ORIGINAL RESEARCH

Ecology and Evolution
Open Access

WILEY

Effects of consumer surface sterilization on diet DNA metabarcoding data of terrestrial invertebrates in natural environments and feeding trials

Ana Miller-ter Kuile 🆔 | Austen Apigo | Hillary S. Young

Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA, USA

Correspondence
Ana Miller-ter Kuile, Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106, USA.
Email: ana.miller-ter.kuile@lifesci.ucsb.edu

Abstract

1. DNA metabarcoding is an emerging tool used to quantify diet in environments and consumer groups where traditional approaches are unviable, including small-bodied invertebrate taxa. However, metabarcoding of small taxa often requires DNA extraction from full body parts (without dissection), and it is unclear whether surface contamination from body parts alters presumed diet presence or diversity.

2. We examined four different measures of diet (presence, rarefied read abundance, richness, and species composition) for a terrestrial invertebrate consumer (the spider *Heteropoda venatoria*) both collected in its natural environment and fed an offered diet item in contained feeding trials using DNA metabarcoding of full body parts (opisthosomas). We compared diet from consumer individuals surface sterilized to remove contaminants in 10% commercial bleach solution followed by deionized water with a set of unsterilized individuals.

3. We found that surface sterilization did not significantly alter any measure of diet for consumers in either a natural environment or feeding trials. The best-fitting model predicting diet detection in feeding trial consumers included surface sterilization, but this term was not statistically significant ($\beta = -2.3$, p-value = .07).

4. Our results suggest that surface contamination does not seem to be a significant concern in this DNA diet metabarcoding study for consumers in either a natural terrestrial environment or feeding trials. As the field of diet DNA metabarcoding continues to progress into new environmental contexts with various molecular approaches, we suggest ongoing context-specific consideration of the possibility of surface contamination.

KEYWORDS
consumptive interactions, contamination, diet analysis, food web, invertebrates, predator–prey interactions

1 | INTRODUCTION

Biological communities and ecosystem function are shaped by interactions between organisms (Hooper et al., 2005). Among the many interaction types, consumptive interactions (including herbivory, predation, and parasitism) can shape the stability of biologically diverse communities (Delmas et al., 2019). Until recently, consumptive interactions were most often measured by visual observations of feeding or by gut dissection or inspection of fecal contents (Baker et al., 2014; Nielsen et al., 2018), which made it challenging or impossible to conduct diet analyses for many consumer groups. Specifically, these diet analyses are not possible for consumers that (a) are too small for dissection and food identification and (b) have feeding habits or food items which make diet visually unidentifiable (Sheppard & Harwood, 2005). This group of consumers, including terrestrial insects, spiders, and other arthropods, form the base of most terrestrial food webs and are integral to maintaining biodiversity and ecosystem functioning in ecosystems worldwide (Wilson, 1987). For these consumer groups, the use of high-throughput sequencing is one of the most promising emerging approaches for determining gut contents. High-throughput sequencing (hereafter referred to as "diet DNA metabarcoding") can identify a suite of diet species at once and provides a comprehensive and efficient method for determining intrapopulation, intraspecific, and interspecific diets (Lucas et al., 2018; Pompanon et al., 2012; Quéméré et al., 2013; Soininen et al., 2015). These methods have already illuminated new interactions and ecological trends in a variety of environments (e.g., host–parasitoid: (Wirta et al., 2014); plant–herbivore: (Kartzinel et al., 2015); host–parasite: (Schnell et al., 2012); and predator–prey: (Toju & Baba, 2018).

As diet DNA metabarcoding methods continue to advance, however, they need to be validated so that the ecological inference made from them is robust. Focusing on the challenges of small organisms where small body size has limited other diet analysis methods, DNA diet analyses are often performed on full organisms or body parts without gut dissection (e.g., Jacobsen et al., 2018; Toju & Baba, 2018). The necessity to use full organisms or body parts increases the possibility of surface contamination, altering detection and species composition of presumed diet items. Surface sterilization, the use of chemical treatments or physical action to remove surface contaminants, is systematically used in other fields to reduce the risk of contamination in DNA metabarcoding datasets (Burgdorf et al., 2014; Zimmerman & Vitousek, 2012). However, surface sterilization has not been systematically used in diet metabarcoding studies. While some fields have developed informed protocols based on decades of research into best practices and study-specific considerations (Brown et al., 2018), the field of diet DNA metabarcoding has not developed a similarly systematic approach (e.g., ethanol: Doña et al., 2019, bleach: Anslan et al., 2016, and no sterilization: Jacobsen et al., 2018; Wirta et al., 2014). The lack of systematic surface sterilization in diet DNA metabarcoding when using full individuals or body parts may be due to the desire to avoid DNA destruction in relatively permeable animal cells (Greenstone et al., 2012). However,

without considering surface sterilization as a treatment for surface contamination, we have limited ability to confidently assign DNA sequences to ingested diet items (Greenstone et al., 2011, 2012; Linville & Wells, 2002).

In this study, we look at the effects of surface sterilization to remove surface contaminants on our understanding of consumer diets where the DNA of full body parts (no internal dissection) is used for diet DNA metabarcoding. Targeting the CO1 gene region, we produced high-throughput sequencing results from the full body parts (opisthosomas without gut dissection) of an invertebrate consumer species (the spider, *Heteropoda venatoria*). We surface sterilized half of the consumers prior to DNA extraction using a series of washes in a 1:10 dilution of bleach (10% commercial bleach) and deionized water; we left the other half of consumers unsterilized. We first determined how surface sterilization to remove contaminants impacts presumed diet from consumers collected in their natural environment, comparing surface sterilized individuals to those which were not surface sterilized, to ask whether surface sterilization influences (a) detection, (b) rarefied abundance, (c) richness, and (d) composition of potential diet items. We then performed a laboratory feeding trial, comparing surface sterilized individuals to those which were not surface sterilized to ask whether surface sterilization influenced (a) detection or (b) rarefied abundance of offered diet items. Exploring these questions in natural and contained settings addresses whether surface contamination alters interpretations of feeding interactions and thus whether it should be incorporated into standard protocols in diet metabarcoding.

2 | MATERIALS AND METHODS

2.1 | Field site and collections

We conducted fieldwork on Palmyra Atoll National Wildlife Refuge, Northern Line Islands, USA (5°53′N, 162°05′W). Palmyra Atoll has a well-characterized species list and is relatively species poor, allowing for relatively complete characterization of consumer and diet items (Handler et al., 2007). We targeted a generalist, active hunting spider species (*Heteropoda venatoria*) because (a) it occurs in high abundance on the atoll and is easy to collect, (b) it is a generalist species that feeds on a wide suite of organisms (including spiders, other invertebrates, and two geckos in the genus *Lepidodactylus*), and (c) it is the only species in its family on the atoll, meaning consumer DNA can be differentiated from potential diet DNA. All individuals were stored individually in sterilized containers (Greenstone et al., 2011).

2.2 | Natural environment consumer collection

In 2015, we collected consumers ($n = 47$) from natural environments, which had fed on available diet items and come into contact with environmental surfaces, to test whether DNA metabarcoding

detects diet DNA effectively. Consumers were collected at night via eye shine while they were actively hunting. We collected the first individuals we observed in each survey period and so they represent the distribution of body size and population demographics of this species that actively hunt in that environment. We froze all individuals at −80°C immediately following collection until surface sterilization and DNA extraction in 2019.

2.3 | Feeding trial consumer setup and feeding

In 2017, we conducted laboratory trials ($n = 26$) to test whether DNA metabarcoding detects DNA from diet items offered in a contained environment. We created feeding environments from one-liter plastic yogurt containers with holes for air transfer and placed one *H. venatoria* in each container. After 12 hr, we placed one large grasshopper (*Oxya japonica*, a likely diet item (Handler et al., 2007)) in each container and left all containers for 24 hr. We then froze (−20°C) each *H. venatoria* that had killed the grasshopper ($n = 25$, consumption was not easily detectable and thus not considered in analyses). We cleaned all containers between trials with 10% bleach solution.

To test surface sterilization's efficacy at removing possible contaminants, we used a surface sterilization treatment (Burgdorf et al., 2014; Schulz et al., 1993) on ~half the consumers for each set: those collected from the natural environment and those subjected to controlled feeding trials. We submerged and stirred each (whole) consumer in 10% commercial bleach by volume for 2 min and washed each in deionized water for 2 min. Similar bleach submersion leads to undetectable DNA degradation in similar soft-exoskeleton consumers (Greenstone et al., 2012; Linville & Wells, 2002). Natural environment consumers (2015) had been frozen at −80°C since collection; we surface sterilized these consumers in a sterilized laminar flow hood in 2019 just before DNA extraction ($n = 22$ surface sterilized, $n = 25$ not surface sterilized; Table 1). We surface sterilized feeding trial consumers (2017) in the laboratory on the atoll in 2017 following freezing at −20°C and then stored each in individual vials of 95% ethanol in a −20°C freezer until DNA extraction (no −80°C freezer was available at the field station that year) ($n = 10$ surface sterilized; $n = 14$ not surface sterilized). Prior to DNA extraction, we dried all samples for 1–3 hr in a sterilized laminar flow hood and then removed the full opisthosoma (containing the hind gut region) using a sterilized scalpel. Between all steps, tools were sterilized with either ethanol and flame (scalpels and forceps) or 10% bleach (surfaces) between handling each individual.

2.4 | DNA extraction and removal of consumer DNA with AMPure XP beads

We extracted DNA from each consumer following a modified CTAB extraction protocol (Fulton et al., 1995). We quantified DNA using a Qubit (Invitrogen) fluorometer with the high sensitivity double-stranded DNA quantification kit. We followed Krehenwinkel et al. (2017) to isolate a proportion of lower molecular weight DNA with AMPure XP beads prior to PCR (Appendix S5, Figure S1). We diluted each DNA sample to 20ng/μl (creating a total sample volume of 40μl), mixed each sample using AMPure XP beads (0.75x bead-to-DNA ratio), and kept the supernatant. With the supernatant, we precipitated the DNA pellets with isopropanol and 5 M potassium acetate and washed DNA pellets with ethanol (Appendix S6). We quantified this cleaned DNA again using a Qubit fluorometer and diluted all samples to 10 ng/μl prior to PCR steps. All DNA pellets were stored in and diluted with TE buffer.

2.5 | PCR amplification, library preparation, and sequencing

We amplified the CO1 gene with general metazoan primers (Krehenwinkel et al., 2017; Leray et al., 2013; Yu et al., 2012; Table 2). We performed all PCR preparation steps in a UV-sterilized biosafety cabinet. We used PCR volumes of 25μl (9μl nuclease free water, 12.5μl GoTaq Green Master Mix (Promega Corp.), 1.25 μl of each of the primers (at 10 mM), and 1 μl of DNA template (at 10 ng/μl)). We ran each sample in duplicate along with duplicated negative samples each PCR run. PCRs are as follows: initial denaturation step at 95°C for 3 min and then 35 cycles of (a) 95°C for 30 s, (b) 46°C for 30 s, and (c) 72°C for 1 min, followed by a final 5 min at 72°C. We cleaned PCR products with AMPure XP beads at a 0.8x bead-to-DNA ratio and resuspended from beads using a 10 mM TRIS buffer.

We attached Illumina index primers with an additional PCR step following standard protocols (Nextera XT Index Kit v2, Illumina, 2019). We combined duplicate samples for which both duplicates successfully amplified and diluted to a concentration of 5 nM. We multiplexed all samples with one negative control and two fungal clone positive controls (GenBank accession numbers: MG840195 and MG840196; Apigo & Oono, 2018; Clark et al., 2016; Toju et al., 2012). We submitted multiplexed samples for sequencing at the University of California, Santa Barbara Biological Nanostructures Laboratory Genetics Core. Samples were run on an

TABLE 1 Sample sizes for successfully extracted and PCR-amplified samples of surface sterilized and unsterilized *Heteropoda venatoria* individuals in the natural environment and feeding trial studies

Environment	Surface sterilized		Unsterilized	
	Extracted	Amplified and sequenced	Extracted	Amplified and sequenced
Natural environment	22	**18**	25	**19**
Feeding trial	10	**8**	14	**11**

Note: Bold numbers indicate final sample sizes for statistical analyses.

TABLE 2 Primers with Illumina overhang adapters (in bold) used to amplify the CO1 region in this study

Primer	Sequence (5′–3′)	Source
mlCOIintF	**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**GGWACWGGWTGAACWGTWTAYCCYCC	Yu et al. (2012)
Fol-degen-rev	**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**TANACYTCNGGRTGNCCRAARAAYCA	Leray et al. (2013)

Illumina MiSeq platform (v2 chemistry, 500 cycles, paired-end reads) with a 15% spike-in of PhiX. Following sequencing, samples were demultiplexed using Illumina's bcl2fastq conversion software (v2.20) at the Core facility. Our full protocol from DNA extraction through submission for Illumina sequencing can be found in Appendix S6.

2.6 | Sequence merging, filtering, and clustering with UNOISE3

We merged, filtered (max ee = 1.0), and denoised (clustered) our sequences around amplicon sequence variants (ASVs) using the UNOISE3 algorithm (unoise3 command in the open-source USEARCH 32-bit version 11.0.667; Edgar, 2016, Appendix S5, Figure S3). Prior to denoising with UNOISE3, we used cutadapt (version 1.18, Martin, 2011) to remove primers from each sequence. We also repeated analyses with the DADA2 algorithm run through R (dada2 package version 1.1.14.0; Callahan et al., 2016) and with a data cleaning step run through BBSplit (Bushnell, 2019) to remove consumer DNA prior to ASV assignment (because ASV assignment is abundance-sensitive). We considered analyses from the UNOISE3 algorithm only because UNOISE3 assigned more sequence reads to positive controls than DADA2 (on average, 3× as many reads per positive control) and the cleaning step paired with either DADA2 or UNOISE3 did not increase potential diet DNA detection (summary and comparisons in Appendices S1 and S2).

We created a list of unique ASVs and a matrix of ASV abundances across samples. We matched ASVs to taxonomies in the GenBank and BOLD databases. For GenBank, we used BLAST (version 2.7.1) with the blastn command for taxonomic assignment of each ASV using the computing cluster at UC Santa Barbara, comparing against the GenBank nucleotide database with an evalue of 0.01 (downloaded on 20 November 2019). We visualized and exported taxonomic alignment using MEGAN Community Edition (version 6.18.0, Huson et al., 2016), using default settings (LCA = naïve, MinScore = 50.0, MaxExpected = 0.01, TopPercent = 10.0, MinSupportPercent = 0.05) and selecting the subtree with all possible diet items for this species (Kingdom: Animalia, Clade: Bilateria). For taxonomies which were not assigned below the order level (n = 24), we submitted each ASV individually to the BLAST Basic Local Alignment Search Tool and assigned them a family based on the best sequence match in the database, given that the top ten database matches were from the same family. For BOLD taxonomic assignment, we used the BOLD IDEngine of the CO1 gene with Species Level Barcode Records (accessed 5–16 February 2020; 3,825,490 Sequences,

216,704 Species, and 95,537 Interim Species in database) to match each ASV list to taxonomies. We combined taxonomic assignments from both programs and discarded taxonomic assignments that were mismatched at the family level or higher (Elbrecht et al., 2017).

2.7 | Detection of potential diet items

For consumers from both natural environment and feeding trials, we asked whether surface sterilization altered detection of potential diet items for each consumer. For natural environment consumers, we examined all potential diet items (which could represent either diet or surface contaminants). For feeding trial consumers, we focused our detection analysis on the offered diet item we provided the consumers in the feeding trial environment (O. japonica, which all consumers were observed to have killed, but not necessarily ingested). We rarefied (McKnight et al., 2019, Appendix S5, Figure S4) based on the sample with the lowest sequencing depth which had been sequenced with 95%+ sampling completeness based on iNEXT (version 2.0.20) interpolation and extrapolation methods (Hsieh & Chao, 2017, 16,004 reads for natural environment and 55,205 reads for feeding trial consumers). We rarefied using the rrarefy() function in the vegan (version 2.5.6) package in R and rarefied the field and laboratory consumers separately.

We then selected all ASVs that matched potential diet items for the natural environment consumers (diet filtered to include all ASVs in the Kingdom: Animalia; Clade: Bilateria, excluding consumer DNA) and just the offered diet item for the feeding trial consumers (including species: *Oxya japonica*, genus: *Oxya*, and family: Acrididae, excluding those which only matched to order). Because the consumer species *H. venatoria* is the only species in the family Sparassidae on Palmyra Atoll, removing consumer DNA meant excluding all ASVs that received a family-level taxonomic assignment of "Sparassidae." As all ASVs received family-level taxonomic assignment, we pooled ASVs that matched at the family level into one taxonomic unit using cumulative read abundance (i.e., all ASVs matched to *diet family A* were pooled into *diet family A* taxonomic unit), a practice common in diet metabarcoding (Kartzinel et al., 2015) and predator–prey interaction (Brose et al., 2019) studies.

2.8 | Statistical analyses

For potential diet detection and rarefied abundance in both sets of consumers (natural environment and feeding trial), we used

generalized linear models to assess the effect of surface sterilization treatment. For prey detection, we used all potential (natural environment) or offered (feeding trial) diet item detection (presence–absence per sample) as the response variable in the full model with surface sterilization as a fixed effect and a binomial distribution. For rarefied diet abundance, we only assessed consumers for which we had detected diet and not those with no diet detection ($n = 33$ of 37 for natural environment; $n = 14$ of 19 for feeding trials). For this model, we treated the number of all potential (natural environment) or offered (feeding trials) diet DNA reads per sample as the response variable, surface sterilization treatment as a fixed effect, total read abundance of the sample (constant across all) as an offset term, and a Poisson or negative binomial distribution (to correct for overdispersion when needed). We assessed differences in per-sample potential diet richness among sterilization treatments for the natural environment consumers using generalized linear models with the number of potential diet items per sample as the response variable (both family-level taxonomic units or ASVs), surface sterilization treatment as the fixed effect and a Poisson or negative binomial distribution (to correct for overdispersion when needed). We assessed differences in potential diet item composition with family-level taxonomic units between surface sterilized and unsterilized consumers using a presence–absence PERMANOVA model fit with a binomial mixed effects model with surface sterilization treatment as a fixed effect, a random intercept term for potential diet item, and a random slope term for surface sterilization treatment. We also assessed ASV composition as a representation of potential prey composition using a canonical correspondence analysis (CCA) with surface sterilization as a predictor variable. We performed these analyses along with multiple other supplementary analyses and approaches, which can be found in the Supplementary Information (Appendices S4 and S5).

For all generalized linear models and mixed models, we performed model selection by comparing the full model (including the fixed effect of surface sterilization treatment) to a null model without this effect. All models were called in the glmmTMB package (version 1.0.0, Brooks et al., 2017) in R (version 3.6.1) We chose the best-fitting model based on size-corrected AIC values (MuMIn package version 1.43.15). For responses for which the best model included the surface sterilization treatment term, we examined the model summary to determine the standardized coefficients (β) and p-value of the significance between marginal means of the levels of the surface sterilization fixed effect. We assessed model fit using diagnostics in the DHARMa package (version 0.2.7), including tests for heteroscedasticity, and for count models (Poisson or negative binomial), zero inflation and overdispersion (Bolker et al., 2009; Zuur et al., 2009). We performed the CCA using the vegan package in R, comparing a model with surface sterilization as a fixed effect to a null model using an ANOVA. All raw data, data cleaning, and data analyses are available online (Miller-ter Kuile, 2020a, 2020b), and model outputs for primary and supplemental models can be found in Appendices S3 and S4.

3 | RESULTS

3.1 | PCR success, sequence merging, filtering, and clustering with UNOISE3 and DADA2

We successfully extracted DNA from 100% of samples ($n = 72$). Amplification success across all samples was 78%, with 56 of 72 initially extracted samples successfully amplified and were thus sequenced (Table 1). Seventy-three percent (128 of 176) of ASVs matched to a taxonomic assignment. Twenty-three percent of the total ASVs corresponded to potential diet items (41 of 176), and eight percent (14 of 176) corresponded to consumer DNA (the remaining 73 ASVs corresponded to nondiet items, including fungi, bacteria, and human DNA). Amplicon sequence variants that matched to the consumer comprised the majority of each sample ($98 \pm 0.6\%$ of rarefied abundance compared to $1.5 \pm 0.6\%$ for potential diet and $0.3 \pm 0.1\%$ for nondiet). Eighty-five percent of the potential diet ASVs received a species-level taxonomic assignment (35 of 41) from either BLAST or BOLD taxonomic assignments, and every potential diet species received a family-level and order-level taxonomic assignment. In MEGAN, the family-level assignments corresponded to 100% coverage results suggesting evidence of no mitochondrial pseudogenes (NUMTs) at the family level (Saitoh et al., 2016). There were no conflicting taxonomic assignments at the family level or higher between the BOLD and BLAST assignments.

3.2 | Detection of potential diet items

We detected potential diet in 89% (33 of 37) of natural environment consumers and the offered diet in 74% (14 of 19) of feeding trial consumers. For natural environment consumers, family-level taxonomic units corresponded to 20 families of potential diet items. The best model for potential diet detection in natural environment consumers was the null model that did not include surface sterilization treatment as a fixed effect (Figure 1, Appendix S4). For feeding trial consumers, one ASV matched to the offered diet (species: *O. japonica*, genus: *Oxya*, and family: Acrididae), and the best model for diet detection included the fixed effect of surface sterilization treatment, though the model without the surface sterilization term was within two AICc values (ΔAICc = 1.59) and the effect of the surface sterilization term was not statistically clear ($\beta = -2.3$; p-value = .07). We detected offered prey in 50% of consumers that had been surface sterilized compared to 91% of those consumers that were not surface sterilized.

3.3 | Proportion of potential diet DNA

For natural environment consumers, potential diet rarefied DNA sequence reads represented 2.0% ($\pm 1.0\%$) of total per-sample DNA sequence abundance (Figure 2). In feeding trial consumers, offered

(a)

(b)
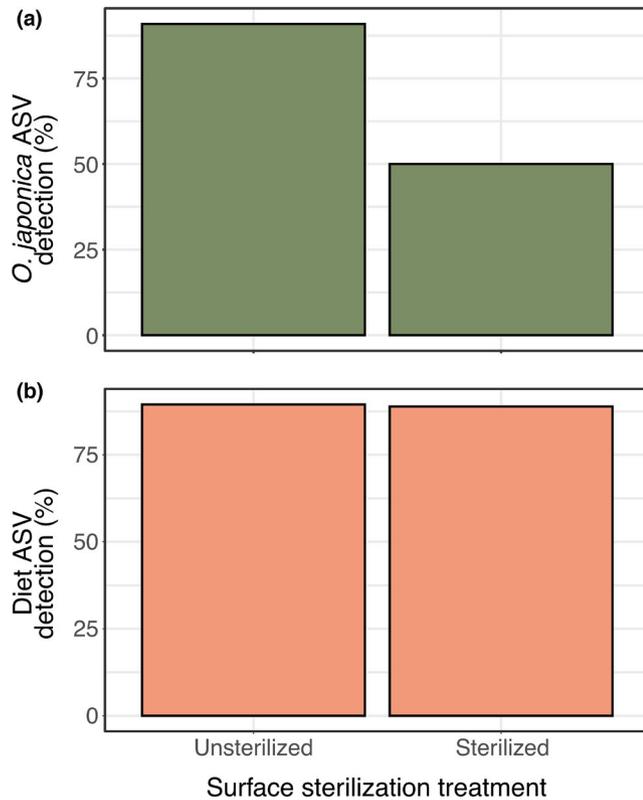
FIGURE 1 (a) Detection of all potential diet DNA in natural environment consumers that were and were not surface sterilized. Detection of diet DNA did not change with sterilization treatment. (b) Detection of offered diet (*Oxya japonica*) DNA in feeding trial consumers that were and were not surface sterilized. While the best-fitting model based on AICc values indicated an effect of surface sterilization treatment (a decrease from 91% without surface sterilization to 50% with surface sterilization), the effect of this term in the model was statistically unclear (*p*-value = .07)

diet DNA sequence reads represented 0.8% (±0.7% *SE*) of total per-sample DNA sequence abundance. For both natural environment and feeding trial consumers, the null models that did not include surface sterilization treatment as a fixed effect were the best models of diet DNA read abundance.

3.4 | Potential diet richness and composition in natural environment consumers

For family-level taxonomic units, potential diet richness per natural environment consumer was an average 2.08 (±0.26 *SE*) families per individual sample, with a maximum of 5 diet families in one consumer diet (Figure 3). Richness of potential diet ASVs for these consumers was similar, with an average of 2.32 (±0.31) potential diet ASVs per sample with a maximum of 7 ASVs in one consumer (Figure 3). The best models for per-sample potential diet richness for both family-level taxonomic units and ASV-level, as well as both family-level PERMANOVA and ASV-level CCA, were the null

(a)

(b)
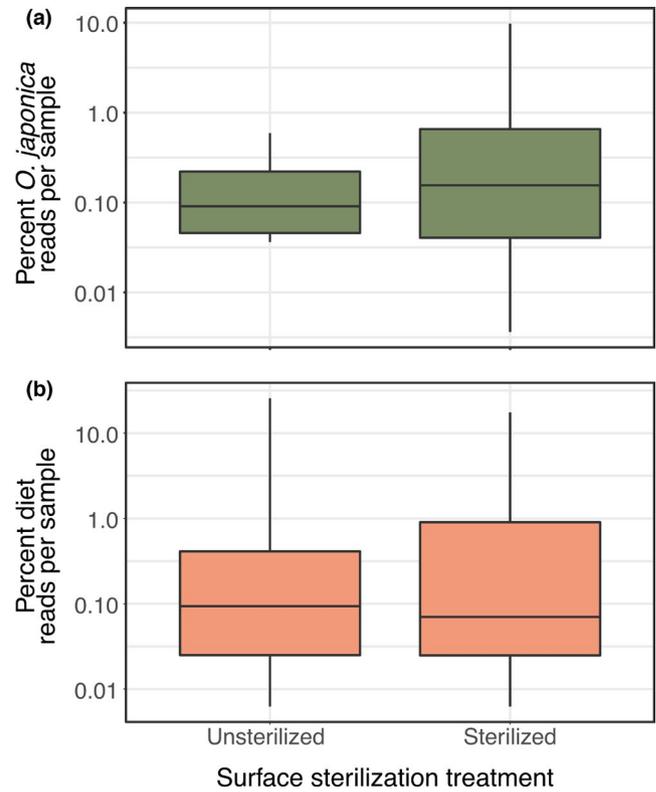
FIGURE 2 Neither the (a) proportion of total potential diet DNA in natural environment consumers or the (b) proportion of offered diet item DNA in feeding trial consumers significantly changed with surface sterilization treatment
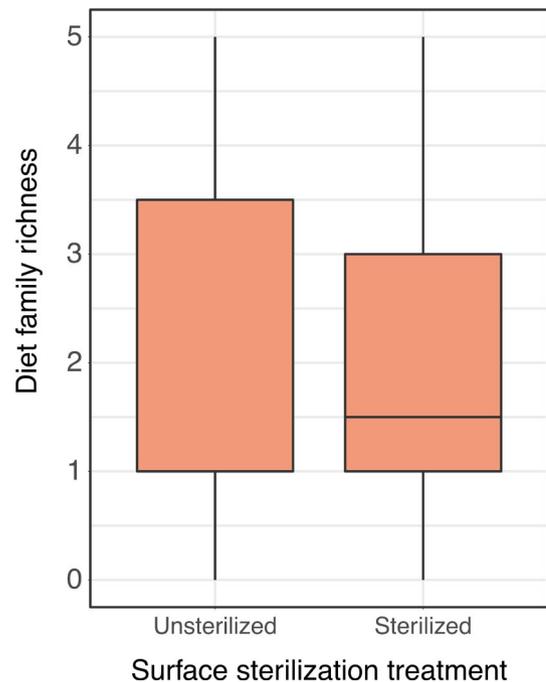
FIGURE 3 In natural environment consumers, surface sterilization did not alter per-sample diet richness of either family-level or ASV-level taxonomic units
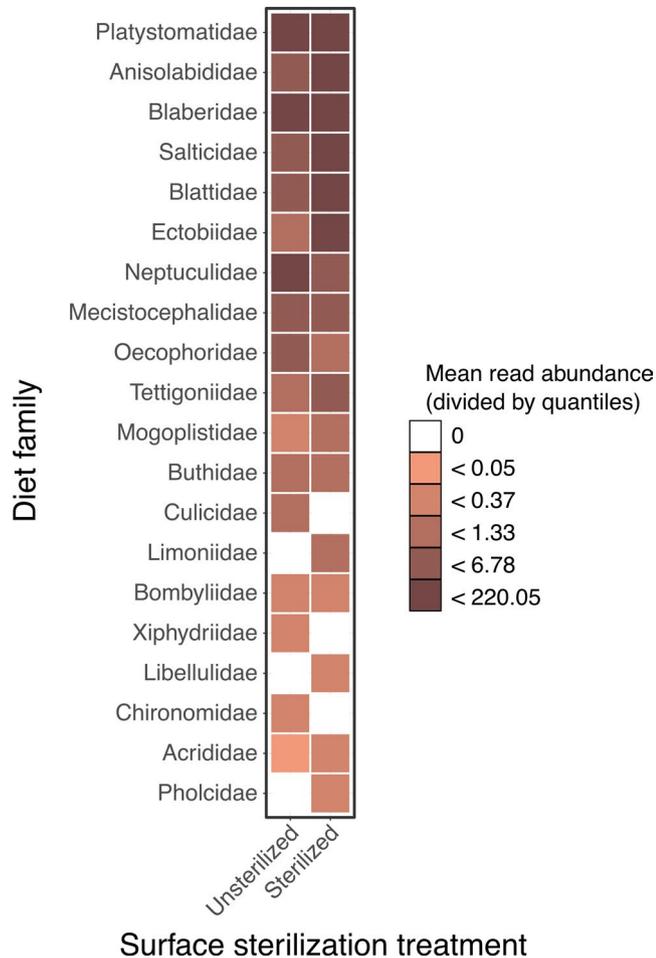
FIGURE 4 For natural environment consumers, surface sterilization did not alter the composition (either with a presence–absence of abundance model) of potential diet items of either family-level taxonomic units or ASV-level taxonomic units. In this figure of family-level taxonomic units by surface sterilization treatment, presence is indicated by a colored box and abundance is indicated by color depth (divided by quartiles due to wide variation in DNA sequence abundance)

metabarcoding. For potential diet richness and composition, in particular, these results did not change when considering potential diet in combined family-level taxonomic units (making them comparable with food web studies in this field, e.g., Brose et al., 2019) and when considering richness of molecular taxonomic units (ASVs). We detected diet across 84% of the total consumers in our study ($n = 47$ of 56), including 20 diet families. Diet DNA metabarcoding has high potential to contribute diet information for small consumers with cryptic feeding habits. Furthermore, it appears that current protocols that do not include surface sterilization steps are sufficient to determine potential diet for these consumers.

The field of diet DNA metabarcoding has not universally adopted surface sterilization practices into common protocols, in particular for studies including DNA extraction of full organisms or body parts without dissection (e.g., Jacobsen et al., 2018; Wirta et al., 2014). We demonstrate that surface sterilization does not seem necessary to avoid contamination effects. The evident lack of the effects of surface contaminants in our study contrasts with obvious surface contaminants that alter ecological interpretations in other fields using high-throughput sequencing to determine community diversity, particularly fungal endophyte studies (Burgdorf et al., 2014). One reason for this difference may be that fungal spores are widespread on and in the surfaces of most environments and organisms (Després et al., 2012) and likely to contaminate studies targeting specific subgroups of these communities. Indeed, even in our dataset, some sequences matched to fungal taxonomies. The fact that these nontarget sequences did not alter our DNA metabarcoding data by hiding target diet DNA, even with the relative rarity of diet DNA compared to consumer DNA (0.006%–26% of each sample), is likely due to differences in biomass of these sources of DNA in our samples and the specificity of our DNA size-selection protocol and PCR primers (Elbrecht et al., 2017; Krehenwinkel et al., 2017). Therefore, our results are promising both in validating the robustness of findings from past diet DNA studies that have not implemented surface sterilization treatments, but also highlight that diet DNA metabarcoding using broad, universal primer sets (e.g., those in this study) is an effective tool even when DNA sequence data contain potential environmental contaminants (Appendix S5, Figure S5).

While we saw no widespread support of the necessity for surface sterilization in our study, a model from the feeding trial that includes surface sterilization performed slightly better than one without this treatment (ΔAICc = 1.59). Thus, it is possible that contained environments may be more prone to contamination than open terrestrial environments. We see this result as an ideal starting point for next steps in validating diet DNA metabarcoding in similar contexts. Specifically, because this study had a relatively limited sample size ($n = 8$ and 11 in each sterilization treatment group) and because we did not confirm ingestion, a similar trial including crossed treatments of sterilization with different forms of diet item contact (e.g., Greenstone et al., 2012) would provide additional evidence of the effects of surface sterilization or surface contamination. Further exploration of these results might reveal that the decision to surface sterilize prior to diet DNA metabarcoding may matter more in some

models which did not include surface sterilization treatment as a predictor (Figure 4, Figure S1). Diet families came from insect, arachnid, and centipede orders (insects: Diptera (5), Dermaptera (1), Blattodea (3), Lepidoptera (3), Orthopotera (3), Hymenoptera (1), Odonata (1); Arachnids: Araneae (2); Scorpiones (1); and Centipedes: Geophilomorpha (1), Figure 4).

4 | DISCUSSION

Surface sterilization does not change diet measures in diet DNA metabarcoding data for the predatory consumer *H. venatoria* in either natural settings or a feeding trial environment, suggesting that surface sterilization is not a necessary step for this consumer. Our results suggest that various measures of diet, including potential diet detection, rarefied abundance, richness, and composition, are not significantly altered by surface sterilizing consumers prior to DNA

environments and experiments than others (e.g., where diet items are in high density or consumers have long handling times (Abrams & Ginzburg, 2000; Samu & Biro, 1993). Furthermore, as earlier studies targeting particular consumer diet pairs explored (e.g., Greenstone et al., 2012), the field of diet DNA metabarcoding is ripe for a comparison of surface sterilization techniques.

Diet DNA metabarcoding can empirically provide diet descriptions for a suite of consumers important to food web ecology and the maintenance of biodiversity on the planet (Stork, 2018).

Characterizing consumptive interactions for small, cryptic species for the first time will build a better picture of nature's complexity and allow ecologists to confidently query how species interactions will change with continued anthropogenic disturbance (Tylianakis et al., 2008). Like any method for determining consumptive interactions in nature, DNA metabarcoding continues to be refined, especially as tools and data emerge (Krehenwinkel et al., 2019; Kvist, 2013). This study builds on past efforts to refine diet DNA metabarcoding by using surface sterilization to pinpoint potential sources of error in diet DNA data. Here, we found that, on the whole, surface sterilization seems unnecessary in two contexts (terrestrial environments and contained feeding trials) when extracting DNA from body parts of invertebrate taxa. Continued context-specific refinement of surface sterilization protocols, and of other steps in diet DNA metabarcoding, will improve the widespread utility of diet DNA metabarcoding across consumer groups and environments.

CONFLICT OF INTEREST
None declared.

AUTHOR CONTRIBUTIONS
Ana Miller-ter Kuile: Conceptualization (equal); data curation (lead); formal analysis (lead); funding acquisition (equal); investigation (lead); methodology (equal); project administration (lead); visualization (lead); writing–original draft (lead); writing–review and editing (lead). **Austen Apigo:** Conceptualization (equal); formal analysis (supporting); investigation (supporting); methodology (equal); project administration (supporting); visualization (supporting); writing–original draft (supporting); writing–review and editing (supporting). **Hillary S. Young:** Conceptualization (supporting); formal analysis (supporting); funding acquisition (equal); investigation (supporting); project administration (supporting); supervision (lead); writing–original draft (supporting); writing–review and editing (supporting).

ORCID
Ana Miller-ter Kuile 🆔 https://orcid.org/0000-0003-2599-5158

REFERENCES
Abrams, P. A., & Ginzburg, L. R. (2000). The nature of predation: Prey dependent, ratio dependent or neither? _Trends in Ecology and Evolution_, _15_(8), 337–341. https://doi.org/10.1016/S0169-5347(00)01908-X

Anslan, S., Bahram, M., & Tedersoo, L. (2016). Temporal changes in fungal communities associated with guts and appendages of Collembola as based on culturing and high-throughput sequencing. _Soil Biology and Biochemistry_, _96_, 152–159. https://doi.org/10.1016/j.soilbio.2016.02.006

Apigo, A., & Oono, R. (2018). _MG840195 and MG840196_.

Baker, R., Buckland, A., & Sheaves, M. (2014). Fish gut content analysis: Robust measures of diet composition. _Fish and Fisheries_, _15_(1), 170–177. https://doi.org/10.1111/faf.12026

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. _Trends in Ecology and Evolution_, _24_(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., & Bolker, B. M. (2017). Modeling zero-inflated count data with glmmTMB. _BioRxiv_, _132753_, https://doi.org/10.1101/132753

Brose, U., Archambault, P., Barnes, A. D., Bersier, L.-F., Boy, T., Canning-Clode, J., Conti, E., Dias, M., Digel, C., Dissanayake, A., Flores, A. A. V., Fussmann, K., Gauzens, B., Gray, C., Häussler, J., Hirt, M. R., Jacob, U., Jochum, M., Kéfi, S., … Iles, A. C. (2019). Predator traits determine food-web architecture across ecosystems. _Nature Ecology and Evolution_, _3_(6), 919–927. https://doi.org/10.1038/s41559-019-0899-x

Brown, S. P., Leopold, D. R., & Busby, P. E. (2018). Protocols for investigating the leaf mycobiome using high-throughput DNA sequenci. In _Plant pathogenic fungi and oomycetes: Methods and protocols_ (Vol. _1848_). https://doi.org/10.1007/978-1-4939-8724-5

Burgdorf, R. J., Laing, M. D., Morris, C. D., & Jamal-Ally, S. F. (2014). A procedure to evaluate the efficiency of surface sterilization methods in culture-independent fungal endophyte studies. _Brazilian Journal of Microbiology_, _45_(3), 977–983. https://doi.org/10.1590/S1517-83822014000300030

Bushnell, B. (2019). *BBMap*.

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

Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2016). GenBank. *Nucleic Acids Research*, 44(D1), D67–D72. https://doi.org/10.1093/nar/gkv1276

Delmas, E., Besson, M., Brice, M.-H., Burkle, L. A., Dalla Riva, G. V., Fortin, M.-J., Gravel, D., Guimarães, P. R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., & Poisot, T. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, 94(1), 16–36. https://doi.org/10.1111/brv.12433

Després, V. R., Huffman, J. A., Burrows, S. M., Hoose, C., Safatov, A. S., Buryak, G., Fröhlich-Nowoisky, J., Elbert, W., Andreae, M. O., Pöschl, U., & Jaenicke, R. (2012). Primary biological aerosol particles in the atmosphere: A review. *Tellus, Series B: Chemical and Physical Meteorology*, 64(1), 15598. https://doi.org/10.3402/tellusb.v64i0.15598

Doña, J., Proctor, H., Serrano, D., Johnson, K. P., van Oploo, A. O., Huguet-Tapia, J. C., Ascunce, M. S. & Jovani, R. (2019). Feather mites play a role in cleaning host feathers: New insights from DNA metabarcoding and microscopy. *Molecular Ecology*, 28(2), 203–218. https://doi.org/10.1111/mec.14581

Edgar, R. C. (2016). UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv*, 081257. https://doi.org/10.1101/081257

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, 7(17), 6918–6926. https://doi.org/10.1002/ece3.3192

Fulton, T. M., Chunwongse, J., & Tanksley, S. D. (1995). Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter*, 13(3), 207–209. https://doi.org/10.1007/BF02670897

Greenstone, M. H., Weber, D. C., Coudron, T. C., & Payton, M. E. (2011). Unnecessary roughness? Testing the hypothesis that predators destined for molecular gut-content analysis must be hand-collected to avoid cross-contamination. *Molecular Ecology Resources*, 11(2), 286–293. https://doi.org/10.1111/j.1755-0998.2010.02922.x

Greenstone, M. H., Weber, D. C., Coudron, T. A., Payton, M. E., & Hu, J. S. (2012). Removing external DNA contamination from arthropod predators destined for molecular gut-content analysis. *Molecular Ecology Resources*, 12(3), 464–469. https://doi.org/10.1111/j.1755-0998.2012.03112.x

Handler, A., Gruner, D., Haines, W., Lange, M., & Kaneshiro, K. (2007). Arthropod surveys on Palmyra Atoll, Line Islands, and insights into the decline of the native tree *Pisonia grandis* (Nyctaginaceae). *Pacific Science*, 61(4), 485–502. https://doi.org/10.2984/1534-6188(2007)61

Hooper, D. U., Chapin, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J. H., Lodge, D. M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A. J., Vandermeer, J., & Wardle, D. A. (2005). Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs*, 75(1), 3–35. https://doi.org/10.1890/04-0922

Hsieh, T. C., & Chao, A. (2017). Rarefaction and extrapolation: Making fair comparison of abundance-sensitive phylogenetic diversity among multiple assemblages. *Systematic Biology*, 66(1), 100–111. https://doi.org/10.1093/sysbio/syw073

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Tappu, R. (2016). MEGAN community edition - Interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Computational Biology*, 12(6), 1–12. https://doi.org/10.1371/journal.pcbi.1004957

Illumina, (2019). Illumina. *Nextera XT DNA Library Prep Reference Guide.*, (May), Document # 15031942 v05.

Jacobsen, R. M., Sverdrup-Thygeson, A., Kauserud, H., & Birkemoe, T. (2018). Revealing hidden insect-fungus interactions; moderately specialized, modular and anti-nested detritivore networks. *Proceedings. Biological Sciences*, 285(1876), 20172833. https://doi.org/10.1098/rspb.2017.2833

Kartzinel, T. R., Chen, P. A., Coverdale, T. C., Erickson, D. L., Kress, W. J., Kuzmina, M. L., Rubenstein, D. I., Wang, W., & Pringle, R. M. (2015). DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proceedings of the National Academy of Sciences*, 112(26), 8019–8024. https://doi.org/10.1073/pnas.1503283112

Krehenwinkel, H., Kennedy, S. R., Adams, S. A., Stephenson, G. T., Roy, K., & Gillespie, R. G. (2019). Multiplex PCR targeting lineage-specific SNPs: A highly efficient and simple approach to block out predator sequences in molecular gut content analysis. *Methods in Ecology and Evolution*, 10(7), 982–993. https://doi.org/10.1111/2041-210X.13183

Krehenwinkel, H., Kennedy, S., Pekár, S., & Gillespie, R. G. (2017). A cost-efficient and simple protocol to enrich prey DNA from extractions of predatory arthropods for large-scale gut content analysis by Illumina sequencing. *Methods in Ecology and Evolution*, 8(1), 126–134. https://doi.org/10.1111/2041-210X.12647

Kvist, S. (2013). Barcoding in the dark?: A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. *Molecular Phylogenetics and Evolution*, 69(1), 39–45. https://doi.org/10.1016/j.ympev.2013.05.012

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & 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), 1–14. https://doi.org/10.1186/1742-9994-10-34

Linville, J. G., & Wells, J. D. (2002). Surface sterilization of a maggot using bleach does not interfere with mitochondrial DNA analysis of crop contents. *Journal of Forensic Sciences*, 47(5), 15532J. https://doi.org/10.1520/jfs15532j

Lucas, A., Bodger, O., Brosi, B. J., Ford, C. R., Forman, D. W., Greig, C., Hegarty, M., Jones, L., Neyland, P. J., & De Vere, N. (2018). Floral resource partitioning by individuals within generalised hoverfly pollination networks revealed by DNA metabarcoding. *Scientific Reports*, 8(1), 1–11. https://doi.org/10.1038/s41598-018-23103-0

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet Journal*, 17(1), 10–12. https://doi.org/10.14806/ej.17.1.200

McKnight, D. T., Huerlimann, R., Bower, D. S., Schwarzkopf, L., Alford, R. A., & Zenger, K. R. (2019). Methods for normalizing microbiome data: An ecological perspective. *Methods in Ecology and Evolution*, 10(3), 389–400. https://doi.org/10.1111/2041-210X.13115

Miller-ter Kuile, A. (2020a). *BioProject: PRJNA639981*.

Miller-ter Kuile, A., Apigo, A., & Young, H. (2020b). Diet DNA metabarcoding data from spiders (Heteropoda venatoria) from Palmyra Atoll (2015-2017) with both individual samples that have and have not been surface sterilized. https://doi.org/10.5061/dryad.gqnk98snc

Nielsen, J. M., Clare, E. L., Hayden, B., Brett, M. T., & Kratina, P. (2018). Diet tracing in ecology: Method comparison and selection. *Methods in Ecology and Evolution*, 9(2), 278–291. https://doi.org/10.1111/2041-210X.12869

Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: Diet assessment using next generation sequencing. *Molecular Ecology*, 21(8), 1931–1950. https://doi.org/10.1111/j.1365-294X.2011.05403.x

Quéméré, E., Hibert, F., Miquel, C., Lhuillier, E., Rasolondraibe, E., Champeau, J., Rabarivola, C., Nusbaumer, L., Chatelain, C., Gautier, L., Ranirison, P., Crouau-Roy, B., Taberlet, P., & Chikhi, L. (2013). A DNA metabarcoding study of a primate dietary diversity and

10 | **WILEY**_Ecology and Evolution MILLER-TER KUILE ET AL.

Open Access

plasticity across its entire fragmented range. *PLoS One*, 8(3), https://doi.org/10.1371/journal.pone.0058971

Saitoh, S., Aoyama, H., Fujii, S., Sunagawa, H., Nagahama, H., Akutsu, M., Shinzato, N., Kaneko, N., & Nakamori, T. (2016). A quantitative protocol for DNA metabarcoding of springtails (Collembola). *Genome*, 59(9), 705–723. https://doi.org/10.1139/gen-2015-0228

Samu, F., & Biro, Z. (1993). Functional response, multiple feeding and wasteful killing in a wolf spider (Araneae: Lycosidae). *European Journal of Entomology*, 90, 471–476.

Schnell, I. B., Thomsen, P. F., Wilkinson, N., Rasmussen, M., Jensen, L. R. D., Willerslev, E., Bertelsen, M. F., & 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

Schulz, B., Wanke, U., Draeger, S., & Aust, H. J. (1993). Endophytes from herbaceous plants and shrubs: Effectiveness of surface sterilization methods. *Mycological Research*, 97(12), 1447–1450. https://doi.org/10.1016/S0953-7562(09)80215-3

Sheppard, S. K., & Harwood, J. D. (2005). Advances in molecular ecology: Tracking trophic links through predator-prey food-webs. *Functional Ecology*, 19(5), 751–762. https://doi.org/10.1111/j.1365-2435.2005.01041.x

Soininen, E. M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., Gussarova, G., Bellemain, E., Hassel, K., Stenøien, H. K., Epp, L., Schrøder-Nielsen, A., Brochmann, C., & Yoccoz, N. G. (2015). Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. *PLoS One*, 10(1), 1–18. https://doi.org/10.1371/journal.pone.0115335

Stork, N. E. (2018). How many species of insects and other terrestrial arthropods are there on earth? *Annual Review of Entomology*, 63, 31–45. https://doi.org/10.1146/annurev-ento-020117-043348

Toju, H., & Baba, Y. G. (2018). DNA metabarcoding of spiders, insects, and springtails for exploring potential linkage between above- and below-ground food webs. *Zoological Letters*, 4(1), 1–12. https://doi.org/10.1186/s40851-018-0088-9

Toju, H., Tanabe, A. S., Yamamoto, S., & Sato, H. (2012). High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PLoS One*, 7(7). https://doi.org/10.1371/journal.pone.0040863

Tylianakis, J., Didham, R., Bascompte, J., & Wardle, D. (2008). Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, 11, 1351–1363. https://doi.org/10.1111/j.1461-0248.2008.01250.x

Wilson, E. O. (1987). The little things that run the world (The importance and conservation of invertebrates). *Conservation Biology*, 1(4), 344–346.

Wirta, H. K., Hebert, P. D. N., Kaartinen, R., Prosser, S. W., Várkonyi, G., & Roslin, T. (2014). Complementary molecular information changes our perception of food web structure. *Proceedings of the National Academy of Sciences*, 111(5), 1885–1890. https://doi.org/10.1073/pnas.1316990111

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3(4), 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

Zimmerman, N. B., & Vitousek, P. M. (2012). Fungal endophyte communities reflect environmental structuring across a Hawaiian landscape. *Proceedings of the National Academy of Sciences of the United States of America*, 109(32), 13022–13027. https://doi.org/10.1073/pnas.1209872109

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, Anatoly, A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R. Mixed effects models and extensions in ecology with R* (Vol. 53). https://doi.org/10.1017/CBO9781107415324.004

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.