**Title:** Effects of surface sterilization on diet DNA metabarcoding data of terrestrial invertebrate consumers in natural environments and feeding trials

**Running title:** Surface sterilization in diet metabarcoding

**Authors:** Ana Miller-ter Kuile1, Austen Apigo1, Hillary S. Young1

**Author Affiliations:**

Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106

**Corresponding Author:** Ana Miller-ter Kuile

**Abstract**

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, and it is unclear if surface contamination from body parts alters presumed diet presence or diversity. We examine 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 individuals surface sterilized in 10% commercial bleach solution followed by deionized water with a set of unsterilized individuals. We found that surface sterilization did not significantly alter any measure of diet for either consumers from a natural environment or in feeding trials. The best-fitting model predicting diet detection in feeding trial consumers included surface sterilization, but this term was not statistically significant (ΔAICc = 1.59 compared to the null model, β = -2.3, p-value = 0.07). 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 in 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, invertebrates, contamination, food web, predator prey interactions, diet analysis

**Introduction**

Biological communities and ecosystem function are shaped by interactions between organisms (J. H. Brown et al., 2001; Hooper et al., 2005; Schleuning et al., 2015). Among the many interaction types, consumptive interactions (including herbivory, predation, and parasitism) can shape the stability of biologically diverse communities (Delmas et al., 2019; Ings et al., 2009). Until recently, these consumptive interactions were most often measured by visual observations of feeding or by gut dissection or inspection of fecal contents (Baker et al., 2014; Duffy & Jackson, 1986; Hyslop, 1980; 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, which includes 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 methods for determining gut contents is one of the most promising emerging approaches. 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 intra-population, 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), 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 the 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 (e.g. fungal endophyte research; 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 (S. P. Brown et al., 2018; J. Hallmann et al., 1997), the field of diet DNA metabarcoding has not developed a similarly systematic approach (e.g. ethanol washes in Doña et al., 2019, bleach washes in Anslan et al., 2016, no sterilization in 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 destruction of DNA 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 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) 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 influences potential 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 1) detection, 2) rarefied abundance, 3) richness, and 4) composition of potential diet items. We then performed a laboratory feeding trial in which we fed consumer individuals a specific diet item, comparing surface sterilized individuals to those which were not surface sterilized to ask whether surface sterilization influenced 1) detection or 2) rarefied abundance of offered diet items. Exploring these questions in both natural and contained settings help address whether surface contamination could alter ecological interpretations of community-scale species interactions and whether surface sterilization needs to be incorporated into standard protocols in this field.

**Materials and Methods**

*Field site and collections*

We conducted this work on Palmyra Atoll National Wildlife Refuge, Northern Line Islands (5º53’ N, 162º05’W). consumer and in DNA metabarcoding dataWe targeted a generalist, active hunting spider species (*Heteropoda venatoria*) as the consumer for this project because a) it occurs in high abundance on the atoll and is therefore easy to collect, b) it is a generalist species that feeds on a wide suite of other organisms (including insects, spiders, other invertebrates, and two vertebrates, geckos in the genus *Lepidodactylus*), c) due to its small size and arachnid feeding habits, there are no viable non-genetic methods of diet analysis, and d) it is the only species in its family on the atoll, meaning consumer DNA can be differentiated from potential diet DNA. We collected consumer individuals during two summers. In 2015, we collected individuals in natural habitats across the atoll. containerAll individuals were collected individually in sterilized collection containers to avoid contamination (Greenstone et al., 2011).

*Natural environment consumer collection*

In 2015, we collected consumers (n = 47) in natural environments in order to test whether DNA metabarcoding would detect potential diet DNA from consumers which fed on available diet items and came into contact with natural environmental surfaces. We collected and froze all individuals in separate containers at -80°C immediately following collection until surface sterilization and DNA extraction in 2019 (see below).

*Feeding trial*

laboratory. environmentscontainer and acontainter, containerscontainerscontainers

*Both natural environment and feeding trial consumers: surface sterilization*

Because we planned to extract DNA from entire body parts (opisthosomas) of consumer individuals (following methods from Krehenwinkel et al., 2017; Macías-Hernández et al., 2018), we wanted to determine whether surface sterilization of *H. venatoria* consumer individuals altered common diet DNA measures (detection, abundance, richness, and composition). We used a surface sterilization treatment to remove possible contaminants from some consumer individuals from both the natural environment and feeding trials while leaving some individuals unsterilized. We used surface sterilization techniques common in other fields of molecular ecology (i.e. plant endophytes Burgdorf et al., 2014; Schulz et al., 1993) by submerging and stirring each full consumer in 10% commercial bleach by volume (0.5% sodium hypochlorite) for 2 minutes and then washing each consumer by submerging and stirring in deionized water for 2 minutes. Similar or longer periods of bleach washing at equal or greater concentrations have led to undetectable DNA degradation in similar soft-exoskeleton consumers (e.g. maggots and beetle nymphs; Greenstone et al., 2012; Linville & Wells, 2002). Natural environment consumers (2015) had been frozen at -80ºC since collection; these consumers were surface sterilized following the bleach wash protocol 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) following the bleach wash protocol in the lab on the atoll in 2017 following freezing at -20°C and then stored each sample in individual vials of 95% ethanol in a -20ºC freezer until DNA extraction because no -80ºC freezer was available at the field station that year (n = 10 surface sterilized; n = 14 not surface sterilized). Prior to DNA extraction, all samples from both 2015 and 2017 were allowed to dry for 1-3 hours in a sterilized laminar flow hood and the opisthosoma was removed from every consumer individual for DNA extraction using a sterilized scalpel in a sterilized laminar flow hood. For all sterilization steps, forceps, scalpels, and laboratory surfaces were sterilized with either ethanol and flame (scalpels and forceps) or 10% bleach (surfaces) between handling each individual.

*DNA extraction and removal of consumer DNA with Ampure XP beads*

We extracted DNA from each *H. venatoria* consumer opisthosoma following a modified CTAB extraction protocol (Fulton et al., 1995). At least twenty-four hours following extraction, we quantified DNA using a Qubit (Invitrogen) fluorometer with the high sensitivity double-stranded DNA quantification kit using 1L of DNA template per reaction. We used methods developed by Krehenwinkel et al., (2017) to isolate a proportion of lower molecular weight DNA with Ampure XP beads prior to PCR. Ampure XP beads preferentially bind to heavier molecules of more intact consumer DNA, leaving the smaller fragments of presumed semi-digested diet DNA in the supernatant (Appendix E, Figure 1). Thus, by keeping the supernatant, we aimed to work with a sample that had a larger proportion of lower molecular weight diet DNA after removing consumer DNA that bound to beads (Krehenwinkel et al., 2017). To do this, we diluted each DNA sample to 20ng/l (creating a total sample volume of 40L), mixed each sample using Ampure XP beads (0.75x bead-to-DNA ratio), and kept the supernatant from this step. With the supernatant, we repeated the CTAB protocol steps for precipitating DNA pellets with isopropanol and 5M potassium acetate and cleaned DNA pellets with ethanol washes (Appendix F). After at least another twenty-four hours, we quantified DNA again using a Qubit fluorometer (following the same methods as above) and diluted all samples to 10ng/L prior to PCR steps. All DNA pellets were stored in and diluted with TE buffer.

*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). The CO1 gene is well-represented in the GenBank sequencing database (Porter & Hajibabaei, 2018). We performed all PCR preparation steps in a UV-sterilized biosafety cabinet. We used a standard desalted primer set tested by Krehenwinkel et al., (2017) for use in diet analyses of invertebrate predatory consumers, including spiders (Table 2). These primers included overhang adapters compatible with the Illumina indexing PCR (Illumina 2009) that immediately followed CO1 amplification.

We amplified the CO1 gene in each sample by PCR in a 25L reaction volume that included 9L nuclease free water, 12.5L GoTaq Green Master Mix (Promega Corp.), 1.25 L of each of the primers (at 10mM), and 1 L of DNA template (at 10ng/L). When DNA concentrations were lower than 10ng/L, we added more DNA to the sample to equal 10ng of total template and reduced the amount of water added. Each sample was run in duplicate until after Illumina indexing PCR, and we ran a duplicated negative sample each PCR run. We ran each reaction with an initial denaturation step at 95°C for 3 minutes, and then 35 cycles of: 1) denaturation at 95°C for 30 seconds, 2) annealing at 46°C for 30 seconds, and 3) elongation at 72°C for one minute. We ended each PCR run with a final elongation step at 72°C for 5 minutes and then held samples at 4°C until placed in a 4°C refrigerator. To remove reaction dimer before attaching Illumina P5/P7 indices, we removed lower molecular weight amplicons (~200 bp) with Ampure XP beads at a 0.8x bead-to-DNA ratio. Samples were re-suspended from beads using a 10mM TRIS resuspension buffer.

We then attached Illumina index primers with an additional PCR step (Nextera XT Index Kit v2). Each total reaction volume was again 25L, with 5L of nuclease free water, 12.5L GoTaq Green Master Mix, 1.25L of each primer (at 10mM), and 5L of PCR product. These were run in a standard PCR protocol for these primers: an initial denaturing step at 95°C for 3 minutes, followed by 10 cycles of: 1) denaturation at 95°C for 30 seconds, 2) annealing at 55°C for 30 seconds, and 3) elongation at 72°C for 30 seconds. We ended each run with a final elongation step at 72°C for 5 minutes and then held samples at 4°C until placed in a 4°C refrigerator.

We verified PCR amplification by visualizing 3-4 L of each PCR product in a 1.5% agarose gel using GelRed (Biotium) at 100V for 30-40 minutes. Gels were visualized with a Bio-Rad Gel Doc XR+ imager using Image Lab 5.0. We kept samples for which both duplicates successfully amplified during the PCR steps and only kept samples on PCR runs in which both negative control duplicates resulted in no product detection by gel electrophoresis. For successful samples, we combined duplicates and mixed with an Ampure XP bead-to-DNA ratio of 0.7x. We determined the average length of the gene region using an Agilent TapeStation with a D1000 ScreenTape System following the standard protocol from the quick start guide. We then quantified these final PCR products using a Qubit fluorometer and a high sensitivity kit with 1L of sample per reaction tube and diluted each sample in 10mM TRIS to a final concentration of 5nM.

We multiplexed all samples along with one negative control and two PCR4-TOPO TA vectors (Invitrogen, Carlsbad, CA, USA) containing the internal transcribed spacer 1 region from two fungal species as 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 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 F.

*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). This ASV denoising approach incorporates sequence abundance, quality, and error rates to cluster reads in high throughput sequencing data into a smaller subset of biological units (Appendix E, Figure 3). 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 chose to consider analyses from the UNOISE3 algorithm only because UNOISE3 assigned more sequence reads to positive controls than DADA2 (on average, 3x 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 A and B).

From the output, we created a list of unique ASVs and a matrix of ASV abundances across samples. We matched ASVs to taxonomies both 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 November 20, 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 February 5-16, 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).

*Detection of potential diet items*

For consumers from both the natural environment and feeding trials, we wanted to know whether surface sterilization altered the detection of potential diet items for each consumer. 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). Samples were rarefied (McKnight et al., 2019) because sequencing depth, or the total number of DNA sequences assigned per sample, can vary considerably in high throughput sequencing runs (Appendix E, Figure 4). We rarefied 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). We rarefied using the rrarefy() function in the vegan (version 2.5.6) package in R. We rarefied to 16,004 reads per sample for the natural environment and 55,205 reads per sample for the feeding trial consumers. We rarefied these separately because samples had been preserved in different ways and for different times, which can have large effects on DNA extraction outcomes (Murphy et al., 2002).

Following rarefying, we selected all ASVs that matched all potential diet items for the natural environment consumers (Kingdom: Animalia; Clade: Bilateria, excluding consumer DNA) and just the offered diet item for the feeding trial consumers (including species: *Oxya japonica*, genus: *Oxya*, family: Acrididae, excluding those which only matched to order: Orthoptera with no lower taxonomies). 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”. All ASVs received a family-level taxonomic assignment, so we chose to combine ASVs at the family level. Furthermore, because family-level taxonomic assignments are common in other diet metabarcoding studies (Kartzinel et al., 2015) and in the field of predator-prey interactions more broadly (Brose et al., 2019), this level allows diet data to be most comparable to studies across environments. We combined family-level taxonomic units by combining ASVs that matched at the family level into one combined taxonomic unit with cumulative read abundance (i.e. all ASVs matched to *diet family A* were combined into one *diet family A* taxonomic unit with cumulative read abundance).

*Abundance of potential diet DNA*

To test whether surface sterilization altered the abundance of DNA representing all potential (natural environment) or offered (feeding trial) diet items, we assessed per sample offered or potential diet DNA abundance for both sets of consumers (natural environment and feeding trial) separately. For this analysis, we used only consumer individuals for which we detected potential or offered diet DNA (33 of 37 for natural environment; n = 14 out of 19 for feeding trials), to test whether contaminants altered potential diet abundance only when potential diet DNA is present.

*Potential diet richness and composition in natural environment consumers*

In addition to allowing detection of potential diet items, DNA metabarcoding also enables the analysis of potential diet communities, allowing explorations of individual-, population-, and species-level diet richness and composition. If surface contaminants alter these metrics, ecological interpretation of these community-level data could be misleading, either by increasing the richness of consumptive interactions attributed to a consumer, or by hiding interactions that occur more rarely or further back in time (e.g. Macías-Hernández et al., 2018; MacKenzie & Kendall, 2002). We determined the potential diet richness and composition in natural environment consumers. For per sample potential diet richness and composition, we performed analyses used both taxonomic units combined at the family level (described above) and with the original number and composition of ASVs matched to potential diet items.

*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 potential prey detection, all potential (natural environment) or offered (feeding trial) diet item detection (presence-absence per sample) was the response variable in the full model with surface sterilization treatment as a fixed effect and a binomial distribution. For diet abundance, we treated the number of all potential (natural environment) or offered (feeding trials; *O. japonica*) 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 also examined the abundance of other potential diet items for the feeding trial consumers as well as DNA which was sequenced but was not diet (e.g. fungi and potential endoparasites) with results in the Supplemental Information (Appendix E, Figure 5). 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. Incorporating a random intercept term for potential diet item combined with a random slope term for surface sterilization treatment allows the effect of surface sterilization treatment to vary by potential diet item, such that some potential diet items may increase in presence with surface sterilization (i.e. hidden by contaminants), while others may decrease in presence (i.e. potential diet item is a contaminant; Zuur et al., 2009). 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 repeated the natural environment consumer potential diet item PERMANOVA with abundance data (Poisson distribution), conducted both presence-absence and abundance based PERMANOVA analyses on all potential diet items (including offered item) for feeding trial consumers, and repeated each analysis using the adonis() function from the vegan package (version 2.5.6) in R (dist = “jaccard” with binary = TRUE for presence/absence and dist = “bray” for abundance; Appendix D, Appendix E, Figures 6-8).

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 (, or the degree of change in the response with every unit change in the predictor variables, with positive or negative values depending on the response direction) 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 heteroskedasticity, and for count models (Poisson or negative binomial), zero inflation and overdispersion (Bolker et al., 2009; Zuur et al., 2009). We performed the CCA analysis 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, 2020b, 2020a), and model outputs for primary and supplemental models can be found in Appendices C and D.

**Results**

*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 sequenced (natural environment: n = 18 surface sterilized, n = 19 unsterilized, feeding trials: n = 8 surface sterilized, n = 11 unsterilized; Table 1). The Illumina MiSeq run yielded 33,332,804 unpaired reads and had a Q30 quality score of 78.03%. After quality filtering and denoising with UNOISE3, 8,029,959 paired-end reads corresponded to 176 ASVs. 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 non-diet items, including fungi, bacteria, and human DNA). ASVs that matched to the consumer comprised the majority of each sample (98 ± 0.6% of rarefied abundance compared to 1.5 ± 0.6% for potential diet and 0.3 ± 0.1% for non-diet). Eighty-five percent of the potential diet ASVs received a species-level taxonomic assignment (35 of 41) from either the 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 family-level assignments corresponded to 100% coverage results (LCA parameters: MinScore = 100, MaxExpected = 0.01, TopPercent = 10.0, MinSupportPercent = 0.05, LCA = naïve) 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.

*Detection of potential diet items*

We detected potential diet in 89% (33 of 37) of natural environment consumers and offered diet (*O. japonica*) 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 D). For feeding trial consumers, one ASV matched to the offered diet (Species: *O. japonica*, Genus: *Oxya*, 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 surface sterilization term was not statistically significant in the full model at an α = 0.05 (surface sterilization parameters: β = -2.3; p-value = 0.07). We detected *O. japonica* in 50% of consumers that had been surface sterilized compared to 91% of those consumers that were not surface sterilized.

*Proportion of potential diet DNA*

For natural environment consumers, potential diet rarefied DNA sequence reads represented 2.0% (± 1.0%) of total per-sample DNA sequence abundance (Figure 2). In feeding trial consumers, offered diet rarefied DNA sequence reads (*O. japonica*) represented 0.8% (± 0.7% SE) of total per-sample DNA sequence abundance. For both the 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.

*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) 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 the family-level taxonomic units and ASV-level were the null models which did not include surface sterilization treatment as a fixed effect. The best models for potential diet composition for family-level taxonomic units (mixed model PERMANOVA) and ASV-level taxonomic units (CCA) also did not include surface sterilization treatment as fixed effects (Figure 4, Supplementary Figure 1).

**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 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 for which diet has been challenging to determine because of small consumer size or 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 in most or all of these types of studies to avoid contamination effects. The evident lack of the effects of surface contaminants in our study contrast 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 (Colston & Jackson, 2016; Després et al., 2012; Philippot et al., 2013) 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 non-target 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, similar to other studies; Krehenwinkel et al., 2017), is likely due to the 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 E, Figure 5).

While we saw no widespread support of the necessity for surface sterilization in our study, because our model results from the feeding trial environment suggested a model including the effect of surface sterilization on diet detection performed slightly better than one without this effect (ΔAICc = 1.59), there is potential that more contained environments may be more prone to contamination than open terrestrial environments. We see this result as an ideal starting point for the next steps in validating diet DNA metabarcoding data in similar contexts. Specifically, because this study had a relatively low sample size (n = 8 and 11 in each treatment group) and because we did not confirm ingestion, repetition of a similar feeding trial including crossed treatments of surface sterilization with different forms of potential diet item contact (e.g. prey ingested, no prey offered, prey contact on outside but no ingestion, similar to Greenstone et al., 2012) would provide additional evidence of the effects of surface sterilization or surface contamination in this more contained environmental context. Further exploration of these results might reveal that the decision to surface sterilize prior to diet DNA metabarcoding may matter more in some environments and experiments than others (e.g. where diet items are in high density or consumers have long handling times; Abrams & Ginzburg, 2000; Jeschke et al., 2002; Samu & Biro, 1993; Scharf et al., 1998). Furthermore, as earlier studies in molecular diet methods 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. This current study was not designed to look for the negative effects of bleach sterilization, for example; thus future work should explicitly explore the potential negative effects of surface sterilization treatments on DNA degradation versus removal due to physical or chemical treatments. Nor was the current study an examination of whether primers that target various gene region lengths (e.g. shorter gene-length arthropod primer sets in Krehenwinkel et al., 2017; Thomsen & Sigsgaard, 2019; Zeale et al., 2011) nor methods used to block consumer DNA (Krehenwinkel et al., 2019) may reveal nuances not detected in this current study.

Diet DNA metabarcoding is providing the first glimpse at comprehensive diet for a suite of consumers important to the field of food web ecology and to the maintenance of biodiversity on the planet (C. A. Hallmann et al., 2017; Nielsen et al., 2018; Wilson, 1987). Being able to determine consumptive interactions for many species and environments for the first time will continue to build a better picture of the complex structure of nature and how natural systems will change with anthropogenic disturbance (Brophy et al., 2017; Dunne et al., 2002; Harvey et al., 2017; Ives et al., 2005; Pilosof et al., 2017; Rudolf, V. H W, Lafferty, 2011; Tylianakis et al., 2008). Like any method for determining consumptive interactions in nature, DNA metabarcoding continues to be refined, especially as new tools and data are continually emerging (Gómez-Rodríguez et al., 2017; Jain et al., 2015; Kvist, 2013; Zinger et al., 2019). This study builds on past efforts to refine the field of 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 to be unnecessary in two contexts (terrestrial environments and feeding trial containers) when extracting DNA from body parts of invertebrate taxa. However, continued context-specific refinements of surface sterilization protocols, along with other steps in diet DNA metabarcoding studies, will continue to increase the validity and widespread utility of diet DNA metabarcoding across consumer groups and environments.

**Acknowledgements**

Field work for this project was funded by the National Science Foundation (DEB #1457371), National Geographic Society, and a Faculty Research Grant from the UC Santa Barbara Academic Senate. We would like to thank field technicians Carina Motta and Michelle Lee for help collecting samples for this project and laboratory technicians Emily Lutz and Tessa Chou for helping prepare genetic samples. We would like to thank the U.S. Fish and Wildlife Service and Palmyra Atoll Research Consortium for supporting field work for this project. We would like to thank Dr. Ryoko Oono for use of her laboratory space and equipment. We acknowledge the use of the Biological Nanostructures Laboratory within the California NanoSystems Institute, supported by the University of California (UC) Santa Barbara and the University of California Office of the President. We especially thank Dr. Jennifer Smith, manager of the Biological Nanostructures Laboratory for her assistance in preparing and troubleshooting our samples. We acknowledge the use of computational facilities at the Center for Scientific Computing (CSC), which was purchased with funds from the National Science Foundation (CNS-1725797) and is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara. We thank D. Orr, E. Forbes, H. Lowman, A. Bui, D. Preston, D. Trovillion, E. Crone, E. Sauer, L. Falke, B. DiFiore, C. Jerde, M. Lee, and R. Ramiro for help in aspects of this manuscript, including statistics, bioinformatics, framing, and editing. We thank XX anonymous reviewers for help revising this manuscript. This is publication number PARC-XXX from the Palmyra Atoll Research Consortium.

**Reference**

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., Skaug, H. J., Maechler, M., & Bolker, B. M. (2017). Modeling Zero-Inflated Count Data With glmmTMB. *BioRxiv*, 132753. https://doi.org/10.1101/132753

Brophy, C., Dooley, Á., Kirwan, L., Finn, J. A., McDonnell, J., Bell, T., Cadotte, M. W., & Connolly, J. (2017). Biodiversity and ecosystem function: making sense of numerous species interactions in multi-species communities. *Ecology*, *98*(7), 1771–1778. https://doi.org/10.1002/ecy.1872

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, J. H., Whitham, T. G., Morgan Ernest, S. K., & Gehring, C. A. (2001). Complex species interactions and the dynamics of ecological systems: Long-term experiments. *Science*, *293*(5530), 643–650. https://doi.org/10.1126/science.293.5530.643

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

Colston, T. J., & Jackson, C. R. (2016). Microbiome evolution along divergent branches of the vertebrate tree of life: what is known and unknown. *Molecular Ecology*, *25*(16), 3776–3800. https://doi.org/10.1111/mec.13730

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., Alex Huffman, J., 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). https://doi.org/10.3402/tellusb.v64i0.15598

Doña, J., Proctor, H., Serrano, D., Johnson, K. P., Oploo, A. O. van, 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

Duffy, D. C., & Jackson, S. (1986). Diet Studies of Seabirds: A Review of Methods. *Colonial Waterbirds*, *9*(1), 1–17.

Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Network structure and biodiversity loss in food webs: robustness increase with connectance. *Ecology Letters*, *5*, 558–567. https://doi.org/10.1046/j.1461-0248.2002.00354.x

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

Gómez-Rodríguez, C., Timmermans, M. J. T. N., Crampton-Platt, A., & Vogler, A. P. (2017). Intraspecific genetic variation in complex assemblages from mitochondrial metagenomics: comparison with DNA barcodes. *Methods in Ecology and Evolution*, *8*(2), 248–256. https://doi.org/10.1111/2041-210X.12667

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

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

Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H., Stenmans, W., Müller, A., Sumser, H., Hörren, T., Goulson, D., & De Kroon, H. (2017). More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE*, *12*(10). https://doi.org/10.1371/journal.pone.0185809

Hallmann, J., Quadt-Hallmann, A., Mahaffee, W. F., & Kloepper, J. W. (1997). Bacterial endophytes in agricultural crops. *Canadian Journal of Microbiology*, *43*(10), 895–914. https://doi.org/10.1139/m97-131

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

Harvey, E., Gounand, I., Ward, C. L., & Altermatt, F. (2017). Bridging ecology and conservation: from ecological networks to ecosystem function. *Journal of Applied Ecology*, *54*, 371–379. https://doi.org/10.1111/1365-2664.12769

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., Setala, 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., Ruscheweyh, H. J., & 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

Hyslop, E. J. (1980). Stomach contents analysis - a review of methods and their application. *Journal of Fish Biology*, *17*(4), 411–429. https://doi.org/10.1111/j.1095-8649.1980.tb02775.x

Ings, T. C., Montoya, J. M., Bascompte, J., Bluthgren, N., Brown, L., Dormann, C., Edwards, F., Figueroa, D., Jacob, U., Jones, J. I., Lauridsen, R. B., Ledger, M. E., Lewis, H. M., Olesen, J. M., Van Veen, F. J. F., Warren, P. H., & Woodward, G. (2009). Ecological networks - Beyond food webs. *Journal of Animal Ecology*, *78*(1), 253–269. https://doi.org/10.1111/j.1365-2656.2008.01460.x

Ives, A. R., Cardinale, B. J., & Snyder, W. E. (2005). A synthesis of subdisciplines: Predator-prey interactions, and biodiversity and ecosystem functioning. *Ecology Letters*, *8*(1), 102–116. https://doi.org/10.1111/j.1461-0248.2004.00698.x

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

Jain, M., Fiddes, I. T., Miga, K. H., Olsen, H. E., Paten, B., & Akeson, M. (2015). Improved data analysis for the MinION nanopore sequencer. *Nature Methods*, *12*(4), 351–356. https://doi.org/10.1038/nmeth.3290

Jeschke, J. M., Kopp, M., & Tollrian, R. (2002). Predator functional responses: Discriminating between handling and digesting prey. *Ecological Monographs*, *72*(1), 95–112. https://doi.org/10.1890/0012-9615(2002)072[0095:PFRDBH]2.0.CO;2

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., 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

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

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*(1), 1. 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

Macías-Hernández, N., Athey, K., Tonzo, V., Wangensteen, O. S., Arnedo, M., & Harwood, J. D. file:///Users/Ana/Downloads/j. 136.-2427. 2005. 01398. x. pd. (2018). Molecular gut content analysis of different spider body parts. *PLoS ONE*, *13*(5), 1–16. https://doi.org/10.1371/journal.pone.0196589

MacKenzie, D. I., & Kendall, W. C. (2002). How should detection probability be incorporated into estimates of relative abundance? *Ecology*, *83*(9), 2387–2393. https://doi.org/10.1890/0012-9658(2002)083[2387:HSDPBI]2.0.CO;2

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBNet Journal*, *17*(1).

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*. NCBI BioProject Database.

Miller-ter Kuile, A. (2020b). *DNA\_Diet\_Methods*. Github Repository.

Murphy, M. A., Waits, L. P., Kendall, K. C., Wasser, S. K., Higbee, J. A., & Bogden, R. (2002). An evaluation of long-term preservation methods for brown bear (Ursus arctos) faecal DNA samples. *Conservation Genetics*, *3*, 435–440.

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

Philippot, L., Raaijmakers, J. M., Lemanceau, P., & Van Der Putten, W. H. (2013). Going back to the roots: The microbial ecology of the rhizosphere. *Nature Reviews Microbiology*, *11*(11), 789–799. https://doi.org/10.1038/nrmicro3109

Pilosof, S., Porter, M. A., Pascual, M., & Kéfi, S. (2017). The multilayer nature of ecological networks. *Nature Publishing Group*, *1*(March), 1–9. https://doi.org/10.1038/s41559-017-0101

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

Porter, T. M., & Hajibabaei, M. (2018). Over 2.5 million COI sequences in GenBank and growing. *PLoS ONE*, *13*(9), 1–16. https://doi.org/10.1371/journal.pone.0200177

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

Rudolf, V. H W, Lafferty, K. D. (2011). Stage structure alters how complexity affects stability of ecological networks. *Ecology Letters*, *14*, 75–79. https://doi.org/10.1111/j.1461-0248.2010.01558.x

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.

Scharf, F. S., Buckel, J. A., Juanes, F., & Conover, D. O. (1998). Predation by juvenile piscivorous bluefish (*Pomatomus saltatrix*): the influence of prey to predator size ratio and prey type on predator capture success and prey profitability. *Canadian Journal of Fisheries and Aquatic Sciences*, *55*(7), 1695–1703. https://doi.org/10.1139/cjfas-55-7-1695

Schleuning, M., Fründ, J., & García, D. (2015). Predicting ecosystem functions from biodiversity and mutualistic networks: An extension of trait-based concepts to plant-animal interactions. *Ecography*, *38*(4), 380–392. https://doi.org/10.1111/ecog.00983

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

Thomsen, P. F., & Sigsgaard, E. E. (2019). Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. *Ecology and Evolution*, *9*(4), 1665–1679. https://doi.org/10.1002/ece3.4809

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., Varkonyi, 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

Zeale, M. R. K., Butlin, R. K., Barker, G. L. A., Lees, D. C., & Jones, G. (2011). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, *11*(2), 236–244. https://doi.org/10.1111/j.1755-0998.2010.02920.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

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., Chariton, A. A., Creer, S., Coissac, E., Deagle, B. E., De Barba, M., Dickie, I. A., Dumbrell, A. J., Ficetola, G. F., Fierer, N., Fumagalli, L., Gilbert, M. T. P., Jarman, S., Jumpponen, A., … Taberlet, P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857–1862. https://doi.org/10.1111/mec.15060

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. In *Mixed Effects Models and Extensions in Ecology with R* (Vol. 53, Issue 9). https://doi.org/10.1017/CBO9781107415324.004

**Data Accessibility**

Raw sequence data are available on GenBank (reviewer link: https://dataview.ncbi.nlm.nih.gov/object/PRJNA639981?reviewer=2k2u1qmrtehqbsmc05vdqivkor) and will be made publicly available following acceptance of this manuscript.

Data and analyses are currently available in a GitHub repository (https://github.com/anamtk/DNA\_Diet\_Methods.git) and will be uploaded to Dryad after acceptance of this manuscript.

**Author Contributions**

AM-tK, AA, and HY conceived the idea for this study. AM-tK collected field samples and conducted mesocosm study. AM-tK and AA designed laboratory analyses for this study. AM-tK performed all lab processing and data analyses for the study. AA and HY provided feedback on data analysis methods. AM-tK led the writing of the manuscript. All authors contributed to editing of the manuscript.

**Figures and Tables**

Table 1: Sample sizes for successfully extracted and PCR amplified samples of surface sterilized and unsterilized *H. venatoria* individuals in the natural environment and feeding trial studies. Bold numbers indicate final sample sizes for statistical analyses.

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

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

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

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 non-significant at a cutoff of α = 0.05 (p-value = 0.07).

****

Figure 2: Neither the a) proportion of total potential diet DNA in natural environment consumers or 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 either of 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). We observed similar results for ASV-level composition (Supplementary Figure).

****

**Supplemental Information:**

Appendix A: Comparisons between UNOISE3 and DADA2 with and without BBSplit.

Appendix B: BBSplit methods.

Appendix C: Model outputs for GLMMs.

Appendix D: Model outputs from supplementary data analyses.

Appendix E: Supplemental figures on DNA extraction and amplification protocols, as well as additional figures of diet composition and diversity for natural-environment and feeding trial consumers.

Appendix F: Full DNA extraction and PCR protocol