Using surface environmental DNA to assess arthropod biodiversity within a forested ecosystem

Michael Allen[1], Julie Lockwood[1], Robert Kwait[1], Anthony Vastano[1], Donnie Peterson[1], Leon Tkacenko[1], Jordan Angle[2], and Benjamin Jaffe[3]

[1]Rutgers The State University of New Jersey
[2]ExxonMobil Upstream Research Company
[3]ExxonMobil Biomedical Sciences Inc

February 22, 2024

Abstract

Terrestrial arthropods are abundant and diverse with outsized ecological and economic importance. Our ability to monitor this diversity is hampered by the variety of sampling techniques and taxonomic expertise required to catalog the species in an area. DNA metabarcoding approaches show promise but have mainly been limited to trapping studies where DNA is extracted from captured individuals. Here we illustrate the promise of terrestrial plant surfaces as reservoirs of environmental DNA (eDNA) that is rich in arthropod biodiversity information. We posit that collection of surface eDNA will enable easier and more rapid arthropod inventories. We collected 40 paired samples using two novel terrestrial surface eDNA sampling techniques – 'roller' tree bark and 'spray' foliage sampling – in a New Jersey, USA pine barrens forest. Metabarcoding using two primer sets (COI and 16S) revealed the presence of 177 arthropod families (from 21 orders), representing 80% of the family-level diversity expected in the area based on accumulation curves. Spray samples revealed more families than roller (148 vs. 126), while the two methods showed distinct, though overlapping, community composition. The two primer sets revealed similar alpha diversity, although they also captured different taxonomic subsets. A more limited comparison of roller and spray sampling with traditional aquatic and soil eDNA samples revealed a greater family diversity in surface samples, especially compared with soil. Our study highlights the value of eDNA metabarcoding surveys for achieving the elusive goal of rapid, cost-effective arthropod inventories, and thus realizing a range of ecological research and management goals.

Title: Using surface environmental DNA to assess arthropod biodiversity within a forested ecosystem

Running title: Arthropod monitoring via surface eDNA

Michael C. Allen[1*], Julie L. Lockwood[1], Robert Kwait[1], Anthony R. Vastano[1], Donnie L. Peterson[1,2], Leon A. Tkacenko[1], Jordan Angle[3] and Benjamin Jaffe[4]

[1]Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ 08902 USA

[2]Current address: Southern Swedish Forest Research Centre, Swedish University of Agricultural Sciences, Alnarp, Sweden

[3]ExxonMobil Upstream Research Company, Spring, Texas

[4]ExxonMobil Biomedical Sciences Inc., Annandale, New Jersey

*Corresponding author: michael.allen@rutgers.edu

ABSTRACT

Terrestrial arthropods are abundant and diverse with outsized ecological and economic importance. Our ability to monitor this diversity is hampered by the variety of sampling techniques and taxonomic expertise required to catalog the species in an area. DNA metabarcoding approaches show promise but have mainly been limited to trapping studies where DNA is extracted from captured individuals. Here we illustrate the promise of terrestrial plant surfaces as reservoirs of environmental DNA (eDNA) that is rich in arthropod biodiversity information. We posit that collection of surface eDNA will enable easier and more rapid arthropod inventories. We collected 40 paired samples using two novel terrestrial surface eDNA sampling techniques – 'roller' tree bark and 'spray' foliage sampling – in a New Jersey, USA pine barrens forest. Metabarcoding using two primer sets (COI and 16S) revealed the presence of 177 arthropod families (from 21 orders), representing 80% of the family-level diversity expected in the area based on accumulation curves. Spray samples revealed more families than roller (148 vs. 126), while the two methods showed distinct, though overlapping, community composition. The two primer sets revealed similar alpha diversity, although they also captured different taxonomic subsets. A more limited comparison of roller and spray sampling with traditional aquatic and soil eDNA samples revealed a greater family diversity in surface samples, especially compared with soil. Our study highlights the value of eDNA metabarcoding surveys for achieving the elusive goal of rapid, cost-effective arthropod inventories, and thus realizing a range of ecological research and management goals.

Keywords: monitoring, metabarcoding, COI, ITS, eDNA, insects

INTRODUCTION

Arthropods are the most abundant and diverse non-microbial organisms on Earth, yet comprehensive information on patterns of richness, endemism and biogeography are lacking (Beng et al., 2016). Arthropods also provide key ecosystem services (e.g., pollination) and can be indicators of ecosystem health (Schulze et al., 2004), thus the ability to monitor their response to management and restoration actions is critical to realizing sustainability goals (Dangles & Casas, 2019; van der Heyde et al., 2022a). Comprehensive surveys of arthropod biodiversity within or across habitats, and through time, have been difficult to execute given the range of sampling methods necessary to capture more than one taxonomic group and the expertise required to identify each individual captured (Porter et al., 2019; Zenker et al., 2019). While sufficient field sampling is still a rate-limiting step for arthropod biodiversity surveys (Porter et al., 2019), the use of DNA metabarcoding on samples of collected arthropods is increasingly viewed as an inexpensive, accurate, and efficient way to identify and categorize taxonomic groups (Ärje et al., 2020). Arthropods, even within a single habitat, such as temperate forests, are hyper-diverse, highly cryptic, and span multiple orders of magnitude in size (Beng et al., 2016). Even when using metabarcoding tools, the suite of species that can be assessed is constrained by the field methods needed to target and collect individuals (e.g., Beng et al., 2016; Porter et al., 2019). The use of sources of DNA left behind by organisms as they move through the environment (eDNA) removes the need for physical, destructive capture of individuals in arthropod surveys, thus representing a substantial advance in arthropod biodiversity collection and assessments (Belle et al., 2019; Thomsen & Sisgaard, 2018; Porter et al., 2019). However, eDNA surveys of arthropods have thus far seen very limited use, primarily via sampling eDNA from soils and water (Belle et al., 2019; van der Heyde et al., 2022b; Porter et al., 2019). Here we provide

evidence that surface eDNA sampling techniques can provide a comprehensive assessment of aboveground forest arthropod biodiversity, and that plant surfaces are more efficient sampling substrates for eDNA surveys than either soil or water.

In the last decade, environmental DNA-based methods have become a widespread and powerful suite of tools to detect the presence of species and to characterize communities within marine and freshwater ecosystems (Beng & Corlett, 2020). The process of collecting and analyzing eDNA using metabarcoding techniques to inventory species within aquatic ecosystems is consistently as cost effective, or more so, than encounter-based sampling with morphology-based species identification (e.g., Balint et al., 2018). eDNA metabarcoding can also circumvent sampling issues associated with cryptic species (e.g., mosquitos; Boerlijst et al., 2019) or habitats that are dangerous or expensive to visit (e.g., deep ocean ecosystems; Closek et al., 2019). These successes have spurred research into novel methods for eDNA biodiversity surveys of terrestrial habitats targeting diverse taxa such as insects (Thomsen & Sisgaard, 2018), mammals (Leempoel et al., 2020), reptiles (Kyle et al., 2022), and birds (Ushio et al., 2018). However, arthropods, especially insects, are vastly under-represented in this body of research (Belle et al., 2019; Jinbo et al., 2022).

Although bulk environmental samples such as soil and water harbor eDNA from soil-dwelling and aquatic arthropods, respectively, they may not contain the DNA from species of aboveground terrestrial habitats such as forest canopy, tree bark, and understory vegetation (Marquina et al., 2019; Oliverio et al., 2018). A solution is to sample above-ground surfaces that arthropods use for feeding or cover, which are likely to accumulate DNA deposited by individuals via their normal activities shedding scales, excrement, exuvia, and saliva (Valentin et al., 2020). Surface eDNA collection methods have been used to survey for targeted pest insects

within forested and agricultural settings, providing substantial boosts in detection probability (Allen et al., 2021; Valentin et al., 2018, 2020), but no such study has employed a metabarcoding approach, which could prove especially useful for biodiversity inventories. Here, we use vegetation surface (tree bark and foliage) eDNA collection techniques and metabarcoding to characterize arthropod diversity in a forested ecosystem and evaluate the sampling effort needed to maximize the number of arthropod taxa identified. We also compare the performance of surface eDNA techniques in capturing arthropod diversity to more established techniques that rely on sampling of arthropod eDNA within bulk substrates such as soil or waterbodies.

MATERIALS AND METHODS

Our study area was a ~ 1.4 ha area of upland pine-deciduous forest within the Colliers Mills Wildlife Management Area, New Jersey, USA. The site lies within the Pinelands National Reserve, the largest expanse of Atlantic coastal pine barrens remaining in North America (> 5000 km^2) and a unique ecosystem classified as a United Nations International Biosphere Reserve due to its unusual fire-adapted flora and fauna, including numerous rare or threatened plants, vertebrates, and arthropods (Boyd, 1991).

Our methods consisted of performing a 'head-to-head' trial of two recently-developed terrestrial eDNA sampling methods – targeting bark and foliage surfaces, respectively (Valentin et al., 2020) – combined with metabarcoding with two primer sets to evaluate these methods as a means of monitoring forest arthropod biodiversity. We also performed a more limited study comparing the performance of the same primers across differing forest substrates, including vegetation surfaces, water, and soil.

Primary study: comparison of two vegetation surface eDNA methods

The primary field study consisted of collecting paired eDNA samples, from tree bark and understory foliage, at 20 sites within a ~ 30 m wide strip of forest along the southeastern bank of Colliers Lake (north site boundary latitude and longitude: 40.0709, -74.4478; south boundary: 40.0686, -74.4475). Sites were 20 m^2 in area, separated by a minimum of 3.5 m, and were centered around clusters of deciduous (*Acer*, *Quercus*, *Nyssa*) or coniferous (*Pinus*, *Juniperus*) trees with surrounding understory vegetation. Collection methods at each site followed Valentin et al. (2020) and included collecting a single pooled tree bark ('roller') sample from one tree per site and a single foliage ('spray') sample from the leaves of understory plants within 3 m of that tree. Samples for the primary study were collected on 29 September 2021.

Tree bark, or 'roller', samples consisted of applying deionized (DI) water to moisten a commercial paint roller and then, with gentle pressure, moving it around the entire surface of the tree bark from the base to a height of ~ 2 m (Valentin et al., 2020; see Appendix S1). At each site, we used the roller to sample one tree > 25 cm in diameter, either deciduous (n = 10 sites) or coniferous (n = 10 sites), following recommended practices to avoid field sample contamination (Valentin et al., 2020; Appendix S1). After sampling, we placed the roller into a sterile bag, added deionized (DI) water until ~ 30-40% of the roller was submerged (~ 130 mL), and massaged the roller for 15 s within the bag to dislodge and suspend DNA into the water (Peterson et al., 2022). Finally, we removed the roller from the bag, and filtered the remaining water through a 5 μm polycarbonate track etched (PCTE) membrane filter housed in a 47 mm plastic filter holder (Whatman Swin-Lok, Cytiva, Marlborough, Massachusetts, USA) using a peristaltic field pump. After all water was filtered, or the filter clogged, we removed the filter from the plastic holder with clean forceps and placed it into a 1.5 ml tube of 100% non-denatured ethanol for transportation to the lab.

Foliage, or 'spray', sampling involved spraying understory vegetation with pressurized DI water using an 18 L backpack sprayer (Allen et al., 2021; Valentin et al., 2020; Appendix S1). We sprayed the foliage evenly as we walked slowly through the site, keeping the nozzle at a distance of ~ 30-50 cm, and collecting as much of the residual water as possible in a 2 L sterilized bucket (~ 300-400 ml per sample), following best practices to avoid field sample contamination (Appendix S1). At each site, we sprayed a ~ 10 m^2 patch of understory vegetation, or roughly 4-8 shrubs and saplings of ~ 1 m in height. We then filtered the collected spray water and processed samples as above.

Substrate comparison study

The smaller 'substrate comparison' study involved collecting samples in the same forest as the primary study, but at fewer sites (n = 5) and with four, not two, collection methods: the two surface methods (roller and spray), plus soil and water. We collected samples for this study on 21 July 2020 (roller, spray, and soil samples) and 28 August 2020 (water samples).

Roller and spray samples were collected using the same sampling procedures and processing methods as the primary study except that each roller sample was collected from three trees (i.e., as one pooled sample) instead of one. One soil sample was collected at each site in which a 50 ml Falcon tube was filled with surface soil (< 1 cm depth) and placed in a cooler at ambient temperature until transported back to the lab within 1-2 h for storage in a -80 ℃ freezer. Only two water samples were collected, one each at the first two sites sampled. Each water sample was collected ~ 2-3 m from the pond edge adjacent to the site using a 3.5 m aluminum grab sampling pole equipped with 1 L PETG bottles. Filled bottles were placed in a cooler and brought back to the lab within 1-2 hours where the water was immediately filtered and processed

as described above. Further details for field equipment and contamination prevention procedures are described in Appendix S1.

DNA extraction and sequencing

We performed all extractions in a dedicated lab free of PCR products. We used a vacuum centrifuge to evaporate ethanol from all filter samples (i.e., roller, spray, and water samples) immediately prior to extraction using the Qiagen DNeasy Blood and Tissue Kit (Qiagen Sciences Inc., Maryland, USA). Most roller samples from the primary study (17 of 20) were suspected to contain PCR inhibitors following extraction as they failed to amplify during the first PCR step (see below). We removed inhibitors from these samples by performing a bead clean-up using Ampure XP magnetic beads (Beckman Coulter Inc., New Jersey, USA) followed by inhibitor removal (Zymo Research, Irvine, California, USA) following manufacturers' instructions. Soil samples were extracted using the MoBio PowerSoil kit (Molecular Biosciences, California, USA). This kit allows for the processing of up to 10 g of soil per sample and includes steps to remove PCR inhibitors and impurities commonly found in soil. One extraction negative was constructed for each batch of extracted samples resulting in 4 extraction negatives for the primary study and 3 for the substrate comparison study. These negatives consisted of PCR-grade water in place of sample and extracted alongside the other samples. A similar negative control was included for each batch of samples run during the inhibitor removal process (n = 3, for primary study only). Following extraction, DNA was quantified using the Qubit High Sensitivity DNA assay (Life Technologies Corporation, California, USA).

Libraries were prepared using the Illumina (2013) two-step PCR metabarcoding protocol. Arthropod DNA was amplified using two different primer sets. The first set, 'ZBJ-Art', amplifies a ~ 160 bp fragment of the COI region and was designed to detect a broad range of

arthropods (Zeale et al., 2011). The second set, 'Coleop_16S', amplifies a ~ 110 bp fragment of the 16S region (Epp et al., 2012). We chose this second set to complement and broaden taxonomic coverage, notably for the beetles (Coleoptera). Hereafter, we refer to these primer sets as 'Zeale' and 'Coleop', respectively.

We performed two rounds of PCR for each primer set: the first to add adapter sequences and the second to add sample specific indices (Illumina, 2013). Three technical replicates per sample were included for each PCR in the primary study, while only one was included in the substrate comparison study. Reaction concentrations for the initial PCR followed Illumina (2013) and had the following cycle parameters: 95℃ for 3 minutes; 35 cycles of 30 s at 95 ℃, 30 s at 57 ℃ (Zeale) or 60 ℃ (Coleop), 1 min at 72 ℃; 5 min at 72 ℃; and hold at 4 ℃. We then pooled the amplicons from each primer set in equimolar ratios for each sample and performed a bead clean-up using Ampure XP magnetic beads. For the second PCR, reaction preparation and cycle parameters followed the Illumina (2013) protocol exactly. Libraries were then purified with a second magnetic bead cleanup and again pooled in equimolar ratios. We included one PCR negative for each PCR plate that was run, and negatives from each primer specific PCR were pooled, resulting in 8 total PCR negative samples for the primary study and 2 for the substrate comparison study.

Bioinformatics

Sequencing was performed at the Princeton University Genomics core using an Illumina MiSeq. Filtering and taxon assignments were made using Cutadapt, OBITools and EcoPCR following OBITools documentation and others (Boyer et al., 2016; Leempoel et al., 2020). First, we aligned forward and reverse reads using OBITools *illuminapairedend* and removed all unpaired reads and reads with join scores < 40. We then removed adapter sequences and bases with

quality scores < 30 using Cutadapt. We assigned reads to samples and separated based on primer pair using the OBITools command *ngsfilter*, removed all unassigned reads, and dereplicated using *obiuniq*. We then removed all sequences having a sequence count of ≤ 30 (or ≤ 10 for the substrate comparison study) or that were < 80 bp in length and used *obiclean* to filter all reads for PCR or sequencing errors (see Table S1).

To create reference databases, we downloaded all invertebrate sequences available from Ensembl (Cunningham et al., 2022) and used EcoPCR to simulate a PCR on the constructed database for each primer pair. We used EcoPCR to assign molecular operational taxonomic units (MOTUs) to sequence reads, only retaining those that we could associate to a known taxon with greater than 95% probability. For all sequences not matched to the Ensembl database, we assigned taxa manually using BLAST (Johnson et al., 2008), only retaining MOTUs with matches above 90%. Following taxon assignment, we further controlled for contamination by removing (subtracting) any counts that appeared in negatives (i.e., extraction, clean-up, PCR, or field equipment negatives; see Appendix S1) from the samples.

Statistical analysis

We used the same procedures, with minor variations, to analyze data from the primary and substrate comparison field studies. We describe the basic steps for both here, noting where the two differ as appropriate.

To compare the sampling efficiency and taxonomic breadth captured by the two primer sets, we used read data from all MOTUs identified as arthropods, performing separate analyses for the primary and substrate comparison study. We excluded one sample in the primary study with low sequencing depth (195 total reads). We then rarefied the primary and substrate comparison datasets to match the sample with the lowest read count (3660 and 770 total reads,

222 respectively) and calculated the rarefied MOTU richness, Simpson evenness index, and Shannon
223 diversity index for each sample and primer set using the R package *vegan* (v. 2.5-7; Oksanen et
224 al., 2020; R Core Team, 2021). Means and 95% confidence intervals (± 2 SE) of these indices
225 were then calculated for each study and primer set for comparison.

226 To compare the arthropod communities captured by the various eDNA collection
227 methods (and around coniferous vs. deciduous trees for the primary study), we retained all
228 arthropod MOTUs that were identified to the family level, the lowest taxonomic classification
229 where we felt that all MOTUs could be reliably identified. We pooled family-level detection and
230 non-detection data from the Zeale and Coleop primers into a single data set, considering a family
231 as 'detected' in a sample if ≥ 1 MOTU from it had > 0 reads by either primer set. We evaluated
232 family-level compositional differences between groups (sampling method or tree type) using two
233 approaches. First, we created heat trees (Foster et al., 2017) that visualize the taxonomic relations
234 among families detected by the various sampling methods, as well as the overall prevalence of
235 families within samples (i.e., % of sites). Second, we used non-metric multidimensional scaling
236 (NMDS; *metaMDS* function) and PERMANOVA (*adonis* function) in *vegan* to statistically
237 evaluate compositional differences in family presence-absence between sample collection
238 methods, and, for the primary study, between coniferous and deciduous samples. We evaluated
239 sampling efficiency of the various eDNA collection methods for adequately capturing family-
240 level arthropod diversity at our study site using accumulation curves (function *specaccum* in
241 *vegan*).

242 Differences in the proportion of taxa that each technique captured may be a function of
243 DNA persistence or transport dynamics within pine-deciduous forests or the behavioral and
244 habitat preferences of the species themselves. Therefore, we used Bayesian multinomial logistic

models to explore the extent to which families with different ecological traits – dietary niche, terrestrial vs. aquatic life mode, and body size – were captured with different frequencies based on sampling method or tree type (coniferous vs. deciduous). We classified each family into one of six dietary guilds, four life modes, and four body size classes (see Fig. 4) using data from multiple sources including peer-reviewed articles, field guides, textbooks, and a published insect trait data set (see Appendix S2). Our models assumed that counts of families within each ecological category in each sample (e.g., the four size categories) arose from a multinomial distribution, and that sampling method and/or tree type (conifer vs. deciduous) affect each multinomial probability on the logit scale as categorical covariates. For the primary study, we used leave-one-out information criterion (LOOIC) to evaluate models with alternate covariate structures, retaining the model with the lowest LOOIC value for inference. We evaluated the results based on posterior distributions of estimated proportions, with point estimates (posterior mean) and 80% or 95% credible intervals. Modeling was performed in Stan via the *brms* package (v. 2.17.0; Bürkner, 2017) using non-informative priors and 6000 draws from the posterior (in 3 chains, each with 2000 warmup and 2000 sampling iterations). Convergence was verified using Gelman-Rubin statistics (rhat < 1.1) and examination of trace plots. All data and code associated with this study are archived and are openly available (Allen et al., 2022).

RESULTS

Three samples in the primary study (all spray samples) and four in the substrate comparison study (two roller and two spray samples) failed to amplify during PCR despite bead clean-up efforts and thus were not sent out for sequencing; these samples were excluded from further

267 analyses. The remaining samples, across both studies, had mean sequencing depths of 37920
268 (5048 SE) to 101088 (5605 SE) reads per sample after all filtering steps (see Tables S1 & S2).
269 MOTU-level richness, evenness, and diversity were similar between the two primer sets,
270 but differences in taxonomic composition were evident. In the primary study, mean [95% CI]
271 rarefied MOTU richness was similar for Zeale vs. Coleop primers (36.5 [29.4, 45.2] vs. 35.9
272 [31.5, 40.9]) as was evenness (0.58 [0.48, 0.67] vs. 0.54 [0.44, 0.64]) and Shannon diversity
273 (1.65 [1.30, 2.00] vs. 1.54 [1.21, 1.87]; Fig. S1). Notable differences in taxonomic composition
274 included that arachnids and springtails were detected only by the Zeale primer set, while
275 centipedes and multiple insect orders (most with low read counts) were detected only by the
276 Coleop primer set (Fig. S1). A parallel analysis of the substrate comparison dataset yielded
277 similar results (Fig. S2).
278 In the primary study, we detected a total of 177 arthropod families representing 14 orders
279 and 4 classes across both roller and spray eDNA sampling techniques (Fig. 1). Accumulation
280 curves indicated that the full family-level richness had not yet been captured in the study area
281 with our 20 samples (Fig. 2). Roller sampling recovered 126 families, or ~ 78% of the
282 extrapolated richness estimate of 162 families (SE = 15; Chao estimator). Spray sampling
283 recovered 148 families or a similar ~ 78% of extrapolated richness (Chao estimator = 190
284 families; SE = 16). Pooling data from the two methods (i.e., viewing them as a combined method
285 of detection) revealed an extrapolated richness of 213 families (SE = 14), of which we detected
286 83% with 20 samples (177 families). In the pooled samples, the rate of family accumulation
287 decreased from 62 families in the first five samples, to 13 in the last five samples (Fig. 2).
288 Multivariate analyses also revealed family-level compositional differences between roller
289 and spray samples in the primary study (PERMANOVA, $F = 3.98$, $df = 1, 34$, $p < 0.001$), and, to

290 a lesser extent, between coniferous and deciduous samples ($F = 1.43$, $df = 1, 34$, $p = 0.039$). This
291 result was reflected in the clear separation between roller and spray samples in the NMDS plot,
292 while little separation was apparent between coniferous and deciduous samples (Fig. 3). In
293 general, compositional differences evident between the two methods were at the family level
294 (Fig. 1) as most of the 21 orders appeared to have roughly similar prevalence in both methods
295 (Fig. 4). Exceptions to this included the lacewings (Neuroptera), which were much more
296 prevalent in roller samples, and the thrips (Thysanoptera), which were more prevalent in spray
297 samples (Fig. 4).

298 In the substrate comparison study, we detected 86 arthropod families (Table S3),
299 including 18 families and one order (Megaloptera) not detected in the primary study. Comparing
300 among substrates, the most families were detected using roller sampling (71 families in 3
301 samples) followed by spray (67 families in 3 samples), water (53 families in 2 samples), and soil
302 sampling (38 families in 5 samples; Fig. 5 and Figs. S3-S5). In that study, roller samples also
303 contained the greatest number of unique families ($n = 10$; i.e., those found only using that
304 method), and were the only samples to detect the order Megaloptera. In comparison, four unique
305 families were detected by soil sampling, three by spray, and two by water (Fig. 5). The highest
306 overlap in families detected occurred between spray and roller sampling (73% overlap; Fig. 5).
307 Water also had relatively high overlap with roller (65%) and spray (67%) sampling, while soil
308 had the lowest pairwise overlap with the other three methods (35-48%; Fig. 5). Statistical
309 analysis of family-level differences in composition among the four substrates (e.g., via
310 PERMANOVA) was precluded by low sample sizes.

311 Out of the 195 families detected across both the primary and substrate comparison
312 studies, only 3 were not detected by a surface eDNA method (spray or roller): Keroplatidae

(fungus gnats), Nycteribiidae (bat flies), and Trogidae (hide beetles), all of which were unique to soil samples.

Multinomial models revealed only minor differences in the proportional representation of different ecological guilds among sampling methods. Notably, in the primary study, roller sampling picked up slightly more predatory families than spray sampling (posterior means: 16% vs. 11%) and more families of small body size (posterior means for the "< 5 mm" class: 42% vs. 35%; see Fig. 4 for extent of credible interval overlap). Results for the substrate comparison study revealed broadly similar proportions of guild membership, but with wider credible intervals (Fig. S6).

DISCUSSION

The high diversity of arthropods has always presented challenges to documenting and describing biodiversity patterns, and there has been limited use of arthropods in monitoring the success of restoration and management designed to increase functional diversity and ecosystem services (van der Heyde et al., 2022a). No one conventional, physical field method is capable of capturing the full range of arthropods that inhabit an ecosystem, necessitating the use of several methods to ensure sufficient taxonomic coverage (e.g., Brunbjerg et al., 2019; Zeale et al., 2011). Even when a range of conventional, physical methods can be used, the volume of individuals captured can be overwhelming in terms of counting and taxonomically classifying these individuals, requiring taxonomic expertise (Beng & Corlett, 2020). Metabarcoding approaches to classifying captured arthropods are increasingly used in biodiversity research and are a clear step forward in producing site-level arthropod biodiversity inventories (e.g., Beng & Corlett, 2020; Fernandes et al., 2018; van der Heyde et al., 2022a). However, these approaches still must employ a range of

field techniques to capture a broad representation of arthropods, and the individuals captured are destructively sampled (e.g., Zeale et al., 2011). Here, we show that surface eDNA sampling is a viable approach to documenting and monitoring arthropod biodiversity in forests, and one that avoids destructive sampling or the deployment of multiple trap designs.

In our primary study, we detected the presence of 177 arthropod families, representing 14 orders and four classes, at a single forested site after minimal investment of field and lab effort. While our sampling effort was limited in spatial and temporal scope, our results suggest that we detected about 80% of the families likely (based on accumulation curves) to be present in the late-summer in this pine-deciduous forest. This result suggests that a more comprehensive survey design that included more sites would detect more arthropod families, perhaps reaching closer to 90% or more of arthropod families occupying this stretch of forest. Our choice to sample in the late summer months reflected the phenology of local arthropod assemblages in that it is in this time of year when adult stages of many species are present and thus likely depositing more DNA on vegetation surfaces. Nevertheless, New Jersey pine-oak forests are highly seasonal ecosystems and we expect that the composition of our samples would certainly change across the spring and summer months. We did not find strong evidence that the families detected differed greatly by tree type (coniferous vs. deciduous) or by feeding guild, life mode, or body size. However, there was enough of a difference in the composition of families detected between approaches (e.g., see Figs. 1-4) that both surface eDNA collection methods, and indeed as many complementary eDNA collection methods as possible, should be employed together within survey designs to ensure maximum taxonomic coverage. Collection of field eDNA samples is consistently more cost-effective than conventional field sampling techniques (e.g., Balint et al., 2018; Smart et al., 2016), and we posit that the same efficiencies would apply to surveying

arthropods in forested ecosystems, making the sampling of taxa from a wide variety of substrates logistically feasible. Lab-based efficiencies in such multi-method eDNA approaches could be achieved, at least in theory, by pooling samples to reduce the amount of lab work required. Our results can inform the design of future survey efforts seeking to track seasonal or annual changes in arthropod communities, or to compare diversity across habitats, including an indication of the families likely to be detected and the sampling effort needed per sampling unit.

Metabarcoding of eDNA collected from vegetation surfaces detected arthropods across a range of feeding guilds and body sizes, indicating the technique was not biased toward detecting only very small species or those that may feed directly on or within vegetation. The range of body sizes that we detected with surface eDNA methods, including families with species that are very small in size, indicates that the techniques perhaps sampled whole individuals that either were transferred onto the roller surface or washed into buckets. If so, these individuals provided sufficient free DNA to be captured within filters and processed accordingly. However, we also identified many large-bodied herbivorous families (e.g., within Lepidoptera and Orthoptera) that were certainly detected based on shed DNA. Additionally, arthropods that sit higher in the food web (e.g., predators, parasitoids) search for prey on plant surfaces and thus may leave their eDNA, and potentially DNA of their prey, on these surfaces. Such transport of DNA by predators or by rainwater (Valentin et al., 2021) may partly explain some of our more surprising findings, ecologically. For example, we detected ground beetles (Carabidae), a primarily ground-dwelling predatory and scavenging group (and one of high interest for ecological monitoring; Hoekman et al., 2017) in a high proportion (65-85%) of samples, indicating either that they either climbed vegetation to feed or their DNA was otherwise transferred there. Similarly, both spray and roller techniques unexpectedly detected bark lice (Psocoptera), which mainly inhabit

tree bark, in roughly equal measure (means: 36-38% of samples for 10 families). Further study of detection dynamics over time in the days surrounding rain events could shed light on such questions related to DNA transport. However, our finding that roller sampling detected fewer taxa than spray sampling (at least in the primary study) may reflect a signal of locally-produced DNA as, in general, there are more foliage feeding than bark associated arthropod species (e.g., Novotny et al., 2010).

Atlantic coastal pine barrens formerly stretched along the east coast of North America from North Carolina to Nova Scotia but are now globally rare ecosystems due to both ancient sea level rise and to modern urban development (Boyd, 1991). Although most remaining pine barren forests are protected, as is the case for our study area, these systems are experiencing rapid changes due to pest insect outbreaks (e.g., southern pine beetle, *Dendroctonus frontalis*), warming temperatures, and altered fire regimes (Dodds et al., 2018; Kretchun et al., 2014). The influence of these factors on animal diversity is mostly unknown, particularly for arthropods, although climate change, in general, is expected to reduce diversity (Halsch et a., 2021). At the same time, synthesis studies have suggested broad-scale insect population declines at regional and even global scales (Forister et al., 2019; Sánchez-Bayo & Wyckhuys, 2019). We show that vegetation surface eDNA techniques may be well-suited for use to document current arthropod biodiversity within these forests, and track changes in composition over time as threats persist or as management and restoration efforts proceed. eDNA is increasingly considered a viable tool for long-term biodiversity monitoring (Fernandes et al., 2018; van der Heyde et al., 2022b), but a full rollout of our approach (or others like it) would first require a deeper exploration of spatial and temporal variation in eDNA detections to determine appropriate survey design. Leveraging existing conventional long-term monitoring schemes for arthropods (e.g., those associated with

405 the National Ecological Observatory Network; Hoekman et al., 2017) may be a fruitful approach
406 for benchmarking nascent eDNA-based ecosystem monitoring programs moving forward. This is
407 especially true for key functional groups, like pollinators, which are considered high priority for
408 management and restoration efforts (Thomsen & Sigsgaard, 2014).

409 Despite the growth of eDNA research over the past decade, there have been very few
410 studies that compare compositional differences in taxa detected across substrates, such as soil,
411 water, or vegetation surfaces, sampled using the same set of DNA primers (van der Heyde,
412 2022b). Although our efforts to compare across substrates were limited by relatively small
413 samples sizes (n = 2-5 samples per substrate), we show that soil sampling, a commonly used
414 eDNA technique for arthropods, detected the fewest families of the four substrates we sampled,
415 despite having the largest number of samples (n = 5). Soil eDNA did detect three families that
416 other substrates did not (fungus gnats, bat flies, and hide beetles), though none of these families
417 predominately inhabit soil. Sampling larger volumes of soil (e.g., see Leempoel et al., 2020)
418 could yield more taxa. Our results indicate that lentic water bodies do serve as a pooling site for
419 eDNA deposited by nearby terrestrial arthropods, as they do for terrestrial birds and mammals
420 (Harper et al., 2019; Ushio et al. 2018). Our water samples, despite their low number (n = 2),
421 allowed detection of a range of terrestrial taxa of diverse body sizes and feeding guilds.
422 However, water sampling revealed few (only 2) unique families, while it missed some significant
423 taxonomic groups that spray, roller, and soil sampling techniques all captured (e.g., an order of
424 springtails, Entomobryomorpha; and mites, Trombidiformes). Both body size and behavior affect
425 DNA deposition rates into the environment (Adams et al., 2019), and it is possible that terrestrial
426 arthropods, many of which have relatively small home ranges and body sizes, may not reach or
427 deposit DNA in water bodies as much as mammals and birds.

Finally, as in all metabarcoding research, our choice of primers likely influenced the accuracy and breadth of taxa we detected (Horton et al., 2017). We used primers to amplify two genomic regions (16S and COI) that are commonly used in arthropod metabarcoding efforts to identify taxa sampled using conventional physical capture techniques (e.g., Malaise traps, pitfall traps). While neither primer set performed decisively better at capturing family-level diversity, the use of a wider selection of primers certainly increased the number of families we detected. In fact, entire classes would have been missed if we omitted either the Zeale or Coleop primers (arachnids and centipedes, respectively). More generally, our eDNA survey was 'blind', meaning that we did not have a target set of species that we wished to detect, nor were reference specimens available as would be the case in a trapping study (Darling et al., 2000). Thus, like other comparable studies (e.g., Thomsen & Sigsgaard, 2014), we are reliant on, and are limited by, the availability of arthropod COI and 16S sequence data in public databases for taxa identification. Although the number of arthropod sequences within these databases is growing quickly, the vast global diversity of arthropod species means that relying on published sequence data will likely always entail some lack of taxonomic resolution and uncertainty (Jinbo et al., 2011). Here we encountered this issue with accurately classifying eDNA sequences to taxonomic levels below family, and we may have even unknowingly missed some family-level detections (i.e., a small number of arthropod MOTUs were only classified to the level of order or higher; see Figs. S1 and S2). A solution to these issues is the de novo construction of a reference sequence database that is built using tissues collected from all species that an eDNA survey is designed to detect (Darling et al., 2020). Such a solution may work well for terrestrial arthropod fit-for-purpose eDNA surveys if the target suite of species is, for example, a set of forest pests, pollinators, or species of conservation concern (Darling et al., 2020). Trap-based arthropod

metabarcoding studies for biodiversity inventories (i.e., those that involve reference specimen collection) can also be used as a springboard toward use of fully trap-free eDNA approaches in defined geographic areas, as local libraries are assembled. However, even completely blind eDNA surveys have value for documenting biodiversity patterns across larger spatial scales (e.g., Porter et al., 2019) or for carefully qualified preliminary surveillance of specific groups, such as forest arthropod pests (Darling et al., 2020).

A little over a decade after the launch of the International Barcode of Life initiative (iBOL), routine DNA-based biodiversity inventories of arthropods, a hugely important and fragile component of Earth's biodiversity, appears increasingly within reach (Jinbo et al., 2011; Sánchez-Bayo & Wyckhuys, 2019). We illustrate the usefulness of tree and plant surface eDNA sampling techniques for arthropod biodiversity assessments, and more generally, the value of eDNA surveys to realizing a range of arthropod research and management goals. The practical advantages of using an eDNA arthropod biodiversity survey are numerous. For example, sampling is simple and non-destructive causing no damage to the species or habitat surveyed, while rare, cryptic, and elusive species can be detected 'sight unseen'. Further, taxonomic identification of species becomes standardized and auditable, and, with a robust reference database, species can be accurately identified no matter what life stages are present at the site (Belle et al., 2019; Bent & Corlett, 2020; Bista et al., 2017). Our results provide proof-of-concept evidence that eDNA surveys are feasible and represent a potentially cost-effective solution for arthropod research, management, and restoration (Darling et al., 2020; Fernandes et al., 2018).

DATA AVAILABILITY STATEMENT

The data and code associated with this study will be openly available in Dryad at http://doi.org/[doi], reference number [reference number].

REFERENCES

Adams, I. C., Hoekstra, A. L., Muell, R. M., & Janzen, J. F. (2019). A brief review of non-avian reptile environmental DNA (eDNA), with a case study of painted turtle (*Chrysemys picta*) eDNA under field conditions. *Diversity*, 11, 50. (doi:10.3390/d11040050)

[dataset] Allen, M. C., Lockwood, J. L., Kwait, R., Vastano, A. R., Peterson, D. L., Tkacenko, L. A., Angle, J. C., & Jaffe, B. D. (2022). Data and code for: "Using surface environmental DNA to assess arthropod biodiversity within a forested ecosystem". *Dryad*, https://doi.org/10.XXXX/XXXX.XXXX.XXXXX

Allen, M. C., Peterson, D. L., Nielsen, A. L., & Lockwood, J. L. (2021). Terrestrial eDNA survey outperforms conventional approach for detecting an invasive pest insect within an agricultural ecosystem. *Environmental DNA*, 3(6), 1102-1112.

Ärje, J., Melvad, C., Jeppesen, M., Madsen, S., Raitoharju, J., Rasmussen, M., Iosifidis, A., Tirronen, V., Gabbouj, M., Meissner, K., & Høye, T. T. (2020). Automatic image-based identification and biomass estimation of invertebrates. *Methods in Ecology and Evolution*, 11, 922-931.

Balint, M., Nowak, C., Marton, O., Pauls, S. U., Wittwer, C., Aramayo, J. L., Shulze, A., Chambert, T., Cocchiararo, B., & Jansen, M. (2018). Accuracy, limitations and cost efficiency of eDNA-based community survey of tropical frogs. *Molecular Ecology Resources*, 18, 1415-1426.

Belle, C. B., Stoeckle, B. C., & Geist, J. (2019). Taxnomic and geographical representation of freshwater environmental DNA research in aquatic conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 29, 1996-2009.

Beng, K. C., Tomlinson, K. W., Shen, X. H., Surget-Groba, Y., Hughes, A. C., Corlett, R. T., & Ferry Slik, J. W. (2016). The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. *Scientific Reports*, 6, 24965.

Beng, K. C., & Corlett, R. T. (2020). Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. *Biodiversity and Conservation*, 29, 2089-2121.

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., Christmas, M., & Creer, S. (2017). Annual time-series analysis of aqueous eDNA reveals ecological relevant dynamics of lake ecosystem biodiversity. *Nature Communications*, 8, 14087.

Boerlijst, S. P., Trimbos, K. B., Van der Beek, J. G., Dijkstra, K. D. B., Van der Hoorn, B. B., & Shrama, M. (2019). Field evaluation of DNA based biodiversity monitoring of Caribbean mosquitoes. *Frontiers in Ecology and Evolution*, 7, Article 240. https://doi.org/10.3389/fevo.2019.00240

Boyd, H.P. (1991). *A field guide to the Pine Barrens of New Jersey*. Plexus Publishing.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). Obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16, 176–182.

Brunbjerg, A. K., Henrik Bruun, H., Brondum, L., Classen, A. T., Dalby, L., Fog, K., Froslev, T. G., Goldberg, I., Johannes Hansen, A., Hansen, M. D. D., Hoye, T. T., Illum, A. A., Laessoe, T., Newman, G. S., Skipper, L., Sochting, U., & Ejrnaes, R. (2019). A systematic survey of regional multi-taxon biodiversity: evaluating strategies and coverage. *BMC Ecology*, 19, Article 43. https://doi.org/10.1186/s12898-019-0260-x

Bürkner, P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1-28. https://doi.org/10.18637/jss.v080.i01

Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M., Bograd, S. J., Hazen, E. L., Field, J. C., & Boehm, A. B. (2019). Marine vertebrate biodiversity and distribution within the Central California Current using environmental DNA (eDNA) metabarcoding and ecosystem surveys. *Frontiers in Marine Science*, 6, Article 732. https://doi.org/10.3389/fmars.2019.00732

Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Ridwan Amode, M., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., … Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, 50, D988–D995.

Dangles, O., & Casas, J. (2019). Ecosystem services provided by insects for achieving sustainable development goals. *Ecosystem Services*, 35, 109-115.

Darling, J. A., Pochon, X., Abbot, C. L., Inglis, G. J., & Zaiko, A. (2020). The risks of using molecular biodiversity data for incidental detection of species of concern. *Diversity and Distributions*, 26, 1116-1121.

Dodds, K. J., Aoki, C. F., Arango-Velez, A., Cancelliere, J., D'Amato, A. W., DiGirolomo, M. F., & Rabaglia, R. J. (2018). Expansion of southern pine beetle into northeastern forests: Management and impact of a primary bark beetle in a new region. *Journal of Forestry*, 116(2), 178-191.

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V. I., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kauserud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, E., & Brochmann, C. (2012). New environmental metabarcodes for analyzing soil DNA: potential for studying past and present ecosystems. *Molecular Ecology*, 21, 1821-1833.

Fernandes, K., van der Heyde, M., Bunce, M., Dixon, K., Harris, R. J., Wardell-Johnson, G., & Nevill, P. G. (2018). DNA metabarcoding – a new approach to fauna monitoring in mine site restoration. *Restoration Ecology*, 26, 1098-1107.

Forister, M. L., Pelton, E. M., & Black, S. H. (2019). Declines in insect abundance and diversity: We know enough to act now. *Conservation Science and Practice*, 1(8), Article e80.

Foster, Z. S. L., Sharpton, T. J., & Grunwald, N. J. (2017). Metacoder: an R package for visualization and manipulation of community taxonomic diversity data. *PLoS Computational Biology*, 13, Article e1005404.

Halsch, C. A., Shapiro, A. M., Fordyce, J. A., Nice, C. C., Thorne, J. H., Waetjen, D. P., & Forister, M. L. (2021). Insects and recent climate change. Proceedings of the National Academy of Sciences, 118(2), Article e2002543117.

Harper, L. R. Lawson Handley, L. Carpenter, A. L., Ghazali, M., Di Muri, C., Macgregor, C. J., Logan, T. W., Law, A., Breithaupt, T., Read, D. S., McDevitt, A. D., & Hänfling, B.

(2019). Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. *Biological Conservation*, 238, Article 108225.

Hoekman, D., LeVan, K. E., Gibson, C., Ball, G. E., Browne, R. A., Davidson, R. L., Erwin, T. L., Knisley, C. B., LaBonte, J. R., Lundgren, J., Maddison, D. R., Moore, W., Niemelä, J., Ober, K. A., Pearson, D. L., Spence, J. R., Will, K., & Work, T. (2017). Design for ground beetle abundance and diversity sampling within the National Ecological Observatory Network. *Ecosphere*, 8(4), Article e01744.

Horton, D. J., Kershner, M. W., & Blackwood, C. B. (2017). Suitability of PCR primers for characterizing invertebrate communities from soil and leaf litter targeting metazoan 18S ribosomal or cytochrome oxidase I (COI) genes. *European Journal of Soil Biology*, 80, 43-48.

van der Heyde, M., Bunce, M., Dixon, K. W., Fernandes, K., Majer, J., Wardell-Johnson, G., White, N. E., & Nevill, P. (2022a). Evaluating restoration trajectories using DNA metabarcoding of ground-dwelling and airborne invertebrates and associated plant communities. *Molecular Ecology*, 31(7), 2172-2188.

van der Heyde, M., Bunce, M., & Nevill, P. (2022b). Key factors to consider in the use of environmental DNA metabarcoding to monitor terrestrial ecological restoration. *Science of the Total Environment*, 848, Article 157617.

Illumina. (2013). *16S Metagenomic sequencing library preparation: Preparing 16S Ribosomal RNA Gene Amplicons Illumina MiSeq System*. Illumina.

Jinbo, U., Kato, T., & Ito, M. (2011). Current progress in DNA barcoding and future implications for entomology. *Entomological Science*, 14(2), 107-124.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Research*, 36, W5-9.

Kretchun, A. M., Scheller, R. M., Lucash, M. S., Clark, K. L., Hom, J., & Van Tuyl, S. (2014). Predicted effects of gypsy moth defoliation and climate change on forest carbon dynamics in the New Jersey Pine Barrens. *PLoS One*, 9(8), Article e102531.

Kyle, K. E., Allen, M. C., Dragon, J., Bunnell, J. F., Reinert, H. K., Zappalorti, R., Jaffe, B. D., Angle, J. C., & Lockwood, J. L. (2022). Combining surface and soil environmental DNA with artificial cover objects to improve terrestrial reptile survey detection. *Conservation Biology*, Article e13939. https://doi.org/10.1111/cobi.13939

Leempoel, K., Hebert, T., & Hadly, E. A. (2020). A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. *Proceedings of the Royal Society Biological Sciences*, 287, Article 20192353. http://doi.org/10.1098/rspb.2019.2353

Marquina, D., Esparza-Salas, R., Roslin, T., & Ronquist, F. (2019). Establishing arthropod community composition using metabarcoding: Surprising inconsistencies between soil samples and preservative ethanol and homogenate from Malaise trap catches. *Molecular Ecology Resources*, 19(6), 1516-1530.

Novotny, V., Miller, S. E., Baje, L., Balagawi, S., Basset, Y., Cizek, L., Craft, K. J., Dem, F., Drew, R. A., Hulcr, J., & Leps, J. (2010). Guild-specific patterns of species richness and host specialization in plant–herbivore food webs from a tropical forest. *Journal of Animal Ecology*, *79*(6), 1193-1203.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, H. H. H., Szoecs, E., & Wagner, H.

(2020). vegan: Community Ecology Package. https://CRAN.R-project.org/package=vegan

Oliverio, A. M., Gan, H., Wickings, K., & Fierer, N. (2018). A DNA metabarcoding approach to characterize soil arthropod communities. *Soil Biology and Biochemistry*, 125, 37-43.

Porter, T. M., Morris, D. M., Basiliko, N., Hajibabaei, M., Doucet, D., Bowman, S., Emilson, E. J. S., Emilson, C. E., Chartrand, D., Wainio-Keizer, K., Seguin, A., & Venier, L. (2019). Variations in terrestrial arthropod DNA metabarcoding methods recovers robust beta diversity but variable richness and site indicators. *Scientific Reports*, 9, Article 18218.

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Sánchez-Bayo, F., & Wyckhuys, K. A. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, 232, 8-27.

Schulze, C. H., Waltert, M., Kessler, P. J., Pitopang, R., Veddeler, D., Mühlenberg, M., Gradstein, S. R., Leuschner, C., Steffan-Dewenter, I., & Tscharntke, T. (2004). Biodiversity indicator groups of tropical land-use systems: comparing plants, birds, and insects. *Ecological Applications*, 14(5), 1321-1333.

Smart, A. S., Weeks, A. R., van Rooyen, A. R., Moore, A., McCarthy, M. A., & Tingley, R. (2016). Assessing the cost-efficiency of environmental DNA sampling. *Methods in Ecology and Evolution*, 7(11), 1291-1298.

Thomsen, P. F., & Sigsgaard, E. E. (2018). Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. *Ecology and Evolution*, 9, 1665-1679.

Ushio, M., Murata, K., Sado, T., Nishiumi, I., Takeshita, M., Iwasaki, W., & Miya, M. (2018). Demonstration of the potential of environmental DNA as a tool for the detection of avian species. *Scientific Reports*, 8, Article 4493.

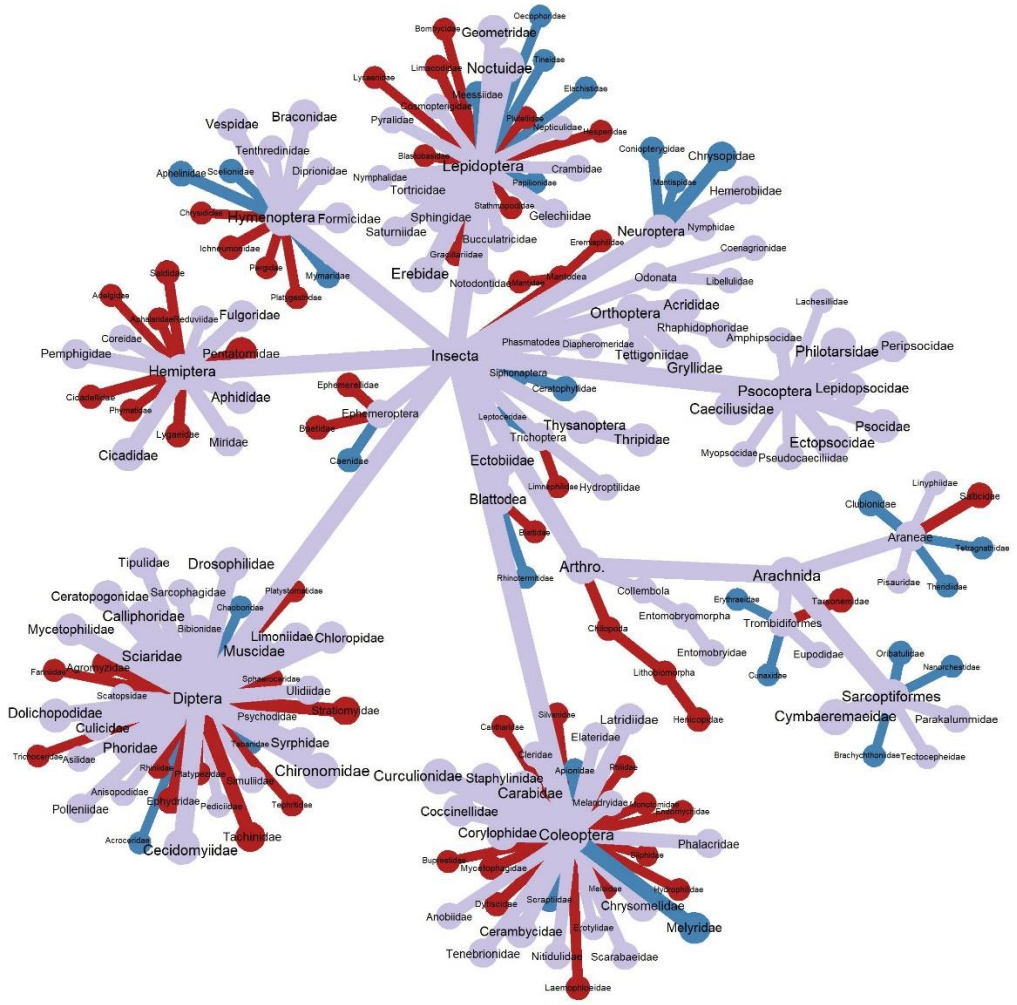
Valentin, R. E., Fonseca, D. M., Nielsen, A. L., Leskey, T. C., & Lockwood, J. L. (2018). Early detection of invasive exotic insect infestations using eDNA from crop surfaces. *Frontiers in Ecology and the Environment*, 16(5), 265-270.

Valentin, R. E., Fonseca, D. M., Gable, S., Kyle, K. E., Hamilton, G. C., Nielsen, A. L., & Lockwood, J. L. (2020). Moving eDNA surveys onto land: strategies for active eDNA aggregation to detect invasive forest insects. *Molecular Ecology Resources*, 20(3), 746-755.

Valentin, R. E., Kyle, K. E., Allen, M. C., Welbourne, D. J., & Lockwood, J. L. (2021). The state, transport, and fate of aboveground terrestrial arthropod eDNA. *Environmental DNA*, 3(6), 1081-1092.

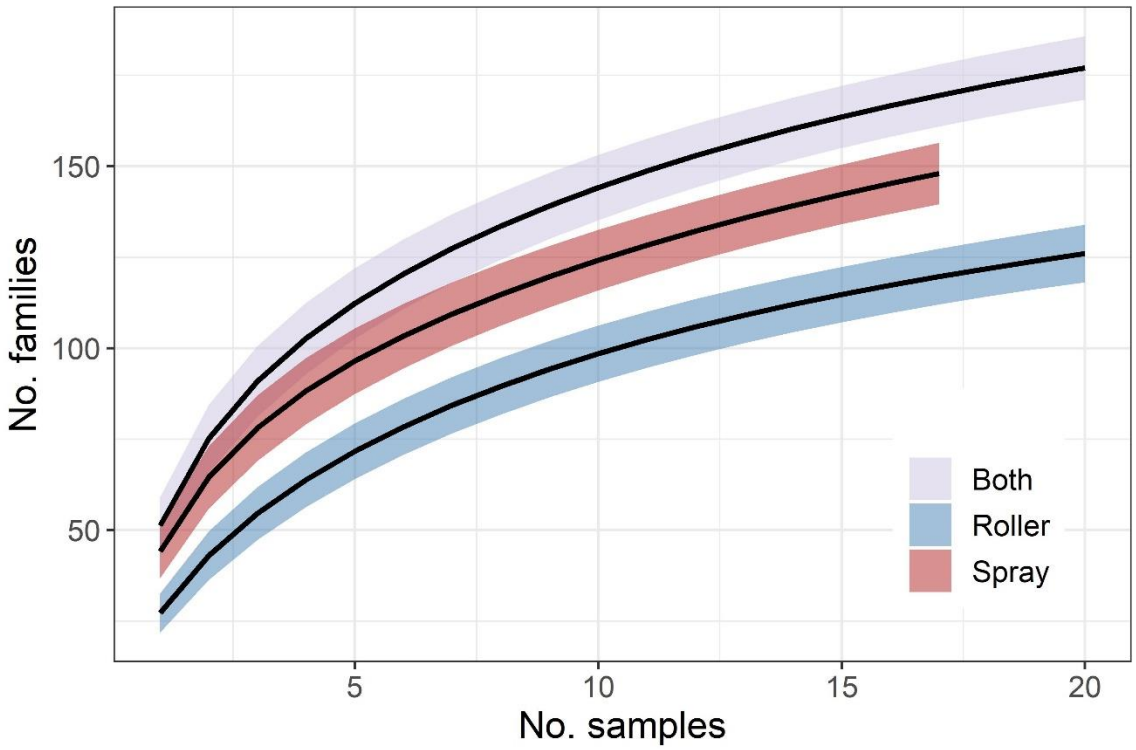
Zeale, M. R. K., Butlin, R. K., Barker, G. L. A., & Lees, D. C. (2011). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, 11, 236-244.

Zenker, M. M., Specht, A., & Fonseca, V. G. (2019). Assessing insect biodiversity with automatic light traps in Brazil: pearls and pitfalls of metabarcoding samples in preservative ethanol. *Ecology and Evolution*, 10, 2352-2366.



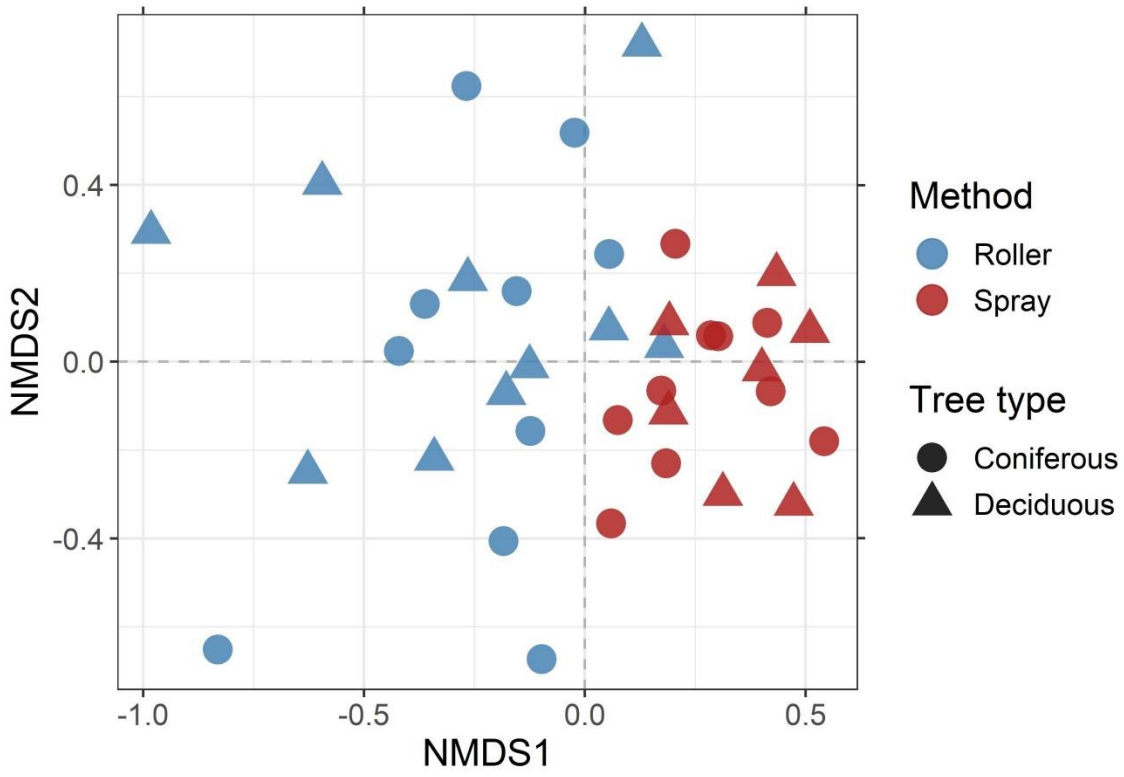
644

645 Figure 1. Taxonomic 'heat tree' depicting the 177 arthropod families detected via DNA
646 metabarcoding and two methods of surface eDNA sampling within a pine-deciduous forest in the
647 Pinelands National Reserve, New Jersey, USA. Families only detected using foliage ('spray')
648 eDNA samples are depicted in red, those only detected in tree bark ('roller') samples are in blue,
649 and those detected by both methods are in purple. The font size and width of the nodes and
650 branches are proportional to the prevalence of each family within the samples (smallest symbols
651 = detected at 5% of sites; largest symbols = detected at 100% of sites).



652

653 Figure 2. Accumulation curves (mean +/- 1 SD) showing that the number of arthropod families
654 detected via metabarcoding of surface eDNA samples increased with increasing numbers of
655 sampling sites within a pine-deciduous forest in the Pinelands National Reserve, New Jersey,
656 USA. The blue shading indicates tree bark ('Roller') samples, the red shading indicates foliage
657 ('Spray') samples, and the purple shading indicates pooled data from these two methods.



658

659 Figure 3. Ordination plot (non-metric multidimensional scaling) showing differences in family-
660 level community composition of metabarcoding samples (presence-absence data) within a pine-
661 deciduous forest in the Pinelands National Reserve, New Jersey, USA. Each point represents a
662 single sample. There was clear evidence for differences between the tree bark ('Roller') vs.
663 foliage ('Spray') eDNA sampling methods (PERMANOVA; R^2 = 10.1%; F = 3.98, df = 1, 34, p
664 < 0.001), but less evidence for a difference between deciduous and coniferous trees (R^2 = 3.6%;
665 F = 1.43, df = 1, 34, p = 0.039).
666

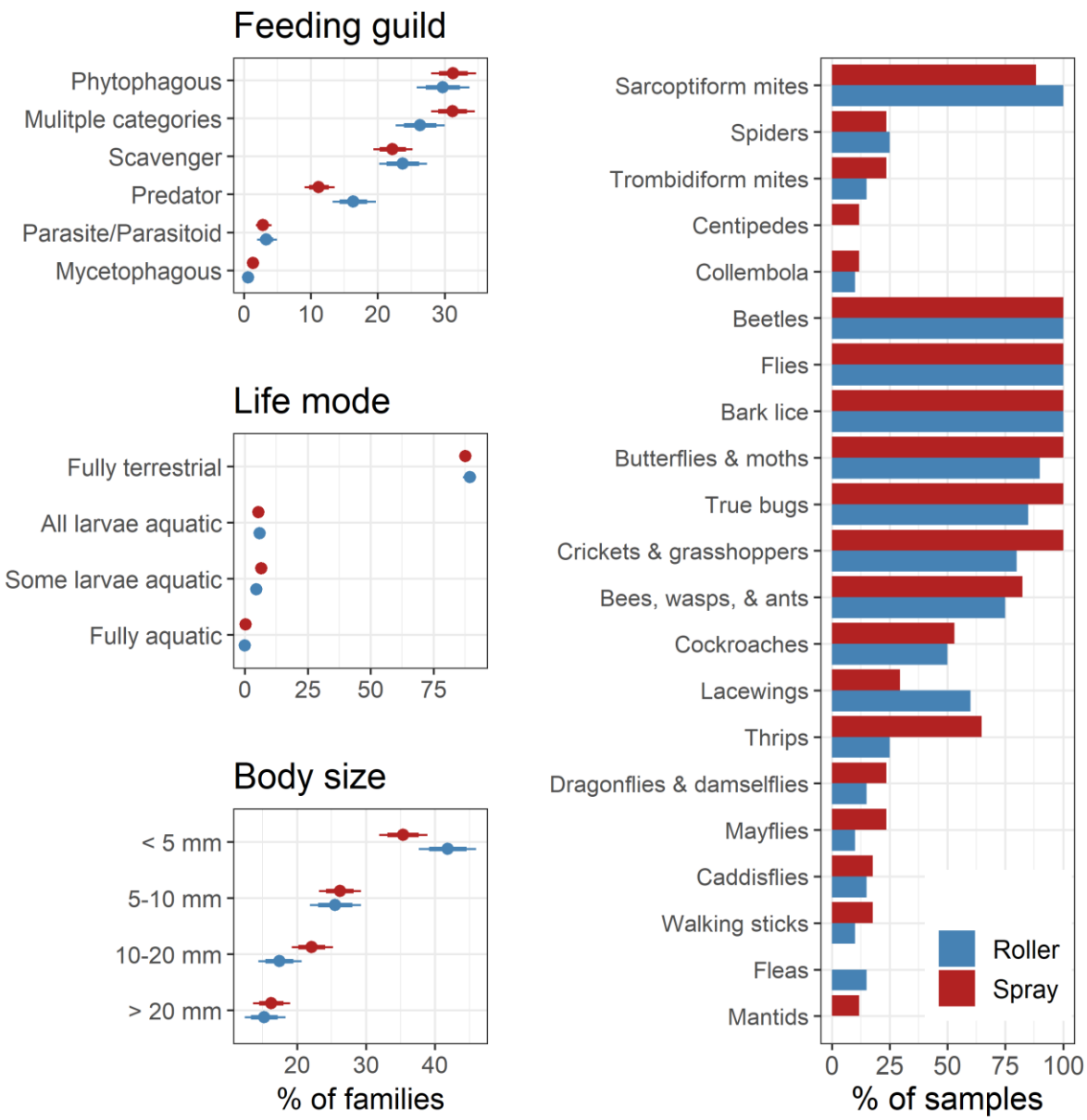
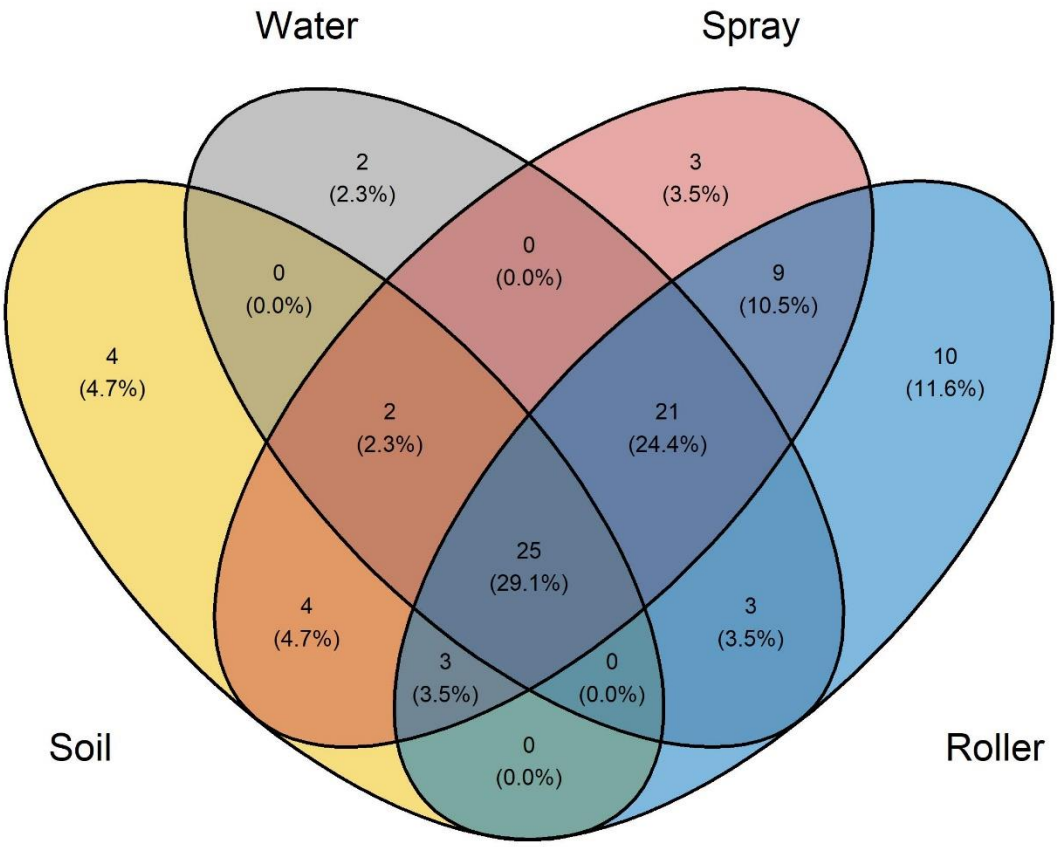


Figure 4. Ecological and phylogenetic differences in arthropod families detected using metabarcoding of tree bark ('Roller') and foliage ('Spray') eDNA collection methods. The x-axis in the left three plots is the estimated % of families detected from each category (feeding guild, life mode, or body size category). Point estimates and 80% and 95% credible intervals are shown from the best-performing Bayesian multinomial model (in each case, the model containing only 'sampling method' performed best: ΔLOOIC = 2.6-4.7). The x-axis in the right plot shows the % of samples in which each Arthropod order was detected (see Table S3 for scientific names).



675

676 Figure 5. Overlap among arthropod families detected with metabarcoding using four different
677 eDNA collection methods that targeted different substrates within a pine-deciduous forest in the
678 Pinelands National Reserve, New Jersey, USA. Note that sample sizes varied among the methods
679 (soil: n = 5 samples, roller: n = 3, spray: n = 3, water: n = 2).