Scalable environmental DNA methods reveal strong associations between landscape-scale forest habitat and insect richness

Rodney T. Richardson[1*], Grace Avalos[2], Cameron Garland[2], Regina Trott[3], Olivia Hager[3], Mark J. Hepner[4], Karen Goodell[5]

[1] Western EcoSystems Technology, Inc., Indianapolis, IN
[2] Department of Entomology, University of Minnesota, Saint Paul, MN
[3] Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, MD
[4] Metamorphic Ecological Research and Consulting, LLC, Alonzaville, VA
[5] Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Newark, OH

* Author for correspondence: rrichardson@west-inc.com

Abstract

1. While aquatic environmental DNA (eDNA) methods have reached relative maturity, the use of terrestrial eDNA is nascent and serious methodological improvements are needed to achieve widespread use. Field-ready applications require eDNA survey methods where samples are easy to collect by inexperienced practitioners, easy to transport between the field and lab, and easy to process thereafter. Here, we demonstrate methods that satisfy these requirements and show strong potential for characterizing diverse terrestrial eDNA samples collected from flower and leaf surfaces.
2. We used novel methods to collect and process 236 flower eDNA samples and 21 leaf surface eDNA samples, obtaining 2,228 Arthropoda eDNA detections spanning 175 families using amplicon sequencing of two genetic markers.
3. Detected taxa were diverse and included numerous groups of conservation concern, such as bees (Hymenoptera; Anthophila, 32 genera spanning 5 families) and Lepidoptera (209 genera from 21 families). Data reveal strong associations between insect community richness and remotely sensed measures of forest habitat, providing a quantitative perspective of relevance to insect conservation.
4. It is increasingly clear that a variety of organisms readily disperse eDNA throughout the environment, supporting the notion that eDNA will be a powerful tool for characterizing species distributions and monitoring at-risk species. However, we conclude that

researchers seeking to characterize fine-scale habitat associations or plant-pollinator interactions using eDNA will need to carefully design studies with appropriate field controls, such as the leaf surface eDNA samples collected here.

Keywords: Invasion biology; Entomology; Conservation; Ecology; Biodiversity

Introduction

In recent decades, North America has experienced considerable shifts in land use, with more than 100,000 km^2 converted into urban developments and row crop agriculture (He et al., 2019; Lark et al., 2020). Areas occupied by short vegetation, including shrublands and grasslands, have decreased by over 200,000 km^2 (Song et al., 2018). Land use changes, and other forms of global environmental change, are intrinsically linked to change in the abundance and geographic distribution of species (Jackson & Sax, 2010). Such changes represent a direct risk to the ecosystem processes that generate natural capital humans depend upon. Accordingly, scientists have argued for rapid investment in the infrastructure and methods needed to better understand the effects of environmental change as well as how they can be managed to limit economic damage and improve ecosystem resilience (Williams, 2022). Attaining these goals requires methods to track shifts in abundance, range and diversity of key biota that underly ecosystem function.

One of the few encouraging attributes of recent land use change includes the reforestation of over 270,000 km^2 across the U.S. and Canada (Song et al., 2018). Forests can help buffer against extreme conditions associated with climate change (Barnes et al., 2024; De Lombaerde et al., 2022) and provide valuable habitat for a wide diversity of wildlife, especially when differential management is used to promote variation in forest age, composition and understory conditions (Hekkala et al., 2023; Hepner et al., 2024; Jeffries et al., 2006; Uhl et al., 2024). Insect community trends in forested systems are difficult to measure, but the maintenance of forest appears to be linked to lower magnitudes of insect loss (Martínez-Núñez et al., 2024; Staab et al., 2023) and, potentially, diversity gains (Ober & Hayes, 2008).

Insects exert major influence over terrestrial ecosystems and are sensitive to environmental change (Jactel et al., 2019). Insects are of broad conservation concern because recent reports suggest insect biomass and richness are declining (Hallmann et al., 2017; Sánchez-

Bayo & Wyckhuys, 2019; Wagner et al., 2021), but efforts to characterize such declines have come to different conclusions about the strength of the evidence and rates of loss (Crossley et al., 2020; Fürst et al., 2023; Rosenheim & Ward, 2020). Meanwhile, there is mounting concern regarding climate change-associated insect pest outbreaks threatening forest health (Bracalini et al., 2024). Disagreement over the magnitude and associated drivers of insect population changes indicates that studies with greater statistical power and improved design are needed. Insects, however, are difficult to monitor because of their high diversity, small size and cryptic variation (van Klink et al., 2022). Even large datasets can fall short of the power needed to link insect compositional change with environmental factors (Tronstad et al., 2022), and the lethal nature of typical designs could compromise vulnerable populations (Tepedino et al., 2015).

In addition to the sheer number of samples needed for large-scale insect monitoring, variation in the properties of different collection methods limits comprehensiveness and comparability across studies (Didham et al., 2020; Montgomery et al., 2021; Portman et al., 2020). Thus, researchers often rely on multiple survey methods to minimize sampling bias, though the extent to which this increases inferential accuracy has received minimal attention. Once samples are collected using traditional methods, the preservation of individual specimens is cumbersome and subsequent identification with taxonomic keys requires high degrees of expertise (Packer et al., 2009). The difficulty of measuring insect population change hampers our ability to stay abreast of important shifts of relevance to insect conservation and forest health (Didham et al., 2020; Knuff et al., 2019; Kuhlman et al., 2021).

Environmental DNA (eDNA) methods have been proposed as a means of circumventing key challenges in the surveillance of terrestrial diversity (Allen et al., 2023; Bell et al., 2024; Bohmann & Lynggaard, 2023). Rapid collection of large numbers of samples can be accomplished by individuals with minimal training using eDNA sampling. These features allow surveys to be distributed across broad communities of scientists. Additionally, samples can be collected easily over short time periods to closely align surveys with the natural life history of target taxa. In this study, we demonstrate the advantages of eDNA sampling for characterizing insect diversity across a broad spatial scale in a heavily forested ecosystem that, like most forests globally, is experiencing the dual threats of climate and land use change. In sequencing collected eDNA samples using genetic markers with previously demonstrated sensitivity across a variety

of insect orders (Avalos et al., 2024), we hypothesized that insect diversity would be positively related to the proportion of forested area and forest tree species diversity at the landscape scale. In addition to evaluating relationships between forest habitat and insect diversity, we provide a rough comparison of the sensitivity of our methods to past insect eDNA efforts.

Methods

Study area and Sample Collection

We conducted a large-scale eDNA survey of the Central Appalachian region of the Eastern United States (U.S.). Central Appalachia is a montane region with large areas of temperate forest and high forest connectivity. Unforested areas in this region tend to include pasture, row-crop agriculture and low-intensity development. From June to September 2022, we conducted 71 site visits across 55 locations spanning 4 states in the Eastern U.S.: Pennsylvania, Maryland, West Virginia and Virginia (Figure 1A). Sites consisted of roadside rights-of-way with abundant floral resources. During each site visit, 1 to 3 surveyors collected flower and leaf surface eDNA samples. Leaf surface samples included leaves from trees or shrubs that were at least 3 meters from prominent floral forage, and which showed no evidence of insect herbivory or visible debris. For each eDNA sample, a single species of flower or leaves were carefully clipped into a sterile quart sized zip-closable plastic bag until the bag was approximately ¼ to ¾ full of loose plant material. Samples were mixed with 25 mL of eDNA preservative (10 percent EtOH v/v, 40 percent propylene glycol v/v and 0.25 percent sodium dodecyl sulfate w/v), mixed gently and allowed to sit for 1 to 5 min before the eDNA-containing preservative rinse was carefully poured into a labelled 50 mL conical vial for transport to the laboratory (Figure 1B). During eDNA collection, researchers used flame-sterilized tools and disposable sterile nitrile gloves to avoid cross contamination between sites or eDNA surveys.

Laboratory eDNA Processing

To purify each field-collected eDNA sample, 10 ul of 11.25 ug/ul yeast tRNA, 15 ul of 0.5 percent w/v linear polyacrylamide (Maixner et al., 2022) and 20 ul of 20 mg/ml glycogen was added to each sample. The concentration of NaCl was adjusted to 0.2 M through addition of 5 M NaCl stock solution. Samples were incubated overnight at –20° C to facilitate aggregation of nucleic acids prior to centrifugation at 7,006 x g for 40 min at 4° C. Sample supernatant was

discarded and the eDNA-containing pellet was air-dried until all visible liquid was evaporated. Pellets were resuspended in 500 ul of lysis buffer (0.5 percent SDS, 10 mM Tris, 1 mM EDTA, 10 mM NaCl, pH 8.0). Resuspended samples were transferred to 2.0 mL microcentrifuge tubes and 5 ul of 20 mg/ml proteinase K was added prior to incubation at 55° C for 1 hour. Each sample was brought to a NaCl concentration of 0.3 M through addition of 5 M stock solution. Samples were incubated on ice for 5 min prior to centrifugation at 20,000 x g for 20 min at 4° C. The eDNA-containing supernatant was transferred to a new 2.0 ml microcentrifuge tube and approximately 2.5 volumes of ethanol were added. Samples were gently mixed by inversion and incubated at –20° C overnight prior to centrifugation at 20,000 x g for 40 min at 4° C. The supernatant was carefully removed, and the eDNA-containing pellets were air dried for approximately 15 minutes prior to being resuspended in 100 ul of nuclease-free water. PCR inhibitors were then removed from the DNA using the Qiagen PowerClean Pro Cleanup Kit. Following purification, Illumina MiSeq amplicon libraries were prepared using a previously-established 3-step PCR-based protocol (Richardson et al., 2019). During library preparation, we used two sets of primers, the 28S primers from Darby et al. (2020), designed for specificity to Anthophila, and the broad arthropod-specific BR2/BF2 primers (Elbrecht & Leese, 2017). During initial BR2/BF2 PCR reactions, 1.6 uL of a 100 uM concentration equimolar mixture of blocking primers was added. These 3’ spacer C3 modified primers were designed to limit fungal byproducts previously observed in past work (Avalos et al., 2024) and are provided in supplemental Table S1. For each marker, we incorporated 15 no-library negative control dual-index combinations during library preparation to quantify critical-mistagging rates during downstream sequencing (Esling et al., 2015). Amplified libraries were purified with the SequalPrep Normalization Plate Kit, pooled equimolarly and subjected to Illumina MiSeq sequencing using a single-end 300 cycle v2 flow cell for 28S and paired-end 500 cycle v2 flow cell for COI.

Taxonomic Annotation of eDNA Data

After sequencing, VSEARCH (Rognes et al., 2016) was used to merge paired-end COI sequences, quality trim single-end 28S sequences and remove priming sites from the 5’ and 3’ ends. During quality trimming, sequences were truncated starting with the first base pair with a Phred score of 18 or lower. Sequences were taxonomically annotated by semi-global VSEARCH

top-hit alignment against custom reference sequence databases trimmed to the amplicon regions of interest using MetaCurator (Richardson et al., 2020), with dependencies HMMER (Eddy, 2011), MAFFT (Katoh et al., 2002), VSEARCH and Taxonomizr (Sherrill-Mix, 2019). During alignment, a minimum query cover of 0.8 was required. Reference databases were constructed using sequencies available through NCBI, downloaded on February 5th, 2023 for 28S and February 2nd, 2023 for COI. Following reference curation, the 28S database consisted of 53,478 sequences representing 1,289 families, 11,440 genera and 28,191 species. COI reference data were relatively more complete, with 606,748 sequences representing 2,129 families, 28,975 genera, and 104,243 species. During alignment, 96 percent identity matches were considered confident genus-level detections for 28S and 95 percent matches were considered genus-level detections for COI. For both markers, 85 percent matches were considered phylum-level matches. Since all Illumina sequencing runs produce low frequency misidentifications when inferring dual-index tags (i.e. critical mistags; Esling et al., 2015), we used an established statistical technique to remove identifications with greater than 0.05 probability of representing a critical mistag-associated detection (Richardson, 2022). Additionally, we removed detections represented by three or fewer sequences in a sample. Computational analysis was performed using the *Ohio Supercomputer Center*, (1987).

Survey Site Landscape Analysis

For each sampling site, remotely sensed maps of land cover and forest species basal areas were used to gather landscape-scale information at 0.5, 1 and 3 km radius scales. Proportional land cover was characterized using the National Land Cover Dataset (NLCD; Dewitz, 2021) and downloaded for our study region using the get_nlcd function in the FedData package (Bocinsky, 2024) on February 5th, 2024. Total forest basal area and species-specific forest basal areas were estimated using a 250 m resolution gridded product based on USFS Forest Inventory and Analysis (FIA) data, MODIS remote sensing of phenology and environmental parameters (Wilson et al., 2012). From these data, we defined three predictor variables for modelling. *DeciduousForest* was represented by the NLCD Deciduous Forest category and *EvergreenForest* was calculated as the sum of the NLCD Evergreen Forest and Mixed Evergreen Forest categories. *ForestDiversity* was estimated by calculating Shannon diversity values from FIA-inferred species-specific basal areas.

Statistical Analysis

Statistical analysis was conducted using R v4.3.2 (R Core Team, 2021). Following taxonomic sequence annotation, we aggregated all detections across the two genetic markers, producing a single table of presence-absence values. Leaf surface eDNA samples were then compared against flower eDNA samples using a Chi-square test to evaluate differences in frequency of eDNA detection per sample and a Wilcoxon Rank Sum test to evaluate differences in genus-level eDNA richness per sample. These tests were conducted for all Insects, as well as Anthophila and Lepidoptera, two taxa enriched for species of conservation concern. Since bee richness and frequency of detection was expected to be greater in flower eDNA relative to leaf surface eDNA samples, one-tailed tests were used when conducting analyses for Anthophila.

To evaluate landscape-scale diversity trends, we analyzed eDNA-derived richness estimates as a function of local forest cover and forest tree species diversity. For these analyses, a starting generalized linear mixed model was constructed by regressing eDNA richness per sample against the proportion of *DeciduousForest*, *EvergreenForest* and *ForestDiversity*. Terms for sampling week (*W*) to account for systematic trends in insect diversity over the sampling period and mean proportion of Arthropoda eDNA sequences (*D*), weighted by the total study-wide detections for each genetic marker to account for how intensively the sample was interrogated during sequencing, were included. To account for repeated measures, a random intercept term for substrate was added (Equation 1). For flower eDNA samples, substrate labels indicated the plant family from which the sample originated, while leaf surface samples were given a single unique factor label. *EvergreenForest* (3-km scale) values were z-scaled to circumvent model convergence issues. A Poisson error distribution was specified, and reverse stepwise selection was conducted where the least significant of all non-significant ($\alpha = 0.05$) features was removed during each modelling round. Feature selection was first conducted with landscape predictors calculated at the 3-km scale, using the Watanabe-Akaike information criterion (WAIC) to select the best-supported model. We then tested the final model using predictors calculated at the 1- and 0.5-km scales. Following selection, each model was tested for overdispersion. Where necessary, confidence values were adjusted using recommendations associated with the lme4 package (Bates et al., 2015). Modelling was conducted for Arthropoda, Hymenoptera and Lepidoptera, all of which exhibited >400 genus-level detections.

Equation 1) $$Richness \sim N(a_{j[i]}, \sigma^2),$$ where

$$a_j \sim N(\beta_0+\beta_1(W)+ \beta_2(D)+\beta_3(DeciduousForest)+\beta_4(EvergreenForest)+\beta_5(ForestDiversity), \sigma^2_{a_j})$$

For *Substrate* j = 1 to j = i.

Results

During field sampling, 236 flower and 21 leaf surface eDNA samples were collected. Following sequencing, 28S and COI libraries were sequenced to mean depths of 38,280 and 32,329 reads per sample, respectively (Figure 2A). The 28S marker showed moderate specificity for Arthropoda amplification, with a median of 10.9 percent of sequences per sample belonging to Arthropoda. However, COI exhibited considerable non-target amplification and a median of only 0.4 percent of sequences per sample represented Arthropoda (Figure 2B), suggesting that the blocking primers employed during amplification were not effective in improving reaction specificity. Similarly, while a median of 4.79 percent of total 28S sequences could be identified to an arthropod genus, only 0.08 percent of COI sequences were identifiable to this level (Figure 2C).

In total, eDNA analysis of the collected samples produced 2,228 Arthropoda detections spanning 543 genera (Figure 3A). A large proportion of detections were represented by the orders Hymenoptera, Lepidoptera and Coleoptera. Hemiptera, Thysanoptera, Diptera and Araneae detections also featured prominently in the data. Similar patterns were observed regarding genus-level diversity of detections (Figure 3B), and prevalence of detection across samples (Figure 3C). Regarding Lepidoptera and Anthophila, groups of heightened conservation concern, eDNA yielded a rich dataset with 558 Lepidoptera detections (21 families spanning 209 genera) 260 bee detections (5 families spanning 32 genera).

In comparing trends across flower surface and leaf surface eDNA, we found no significant differences in the frequency of detection (P = 0.840) or genus-level richness per sample (P = 0.900) when analyzing all Arthropoda detections (Figure 4A). Similarly, we found no effects of substrate type when comparing frequency of detection (P = 0.754) or genus-level richness per sample (P = 0.489) when evaluating Lepidoptera (Figure 4B). However, we found an increased frequency of detection (P = 0.028) and increased genus richness per sample (P = 0.009) when evaluating results for Anthophila (Figure 4C).

During analysis of the associations between total arthropod richness and landscape-scale measures of forest habitat, initial model selection using a 3-km radius scale for forest habitat covariates resulted in the removal of *EvergreenForest* and yielded the model shown in Equation 2, where all predictors were found to be significant at an α of 0.05. When assessing the optimal landscape scale for modelling arthropod richness, the 3-km radius estimates were more strongly supported relative a 0.5-km (ΔWAIC = -14.50) or 1-km (ΔWAIC = -10.19) radius. There was significant evidence in support of all four predictors within this final model: *W* (β = -0.040, *P* = 0.012), *D* (β = 0.261, *P* < 4.10e-12), *DeciduousForest* (Figure 5A, β = 0.857, *P* = 0.007), *ForestDiversity* (Figure 5B, β = 0.754, *P* = 0.018). Additional evaluation of this model revealed strong predictive power (Nakagawa and Schielzeth's Conditional pseudo-R^2 = 0.721) and no evidence of variance inflation due to multi-collinearity (VIF < 1.05 for all predictors).

Equation 2) $$Richness \sim N(a_{j[i]}, \sigma^2), \text{ where}$$

$$a_j \sim N(\beta_0+\beta_1(W)+ \beta_2(D)+\beta_3(DeciduousForest)+\beta_4(ForestDiversity), \sigma^2_{a_j})$$

For *Substrate* j = 1 to j = i.

At the order level, analysis of Lepidoptera richness values revealed some of the largest effect sizes in the data. Model selection at a 3-km radius scale resulted in the removal of *EvergreenForest*, yielding the model shown in Equation 2. When assessing the optimal landscape scale for modelling Lepidoptera richness values, the 3-km radius estimates were more strongly supported relative to the 0.5-km (ΔWAIC = -57.38) or 1-km (ΔWAIC = -31.33). In the final model, *W* (β = 0.055, *P* = 0.055), *D* (β = 0.359, *P* = 2.40e-7), *DeciduousForest* (Figure 5C, β = 1.858, *P* = 0.001) and *ForestDiversity* (Figure 5D, β = 2.43, *P* = 2.80e-5) were all well-supported predictors of Lepidoptera richness within eDNA samples and follow-up evaluations revealed strong predictive power (Conditional pseudo-R^2 = 0.550) and no evidence of variance inflation (VIF < 1.05 for all predictors).

Modelling of Hymenoptera richness values resulted in the removal of *DeciduousForest* and *ForestDiversity*, yielding the model shown in Equation 3. Subsequent analysis supported a 0.5-km landscape scale relative to 1-km (ΔWAIC = -1.43) or 3-km (ΔWAIC = -2.95) scale. There was strong support for all three terms, *W* (β = -0.060, *P* = 9.30e-4), *EvergreenForest* (β = -0.859, *P* = 3.06e-3) and *D* (β = 0.290, *P* = 1.50e-11), and no evidence of variance inflation (VIF < 1.04 for all predictors).

Equation 3) $$Richness \sim N(a_{j[i]}, \sigma^2), \text{ where}$$

$$a_j \sim N(\beta_0+\beta_1(W)+\beta_2(D)+\beta_3(EvergreenForest), \sigma^2_{a_j})$$

For *Substrate* j = 1 to j = i.

Given the strength of associations between forest habitat covariates and Lepidoptera richness, an additional round of model selection was conducted to evaluate whether insect richness values were related to forest habitat when Lepidoptera detections were excluded. As previously described, using Equation 1 as a starting model with forest covariates estimated at a 3-km radius, reverse step-wise selection resulted in the removal of *ForestDiversity* and *EvergreenForest*, yielding the model shown in Equation 4. Evaluation of scale revealed that a 0.5-km radius was more strongly supported relative to a 1-km (ΔWAIC = -7.93) or 3-km (ΔWAIC = -6.48) radius for the remaining forest habitat covariate. Within the final model, predictors *W* (β = -0.076, *P* = 7.10e-7), *D* (β = 0.224, *P* < 3.80e-10) and *DeciduousForest* (β = 0.474, *P* = 0.02) were all significant predictors of non-Lepidoptera arthropod richness.

Equation 4) $$Richness \sim N(a_{j[i]}, \sigma^2), \text{ where}$$

$$a_j \sim N(\beta_0+\beta_1(W)+\beta_2(D)+\beta_3(DeciduousForest), \sigma^2_{a_j})$$

For *Substrate* j = 1 to j = i.

Discussion

Our study developed novel insect eDNA methods for terrestrial systems, where biodiversity is difficult to assess at large scales. Using eDNA, we show that insect richness patterns were strongly predicted by surrounding landscape features including forest extent, both deciduous and mixed evergreen, as well as overall tree species diversity. These results provide a quantitative perspective on the current value of forest ecosystems for arthropod conservation within Central Appalachia. Our conclusions also highlight a potential mechanism through which past and future alterations to forested systems may influence arthropod conservation outcomes. This work will improve our capacity to track how anthropogenic change influences insect populations and guide insect decline mitigation efforts.

Implementation of eDNA methods resulted in high densities of detection and demonstrated clear capacity for characterizing arthropod-habitat associations. With that said, future efforts are warranted to continue improving eDNA applications. Consistent with past studies (Avalos et al., 2024; Jones et al., 2024), our findings highlight the powerful impact

marker choice can have on breadth of detection. Even though there were approximately 2.5 times as many arthropod genera represented in the COI database relative to 28S, we detected 3 times as many arthropod genera using 28S relative to COI. This result is consistent with the observation that the proportion of 28S sequences belonging to Arthropoda was approximately an order of magnitude greater than that of COI. Ultimately, the BR2/BF2 COI primers showed minimal specificity for Arthropoda. Given the heightened power of the COI marker during species delimitation, development of COI primers with improved selectivity for Arthropoda would considerably improve methodological performance.

Relative to past arthropod eDNA studies, this work represents the largest sample size conducted to date (Table 1), reflecting the ease with which large numbers of samples can be collected and processed using the methods we describe. Unsurprisingly, this sampling effort resulted in more arthropod genera detected relative to past efforts. To produce a roughly standardized estimate of detection sensitivity, we divided the total arthropod genera detected in past arthropod eDNA studies by the square root of the sample size to account for the power-law relationship between sample size and species accrual (Macarthur & Wilson, 1967). In our data, this approach was strongly supported by a species accumulation analysis, where accumulated richness was linearly related to square root-transformed sample size ($R^2 = 0.99$, $P < 0.001$, OLS Regression). By this metric, our study, as well as two others (Allen et al., 2023; Stothut et al., 2024), stand out as highly sensitive implementations of arthropod eDNA surveys. Variation in performance across past works will likely guide future eDNA efforts towards greater levels of sensitivity. Of course, readers should note that these comparisons are only approximately standardized given differences in marker selection, bioinformatic processing and baseline sample diversity.

Moving beyond coarse landscape categorizations, eDNA methods show promise for revealing the impacts of a broad range of environmental changes in global forest ecosystems. In the last century, eastern U.S. forests have undergone substantial changes that are expected to alter insect habitat suitability. Fire suppression has contributed to mesophication, driving turnover from oak and hickory forests to maple and beech-dominated stands (Alexander et al., 2021; Nowacki & Abrams, 2008). Invasive species introductions have resulted in the functional eradication of previously abundant species, including American chestnut *(Castanea dentata*;

Hepting, 1974) and most species of ash (*Fraxinus* spp.; Herms & McCullough, 2014). Previous assessments suggest these trends have negatively impacted insect diversity (Sierzega & Eichholz, 2019; Wagner & Todd, 2016), yet our data indicate that Central Appalachian forests continue to exhibit heighted insect conservation potential. The Central Appalachian region remains one of the most intact, high-diversity forest systems in eastern U.S., with greater fire frequency and lower levels of mesophication relative to the broader region (Lafon et al., 2005; Nowacki & Abrams, 2008) and follow-up studies comparing these forests to those of the eastern U.S. more broadly are warranted.

Resolution of disagreements about the magnitude and complexity of insect decline requires monitoring to be conducted at large scales. We argue that eDNA methods could be central to achieving this goal. Large-scale studies would ideally be conducted using probability sampling techniques (Deming, 1975; Firth & Bennett, 1998) within an occupancy modelling framework to account for the potential of systematic bias in sampling location or detectability (MacKenzie et al., 2006). The cost-efficient and easily democratized nature of our methods increases the feasibility of scaling insect monitoring to meet these demands. Further validation of the accuracy and limits of eDNA methods against alternate sampling protocols is warranted, and we expect to see continued improvements to the sensitivity and comprehensiveness of future eDNA implementations.

Acknowledgements

Efforts were primarily supported by a DoD ESTCP Grant to RTR and KG (Project RC22-B5-7373). Additional support was provided by Western EcoSystems Technology, Inc. and Metamorphic Ecological Research and Consulting, LLC. We thank USFWS, USFS, VA DCR-DNH, MD Forest Service, MD Park Service, MD DNR, WV DNR, WV State Parks and PA Game Commission. The authors thank Katherine Turo and John Lloyd for helpful manuscript reviews.

Supporting Data

All sequence data and relevant summary tables needed to reproduce these analyses will be made publicly available through zenodo.org upon publication.

References

Alexander, H. D., Siegert, C., Brewer, J. S., Kreye, J., Lashley, M. A., McDaniel, J. K., Paulson, A. K., Renninger, H. J., & Varner, J. M. (2021). Mesophication of Oak Landscapes: Evidence, Knowledge Gaps, and Future Research. BioScience, 71(5), 531–542.

Allen, M. C., Lockwood, J. L., Kwait, R., Vastano, A. R., Peterson, D. L., Tkacenko, L. A., Kisurin, A., Stringham, O., Angle, J., & Jaffe, B. (2023). Using surface environmental DNA to assess arthropod biodiversity within a forested ecosystem. Environmental DNA, 5(6), 1652–1666.

Avalos, G., Trott, R., Ballas, J., Lin, C. H., Raines, C., Iwanowicz, D., Goodell, K., & Richardson, R. T. (2024). Prospects of pollinator community surveillance using terrestrial environmental DNA metagenetics. Environmental DNA, 6(1), e492.

Barnes, M. L., Zhang, Q., Robeson, S. M., Young, L., Burakowski, E. A., Oishi, A. Christopher., Stoy, P. C., Katul, G., & Novick, K. A. (2024). A Century of Reforestation Reduced Anthropogenic Warming in the Eastern United States. Earth's Future, 12(2), e2023EF003663.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 1, JSSv067i01.

Bell, K. L., Campos, M., Hoffmann, B. D., Encinas-Viso, F., Hunter, G. C., & Webber, B. L. (2024). Environmental DNA methods for biosecurity and invasion biology in terrestrial ecosystems: Progress, pitfalls, and prospects. Science of The Total Environment, 926, 171810.

Bocinsky, R. (2024). FedData: Functions to Automate Downloading Geospatial Data Available from Several Federated Data Sources. https://CRAN.R-project.org/package=FedData

Bohmann, K., & Lynggaard, C. (2023). Transforming terrestrial biodiversity surveys using airborne eDNA. Trends in Ecology & Evolution, 38(2), 119–121.

Bracalini, M., Balacenoiu, F., & Panzavolta, T. (2024). Forest health under climate change: Impact of insect pests. iForest - Biogeosciences and Forestry, 17(5), 295–299.

Crossley, M. S., Meier, A. R., Baldwin, E. M., Berry, L. L., Crenshaw, L. C., Hartman, G. L., Lagos-Kutz, D., Nichols, D. H., Patel, K., Varriano, S., Snyder, W. E., & Moran, M. D. (2020). No net insect abundance and diversity declines across US Long Term Ecological Research sites. Nature Ecology & Evolution, 4(10), 1368–1376.

Darby, B., Bryant, R., Keller, A., Jochim, M., Moe, J., Schreiner, Z., Pratt, C., Euliss, N. H., Jr., Park, M., Simmons, R., & Otto, C. (2020). Molecular sequencing and morphological identification reveal similar patterns in native bee communities across public and private grasslands of eastern North Dakota. PLOS ONE, 15(1), e0227918.

De Lombaerde, E., Vangansbeke, P., Lenoir, J., Van Meerbeek, K., Lembrechts, J., Rodríguez-Sánchez, F., Luoto, M., Scheffers, B., Haesen, S., Aalto, J., Christiansen, D. M., De Pauw, K., Depauw, L., Govaert, S., Greiser, C., Hampe, A., Hylander, K., Klinges, D., Koelemeijer, I., … De Frenne, P. (2022). Maintaining forest cover to enhance temperature buffering under future climate change. Science of The Total Environment, 810, 151338.

Deming, W. E. (1975). On Probability as a Basis for Action. The American Statistician, 29(4), 146–152.

Dewitz, J. (2021). National Land Cover Database (NLCD) 2019 Products (ver. 3.0, February 2024). U.S. Geological Survey data release.

Didham, R. K., Basset, Y., Collins, C. M., Leather, S. R., Littlewood, N. A., Menz, M. H. M., Müller, J., Packer, L., Saunders, M. E., Schönrogge, K., Stewart, A. J. A., Yanoviak, S. P.,

& Hassall, C. (2020). Interpreting insect declines: Seven challenges and a way forward. Insect Conservation and Diversity, 13(2), 103–114.
Eddy, S. R. (2011). Accelerated Profile HMM Searches. PLOS Computational Biology, 7(10), e1002195.
Elbrecht, V., & Leese, F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. Frontiers in Environmental Science, 5, 11.
Esling, P., Lejzerowicz, F., & Pawlowski, J. (2015). Accurate multiplexing and filtering for high-throughput amplicon-sequencing. Nucleic Acids Research, 43(5), 2513–2524.
Firth, D., & Bennett, K. E. (1998). Robust models in probability sampling. Journal of the Royal Statistical Society: Series B, 60(1), 3–21.
Fürst, J., Bollmann, K., Gossner, M. M., Duelli, P., & Obrist, M. K. (2023). Increased arthropod biomass, abundance and species richness in an agricultural landscape after 32 years. Journal of Insect Conservation, 27(2), 219–232.
Gamonal Gomez, N., Sørensen, D. H., Chua, P. Y. S., & Sigsgaard, L. (2023). Assessing flower-visiting arthropod diversity in apple orchards through metabarcoding of environmental DNA from flowers and visual census. Environmental DNA, 5(1), 117–131.
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), e0185809.
He, C., Liu, Z., Gou, S., Zhang, Q., Zhang, J., & Xu, L. (2019). Detecting global urban expansion over the last three decades using a fully convolutional network. Environmental Research Letters, 14(3), 034008.
Hekkala, A.-M., Jönsson, M., Kärvemo, S., Strengbom, J., & Sjögren, J. (2023). Habitat heterogeneity is a good predictor of boreal forest biodiversity. Ecological Indicators, 148, 110069.
Hepner, M. J., Orcutt, E., Price, K., Goodell, K., Roulston, T., Jean, R. P., & Richardson, R. T. (2024). Montane Central Appalachian forests provide refuge for the critically endangered rusty patched bumble bee (Bombus affinis). Forest Ecology and Management, 556, 121751.
Hepting, G. H. (1974). Death of the American Chestnut. Journal of Forest History, 18(3), 61–67.
Herms, D. A., & McCullough, D. G. (2014). Emerald Ash Borer Invasion of North America: History, Biology, Ecology, Impacts, and Management. Annual Review of Entomology 59(59), 13–30.
Jackson, S. T., & Sax, D. F. (2010). Balancing biodiversity in a changing environment: Extinction debt, immigration credit and species turnover. Trends in Ecology & Evolution, 25(3), 153–160.
Jactel, H., Koricheva, J., & Castagneyrol, B. (2019). Responses of forest insect pests to climate change: Not so simple. Current Opinion in Insect Science, 35, 103–108.
Jeffries, J. M., Marquis, R. J., & Forkner, R. E. (2006). Forest Age Influences Oak Insect Herbivore Community Structure, Richness, and Density. Ecological Applications, 16(3), 901–912.
Johnson, M. D., Katz, A. D., Davis, M. A., Tetzlaff, S., Edlund, D., Tomczyk, S., Molano-Flores, B., Wilder, T., & Sperry, J. H. (2023). Environmental DNA metabarcoding from flowers reveals arthropod pollinators, plant pests, parasites, and potential predator–prey

interactions while revealing more arthropod diversity than camera traps. Environmental DNA, 5(3), 551–569.
Jones, D. N., Augustine, B. C., Hutchins, P., Birch, J. M., Yamahara, K. M., Jensen, S., Richardson, R., Trott, R., Campbell, J. R., Barnhart, E., & Sepulveda, A. J. (2024). Autonomous samplers and environmental DNA metabarcoding: Sampling day and primer choice have greatest impact on fish detection probabilities. Metabarcoding and Metagenomics, 8, e122375.
Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research, 30(14), 3059–3066.
Kestel, J. H., Bateman, P. W., Field, D. L., White, N. E., Lines, R., & Nevill, P. (2023). eDNA metabarcoding of avocado flowers: 'Hass' it got potential to survey arthropods in food production systems? Molecular Ecology Resources, 23(7), 1540–1555.
Knuff, A. K., Winiger, N., Klein, A.-M., Segelbacher, G., & Staab, M. (2019). Optimizing sampling of flying insects using a modified window trap. Methods in Ecology and Evolution, 10(10), 1820–1825.
Kuhlman, M. P., Burrows, S., Mummey, D. L., Ramsey, P. W., & Hahn, P. G. (2021). Relative bee abundance varies by collection method and flowering richness: Implications for understanding patterns in bee community data. Ecological Solutions and Evidence, 2(2), e12071.
Lafon, C. W., Hoss, J. A., & Grissino-Mayer, H. D. (2005). The Contemporary Fire Regime of the Central Appalachian Mountains and its Relation to Climate. Physical Geography, 26(2), 126–146.
Lark, T. J., Spawn, S. A., Bougie, M., & Gibbs, H. K. (2020). Cropland expansion in the United States produces marginal yields at high costs to wildlife. Nature Communications, 11(1), 4295.
Macarthur, R. H., & Wilson, E. O. (1967). The Theory of Island Biogeography. Princeton University Press.
MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2006). Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence. USGS Publications Warehouse.
Maixner, F., Mitterer, C., Jäger, H. Y., Sarhan, M. S., Valverde, G., Lücker, S., Piombino-Mascali, D., Szikossy, I., Molnár, E., Pálfi, G., Pap, I., Cipollini, G., & Zink, A. (2022). Linear polyacrylamide is highly efficient in precipitating and purifying environmental and ancient DNA. Methods in Ecology and Evolution, 13(3), 653–667.
Martínez-Núñez, C., Gossner, M. M., Maurer, C., Neff, F., Obrist, M. K., Moretti, M., Bollmann, K., Herzog, F., Knop, E., Luka, H., Cahenzli, F., & Albrecht, M. (2024). Land-use change in the past 40 years explains shifts in arthropod community traits. Journal of Animal Ecology, 93(5), 540–553.
Montgomery, G. A., Belitz, M. W., Guralnick, R. P., & Tingley, M. W. (2021). Standards and Best Practices for Monitoring and Benchmarking Insects. Frontiers in Ecology and Evolution, 8.
Newton, J. P., Bateman, P. W., Heydenrych, M. J., Kestel, J. H., Dixon, K. W., Prendergast, K. S., White, N. E., & Nevill, P. (2023). Monitoring the birds and the bees: Environmental DNA metabarcoding of flowers detects plant–animal interactions. Environmental DNA.

Nowacki, G. J., & Abrams, M. D. (2008). The Demise of Fire and “Mesophication” of Forests in the Eastern United States. BioScience, 58(2), 123–138.
Ober, H. K., & Hayes, J. P. (2008). Influence of forest riparian vegetation on abundance and biomass of nocturnal flying insects. Forest Ecology and Management, 256(5), 1124–1132.
Ohio Supercomputer Center. (1987). [Computer software]. http://osc.edu/ark:/19495/f5s1ph73
Packer, L., Gibbs, J., Sheffield, C., & Hanner, R. (2009). DNA barcoding and the mediocrity of morphology. Molecular Ecology Resources, 9, 42–50.
Portman, Z. M., Bruninga-Socolar, B., & Cariveau, D. P. (2020). The state of bee monitoring in the United States: A call to refocus away from bowl traps and towards more effective methods. Annals of the Entomological Society of America, 113(5), 337–342.
R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/
Richardson, R. T. (2022). Controlling critical mistag-associated false discoveries in metagenetic data. Methods in Ecology and Evolution, 13(5), 938–944.
Richardson, R. T., Curtis, H. R., Matcham, E. G., Lin, C.-H., Suresh, S., Sponsler, D. B., Hearon, L. E., & Johnson, R. M. (2019). Quantitative multi-locus metabarcoding and waggle dance interpretation reveal honey bee spring foraging patterns in Midwest agroecosystems. Molecular Ecology, 28(3), 686–697.
Richardson, R. T., Sponsler, D. B., McMinn-Sauder, H., & Johnson, R. M. (2020). MetaCurator: A hidden Markov model-based toolkit for extracting and curating sequences from taxonomically-informative genetic markers. Methods in Ecology and Evolution, 11(1), 181–186.
Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. PeerJ, 4, e2584.
Rosenheim, J. A., & Ward, P. S. (2020). Insect diversity over 36 years at a protected Sierra Nevada (California) site: Towards an evaluation of the insect apocalypse hypothesis. Ecological Entomology, 45(6), 1490–1494.
Sánchez-Bayo, F., & Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. Biological Conservation, 232, 8–27.
Sherrill-Mix, S. (2019). taxonomizr: Functions to Work with NCBI Accessions and Taxonomy. https://CRAN.R-project.org/package=taxonomizr
Sierzega, K. P., & Eichholz, M. W. (2019). Linking conservation implications of modified disturbance regimes, plant communities, plant associations, and arthropod communities. Oecologia, 189(1), 267–277.
Song, X.-P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., & Townshend, J. R. (2018). Global land change from 1982 to 2016. Nature, 560(7720), 639–643.
Staab, M., Gossner, M. M., Simons, N. K., Achury, R., Ambarlı, D., Bae, S., Schall, P., Weisser, W. W., & Blüthgen, N. (2023). Insect decline in forests depends on species’ traits and may be mitigated by management. Communications Biology, 6(1), 338.
Stothut, M., Kühne, D., Ströbele, V., Mahla, L., Künzel, S., & Krehenwinkel, H. (2024). Environmental DNA metabarcoding reliably recovers arthropod interactions which are frequently observed by video recordings of flowers. Environmental DNA, 6(3), e550.
Tepedino, V. J., Durham, S., Cameron, S. A., & Goodell, K. (2015). Documenting bee decline or squandering scarce resources. Conservation Biology, 29(1), 280–282.

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.

Tronstad, L., Bell, C., & Crawford, M. (2022). Choosing collection methods and sample sizes for monitoring bees. Agricultural and Forest Entomology, 24(4), 531–539.

Uhl, B., Schall, P., & Bässler, C. (2024). Achieving structural heterogeneity and high multi-taxon biodiversity in managed forest ecosystems: A European review. Biodiversity and Conservation.

van Klink, R., August, T., Bas, Y., Bodesheim, P., Bonn, A., Fossøy, F., Høye, T. T., Jongejans, E., Menz, M. H. M., Miraldo, A., Roslin, T., Roy, H. E., Ruczyński, I., Schigel, D., Schäffler, L., Sheard, J. K., Svenningsen, C., Tschan, G. F., Wäldchen, J., … Bowler, D. E. (2022). Emerging technologies revolutionise insect ecology and monitoring. Trends in Ecology & Evolution, 37(10), 872–885.

Wagner, D. L., Grames, E. M., Forister, M. L., Berenbaum, M. R., & Stopak, D. (2021). Insect decline in the Anthropocene: Death by a thousand cuts. Proceedings of the National Academy of Sciences, 118(2), e2023989118.

Wagner, D. L., & Todd, K. J. (2016). New Ecological Assessment for the Emerald Ash Borer: A Cautionary Tale About Unvetted Host-Plant Literature. American Entomologist, 62(1), 26–35.

Williams, J. W. (2022). RAD: A Paradigm, Shifting. BioScience, 72(1), 13–15.

Wilson, B. T., Lister, A. J., & Riemann, R. I. (2012). A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. Forest Ecology and Management, 271, 182–198.

Tables

Table 1: Comparison of current peer-reviewed implementations of arthropod eDNA monitoring.

Study	Sample size (N)	Mean sequencing depth	Fold difference in sequencing depth relative to this study	Total arthropod genera	Total genera / sqrt(N)
This study, both markers	257	70,608	1.0X	543	33.87
This study, 28S only	257	38,280	0.5X	419	26.14
Allen et al. (2023)	40	174,911	2.5X	353	55.81
Stothut et al. (2024)	31[a]	32,073	0.5X	156	28.02
Thomsen & Sigsgaard (2019)	57	5,029,448[b]	71.2X	113	14.97
Avalos et al. (2024)	27	73,447	1.0X	58	11.16
Johnson et al. (2023)	37[a]	2,779,670	39.4X	68	11.18
Newton et al. (2023)	30[a]	35,814[b]	0.5X	47	8.58
Kestel et al. (2023)	80	192,080	2.7X	33	3.69
Gamonal Gomez et al. (2023)	60	22,049	0.3X	10	1.29

[a] Some samples failed to amplify.
[b] Reported sequencing depth does not include sequences which were removed during quality filtering and is likely approximately 50 percent of actual sequencing depth.

Figures

Figure 1: Map of eDNA sampling locations (A), and photos illustrating the sample collection and preservation methods used (B).

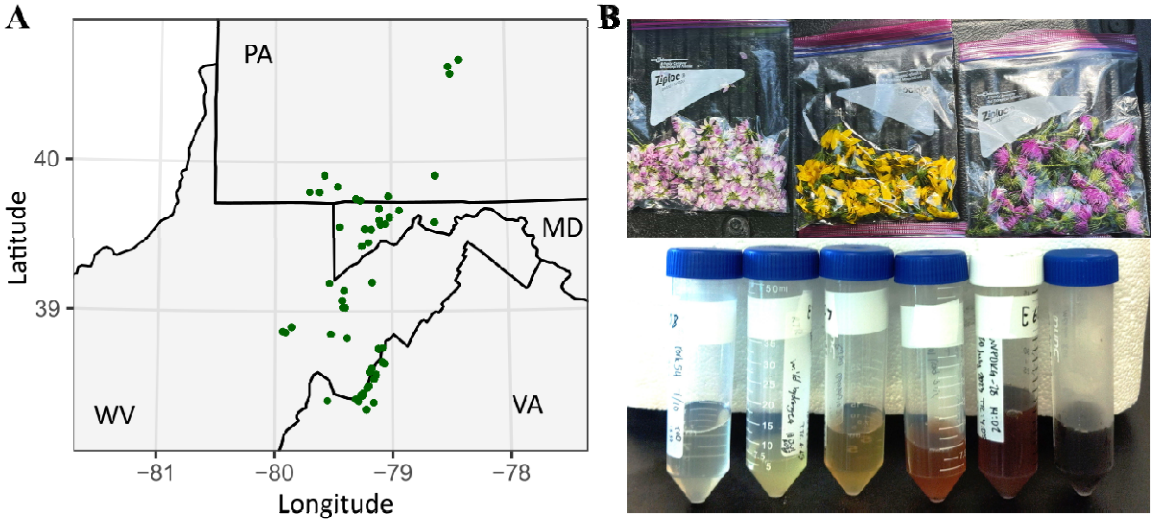


Figure 2: Analysis of per sample sequencing depths achieved (A), proportion of sequences identified as Arthropoda (B) and proportion of sequences identified to an arthropod genus (C) for both markers processed in this study.

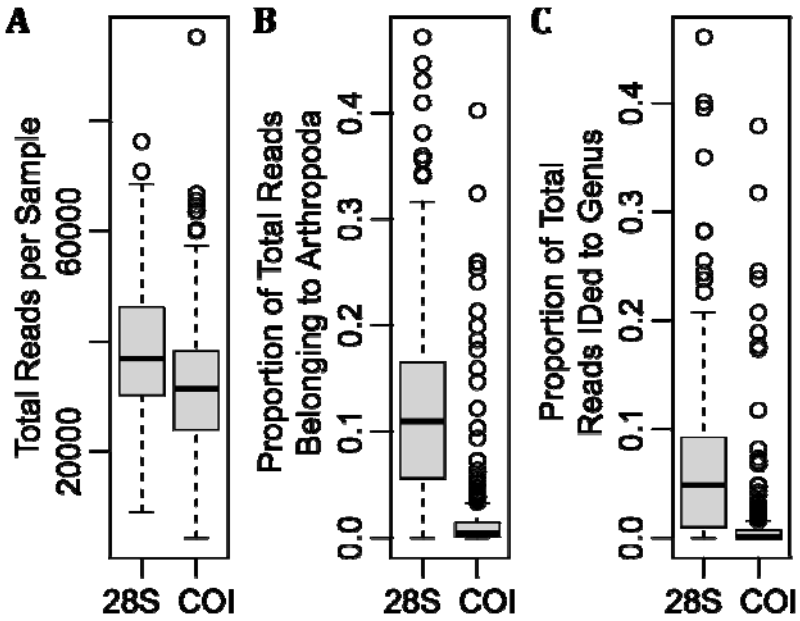


Figure 3: Summaries of study-wide genus-level eDNA detections (A), genus-level richness (B) and detection prevalence (C) across all arthropod orders detected, partitioned by flower surface and leaf surface samples.

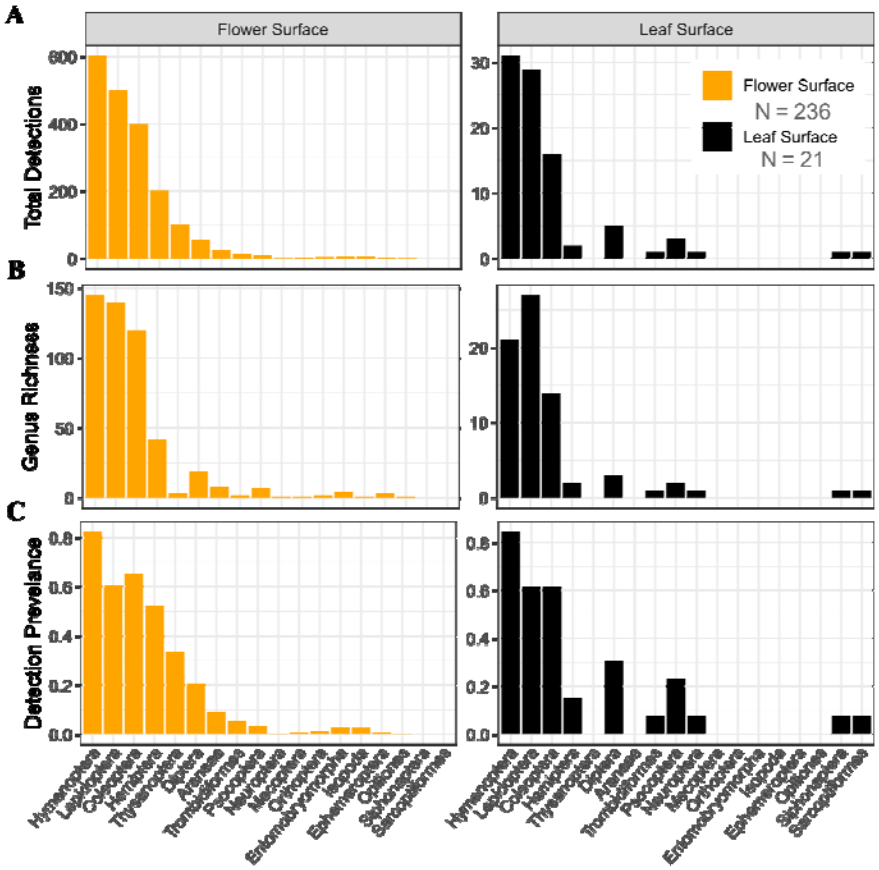


Figure 4: Comparison of flower and field negative eDNA samples. No statistical differences were observed in terms of frequency of detection of genus-level detection richness when conducting analysis for all Arthropoda (A) or all Lepidoptera (B). However, flower eDNA samples exhibited a significantly greater proportion of Anthophila (bee)-positive samples and greater richness of bee genera per sample relative to field negative controls (C).

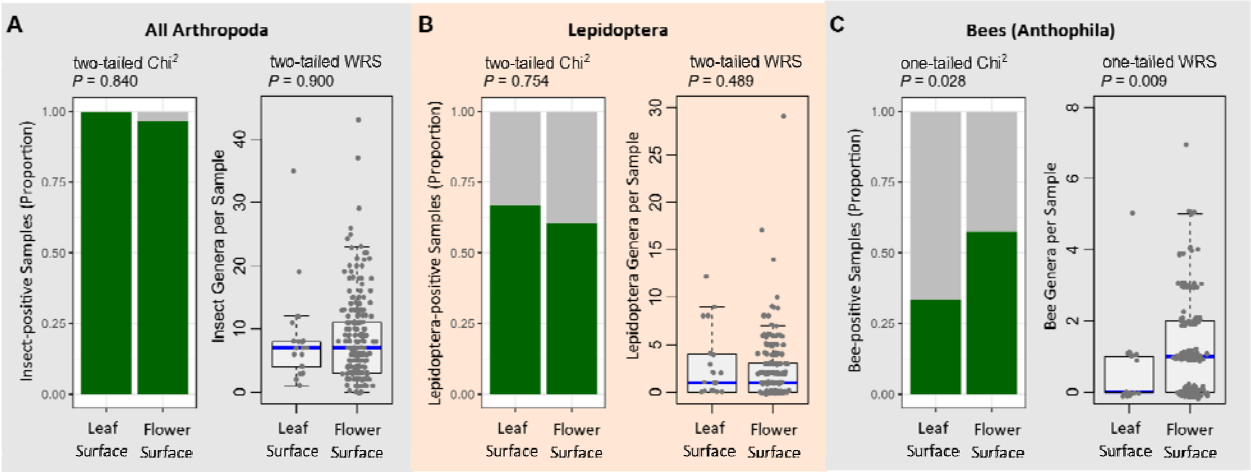
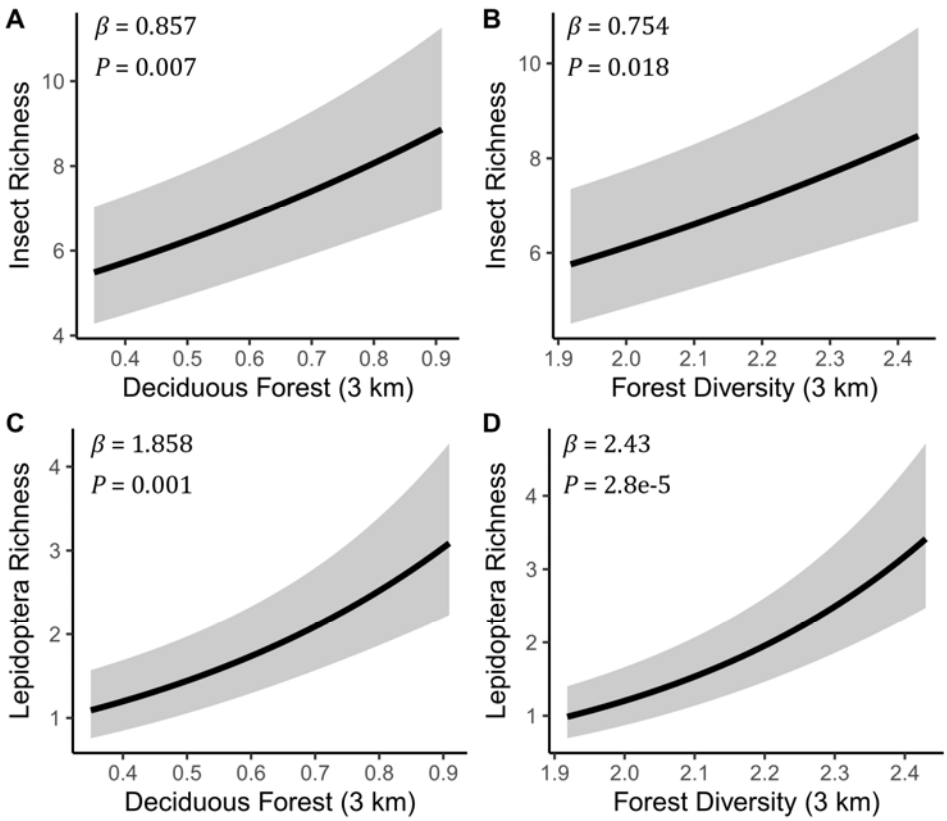


Figure 5: Partial effect plots of the best-fit relationships between eDNA-inferred sample richness values and measures of forest extent and diversity. Total insect richness was found to be significantly associated with the proportion of deciduous forest (A) and forest Shannon diversity (B). Among individual insect orders, Lepidoptera richness exhibited similar, yet stronger, associations with the proportion of deciduous forest (C) and forest Shannon diversity (D). The final WAIC-selected model for both total arthropod richness and Lepidoptera richness values is shown in Equation 2.



Supplemental Table S1: Blocking primers designed to suppress non-target fungal amplification.

Primer Name	Sequence
FUN_BLCKNG323F	GTTTAATMCTATTCTTATTTGCWGC[SpcC3]
FUN_BLCKNG326F	TAAYYCTACTACTWGCTAGYGC[SpcC3]
FUN_BLCKNG398F	CAGGWVTACAAAGYCACWSWGG[SpcC3]
FUN_BLCKNG503F	TMCTRAAYATGMGAIIYYCWGG[SpcC3]