Revisiting the effect of PCR replication and sequencing depth on biodiversity metrics in environmental DNA metabarcoding

Sabrina Shirazi[1], Rachel Meyer[1], and Beth Shapiro[1]

[1]University of California Santa Cruz

May 30, 2021

Abstract

Environmental DNA (eDNA) metabarcoding is a common tool for measuring and cataloguing biodiversity, yet standard methodological approaches to generate metabarcoding data sets have yet to emerge, in part due to challenges understanding the biological and technical biases that affect eDNA profiles. Here, we explore how two experimental choices – depth of sequencing of PCR amplicon libraries and the number of PCR replicates – influence estimates of α and β diversity. We extracted DNA from six soil samples from three ecologically distinct locations, performed 24 PCR replicates from each using two common metabarcodes, and sequenced each to an average depth of 83,898 reads. We found PCR replicates are consistent in composition and relative abundance of abundant taxa, allowing differentiation of samples and sites. However, rare taxa were unique to one or a few replicates, suggesting that even large numbers of experimental replicates may be insufficient to catalogue biodiversity fully. We recommend that to differentiate sites, separately sequencing only a minimum of two PCR replicates to a depth that allows 1,000 reads identified to taxa, is sufficient to differentiate sites. We also conclude that metabarcoding is impractical for exhaustive taxonomic inventory and, because rare taxa are not amplified consistently, taxonomic tallies that rely on consensus among replicates artificially lower richness estimates. These findings provide new considerations for eDNA experimental design and data interpretation.

1- Introduction

Environmental DNA (eDNA) refers to DNA shed by organisms into their environments (Taberlet et al., 2018), such as soil and other sediments (e.g.- Deveautour et al., 2018; Ficetola et al., 2018), water (e.g.- DiBattista et al., 2017; Bista et al., 2017), and air (Kraaijeveld et al., 2015). Although shotgun sequencing of sedimentary DNA is becoming increasingly common (e.g. Pederson et al., 2016, Graham et al., 2016), most eDNA studies to date have targeted either single taxa with species-specific PCR ("barcoding"; e.g.- Baker et al., 2018; Franklin et al., 2019) or phylogenetically diverse taxa using "universal" PCR primers that bind to conserved regions flanking barcode loci ("metabarcoding"; e.g.- Taberlet et al., 2012; Valentini et al., 2016). Metabarcoding has been used, for example, to test hypotheses about biotic and abiotic drivers of changes in community composition (Erlandson et al., 2018; Deveautour et al., 2018; Yan et al., 2018, Giguet-Covex et al., 2014, Epp et al., 2010), including over tens of thousands of years (Willerslev et al., 2003, 2014; Parducci et al., 2012), and to characterize extant eukaryotic diversity (Lallias et al., 2015; Leray and Knowlton, 2016). Because eDNA can be collected non-invasively, eDNA is also a promising tool for studying rare, cryptic, or endangered species (Franklin et al., 2019; Laramie et al. 2015; Schnell et al., 2012), and for tracking early advances of invasive species (Klymus et al., 2017; Xia et al., 2018; Sjogren et al., 2017).

The broad application of eDNA approaches across environmental science, in combination with recent advances in DNA processing and sequencing technologies, has precipitated rapid growth of eDNA as a research field (Pederson et al. 2015; Garlapati et al., 2019). Indeed, several interdisciplinary initiatives now use eDNA to characterize biodiversity across spatial and temporal scales, such as CALeDNA (Meyer et al., 2019), DNAquanet (https://dnaqua.net/), and Maine-eDNA (https://umaine.edu/edna/). These initiatives

1

have shown repeatedly that DNA detection in eDNA samples varies in both specificity and replicability (Garlapati et al., 2019; Cristescu and Hebert, 2018) and the causes of this variation have yet to be fully characterized.

Technical biases can be introduced into eDNA experiments during field sampling, laboratory processing, and bioinformatics (Pederson et al., 2015). These biases have the potential to influence resulting biodiversity profiles, and are difficult to decouple from both true differences among organisms in DNA deposition rates and taphonomic processes that drive variation in long-term DNA survival (Taberlet et al., 2018). DNA isolation protocols, for example, are known to influence which organisms are detected in a sample. Deiner et al. (2018) reported a three-fold change in observed α diversity depending on filter material, pore size, and chemical extraction, including the absence under some extraction conditions of some taxa known to be present. Choice of polymerase for PCR can also influence biodiversity estimates. Nichols et al (2017) showed that the composition and relative abundance of taxa detected in the PCR amplicon pool changed during PCR amplification as some polymerases biased the amplicon pools toward sequences with a particular GC content. Differences in amplicon length, templates secondary structures, and base mismatches at the PCR primer binding site also affect binding and copying efficiency during PCR (Fonesca et al., 2012; Elbrecht & Leese, 2015; Krehenwinkel et al., 2017) also affect binding and copying efficiency during PCR, skewing post-PCR taxon composition and relative abundance estimates (Pawluczyk et al., 2015). As the number of eDNA data sets grows, and along with that the possibility for comparative analysis across data sets, the need to understand and mitigate these many potential biases grows (Braukmann et al., 2019; Ruppert et al., 2019).

As one approach to mitigating these potential biases, many eDNA studies include one or more controls as part of their experimental design. These controls are intended to quantify some component of potential technical biases prior to proceeding with biodiversity analyses. For example, samples taken in the field as replicates can be used to estimate and account for spatial variation in DNA deposition and survival (Andersen et al., 2012; Ficetola et al., 2015). Incorporating negative controls (experimental replicates that include no sample) during both DNA extraction and PCR can track potential laboratory-introduced contaminants or other errors. Incorporation positive controls comprising mixtures of organismal tissue or extracted DNA can confirm that protocols are working as expected and quantify bias (Port et al., 2016; Olds et al., 2016). After data are generated, bioinformatic approaches can also be implemented to detect and mitigate the influence of experimental biases. For example, site occupancy modeling incorporates variation among PCR replicates to overcome imperfect detection (Ficetola et al., 2015), and species richness curves can be estimated to assess whether sufficient replicates have been performed to detect all or most of the taxa preset in a sample (Ficetola et al., 2015; Lundberg et al., 2013; Beentjes et al., 2019).

Variation among PCR replicates is expected in eDNA metabarcoding (Beentjes et al 2019; Ficetola et al. 2008). While much of this variation results from random sampling of low abundance taxa (Leray & Knowlton 2017), variation can also arise due to errors including contamination by exogenous DNA and the accumulation of replication and sequencing errors. To capture this variation, eDNA experimental designs often include one or more replicate PCRs for each DNA extract. Replicate PCRs reduce the effect of stochastic amplification and make it possible to detect potential outlier PCRs (Robasky et al 2014; Leray & Knowlton, 2017), which we define as a PCR amplicon pool that contains a different richness or composition of taxa than other replicates. Most eDNA studies perform from one (e.g. Deveautour et al., 2018; Erlandson et al., 2018) to three PCR replicates (e.g.Yamamoto et al., 2017; Beentjes et al., 2019; Browne et al., 2020), although more PCR replicates are sometimes performed when working with ancient samples (e.g. 8 replicates: Ficetola et al, 2018, Clarke et al., 2019; 15 replicates: Stahlschmidt et al., 2019). However, while PCR replication is common in eDNA research, there is no consensus as to how this mechanism should be applied as a control. Most studies generate replicate PCRs and then pool them prior to sequencing, rather than sequencing them separately (e.g. Lanzen et al., 2017; Smith and Peay 2014; Lin et al., *submitted*), in part because it is cost prohibitive (Fonseca 2018). Some studies that sequence PCR replicates separately discard taxa found in only a single replicate as an approach to ruling out contaminants (De Barba et al., 2014; Giguet-Covex et al., 2014; Hope et al., 2014), or require a taxon or haplotype to be present in all replicates to be included in

downstream analysis (Taberlet et al., 2018; Tsuji et al., 2019). While variation due to errors can artificially inflate biodiversity estimates (Zepeda-Mendoza 2016), discarding taxa found in few replicates may also remove biodiversity that is genuinely present in a sample (Leray & Knowlton, 2017).

Another experimental variable to consider when performing PCR replication is read sampling depth, or how deeply each PCR amplicon pool is sequenced. In published eDNA studies, read sampling depth per PCR replicate tends to be between 1,000 reads (Krehenwinkel et al., 2017) and 25,000 reads (e.g. Lanzen et al 2017, Schnell et al., 2018; Stat et al., 2017), with some exceptions (e.g. >100,000 reads per replicate. Leempoel et al., 2020). Previous studies investigating the role of PCR replication and read sampling depth in eDNA have had conflicting results. In exploring metabarcoding data generated using the Ion Torrent sequencing platform, Murray, Coghlan, and Bunce (2015) found that increasing read sampling depth did not necessarily increase the likelihood of detection of low abundance taxa. Smith and Peay (2014), alternatively, found that increasing read sampling depth decreased dissimilarity between PCR replicates PCR replication but that this had little effect on estimates of either α or β diversity.

Several previous studies have explored the influence of PCR replicates on metabarcoding-based biodiversity estimates, also with conflicting results. Smith and Peay (2014), for example, compared diversity estimates generated from PCR amplicons pools comprising 1, 2, 4, 8 or 16 PCR replicates and found that sequencing platform and sequencing depth both affected recovered taxonomic profiles, but that the number of PCR replicates did not necessarily increase the number of observed taxa. Ficetola et al. (2015), using both simulated and empirical data, found that the number of PCR replicates necessary to observe all taxa in an extract depends on the site, and that the probability of detection of a given taxon increases with that taxon's relative abundance. More recently, Alberdi et al. (2017) compared biodiversity estimates from metabarcoding data amplified in triplicate from 54 bat fecal samples. They observed different taxonomic profiles depending on which subsets of the PCR replicates per sample they analyzed and that increasing read sampling depth increased dissimilarity between PCR replicates. While these results hint that more PCR replicates is better than fewer PCR replicates in assessing total diversity, how many PCR replicates should be performed remains an open question, as does the interaction and potential trade-off between PCR replication and increasing read sampling depth.

Here, we perform 24 replicate PCRs for two commonly used metabarcodes, the *Internal Transcribed Spacer* (*ITS*) for fungi (*ITS1*) and for plants and algae (*ITS2*), to explore how PCR replication and read sampling depth influence metabarcoding-based biodiversity estimates at six ecologically distinct sites. We address explicitly the detection of rare taxa, inference of community composition, site differentiation based on taxon composition, and the detection and prevalence of PCR outliers. Our data provide two key insights for eDNA metabarcoding experimental design. First, we find that abundant taxa are common among PCR replicates and that these taxa, and therefore few PCR replicates, are sufficient to define site uniqueness. Second, we observe that rare taxa most often appear in only one or a few replicates, and alter significantly richness estimates among replicates. These results suggest that metabarcoding may be insufficient to characterize fully the alpha biodiversity at any site, even with large numbers of replicates, but can be sufficient – even with low read sampling depth and few replicates – to characterize beta diversity.

2- Materials and Methods

2.1- Soil Collection

We collected a total of six soil samples from three environmentally distinct locations: two samples from St Paul Island, Alaska, USA (StP.1: 57.136074, -170.82537; StP.2: 57.10577, -170.10563), and four samples from two sites in California, USA. The California samples include two from Fort Ord Natural Reserve in Marina (FO.1, an open sand dune: 36.68448, -121.77731; FO.2, a chaparral ecosystem: 36.68301, -121.78071), and two from and two from Younger Lagoon in Santa Cruz (YL.1, the basin of a coastal lagoon: 36.950081, -122.066756; YL.2, a grassland coastal terrace: 36.949314, -122.063575).

We designed and implemented field sampling protocols to minimize the risk of cross-contamination between soil samples. At each site, we wore clean gloves and used a clean trowel to collect soil from 2-6" below the

surface, which we deposited into sterile 50mL falcon tubes. We sterilized trowels between sites serially with bleach and rinsed with ultraPure water and ethanol. Alaskan samples were kept cool during transfer to California and then frozen. Californian samples were frozen at -20@C immediately after collection.

2.2- DNA Extraction, Amplification, and Sequencing

We processed each soil sample in the eDNA processing center of the UCSC Paleogenomics Lab following strict protocols to avoid contamination. Researchers working in the eDNA lab, which is isolated from PCR amplification products, wear sterile suits, face masks, hair nets, and gloves so as to minimize the possibility of contamination. Prior to DNA extraction, we homogenized the soil samples in the 50mL falcon tube in which they were collected so as to break up any larger pieces of sediment. Next, we removed any identifiable plant matter (leaves and roots) with sterile forceps. We then subsampled 0.25g aliquots of sediment for DNA extraction. We extracted DNA in duplicate for each of the six soil samples using the Qiagen PowerSoil kit and protocol (Qiagen, Germantown MD, USA), following the manufacturer's protocol except for the final elution step, where we substituted the C6 elution buffer with Tris-EDTA-Tween buffer. We included one negative extraction control without soil.

We selected the *ITS* gene in plants (PITS) and fungi (FITS) as universal barcodes for metabarcoding. For PITS, we used primers described by Yao et al. 2010 (ITS-S2F - ATGCGATACTTGGTGTGAAT and ITS-S3R - GACGCTTCTCCAGACTACAAT) and for FITS, we used primers from White et al. 1990 (ITS5-forward - GGAAGTAAAAGTCGTAACAAGG) and Epp et al. 2012 (5.8S_fungi - reverse - CAAGAGATC-CGTTGTTGAAAGTT). The expected amplicon length was 450 base pairs (bp) for PITS and 300 bp for FITS.

We used quantitative PCR (qPCR) to assess PCR inhibition in each extract and determine the appropriate number of PCR cycles for metabarcoding (as recommended in Murray, Coghlan, & Bunce, 2015). We performed qPCR with the Qiagen Multiplex PCR Master Mix and spiked SYBR Green 1 Dye (12.5uL Qiagen MM, 2uL of each 2uM primer, 0.6uL 1:2000 dilution SYBR Green 1 Dye, 5.9uL water, and 2uL extracted DNA). The Qiagen Multiplex PCR Master Mix introduces the least GC amplification bias of tested polymerases (Nichols et al. 2017). We amplified each extract in triplicate with PITS and FITS primers. For each replicate, we set up a serial dilution of 1:0, 1:1, and 1:3 extract to water proportions of the 2uL DNA extract input, and compared the qPCR Ct values across the dilution series. We observed no inhibition in any of the eDNA extracts, and decided to proceed with undiluted extracts. To avoid overamplification, we determined the optimal number of PCR cycles for each extract and PCR primer as the cycle after which the exponential amplification phase ended. We amplified extracts with PITS for 26-31 cycles and with FITS for 18-21 cycles.

We followed a '2-step' protocol to build amplicon sequencing libraries (Nichols et al. 2017). The same reagent set up was used for metabarcoding PCR as for qPCR, but with the appropriate number of cycles and without SYBR Green 1 dye. The amplification primers included 5' overhang Illumina (Illumina, San Diego, CA, USA) TruSeq adapter sequences, which allowed us to perform the indexing PCR directly after metabarcoding. For each of the six extracts we performed 24 replicate PCRs with PITS and 24 PCR replicates with FITS. We amplified four PITS and four FITS PCR replicates from the extraction negative control (no sediment) and added two additional PCR negative controls (no extract) for each primer.

Following metabarcoding PCR, we purified amplicon pools with SPRI beads (Beckman, Indianapolis, IN, USA). Next, we indexed all PCR products individually using Kapa Hifi (Roche, Pleasanton, CA, USA) to add eight base-pair dual indices and Illumina sequencing adapters to our amplicon pools (12.5uL Kapa Hifi, 5.5uL water, 1uL of each 10uM forward and reverse index, and 5uL purified PCR product), followed by a second SPRI bead clean. We used unique combinations of dual indices for each PCR replicate, though each individual index was used multiple times across different amplicon pools. We then quantified the concentration of DNA in the purified amplicon pools with a Nanodrop (Thermo Scientific, Waltham, MA, USA). We used these estimates of DNA concentration to pool the PCR amplification products into equimolar ratios in two pools, one for PITS and one for FITS. We then quantified the two pools with a Qubit fluorometer

(Thermo Fisher, Waltham, MA, USA) and estimated average fragment sizes with a fragment analyzer.

To detect potential index swapping (the incorrect assignment of sequences to an index as a result of blurring of index signals between adjacent clusters, van der Valk et al., 2019) during sequencing, we amplified the PITS metabarcode from an extract of spiral ginger(*Costus pulverulentus*), provided to us by Kathleen Kay's lab at UCSC. The Kay lab collected samples from La Selva, Costa Rica under permit R-056-2019-OT-CONAGEBIO and extracted floral *C. pulverulentus* tissue with the Qiagen plant mini kit. Spiral ginger is native to the neotropics and not found in either California or Alaska. We generated three replicate PCR amplicon libraries from the spiral ginger extract following the 2-step protocol described above. Following sequencing and demultiplexing, we estimated the rate of index hopping based on the number of non-ginger sequences assigned to the ginger dual indices.

We pooled and sequenced 308 sediment and three spiral ginger libraries (a total of 311 libraries) on an Illumina Miseq using v3 chemistry and a 2x300 approach. We targeted 100,000 reads per FITS library and 50,000 reads per PITS library, based on the anticipated higher taxonomic diversity amplified by FITS and the anticipated higher discard rate of FITS-amplified sequences due to the incompleteness of the fungal taxonomy databases.

2.3- Data Processing

We used the *Anacapa Toolkit* (Curd et al, 2019) to process the resulting data. *Anacapa* is a metabarcoding data processing pipeline that enables simultaneous data processing of multiplexed barcodes and libraries. Within *Anacapa* , we trimmed the TruSeq adapters using *cutadapt* (Martin, 2011), removed bases with Q-scores below 35 with the *FastX-Toolkit* (Gordon and Hannon, 2010), and then trimmed, again using cutadapt. We trimmed the first 5 bases on the 5' end of the forward read and the first 10 bases of the 5' end of the reverse read for the PITS data set. We trimmed 40 bases off the 5' end of the forward read and the first 50 bases of the 5' end of the reverse read for the FITS data set. We used *dada2* (Callahan et al. 2016) to merge the forward and reverse reads, remove chimeric sequences, and identify amplicon sequence variants (ASVs). We performed ASV assignment to taxa via global and local alignment using*Bowtie2* (Langmead and Salzberg, 2012) to PITS and FITS*CRUX* reference databases released with Curd et al. 2019. The top 100 hits of the *Bowtie2* alignment to reference were bootstrapped with *BLCA* (Gao et al, 2017) to assign each ASV to a taxon and provide uncertainty estimates. We used a 60% bootstrap confidence threshold of taxonomic assignment, as suggested in the *Anacapa*documentation (Curd et al, 2019). *Anacapa* output two taxonomy tables (one per amplicon) formatted as matrices of the number of reads from each PCR replicate assigned to a given taxa.

We used the negative PCR and DNA extraction controls to detect bioinformatically and remove contaminants and the positive ginger control to infer the rate of barcode swapping. We converted taxonomy tables and the PCR replicate-associated metadata to *phyloseq* (v. 1.22.3; McMurdie and Holmes 2013) objects using *Ranacapa*(Kandlikar et al., 2018). We then used the R package *decontam*(v1.1.0; Davis et al., 2018) with default settings to compare extraction and PCR negative controls to PCR replicates amplified from sediment, and looked bioinformatically for any index hopping between spiral ginger and soil-derived libraries.

To generate data sets for hypothesis testing, we next rarefed our two decontaminated taxonomy tables to generate PCR replicate sets of equal read sampling depth, and then applied a minimum read cutoff to account for low level contamination, PCR error, sequencing error, and undetected index hopping. We used the rarefy_even_depth() function of*phyloseq* to rarefy our data at depths of every thousand between 1,000 to 20,000 (ex. 1k, 2k, 3k. . .). As we increased rarefaction depth, some libraries that were sequenced less deeply dropped out of the analysis. Following rarefaction, we generated three data sets for each rarified library in which we applied minimum read cutoffs of 2, 5, and 10 reads using a custom R script that turns any value below these values to 0. Increasing the minimum read cutoff is more conservative, though may remove the ability to track low abundance taxa. Rarefaction results at different read depths were plotted as taxon accumulation curves at minimum read cutoff of 5 reads, and to account for stochasticity in rarefaction, the average of 25 rarefactions per PCR replicate were plotted. All raw and rarefied taxonomy

tables (at rarefaction depths 1,000 to 20,000, with minimum read cutoffs of 2, 5, and 10) are available at *https://github.com/sashirazi/eDNA-PCR-Project/*.

We rarefied datasets prior to estimating various Alpha diversity metrics. We used the R package *iNEXT* for richness extrapolation. We implemented iNEXT with q=0, datatype="abundance", knots=40, se=TRUE, conf=0.95, nboot=50. With iNext, we extrapolated replicates to twice their true read sampling depth and recorded the extrapolated diversity statistic. We performed outlier tests on extrapolated observed richness by identifying points that fall outside values of 1.5 times the interquantile range. We calculated observed richness, the Shannon diversity index (Shannon 1948), and Simpson index (Simpson 1949) with the *vegan* package in R. While observed α diversity considers only taxon presence, the Shannon and Simpson's estimators consider both the relative abundance (RA) of taxa within a sample in addition to taxon presence. We then performed two-sided t-tests and chi-square tests in R*stat* .

We performed statistical tests for beta diversity using*MicrobiomeSeq* in R that draws on the *vegan* and*phyloseq* packages. We plotted taxon relative abundance barplots and calculated local contribution to beta diversity (LCBD) using the Canberra β diversity measure (a measure within the Bray-Curtis β diversity family of estimators). The LCBD test calculates the contribution to total β diversity estimated for a particular site from each PCR replicate. We considered replicates as outliers when p<0.05.

3- Results

3.1- Data summary

We generated an average of 78,809 PITS (range: 9,352-282,57; Table S1) and 88,987 FITS (range: 15,409-382,888; Table S2) sequences for each of our 288 amplicon libraries (24 PCR replicates for each of six extracts, two primer sets). Following adapter removal and quality trimming, we retained an average of 37,640 PITS reads (range: 6,148- 166,279; Table S1) and 63,436 FITS reads (range: 12,360-323,310; Table S2) per PCR replicate. We were unable to assign to taxa an average of 2.3% of PITS and 76.1% of FITS quality trimmed reads. The large number of unassignable reads in the FITS data are probably the result of reference database incompleteness, which is a known problem for this taxonomic group and one that we attempted to mitigate by increasing the target depth of sequencing for our FITS libraries.

We found little evidence of contamination introduced during sample processing, and no evidence of index hopping between libraries during sequencing. Of the eight extraction negative control libraries and four PCR negative control libraries per primer that each had between 2,684 and 169,767 reads (Table S1 and S2), species were not shared between control and true samples. We found no evidence of contamination in our PITS data set. The FITS extraction negative control libraries contained a maximum of 11 reads that matched an "unidentified environmental" fungus. We removed all reads from the PCR amplicon libraries that were assigned to this "unidentified environmental" fungus. The PCR negative control libraries generated using both the PITS and FITS primers included only primer dimer chimeras. The three spiral ginger libraries used to track index hopping generated 52,675-299,400 sequences. All ginger sample sequences assigned to plant taxa aligned to *Costus pulverulentus* , and no sequences assigned to barcodes for libraries other than the ginger sample libraries were assigned to *Costus* , indicating there was no index hopping.

3.2- Τηε ινφλυενςε οφ ρεαδ σαμπλινγ δεπτη ον ταξον αςςυμυλατιον ανδ α διερσιτψ

Read sampling depth strongly influences the number of taxa counted in an individual PCR replicate (observed alpha diversity). Taxon accumulation curves (Figure 1) show that for all PCR replicates, observed α diversity increases with read sampling depth across the range of depths used. On average across PCR replicates and sites, a read depth of 1,000 included 51% (PITS) and 33% (FITS) of the observed α diversity compared to that detected with 20,000 reads. A read depth of 5,000 retained 77% (PITS) and 64% (FITS) of taxa detected at 20,000 reads, and a read depth of 10,000 retained 89% (PITS) and 82% (FITS) of taxa detected with 20,000 reads**.** In all sites, the PITS curves surpass the inflection point where slope begins to decrease (asymptote) at a sampling depth under 5000 reads. In FITS, the inflection point is less visually observable and more taxa continue to accumulate at high read depth (Figure 1). Nonetheless, we observed that in both

6

PITS and FITS datasets, the maximum observed richness values vary considerably across PCR replicates.

Because the maximum observed richness could be influenced by read sampling depth, we explored PCR replicate richness variation using asymptote values extrapolated to twice the original depth with *iNEXT* . We found Observed diversity estimates were rarely normally distributed and variance was high, with up to five replicates per group being outliers from the mean (Table S3). After outlier removal, PCR replicate richness at the extrapolated asymptote still exhibited multiple fold differences in PITS and standard deviations equivalent to up to 30% of the maximum richness of the group (Table 1). We found the highest fold differences in observed richness in the PITS data set from YL.1 (Figure 1e), a site situated within a marine lagoon at a location that is regularly inundated with both marine water and stream runoff. We expect these physical processes increase true taxon richness and possibly heterogeneity within the environmental samples. We observed fewer outliers in extrapolated richness for FITS, with only up to two outliers per group, but that variation was high, with standard deviations up to 21% of the maximum richness of the group. We observed the highest fold differences in observed richness in FITS at YL.1 and FO.2.

Increasing the read sampling depth from 1,000 to 10,000 reads resulted in an average 1.8-fold increase in observed α diversity for PITS and 2.4-fold increase for FITS (Table S4). This increase in observed α diversity with depth was significant for all six sites and both markers (Table S4). Shannon diversity did not significantly increase with read sampling depth for five of the six sites in the PITS data set, but did significantly increase with all FITS data set (Table S4). Simpson diversity did not significantly increase with read sampling depth for any PITS data set, but did for five of the six sites in the FITS data set (Table S4).

3.3 Within-site variation composition and relative abundance (RA)

We next explored similarity between PCR replicates by comparing the RA of taxa detected in them. Qualitatively, for both PITS and FITS, within-site PCR replicates appear similar to each other and each site is readily distinguishable from other sites (Figure 3). Only one replicate appeared to have a different community profile, and this was in PITS results from the YL.1 site (Figure 3). Analysis of 'local contribution to beta diversity' (LCBD) provides a quantitative examination of differences in composition and RA. At a read sampling depth of 5,000 and with a five read minimum cutoff, LCBD statistics identified several PCR outliers in the PITS data set (11 PCR replicates from the YL.1 site and 2 PCR replicates from the FO.2 site; Table S5) and one outlier in the FITS data set (1 outlier replicate from the FO.2 data set; Table S5).

To explore how read sampling depth and minimum read cutoff influence LCBD, we repeated these analyses at all three read sampling depths (1,000, 5,000, and 10,000 reads) with minimum read cutoffs of two, five, or ten reads (Table S5), and performed Chi-squared tests for significant differences among groups. The number of LCBD-based outliers increased significantly with higher read sampling depth in the FITS data set (p= 3.861E-15), but not in the PITS data set (p=0.25). We found no significant effect of minimum read cutoff for either the PITS (p=0.71) or FITS (p=0. 79) data set, suggesting that low abundance taxa, which are most likely to be impacted by changing the minimum read cutoff, are not causing outliers. For both the PITS and FITS data set, we found that site itself has a significant effect on the number of observed PCR outliers (both p <2.2e-16).

3.4 Beta diversity between sites and replicates

We evaluated inter and intra site-based estimates of β diversity using the Jaccard metric, which weighs all observed taxa equally. At both 1,000 and 10,000 read sampling depths, we found that PCR replicates in both PITS and FITS results cluster by site and that sites from the same geographic area cluster near each other (Figure 4). None of the identified PCR replicate outliers reduced the ability to differentiate sites based on the PCoA. While read sampling depth did not affect dispersion in the PCoA between PCR replicates for either PITS or FITS, increasing read sampling depth changed the position of some sites relative to each other in the ordination (Figure 4). For example, the two Fort Ord sites all had similar PITS taxon composition and could not be distinguished from each other at either read sampling depth. However, the two St.P sites could be differentiated at both read sampling depths for PITS but only at the lower sequencing read depth in FITS. We next calculated β diversity using the Bray-Curtis estimator (Bray and Curtis, 1957), which considers RA,

and compared the resulting PCoA plots to those generated using the Jaccard estimator (Figure S1; Figure 4). The plots were similar overall, with slightly closer clustering among replicates at the two St.P sites in the PITS data set. The similarity between results using Jaccard and Bray-Curtis estimates suggests that low abundance taxa may not influence strongly estimates of β diversity.

3.5- PCR replication and taxon accumulation.

To assess the extent of low abundance and possibly unique taxa in PCR replicates, we calculated increases in α diversity as PCR replicates are added to a combined data set. Although the order in which PCR replicates are added will not influence cumulative α diversity, the trend to this endpoint will vary. We therefore bootstrapped the analysis 100 times and plotted the mean. Notably, after the addition of all 24 PCR replicates, we did not observe a plateau in species richness, indicating that even this large number of PCR replicates was insufficient to fully sample the diversity of taxa within the DNA extract (Figure 5).

We next calculated the number of PCR replicates per sample needed to reach the point at which the taxon accumulation curve is saturated, or at which it increases by fewer than one taxon on average (based on our bootstrapped analysis) when another PCR replicate is added (Table 2). We performed this analysis at different sequencing read depths and read cutoffs. The number of added PCR replicates necessary to achieve saturation of the taxon accumulation curve varied between sites, sampling read depth, and minimum read cutoff, although fewer replicates were necessary to reach saturation at higher read cutoff (Table 2). Increasing the rarefaction read depth surprisingly increases the number of replicates required (Table 2).

We then plotted histograms of the frequency of taxa detected across PCR replicates (Figure 6). Most taxa are either singletons (present in only one PCR replicate) or occur in all PCR replicates. Based on PITS data, singletons did not appear to be sequencing artefacts because out of the 70 singleton species found, only eleven occurred within the same genus as another species found at high frequency (found in at least 20 replicates) (see *Chlamydomonas* ; Table S1). To evaluate if singleton taxa were also low relative abundance taxa, we plotted the relationship between a taxon's within-replicate sequence abundance and its frequency across replicates (Figure 7). For all read depths and minimum read cutoffs, we find a significant positive correlation (Figure 7), as indicated with a fitted linear model (PITS: p<2e-16, T=24.73, adjusted r^2= 0.7324; FITS: p<2e-16, T=39.91, adjusted r^2=0.8219). This indicates that taxa that occur at low sequence abundance within PCR replicates also occur less frequently across replicates, and that taxa that are abundant within PCR replicates are more likely to occur in all PCR replicates. In PITS results, only when a taxon's relative abundance is over roughly 10% does it occur in most replicates (Figure 7). In FITS results, only when a taxon is over 1% does it occur in most replicates. Most taxa in soil and sediments were at below 1% relative abundance (Figure 7).

4. Discussion

We explored the impact of the number of PCR replicates and read sampling depth, two common parameters in eDNA experimental design, on estimates of taxonomic diversity. Using eDNA extracts from six sites at three ecologically and geographically distinct locations, we performed 24 PCR replicates for each of metabarcodes: plant ITS (PITS) and fungal ITS (FITS). We then analyzed these replicates by compiling data sets that included different read sampling depths and minimum read cutoffs. We find that PCR replicates are consistent in the composition (Figure 7) and relative abundance (RA) (Figures 3 and S1) of high abundance taxa, but inconsistent in recovery of low abundance taxa, and that even large numbers of PCR replicates are insufficient to fully characterize diversity at any site.

When considering only high abundance taxa, our PCR replicates produced community profiles that distinguished sites from each other, even sites that are geographically proximate and presumably similar in community composition (Figure 4). The majority of high abundance taxa were detected in all 24 PCR replicates, excluding outlier PCRs. This result provides empirical support for the modeling-based prediction by Ficetola et al (2015) that PCR replicates will consistently detect taxa that have high "detection probability", which they define as taxa present in high abundance relative to other taxa at a site. We also observed that the number of reads assigned to a taxon within a PCR was positively correlated to the frequency with

which that taxon was observed across PCR replicates (Figures 3, 7), which was also reported by Smith and Peay (2014). Together, these results confirm that community profiles based on high abundance taxa are replicable among PCRs and capable of distinguishing sites. Minimal PCR replication is therefore necessary to characterize sites using β diversity statistics that derive from high abundance taxa.

While high abundance taxa were recovered consistently among our PCR replicates, low abundance taxa were not (Figures 1, 2, 6). Low abundance taxa rarely occurred in PCR replicates; we observed some low abundance taxa in several PCR replicates but most in only a single replicate. Unsurprisingly, this stochasticity in recovery of low abundance affected biodiversity statistics that rely on raw taxon counts, such as α diversity. While this observation has been reported previously (e.g. Beentjes et al., 2019; Ficetola et al., 2008) our results highlight how the problem can be exacerbated by shallow sequencing read depths and low minimum read cutoffs. Specifically, we find several fold differences in maximum richness and rarefied richness among replicates depending on what values we selected for read depth and minimum read cutoff. While the stochasticity in recovery of low abundance taxa poses challenges in interpretation of some biodiversity statistics, it tends not to influence β diversity between sites measured as either presence-absence (Figure 4) or RA (Figure S1), or on position with a PCoA.

At our six sites, 24 PCR replicates were not sufficient to detect all rare taxa and therefore stabilize the species accumulation curves (Figure 5). This result supports previous observations that using different numbers of PCR replicates will alter taxonomic profiles (Alberdi et al. 2017; Murray, Coghlan, and Bunce 2015). Intriguingly, Smith and Peay (2014) reported the opposite conclusion: that increasing the number of PCR replicates does not influence α diversity. As is common in eDNA studies, Smith and Peay amplify each of their PCRs over 30 cycles, whereas we estimated the optimal number of cycles for each reaction separately using qPCR, following Murray, Coghlan, and Bunce (2015). Overamplification of PCR amplicon pools can reduce the complexity of the amplicon pool as read "species" that replicate more efficiently outcompete others that replicate less efficiently (Nichols et al., 2017). Consequently, taxa that are least efficiently amplified will become increasingly rare and may not be observed, in particular at low read sampling depths.

We observed most singleton taxa in only one PCR replicate (Figure 6). This finding supports the conclusion by Leray and Knowlton (2017) that random sampling of rare taxa across PCR replicates accounts for most of the variation between PCR replicates. Increasing read sampling depth did not reduce the number of replicates that were required to stabilize the taxon accumulation curve (Table 2). However, increasing the minimum read cutoff did reduce the number of PCR replicates necessary to stabilize the curve (Table 2), presumably by removing many of the low abundance taxa from each data set such that only the high abundance taxa, most of which were common to each PCR, remained.

We found that increasing the read sampling depth significantly increased the number of taxa detected at each of our sites (Figures 1, 2; Table S4). As many eDNA studies and consortia sequence amplicon pools to the shallowest of our depths (1,000 reads), this result has implications for how biodiversity estimates based on these published data sets can be interpreted and compared. The impact of this parameter choice depends on how the data are analyzed. For example, we estimated significantly higher observed α diversity at a depth of 10,000 reads than at a depth of 1,000 reads across all sites, but found no difference between read depths when α diversity was calculated using the Shannon or Simpson metrics, which underweight low abundance taxa compared to common taxa (Hsieh et al., 2016). The significant increase in α diversity that we observed is in contrast to Murray, Coghlan, and Bunce (2015), who found that sampling depth per PCR replicate did not necessarily increase detection of low abundance taxa. This difference may be due to the use by Murray, Coghlan, and Bunce of Ion Torrent rather than Illumina sequencing technology, as the higher error profiles generated by the Ion Torrent platform require more stringent removal of rare taxa (Salipante et al., 2014). Because sites will vary in the amount of total diversity present, taxon accumulation curves such as those in Figure 1 may be useful in determining the appropriate read sampling depth for a given site.

Our results also reiterate the need to consider the physical and ecological setting during eDNA experimental design (Anderson et al., 2012; Ficetola et al. 2015). We observed the most variation in observed α diversity among PCR replicates in the PITS dataset at YL.1 (Figure 1e and 2), a lagoon basin into which water

and wind carries and deposits DNA-containing materials from the surrounding environment. The constant influx of DNA from the surrounding habitats may explain why amplicon pools from this site include many low abundance taxa. Although these low abundance taxa have little effect on β diversity estimates, they are contributing members of local communities. Metabarcoding may therefore be particularly inefficient tool for estimating and comparing α diversity at sites with high biological turnover or input.

Because we sequenced each PCR replicate individually, we were also able to explore the rate of occurrence and potential impact of PCR outliers, which we define as PCR amplicon pools that differ significantly in either composition or relative abundance of taxa compared to other replicates from the same eDNA extract. We found PCR outliers to be more common at sites with high diversity, like YL.1. Increasing read sampling depth also increased the frequency of PCR outliers, but only for the FITS data sets (Figure 4), possibly because of the higher taxonomic diversity among low abundance taxa recovered by this metabarcode. While we are unable to determine the precise cause of outlier PCRs, we note that they are only observable as outliers if more than one PCR replicate is performed. This rationale is often used by groups that perform three PCR replicates per sample (Taberlet et al., 2018), which allows disambiguation between an outlier and a "normal" PCR without additional laboratory work.

Given our results, we present the following conclusions, which can serve as recommendations for experimental design in eDNA metabarcoding experiments:

1. **PCR Replication:** A single PCR often captures the diversity of common taxa at a site and allows sites to be differentiated based on these common taxa. However, because outlier PCRs are a possibility, a minimum of two PCR replicates is recommended. When multiple PCR replicates are performed, the LCBD statistic can be used to identify PCR outliers by quantifying replicate uniqueness.
2. **Read sampling depth:** Increasing sequencing read depth increases the chance that low abundance taxa are recovered from within the amplicon pool. However, because PCR replicates vary in taxonomic composition, exhausting the sequence complexity of an amplicon pool through deep sequencing is not the same as exhausting the sequence complexity of a DNA extract. Variation between PCR replicates in taxonomic composition or relative abundance does not diminish with increased sequencing read depth.
3. **Minimum read cutoff:** Higher minimum read cutoffs remove low abundance taxa from a PCR amplicon pool. Removed taxa will include both low abundance contaminants and low abundance authentic taxa. As such, the minimum read threshold may influence α diversity but is less likely to influence β diversity.

5- Conclusion

Here, we provide a thorough examination of the influence of PCR replication and read sampling depth on two common measures of diversity in environmental DNA metabarcoding research: α diversity and β diversity. We find that metabarcoding PCR robustly estimates measures of diversity that rely on high abundance taxa, and that low numbers of PCR replicates are sufficient to distinguish sites from each other. However, we also find that low abundance taxa tend to occur in only one or a few replicate PCRs, and often require deeper sequencing of PCR amplicon pools to be counted, in particular when read cutoff thresholds are high. Importantly, the challenges of counting rare taxa are so great, in particular in environmentally diverse sites, that even large numbers of PCR replicates and relatively deep sequencing are insufficient to guarantee that they will be recorded. Experiments aiming to catalogue diversity should consider using replicate PCRs, replicate extracts, and even replicate field sampling. While this problem necessarily limits the utility of environmental DNA as a mechanisms to catalogue the full diversity of taxa present at a given site, it also shows the potential power of eDNA to recover even the rarest of taxa, although the authenticity of exceptional taxa may need to be explored using a different enrichment approach, such as hybridization capture or targeted amplification.

Acknowledgments

References

Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-Madsen, H., Kjær, K. H., . . . Willerslev, E. (2012). Meta-barcoding of "dirt" DNA from soil reflects vertebrate biodiversity. Molecular Ecology, 21(8), 1966–1979. *https://doi.org/10.1111/j.1365-294X.2011.05261.x*

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2017). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution* , *9* (1), 134–147. https://doi.org/10.1111/2041-210X.12849

Baker, C. S., Steel, D., Nieukirk, S., & Klinck, H. (2018). Environmental DNA (eDNA) From the Wake of the Whales: Droplet Digital PCR for Detection and Species Identification. Frontiers in Marine Science, 5, 133. *https://doi.org/10.3389/fmars.2018.00133*

Beentjes, K. K., Speksnijder, A. G. C. L., Schilthuizen, M., Hoogeveen, M., & Van Der Hoorn, B. B. (2019). The effects of spatial and temporal replicate sampling on eDNA metabarcoding. PeerJ, 2019(7), 1–18. *https://doi.org/10.7717/peerj.7335*

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., . . . Creer, S. (2017). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. Nature Communications, 8, 1–11. *https://doi.org/10.1038/ncomms14087*

Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Elbrecht, V., Steinke, D., Ratnasingham, S., . . . Hebert, P. D. N. (2019). Metabarcoding a diverse arthropod mock community. Molecular Ecology Resources, 19(3), 711–727. doi: 10.1111/1755-0998.13008

Bray, J. R. & Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27, 325–349.

Browne, L., Mead, A., Horn, C., Chang, K., Celikkol, Z.A., Henriquez, C.L., Ma, F., Beraut, E., Meyer, R.S., Sork, V.L.. (2020). Experimental DNA Demethylation Associates with Changes in Growth and Gene Expression of Oak Tree Seedlings. G3: Genes, Genomes, Genetics, 10(3), 1019-1028.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods, 13(7), 581–583. *https://doi.org/10.1038/nmeth.3869*

Clarke, C.L., Edwards, M.E., Gielly, L., Ehrich, D., Hughes, P.D.M., Morozova, L.M., Haflidason, H., Mangerud, J., Svendsen, J.I., Alsos, I.G.. (2019). Persistence of arctic-alpine flora during 24,000 years of environmental change in the Polar Urals. Scientific Reports 9, Article No. 19613

Cristescu, M. E., & Hebert, P. D. (2018). Uses and Misuses of Environmental DNA in Biodiversity Science and Conservation. *Annual Review of Ecology, Evolution, and Systematics* , *49* (1), 209–230. doi: 10.1146/annurev-ecolsys-110617-062306

Curd, E. E., Gold, Z., Kandlikar, G. S., Gomer, J., Ogden, M., O'Connell, T., . . . Meyer, R. S. (2019). Anacapa Toolkit : an environmental DNA toolkit for processing multilocus metabarcode datasets. Methods in Ecology and Evolution, 2041–210X.13214. *https://doi.org/10.1111/2041-210X.13214*

Davis, N. M., Proctor, Di. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome, 6(1), 1–14. *https://doi.org/10.1186/s40168-018-0605-2*

De Barba M, Miquel C, Boyer F *et al* . (2014) DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources* , 14, 306– 323.

Deiner K, Lopez J, Bourne S, Holman LE, Seymour M, Grey EK, Lacoursière-Roussel A, Li Y, Renshaw MA, Pfrender ME, Rius M, Bernatchez L, Lodge DM (2018) Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. Metabarcoding and Metagenomics 2: e28963.*https://doi.org/10.3897/mbmg.2.28963*

Deveautour, C., Donn, S., Power, S. A., Bennett, A. E., & Powell, J. R. (2018). Experimentally altered rainfall regimes and host root traits affect grassland arbuscular mycorrhizal fungal communities. Molecular Ecology, 27(8), 2152–2163.*https://doi.org/10.1111/mec.14536*

DiBattista, J. D., Coker, D. J., Sinclair-Taylor, T. H., Stat, M., Berumen, M. L., & Bunce, M. (2017). Assessing the utility of eDNA as a tool to survey reef-fish communities in the Red Sea. Coral Reefs, 36(4), 1245–1252.*https://doi.org/10.1007/s00338-017-1618-1*

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, e0130324. https://doi.org/10.1371/journal.pone.0130324

Epp, L. S., K. R. Stoof, M. H. Trauth, and R. Tiedemann. (2010). Historical genetics on a sediment core from a Kenyan lake: intraspecific genotype turnover in a tropical rotifer is related to past environmental changes. Journal of Paleolimnology 43, 939–954.

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(8), 1821–1833. https://doi.org/10.1111/j.1365-294X.2012.05537.x

Erlandson, S., Wei, X., Savage, J., Cavender-Bares, J., & Peay, K. (2018). Soil abiotic variables are more important than Salicaceae phylogeny or habitat specialization in determining soil microbial community structure. Molecular Ecology, 27(8), 2007–2024.*https://doi.org/10.1111/mec.14576*

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. *Biology Letters* , 4, 423– 425.

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., . . . Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources* , 15, 543–556. https://doi-org.oca.ucsc.edu/10.1111/1755-0998.12338

Ficetola, G. F., Poulenard, J., . . . Arnaud, F.. (2018). DNA from lake sediments reveals long-term ecosystem changes after a biological invasion. Science Advances, 4(5), eaar4292.*https://doi.org/10.1126/sciadv.aar4292*

Fonseca, V. G., Nichols, B., Lallias, D., Quince, C., Carvalho, G. R., Power,, D. M., & Creer, S. (2012). Sample richness and genetic diversity as drivers of chimera formation in nSSU metagenetic analyses. Nucleic Acids Research, 40, e66. https://doi.org/10.1093/nar/gks002

Fonseca, V. G. (2018). Pitfalls in relative abundance estimation using eDNA metabarcoding. Molecular Ecology Resources, 18(5), 923–926.*https://doi.org/10.1111/1755-0998.12902*

Franklin, T. W., McKelvey, K. S., Golding, J. D., Mason, D. H., Dysthe, J. C., Pilgrim, K. L., . . . Schwartz, M. K. (2019). Using environmental DNA methods to improve winter surveys for rare carnivores: DNA from snow and improved noninvasive techniques. Biological Conservation, 229(October 2018), 50–58.*https://doi.org/10.1016/j.biocon.2018.11.006*

Gao, X., Lin, H., Revanna, K., & Dong, Q. (2017). A Bayesian taxonomic classification method for 16S rRNA gene sequences with improved species-level accuracy. BMC Bioinformatics, 18(1), 247. https://doi.org/10.1186/s12859-017-1670-4

Garlapati, D., Charankumar, B., Ramu, K., Madeswaran, P., & Ramana Murthy, M. V. (2019). A review on the applications and recent advances in environmental DNA (eDNA) metagenomics. Reviews in Environmental Science and Biotechnology, 18(3), 389–411.*https://doi.org/10.1007/s11157-019-09501-4*

Giguet-Covex C, Pansu J, Arnaud F *et al* . (2014) Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nature Communications* , 5, 3211.

Gordon A. & Hannon GJ.. (2010). FastX Toolkit. Available:*http://hannonlab.cshl.edu/fastx_toolkit/*

Graham, R.W., Belmecheri, S., Choy, K., Culleton, B.J., Davies, L.J., Froese, D., Heintzman, P.D., Hritz, C., Kapp, J.D., Newsom, L.A., Rawcliffe, R., Saulnier-Talbot, E., Shapiro, B., Wang, Y., Williams, J.W., Wooller, M.J.. (2016). Timing and causes of mid-Holocene mammoth extinction on St. Paul Island, Alaska. Proceedings of the National Academy of Sciences USA 113, 9310–9314.

Hope PR, Bohmann K, Gilbert MTP, Zepeda-Mendoza M, Razgour O, Jones G. (2014). Second generation sequencing and morphological faecal analysis reveal unexpected foraging behaviour by Myotis nattereri (Chiroptera, Vespertilionidae) in winter. Front Zool.11(1):39.

Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). Methods in Ecology and Evolution, 7(12), 1451–1456. doi: 10.1111/2041-210x.12613

Kandlikar, G. S., Gold, Z. J., Cowen, M. C., Meyer, R. S., Freise, A. C., Kraft, N. J. B., . . . Curd, E. E. (2018). ranacapa: An R package and Shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. F1000Research, 7, 1734.*https://doi.org/10.12688/f1000research.16680.1*

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), 1–24.*https://doi.org/10.1371/journal.pone.0177643*

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

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. Scientific Reports, 7, 17668. https://doi.org/10.1038/s41598-017-17333-x

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357–359. https://doi.org/10.1038/nmeth.1923

Laramie, M. B., Pilliod, D. S., & Goldberg, C. S. (2015). Characterizing the distribution of an endangered salmonid using environmental DNA analysis. Biological Conservation, 183, 29–37.*https://doi.org/10.1016/j.biocon.2014.11.025*

Lallias, D., Hiddink, J. G., Fonseca, V. G., Gaspar, J. M., Sung, W., Neill, S. P., . . . Creer, S. (2015). Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems. ISME J, 1208-1221. doi:10.1038/ismej.2014.213

Lanzen, A., Lekang, K., Jonassen, I., Thompson, E. M., & Troedsson, C. (2017). DNA extraction replicates improve diversity and compositional dissimilarity in metabarcoding of eukaryotes in marine sediments.*PLoS ONE* , *12* (6), 1–18. https://doi.org/10.1371/journal.pone.0179443

Leempoel K, Hebert T, Hadly EA. (2020). A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. Proc. R. Soc. B 287: 20192353. http://dx.doi.org/10.1098/rspb.2019.2353

Leray, M., & Knowlton, N. (2016). Censusing marine eukaryotic diversity in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* , *371* (1702), 20150331.

https://doi.org/10.1098/rstb.2015.0331

Leray M., & Knowlton N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. PeerJ 5:e3006*https://doi.org/10.7717/peerj.3006*

Lin, M., Simons, A.L., Curd, E., Harrigan, R., Schneider, F., . . . Meyer, R.. A Biodiversity Composition Map of California Derived from Environmental DNA Metabarcoding and Earth Observation. *Ecological Applications.* Submitted.

Lundberg, D. S., Yourstone, S., Mieczkowski, P., Jones, C. D., & Dangl, J. L. (2013). Practical innovations for high-throughput amplicon sequencing. *Nature Methods* , *10* (10), 999–1002. doi: 10.1038/nmeth.2634

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal, 17(1), 10. https://doi.org/10.14806/ej.17.1.200

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE, 8(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Meyer, R., E. E. Curd, T. Schweizer, Z. Gold, D. Ramos Ruiz, S. Shirazi, G. Kandlikar, W. Kwan, M. Lin, A. Friese, J. Moberg-Parker, M. Ramos Munguia, B. Shapiro, J. Sexton, L. Pipes, A. Vedrenne Garcia, M. Mejia Palacios, E. Aronson, T. Moore, R. Nielsen, H. Lewin, P. Barber, J. Wall, N. Kraft, and R. Wayne. (2019). "The California Environmental DNA 'CALeDNA' Program." California Agriculture.

Murray DC, Coghlan ML, Bunce M. (2015). From Benchtop to Desktop: Important Considerations when Designing Amplicon Sequencing Workflows. PLoS ONE 10(4): e0124671.*https://doi.org/10.1371/journal.pone.0124671*

Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., . . . Shapiro, B. (2017). Minimizing polymerase biases in metabarcoding. Molecular Ecology Resources.*https://doi.org/10.1111/1755-0998.12895*

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., . . . Lamberti, G. A. (2016). Estimating species richness using environmental DNA. Ecology and Evolution, 6(12), 4214–4226.*https://doi.org/10.1002/ece3.2186*

Parducci, L., Jorgensen, T., Tollefsrud, M.M., Elverland, E., Alm, T., Fontana, S.L., Bennett, K.D., Haile, J., Matetovici, I., Suyama, Y., Edwards, M.E., Andersen, K., Rasmussen, M., Boessenkool, S., Coissac, E., Brochmann, C., Taberlet, P., Houmark-Nielsen, M., Larsen, N.K., Orlando, L., Gilbert, M.T.P., Kjaer, K.H., Alsos, I.G., Willerslev, E.. (2012). Glacial Survival of Boreal Trees in Northern Scandinavia. Science 335, 1083–1086.

Pawluczyk, M., Weiss, J., Links, M. G., Egana Aranguren, M., Wilkinson, M. D., & Egea-Cortines, M. (2015). Quantitative evaluation of bias in PCR amplification and next-generation sequencing derived from metabarcoding samples. Analytical and Bioanalytical Chemistry, 407(7), 1841–1848.*https://doi.org/10.1007/s00216-014-8435-y*

Pedersen, M.W., Overballe-Petersen, S., Ermini, L., Sarkissian,C.D., Haile, J., Hellstrom, M., Spens, J., Thomsen, P.F., Bohmann, K., Cappellini, E., Schnell, I.B., Wales, N.A., Caroe, C.,Campos, P.F., Schmidt, A.M.Z., Gilbert, M.T.P., Hansen, A.J.,Orlando, L., Willerslev, E.. (2015). Ancient and modern environmental DNA. Philosophical Transactions of the Royal Society B 370, Article No. 20130383.

Pedersen, M.W., Ruter, A., Schweger, C., Friebe, H., Staff, R.A., Kjeldsen, K.K., Mendoza, M.L.Z., Beaudoin, A.B., Zutter, C., Larsen, N.K., Potter, B.A., Nielsenlo, R., Rainville, R.A., Orlando, L., Meltzer, D.J., Kjaer, K.H., Willerslev, E. (2016). Postglacial viability and colonization in North America's ice-free corridor. Nature 537, 45–49.

14

Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary, P. R., Litvin, S. Y., Nickols, K. J., . . . Kelly, R. P. (2016). Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. Molecular Ecology, 25(2), 527–541.*https://doi.org/10.1111/mec.13481*

Robasky K, Lewis NE, Church GM. (2014). The role of replicates for error mitigation in next-generation sequencing. Nat Rev Genet 15: 56–62. pmid:24322726

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation, 17. doi: 10.1016/j.gecco.2019.e00547

Salipante, S. J., Kawashima, T., Rosenthal, C., Hoogestraat, D. R., Cummings, L. A., Sengupta, D. J., . . . Hoffman, N. G. (2014). Performance comparison of Illumina and Ion Torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling.*Applied and Environmental Microbiology* , *80* (24), 7583–7591.*https://doi.org/10.1128/AEM.02206-14*

Schnell, I. B., Thomsen, P. F., Wilkinson, N., Rasmussen, M., Jensen, L. R. D., Willerslev, E., . . . Gilbert, M. T. P. (2012). Screening mammal biodiversity using DNA from leeches. *Current Biology* ,*22* (8), R262–R263. https://doi.org/10.1016/j.cub.2012.02.058

Schnell, I. B., Bohmann, K., Schultze, S. E., Richter, S. R., Murray, D. C., Sinding, M.-H. S., . . . Gilbert, M. T. P. (2018). Debugging diversity - a pan-continental exploration of the potential of terrestrial blood-feeding leeches as a vertebrate monitoring tool. Molecular Ecology Resources, 18(6), 1282–1298.*https://doi.org/10.1111/1755-0998.12912*

Shannon, C. E.. (1948). A mathematical theory of communication. The Bell System Technical Journal, 27, 379–423 and 623–656.

Simpson, E. (1949). Measurement of Diversity. Nature 163, 688. doi:10.1038/163688a0

Sjogren, P., Edwards, M.E., Gielly, L., Langdon, C.T., Croudace, I.W., Merkel, K.F., Fonville, T., Alsos, I.G.. (2017). Lake sedi- mentary DNA accurately records 20th Century introductions of exotic conifers in Scotland. New Phytologist 213, 929–941

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

Stahlschmidt, M. C., Collin, T. C., Fernandes, D. M., Bar-Oz, G., Belfer-Cohen, A., Gao, Z., . . . Pinhasi, R. (2019). Ancient Mammalian and Plant DNA from Late Quaternary Stalagmite Layers at Solkota Cave, Georgia. Scientific Reports, 9(1), 1–10.*https://doi.org/10.1038/s41598-019-43147-0*

Stat, M., Huggett, M.J., Bernasconi, R. . . Bunce, M.. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. Sci Rep 7, 12240*https://doi.org/10.1038/s41598-017-12501-5*

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology, 21(8), 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Environmental DNA: For biodiversity research and monitoring. Oxford, UK: Oxford University Press.*https://doi.org/10.1093/oso/9780198767220.001.0001*

Tsuji, S., Miya, M., Ushio, M., Sato, H., Minamoto, T., & Yamanaka, H. (2019). Evaluating intraspecific genetic diversity using environmental DNA and denoising approach: A case study using tank water.*Environmental DNA* , *2* (1), 42–52. https://doi.org/10.1002/edn3.44

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

van der Valk, T., Vezzi, F., Ormestad, M., Dalen, L., & Guschanski, K. (2019). Index hopping on the Illumina HiseqX platform and its consequences for ancient DNA studies. Molecular Ecology Resources, https://doi.org/10.1111/1755-0998.13009

White, T. J., Bruns, T., Lee, S., & Taylor, J. W. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR Protocols: A Guide to Methods and Applications. Academic Press, 64(February), 315–322. https://doi.org/citeulike-article-id:671166

Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T.P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D.A., Coo- per, A., (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science 300, 791–795.

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M.E., Lorenzen, E.D., Vestergard, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L.S., Pear- man, P.B., Cheddadi, R., Murray, D., Brathen, K.A., Yoccoz, N., Binney, H., Cruaud, C., Wincker, P., Goslar, T., Alsos, I.G., Bellemain, E., Brysting, A.K., Elven, R., Sonstebo, J.H., Murton, J., Sher, A., Rasmussen, M., Ronn, R., Mourier, T., Cooper, A., Austin, J., Moller, P., Froese, D., Zazula, G., Pompanon, F., Rioux, D., Niderkorn, V., Tikhonov, A., Savvinov, G., Roberts, R.G., MacPhee, R.D.E., Gilbert, M.T.P., Kjaer, K.H., Orlando, L., Brochmann, C., Taberlet, P.. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. Nature 506, 47–51

Xia, Z., Zhan, A., Gao, Y., Zhang, L., Haffner, G. D., & MacIsaac, H. J. (2018). Early detection of a highly invasive bivalve based on environmental DNA (eDNA). Biological Invasions, 20(2), 437–447.*https://doi.org/10.1007/s10530-017-1545-7*

Yan, D. F., Mills, J. G., Gellie, N. J. C., Bissett, A., Lowe, A. J., & Breed, M. F. (2018). High-throughput eDNA monitoring of fungi to track functional recovery in ecological restoration. Biological Conservation, 217, 113–120.*https://doi.org/10.1016/j.biocon.2017.10.035*

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., . . . Chen, S. (2010). Use of ITS2 Region as the Universal DNA Barcode for Plants and Animals. PLoS ONE, 5(10), e13102. https://doi.org/10.1371/journal.pone.0013102

Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., . . . Miya, M. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. Scientific Reports, 7(December 2016), 1–12.*https://doi.org/10.1038/srep40368*

Zepeda-Mendoza, M.L., Bohmann, K., Carmona Baez, A., Gilbert, T.. (2016). DAMe: a toolkit for the initial processing of datasets with PCR replicates of double-tagged amplicons for DNA metabarcoding analyses.*BMC Res Notes* 9, 255 https://doi.org/10.1186/s13104-016-2064-9

Data Accessibility

Our raw data and metadata are available at*https://github.com/sashirazi/eDNA-PCR-Project/*and will be archived in Dryad and made publicly available upon acceptance for publication.
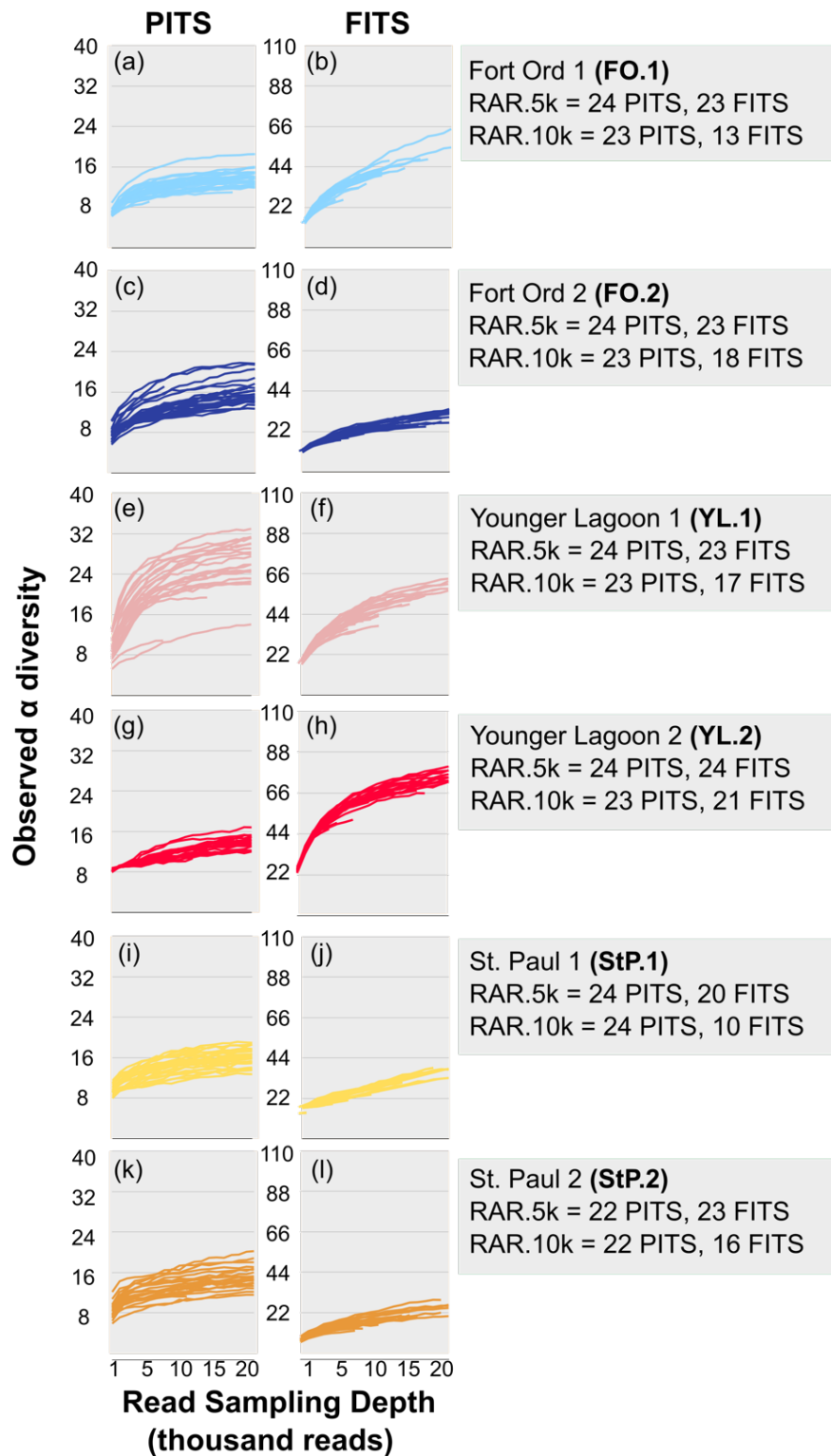
Tables and Figures

Figure 1: Taxon accumulation curve tracking observed number of taxa identified among PCR replicates as a function of read sampling depth for the PITS (a-k) and FITS (j-i) data sets. Plots are created from

17

the average of 25 rarefactions at each round thousand read sampling depth between 1000 and 20,000 reads. After each round of rarefaction, we removed taxa present as fewer than five reads. Each line represents one PCR replicate. Termination of a line prior to the 20,000 read sampling depth denotes missing data. RAR = Retained after rarefaction, referring to the number of PCR replicates retained in analysis following rarefaction to read sampling depths of 5,000 and 10,000 reads. All 24 replicates for each of the six sites and both primers had sufficient data at 1,000 reads to be included in analysis (RAR.1k = 24 for all sites and amplicons).
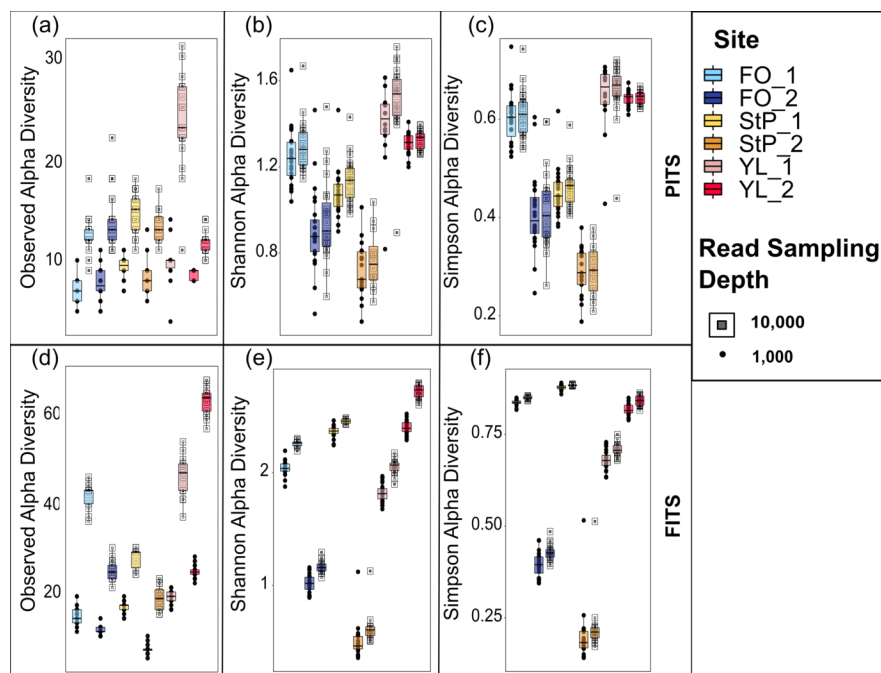

Figure 2. A comparison of observed (a & d), Shannon (b & e), and Simpson (c & f) α diversity measured with read sampling depths of 1,000 (squares) and 10,000 (circles). Taxa present as fewer than five reads were removed prior to analysis. Bars are colored by site and each dot represents a single PCR replicate.
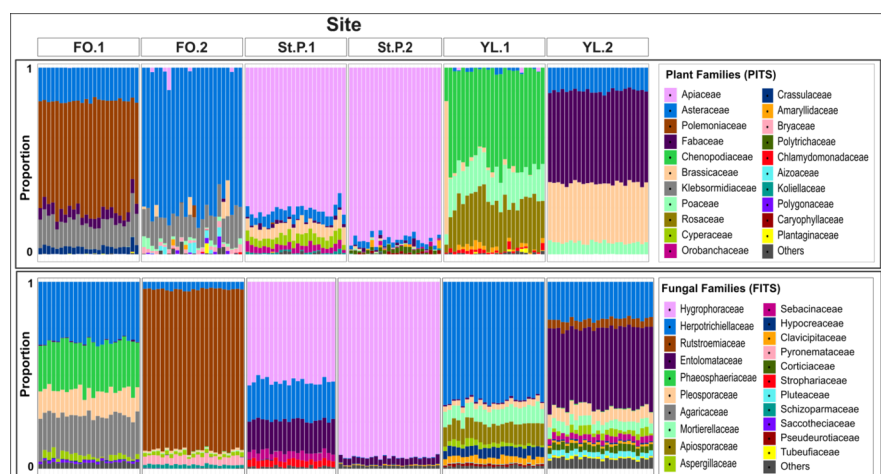

Figure 3: Relative abundance of (a) plant and (b) fungal families detected with 5,000 reads and a five read

18

minimum cutoff. Each bar represents one PCR replicate. PCR replicates are grouped by site. Only the 20 most abundant families are included here.
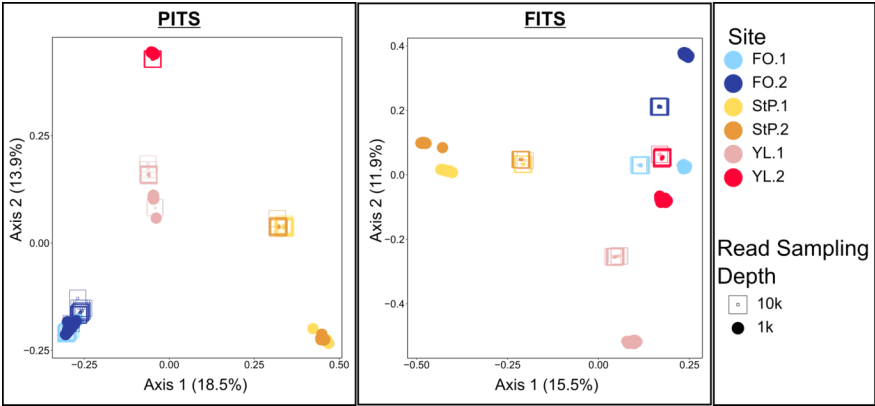

Figure 4. PCoA on Jaccard measure of β diversity for PITS and FITS datasets with a five read minimum cutoff. Each point represents one PCR replicate. Colors indicate site and shapes indicate sequencing depth. The values on each axis represent the percentage of variation attributable to that axis.
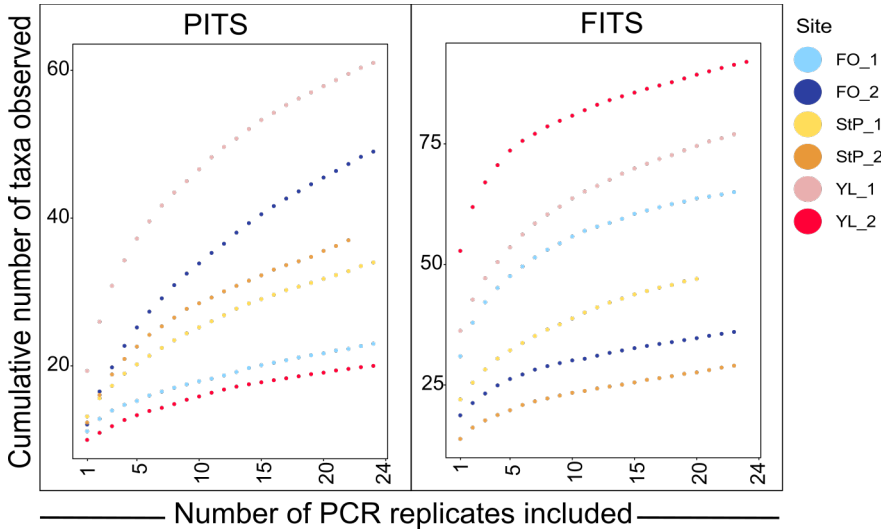

Figure 5. Rarefaction curves describing the cumulative number of taxa detected with increasing number of PCR replicates, each sampled to a read depth of 5000 reads and using a minimum read cutoff of five. Each line reflects the average of 100 bootstraps in which the order at which individual replicates were added was shuffled.
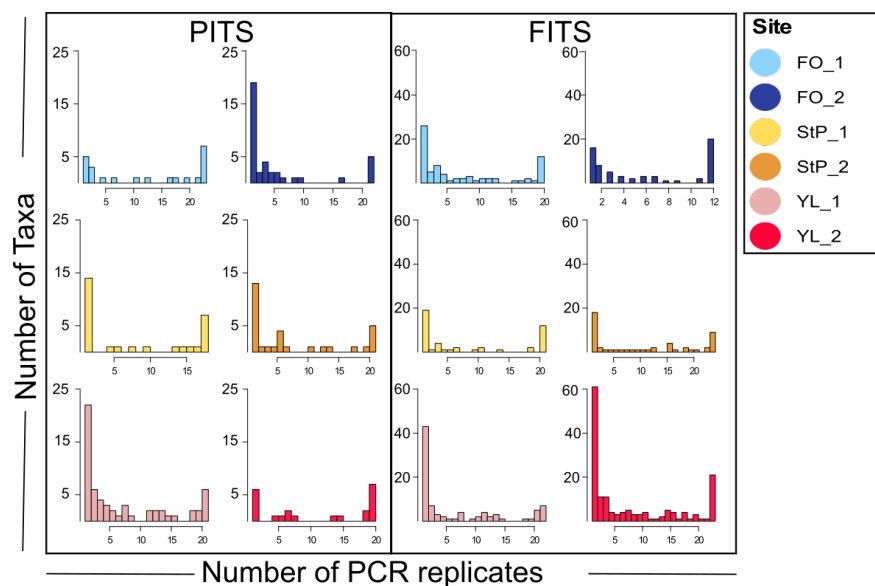
Figure 6. Histograms describing the frequency of individual taxa detected across PCR replicates, each sampled to a read depth of 5000 reads and using a minimum read cutoff of five, out of the total 24 replicates. The right-most bar is a count of taxa present in all replicates while the left-most bar is a count of taxa present in only one replicate. Increasing the minimum read cutoff to ten reduced the number of singleton taxa, while decreasing the minimum read cutoff to two increased the number of singleton taxa (data not shown).
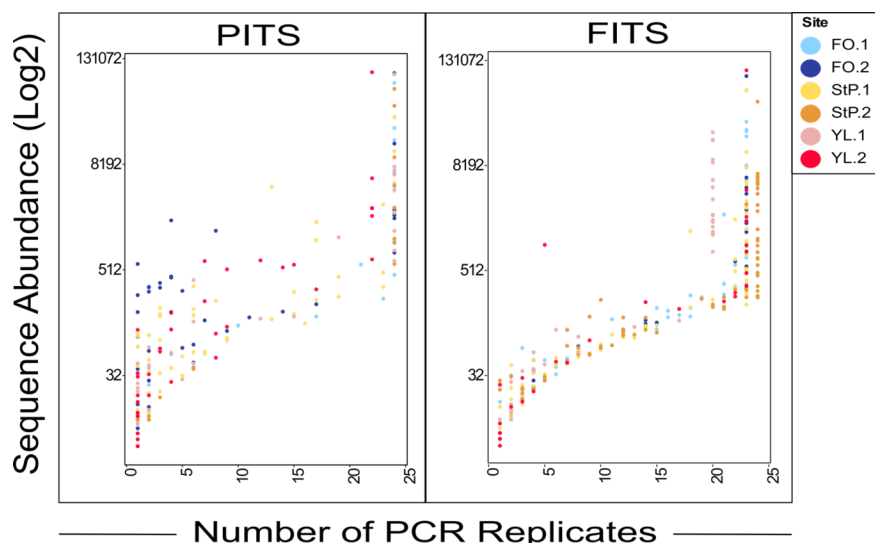

Figure 7. The number of PCRs in which a particular taxon is observed vs the average abundance of that taxon within each PCR. Each data set comprises 5000 subsampled reads and incorporates a minimum read cutoff of five. Sequence abundance is plotted as log-transformed counts of the number of reads per PCR assigned to a particular taxon, averaged across the PCRs in which that taxon is observed. We find a positive correlation between the number of PCR replicates in which a taxon is observed. Each dot represents an individual taxon and is colored according to site.

20

SITE	MIN	MAX	STDEV
Plant *ITS2* (PITS)			
FO1	21.2	48.3	5.8
FO2	2.0	68.3	18.6
YL1	2.0	122.8	31.4
YL2	3.0	49.8	10.5
StP15	1.0	51.4	15.2
StP20	5.0	56.2	12.9
Fungal *ITS1* (FITS)			
FO1	46.4	111.2	14.9
FO2	24.5	106.7	22.2
YL1	48.6	156.1	27.9
YL2	97.0	199.8	24.4
StP15	30.7	78.4	11.6
StP20	34.2	75.2	12.0

Table 1. Variation in extrapolated richness among PCR replicates after outlier removal.

Read Sampling Depth	Minimum read cutoff	FO.1	FO.2	YL.1	YL.2	StP.1	StP.2
Plant ITS2 (PITS)							
5k	2	5	20	17	4	8	10
	5	3	17	15	1	8	9
	10	2	10	12	1	3	8
10k	2	9	17	>23	5	15	19
	5	4	13	19	3	10	18
	10	3	9	16	1	9	9
Fungal ITS1 (FITS)							
5k	2	>23	>23	>23	>24	>20	19
	5	11	7	15	12	12	6
	10	9	2	6	9	4	2
10k	2	>13	>18	>17	>21	>10	>16
	5	>13	>18	>17	>21	>10	>16
	10	>13	4	12	9	7	3

Table 2: Number of PCR replicates required to reach the point at which taxon accumulation curve (shown in Figure 5) increases by less than one taxon with the addition of another PCR replicate. We observe variation by site and marker, and a common trend of decreasing replicates required with an increase in read cutoff. '>X' denotes that greater than the maximum number of replicates X retained after rarefaction is needed to suffice this point.