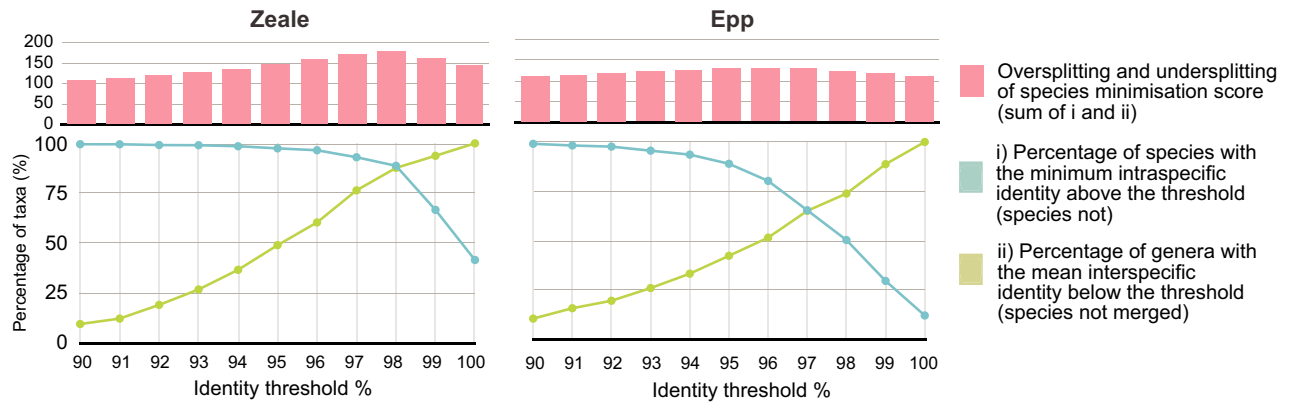
(c) Pairwise similarity values between families within different insect orders

(d) Pairwise similarity values between insect orders

FIGURE 6 Within- and between-taxa similarity values at different taxonomic levels. (a) Histograms showing the minimum genetic similarity between individuals within species (blue) and the mean similarity between species within genera (green). (b) Oversplitting and undersplitting patterns of species under different similarity thresholds. Red vertical bars show the sum of the percentage of species not split and not merged at each threshold value. (c) Pairwise similarity values between families within the eight best represented insect orders in the samples. (d) Pairwise similarity values between orders within the class Insecta

single primer set that is designed to be universal for this group (e.g. Hope et al., 2014). However, as other authors have shown (e.g. Clarke et al., 2014; Piñol, Mir, Gomez-Polo, & Agustí, 2014) we show that even when a single taxonomic group is targeted, the use of a single

primer set can yield biased results. For instance, although the Zeale primer set was originally designed to amplify orthopterans (Zeale et al., 2011), in our study, only one PCR replicate of a single sample amplified orthopterans, while the Epp primers detected them in all

three PCR replicates of multiple samples (Figure 2d). In the light of this, our results indicate that when eDNA extracts contain orthopteran DNA as well as DNA from other taxa for which the Zeale primers have higher amplification efficiency (e.g. Diptera or Lepidoptera; Clarke et al., 2014), the PCR amplification success of orthopterans decreases. This reasoning could also apply to other orders that were detected by some primers and not by others in our study (Figure 2d). Our observations highlight that the use of multiple primers targeting the same taxonomic group minimizes the effect of individual primer sets' biases and increases the taxonomic coverage, which yields a more complete picture of the diversity in the samples.

Selecting appropriate metabarcoding primers is not a simple task. Apart from reference database and taxonomic coverage, amplification of non-target templates should also be considered when selecting or designing metabarcoding primers. For instance, in this study, the Clarke, Epp, and Leray primers amplified DNA from the bat species (Figure 2). Using the Epp primers, only a tiny fraction of sequences and OTUs were assigned to the predator, while >60% of the sequences generated by the Leray primers, which are designed to cover the entire metazoan diversity, were clustered into OTUs assigned to the predators. Amplification of predator DNA is an issue that is seldom considered in either simulated *in silico* analyses or *in vitro* studies using mock communities. Our results demonstrate, however, that the performance of a primer set on real eDNA extracts can radically differ from that predicted from such studies. The Clarke primers have been reported useful for DNA metabarcoding of insects based on *in silico* and *in vitro* mock eDNA studies (Clarke et al., 2014; Elbrecht et al., 2016), but in the present study we observe that when they are used on real eDNA extracts, non-target DNA is PCR amplified to such extent that it prevents detection of the taxonomic group that the primers were designed to amplify, namely insects (Figure 2a, b). Hence, instead of relying solely on the results of *in silico* and *in vitro* analyses, we encourage researchers to perform pilot studies using a subset of their environmental samples to test the performance of the selected primers in a real setting and to make a decision on whether more than one primer set is needed to detect the desired diversity.

4.2 | Should I use 16S rRNA rather than COI primers?

Several studies have suggested that 16S rRNA markers provide broader taxonomic coverage and less biased results than the widely used COI region (e.g. Clarke et al., 2014; Deagle et al., 2014), although contradictory conclusions have been reported (Clarke, Beard, Swadling, & Deagle, 2017; Elbrecht & Leese, 2017). In our study, the Epp (16S) primers did exhibit broader taxonomic range within the arthropods (Figure 2d), although this difference could be related to the shorter marker size of the Epp primers. The analysis of the genetic dissimilarity between reference sequences revealed that the Epp primers can reliably perform order-level assignments (Figure 6d). Hence, considering their broader taxonomic coverage, they might complement the information obtained from the Zeale (COI) primers. However, the reliability of species-level identifications using the Epp

primers is much lower due to the large overlap between the within-species and between-species dissimilarity levels (Figure 6a). This also complicates the selection of an appropriate OTU clustering threshold, because some species will be oversplit, while others will be undersplit regardless of the selected threshold (Figure 6b). This is the likely reason why the Zeale and Epp primers produced so different species-level taxonomic profiles in our study (Figure 2c). Whether other 16S primers, targeting different regions and amplicon sizes, will exhibit similar issues should be empirically tested before opting for using 16S primers to replace or complement COI primers, since the performance of a marker can differ considerably depending on the primers used to amplify it (Elbrecht & Leese, 2017).

4.3 | Is using multiple PCR replicates per sample beneficial?

There were only small amounts of target DNA in the shotgun sequenced bat faecal samples. This is consistent with the levels found in other faecal samples in other studies (e.g. Srivathsan, Ang, Vogler, & Meier, 2016). Low levels of template DNA is a cause of PCR stochasticity in metabarcoding studies (Murray et al., 2015). In concordance with this, when we compared the taxonomic output of the three PCR replicates of each sample, we observed considerable disparity between PCR replicates (Figure 5), suggesting that neither of a sample's individual PCR replicates correctly characterized the true diversity. These differences could be the result of PCR stochasticity and/or accumulation of PCR and sequencing errors, two inherent sources of distortion in metabarcoding that can be minimized by using PCR replicates.

The sequences originating from a sample's PCR replicates can be processed using different bioinformatic filtering strategies (see Figure 1e), which should correspond to a tradeoff between keeping true sequences, while removing erroneous ones. Most metabarcoding studies that use multiple PCR replicates per sample rely on an additive strategy (e.g. Burgar et al., 2014; Leray & Knowlton, 2015), which as our results demonstrate, offsets the variability in individual PCR replicates and thereby maximizes diversity detection (Figure 4a). However, while this approach increases the probability of detecting rare taxa, it also artificially inflates the detected diversity by incorporating artefactual sequences (Flynn et al., 2015). A less commonly implemented strategy is to restrictively process PCR replicates (e.g. Hope et al., 2014). This strategy removes artefactual sequences, although it also implies loss of some real sequences. When comparing these two PCR replicate filtering strategies, as expected we observed that the restrictive approaches drastically decreased the detected diversity (Figure 4b) and the probability of detecting false positives. Restrictive filtering of PCR replicates is also useful for decreasing effects of potential contamination during PCR amplification. Our results suggest that strict restrictive PCR replicate filtering is useful when a high certainty of sequences is needed, and when it is not necessary to detect taxa that are only represented by few sequences. However, in cases where PCR stochasticity is high and many target taxa are present of which some has low template amount, the entire target

diversity in the sample might not be amplified in all PCR replicates of a sample (Figure 5b). Thus, in such cases, adopting a strict restrictive PCR filtering approach might entail a loss of real diversity. Conversely, if a relaxed approach of restrictive PCR replicate filtering is employed (e.g. 2/3), it can balance the positive aspects of the additive (increased detection of diversity) and strict restrictive (increased removal of erroneous sequences) approaches.

4.4 | Does deeper sequencing increase diversity detection?

In metabarcoding studies, a minimum sequencing depth is certainly needed to correctly characterize diversity of environmental samples (Smith & Peay, 2014). However, our results show that increasing the sequencing depth to very high levels does not ensure that the diversity will be better characterized. Firstly, we observed that increasing sequencing depth systematically raised the dissimilarity between PCR replicates (Figure 5a), indicating that increasing sequencing depth raises the amount of artefactual sequences. This in turn artificially inflates the detected diversity unless measures are taken to remove the erroneous sequences by adjusting bioinformatic processing parameters to the increased sequencing depth (Figure 5c), e.g. increasing absolute copy number thresholds or using relative thresholds. Secondly, even when applying restrictive PCR filtering measures to get rid of artefactual sequences, PCR replicates of most samples did not become entirely identical through increasing sequencing depth (Figure 5b). Although it has been proposed that sequencing depth, rather than PCR replication, improves metabarcoding results (Smith & Peay, 2014), our study shows that when metabarcoding analysis is used on complex environmental samples like faecal extracts, sequencing depth alone cannot counteract the effects of PCR stochasticity and primer biases. Hence, if researchers want to increase diversity detection in their samples then in addition to increasing sequencing depth, they need to consider primer choice, optimize their PCRs and include PCR replicates for each sample.

4.5 | Which minimum copy number threshold should I use?

One of the most broadly employed strategies to remove artefactual sequences is to discard sequences with copy numbers under a certain threshold (Giguet-Covex et al., 2014). The rationale behind this strategy is that the lower the copy number of a sequence, the higher the probability that it is an artefact (Leray & Knowlton, 2017). However, in our datasets around 10% of singletons appeared in multiple PCR replicates (Figure 4e), suggesting that some might be real sequences. Hence, when setting copy number thresholds researchers can opt for a high one and thereby gain reliability in the detected diversity by removing errors at the expense of losing real diversity. Alternatively, researchers can opt for a low threshold, which will allow detection of taxa in low copy numbers, at the expense of retaining artefactual sequences. We agree with other authors that the first strategy is more appropriate for ecological

inference (Flynn et al., 2015; Kunin, Engelbrektson, Ochman, & Hugenholtz, 2010). However, removing sequences with low copy numbers is a major decision, indeed our analyses depicted it as the parameter that introduces the largest variation in the final results (Figure 3). Most authors opt to remove singletons (Burgar et al., 2014), although higher thresholds are becoming increasingly common (Leray & Knowlton, 2017). Our results show that removing singletons alone reduces artefactual sequences, but that more measures are needed to completely overcome the problem. Furthermore, our results demonstrate that the appropriate copy number threshold will depend on other factors such as sequencing depth (Figure 5c), thus relative thresholds (e.g. 0.01% of total reads) can be more appropriate than absolute copy number thresholds (Elbrecht & Leese, 2017).

4.6 | Should I remove chimeric sequences?

If chimeric sequences formed during PCR amplification are not detected and removed, they might erroneously inflate the richness and diversity measurements of the results (Bjørnsgaard Aas et al., 2016). In our study, chimera removal reduced the detected diversity (Figure 3), although the effect was small compared to the rest of the analysed parameters. Most sequences identified as chimeric had low copy numbers (Figure S5), thus when a high copy number threshold was used, the effect of removing chimeras was considerably reduced (Figure 4f). Identifying and removing chimeras is thus a recommendable step, although it is unlikely to change the results considerably if other measures to counteract the effect of PCR and sequencing errors are taken.

4.7 | Which OTU clustering threshold should I use?

Most metabarcoding studies use a 97% similarity threshold for OTU clustering, which is indeed the default value of most clustering algorithms (e.g. qiime *pic_otus*, OBITools *sumaclust*). The rationale behind a 97% threshold is that, on average, the minimum dissimilarity between species for most markers targeted by metabarcoding primers (e.g. COI, 16S) is around 3% (Hebert, Ratnasingham, & deWaard, 2003). Mean interspecific dissimilarity does, however, vary across taxa and target sequences (Figure 6). Consequently, so does the most appropriate threshold for defining OTUs. We show that the results of a metabarcoding study can vary considerably with OTU clustering thresholds (Figure 4g). Hence, we encourage researchers to select OTU clustering thresholds by empirical means and adjusted to the marker and target taxa, in order to obtain diversity measurements and taxonomic identifications better fitted to the actual sample.

4.8 | What is the most appropriate taxonomy assignment strategy?

Many approaches can be used to assign taxonomy to OTUs detected in a metabarcoding study. Some authors argue that for COI primers, identifications below 98% identity might be error prone (Clare, Barber,

Sweeney, Hebert, & Fenton, 2011), thus taxonomic assignments should be limited to matches above that threshold. However, such fixed approaches, might result in a low taxonomy assignment success (Figure 2a), more so if the target taxa are not well characterized in reference databases. While species-level identification is essential in some studies (e.g. Comtet et al., 2015), in most cases, incorporating higher taxonomic assignments in addition to species-level identifications allows increasing the ecological inference of the study. In those cases, multi-level taxonomic assignments that assign higher taxonomic levels to lower identities might be useful. Multiple algorithms and approaches exist to perform such taxonomic assignments (e.g. MEGAN, SAP), but our analysis with reference sequences from public databases showed that multi-level taxonomic assignments need to be performed carefully. We observed that the pairwise identities between arthropod families are highly variable, and values above 95% were detected between multiple families (Figure 6c). In contrast, all pairwise identities between insect orders were below 93% identity (Figure 6d). This means that whereas order-level assignments can be made confidently for BLAST identities above 93%, family level assignments are error prone. Because this cutoff can differ between taxonomic groups, but also targeted genomic markers, we believe researchers should analyse the genomic variability between target taxa and markers to choose the identity thresholds that allow confident taxonomic identifications based on empiric data.

5 | CONCLUSIONS

We demonstrate that the diversity detected in metabarcoding studies can drastically change according to the laboratory set-up and the different parameters and thresholds employed during the bioinformatic workflow. While it is likely that none of the approaches employed perfectly reflects reality, it is clear that certain choices critically increase the reliability of the results. Thus, we encourage researchers to acknowledge the benefits as well as potential biases and limitations of the different set-ups when designing a metabarcoding study, and adjust the interpretation of the data to the level of uncertainty. The differences we observed between markers and primer sets also highlight the importance of adjusting parameters to the marker region and taxonomic range based on empirical data rather than relying on general rules of thumb or standard settings of available software. Finally, if metabarcoding is to reliably answer biological questions, we believe that continuous technical refinement and illumination of strengths, weaknesses, and biases are needed in order to ensure unbiased, standardized and optimal detection of true diversity.

AUTHORS' CONTRIBUTIONS

A.A., O.A., M.T.P.G. and K.B. designed the study. A.A. and O.A. participated in the fieldwork and performed the laboratory work. A.A. analysed the data. A.A., O.A. and K.B. wrote the manuscript. All authors contributed to the final version of the manuscript.

DATA ACCESSIBILITY

Raw data are available on Dryad Digital Repository https://doi.org/10.5061/dryad.g14k2 (Alberdi, Aizpurua, Gilbert, & Bohmann, 2017).

REFERENCES

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2017). Data from: Scrutinizing key steps for reliable metabarcoding of environmental samples. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.g14k2

Bjørnsgaard Aas, A., Davey, M. L., & Kauserud, H. (2016). ITS all right mama: Investigating the formation of chimeric sequences in the ITS2 region by DNA metabarcoding analyses of fungal mock communities of different complexities. *Molecular Ecology Resources*, *17*, 730–741. https://doi.org/10.1111/1755-0998.12622

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 & Evolution*, *29*, 358–367.

Burgar, J. M., Murray, D. C., Craig, M. D., Haile, J., Houston, J., Stokes, V., & Bunce, M. (2014). Who's for dinner? High-throughput sequencing reveals bat dietary differentiation in a biodiversity hotspot where prey taxonomy is largely undescribed. *Molecular Ecology*, *23*, 3605–3617.

Chariton, A. A., Court, L. N., Hartley, D. M., Colloff, M. J., & Hardy, C. M. (2010). Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. *Frontiers in Ecology and the Environment*, *8*, 233–238.

Clare, E. L., Barber, B. R., Sweeney, B. W., Hebert, P. D. N., & Fenton, M. B. (2011). Eating local: Influences of habitat on the diet of little brown bats (*Myotis lucifugus*). *Molecular Ecology*, *20*, 1772–1780.

Clare, E. L., Chain, F. J. J., Littlefair, J. E., & Cristescu, M. E. (2016). The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data. *Genome*, *59*, 981–990.

Clare, E. L., Symondson, W. O. C., Broders, H., Fabianek, F., Fraser, E. E., MacKenzie, A., … Reimer, J. P. (2014). The diet of *Myotis lucifugus* across Canada: Assessing foraging quality and diet variability. *Molecular Ecology*, *23*, 3618–3632.

Clarke, J. C., Beard, J. M., Swadling, K. M., & Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution*, *7*, 873–883.

Clarke, L. J., Soubrier, J., Weyrich, L. S., & Cooper, A. (2014). Environmental metabarcodes for insects: In silico PCR reveals potential for taxonomic bias. *Molecular Ecology Resources*, *14*, 1160–1170.

Comtet, T., Sandionigi, A., Viard, F., & Casiraghi, M. (2015). DNA (meta)barcoding of biological invasions: A powerful tool to elucidate invasion processes and help managing aliens. *Biological Invasions*, *17*, 905–922.

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, 306–323.

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, 20140562.

Elbrecht, V., & Leese, F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Sciences*, 5, 11.

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

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., ... Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Molecular Ecology*, 21, 1821–1833.

Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., ... Lodge, D. M. (2016). Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources*, 16, 29–41.

Flynn, J. M., Brown, E. A., Chain, F. J. J., MacIsaac, H. J., & Cristescu, M. E. (2015). Toward accurate molecular identification of species in complex environmental samples: Testing the performance of sequence filtering and clustering methods. *Ecology and Evolution*, 5, 2252–2266.

Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, 13, 851–861.

Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P.-J., Griggo, C., Gielly, L., ... Taberlet, P. (2014). Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nature Communications*, 5, 3211.

Hebert, P. D. N., Ratnasingham, S., & deWaard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B*, 270, S96–S99.

Hope, P. R., Bohmann, K., Gilbert, M. T. P., Zepeda-Mendoza, M. L., Razgour, O., & Jones, G. (2014). Second generation sequencing and morphological faecal analysis reveal unexpected foraging behaviour by *Myotis nattereri* (Chiroptera, Vespertilionidae) in winter. *Frontiers in Zoology*, 11, 39.

Judo, M. S., Wedel, A. B., & Wilson, C. (1998). Stimulation and suppression of PCR-mediated recombination. *Nucleic Acids Research*, 26, 1819–1825.

Kulczynski, S. (1927). Die Pflanzenassociationen der Pienenen. *Bulletin International de L'Academie Polonaise des Sciences et des letters, classe des sciences mathematiques et naturelles, Serie B, Supplement II*, 2, 57–203.

Kunin, V., Engelbrektson, A., Ochman, H., & Hugenholtz, P. (2010). Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*, 12, 118–123.

Leray, M., & Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 2076–2081.

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

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., ... Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, 10, 34.

Munch, K., Boomsma, W., Huelsenbeck, J. P., Willerslev, E., & Nielsen, R. (2008). Statistical assignment of DNA sequences using Bayesian phylogenetics. *Systematic Biology*, 57, 750–757.

Murray, D. C., Coghlan, M. L., & Bunce, M. (2015). From benchtop to desktop: Important considerations when designing amplicon sequencing workflows. *PLoS ONE*, 10, e0124671.

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, 819–830.

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, 1931–1950.

Pornon, A., Escaravage, N., Burrus, M., Holota, H., Khimoun, A., Mariette, J., ... Andalo, C. (2016). Using metabarcoding to reveal and quantify plant-pollinator interactions. *Scientific Reports*, 6, 27282.

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System(http://www.barcodinglife.org). *Molecular Ecology Notes*, 7, 355–364.

Robasky, K., Lewis, N. E., & Church, G. M. (2014). The role of replicates for error mitigation in next-generation sequencing. *Nature Reviews. Genetics*, 15, 56–62.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13, 2498–2504.

Smith, D. P., & Peay, K. G. (2014). Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PLoS ONE*, 9, e90234.

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.

Zeale, M., Butlin, R. K., & Barker, G. (2011). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology*, 11, 236–244.

Zepeda-Mendoza, M. L., Bohmann, K., Carmona Baez, A., & Gilbert, M. T. P. (2016). DAMe: A toolkit for the initial processing of datasets with PCR replicates of double-tagged amplicons for DNA metabarcoding analyses. *BMC Research Notes*, 9, 255.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.