ABSTRACT

Title of Thesis:

REFINING METAGENETIC ENVIRONMENTAL DNA TECHNIQUES FOR SENSITIVE BEE COMMUNITY MONITORING

Grace Avalos, Master of Science, 2023

Thesis Directed By:

Rodney Richardson, Assistant Research Professor, Appalachian Laboratory, University of Maryland Center for Environmental Science

Molecular taxonomic detection is now widespread across the sciences, because of advances in direct PCR, improved marker selection, and increases in sequencing throughput. Facilitated by these advances in sequencing, methodological sensitivity of sample identification has improved substantially. Metagenetic techniques to infer what species are present in a sample by sequencing unknown samples and comparing them to known references has the potential to advance our understanding of biodiversity. Metagenetic analysis of environmental DNA (eDNA) represents a novel, non-lethal method for characterizing floral-associated arthropod communities. Diverse arthropod assemblages interact with flowers, and floral surfaces have been shown to harbor arthropod DNA. We performed metagenetic sequencing on eDNA isolated from flower samples and honey bee-collected pollen samples using multiple markers and compared the frequency and taxonomic breadth of eDNA detections across these genetic markers and substrate types. Understanding which markers and substrates are most effective for eDNA characterization of floral-associated arthropod communities will guide future research and enable low-risk detection of threatened or endangered arthropods.

REFINING METAGENETIC ENVIRONMENTAL DNA TECHNIQUES FOR SENSITIVE BEE COMMUNITY MONITORING

by

Grace Avalos

Thesis submitted to Faculty of the Graduate School of the

University of Maryland, College Park, in partial fulfillment

of the requirements for the degree of

Master of Science

2023

Advisory Committee:

Assistant Research Professor Rodney T. Richardson, Chair

Associate Director for Research and Professor Matthew C. Fitzpatrick

Associate Research Professor Tsvetan R. Bachvaroff

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to Environmental DNA Monitoring of Pollinators

The development of molecular taxonomic identification

The sequencing of deoxyribonucleic acid, or DNA, as a means of identifying the taxonomic origins of a sample represents a widespread application across the biological sciences, from criminal forensics (Norrgard, 2008) to lepidopteran systematics (Hebert et al., 2010). In the past two decades, advances in next-generation sequencing (NGS) platforms have revolutionized these methods, resulting in decreasing costs and increasing throughput and accuracy. Aided by the development of polymerase chain reaction (PCR; Mullis & Faloona, 1987) and universal primers for various branches of the tree of life (Hebert et al., 2003), the direct sequencing of taxonomically-informative PCR products is now readily applicable to numerous areas of ecological research (Townson et al., 1999; Taberlet et al., 2012; Pereira et al., 2008). One such application, DNA barcoding, as first defined by Hebert et al. (2003), involves the use of a specific genetic marker to identify the taxonomic origin of a single specimen or sample. The method typically employs Sanger sequencing to generate a singular DNA sequence reference for precise identification. In contrast, DNA metabarcoding expands upon this concept by employing high-throughput amplicon sequencing to analyze diverse samples. This approach allows for the bulk identification of numerous species within a single sample, aiming to estimate species composition. Metagenetics, within the context of molecular biology, extends beyond metabarcoding by utilizing high-throughput amplicon sequencing with multiple markers to analyze diverse samples. This broader approach enables the taxonomic classification of communities detected within one or multiple samples, providing a more comprehensive understanding of the genetic diversity present in environmental or biological samples.

These molecular taxonomy methods are particularly useful when handling samples that would ordinarily be hard to identify, such as pollen samples (Rowney et al., 2021), cryptic species, or those that are too degraded for identification (e.g. stomach contents or fecal matter). Altogether, high-throughput richness and distribution assessments can facilitate research on a wide range of taxa while enabling relatively greater sampling intensity.

Versatility of environmental DNA and its terrestrial usage

Environmental DNA, or eDNA, is genetic material shed from organisms into the environment which can be isolated and characterized using modern genetic methods (Bienert et al., 2012). By conducting metabarcoding on eDNA, researchers can characterize trace eDNA to perform high-throughput taxonomic assessments from environmental samples. eDNA can exist in many different substrates, such as soil, water, or sediments in the environment (Creer et al., 2016). The lifespan of molecularly usable eDNA depends on characteristics of the organism such as age, feeding activity, physiology, life history, and space use (Barnes & Turner, 2016; Goldberg et al., 2016; Hering et al., 2018). Abiotic factors, such as temperature, pH, or UV-exposure can also influence the lifespan of molecularly usable eDNA (Creer et al., 2016). Historically, eDNA has been applied most frequently in aquatic environments, where the ease of taking large volumes of water samples and multiple field replicates has allowed for numerous studies characterizing fish community compositions across marine and freshwater environments (Jerde et al., 2011; Thomsen et al., 2012). One of the first demonstrated uses of eDNA for a terrestrial species was a 2008 single-species detection study in which environmental DNA in water samples was used to target an invasive species, the American bullfrog, from laboratory as well as field experiments in France (Ficetola et al., 2008). Plant and animal DNA left in the environment by rare species may

represent one of the few pieces of evidence about their presence within a study site, and as such, eDNA can serve several uses for research on these taxa.

In terrestrial systems, eDNA can be used for multiple ecological applications. For surveying low-density populations, eDNA can be used to detect groups that are visually evasive or rare, including endangered species. In situations where detection is difficult using traditional sampling methods, eDNA is a potential avenue toward improving detection sensitivity, aiding in the characterization of current species distributions. For protected species in particular, the minimal disturbance and non-lethality of eDNA methods represents an advantage over traditional surveys.

eDNA is also useful for the detection of exotic or invasive species. A recent example of this application involved the use of eDNA surveys to predict occupancy and detection rates of *Lycorma delicatula*, an invasive ecological pest also known as the spotted lanternfly (Allen et al., 2021). Detection of eDNA fragments within ecosystems can be used to solve forensic questions or those related to ecosystem conditions. eDNA can be used by wildlife managers to detect whether invasive species are present in an area (Morisette et al., 2021). For suspected poaching or illegal wildlife trade, wildlife law enforcement have addressed the need for DNA forensic tools, and eDNA has the potential to provide information for such situations. Recent work on the Eurasian badger, a legally protected species in the UK, provides an example of this (Ogden et al., 2008).

Usage of eDNA is not restricted to gathering current-day data and can be used to improve our understanding of current and trophic interactions. For example, ancient eDNA contained in permafrost soils has been used to reconstruct plant and animal communities of the past (Sønstebø

et al., 2010) and similar inferences can be gained from the large proportions of ancient organisms lacking fossilized remains but which instead left extracellular DNA traces within sediments (Pederson et al., 2015). With these methods, the need to traditionally identify the remains off degraded samples, such as feces, stomach contents, or sediment particles can be avoided by utilizing these molecular techniques which have the potential to reduce bias and improve taxonomic resolution.

The most commonly used applications for eDNA are community-wide surveys designed to characterize patterns of biodiversity and population dynamics. While we do not expect eDNA methods to supplant traditional techniques, eDNA approaches represent a promising method for surveying broad groups of taxa quickly and cost-effectively. The taxonomic information returned by sequencing eDNA can span many taxa, depending on the markers and primers used, while a taxonomic specialist is usually limited to knowledge of a specific taxonomic clade. This often results in biodiversity assessments being implemented at coarser taxonomic levels than ideal when using traditional techniques (Elbrecht et al., 2017). For population assessments, some studies have found that rank abundance corresponds to biomass within eDNA data (Bohmann et al., 2014). One study found that by taking eDNA samples from western Lake Erie clams and snails placed into aquaria, sequenced read abundances reflected relative measured abundances (Klymus, Marshall, & Stepien. 2017).

eDNA methods can be subject to contamination and bioinformatic concerns. For this reason, the incorporation of appropriate controls into study designs is imperative. Both negative field and lab controls lend themselves to ensure valid sampling and handling has occurred, and to identify

potential contamination (Goldberg et al., 2016). From a bioinformatic perspective, there are several concerns that can arise along the workflow, some of which include false positive and false negative assignments. Cross-validation techniques, which involve splitting datasets into training and testing sets, to be used for evaluation of classification performance, can aid researchers in assessing the accuracy of bioinformatic analyses, reducing the risk of false positive and false negative results in eDNA assignments. Additionally, continuous efforts to improve reference sequence databases, encompassing a comprehensive array of taxa and genetic markers, can minimize bioinformatic challenges and improve the reliability of eDNA results. Researchers should remain vigilant in implementing studies with robust designs and employing rigorous bioinformatic methods, recognizing that addressing these challenges is integral to ensuring the validity of eDNA-based biodiversity assessments.

Terrestrial invertebrate and arthropod monitoring with eDNA

As terrestrial eDNA applications have evolved, so have the substrates from which eDNA samples have been collected. Early on, soil, water, and sediments represented the most common substrates for collection of eDNA. More recently, eDNA has been collected from a wide array of substrate types, such as vegetation surfaces (Thomsen & Sigsgaard., 2019; Valentin et al., 2018), tea leaves (Krehenwinkel et al., 2022), and even airborne samples (Lynggaard et al., 2022). It is known that in microbial community studies, the DNA extraction methods impact the community composition detected (Delmont et al., 2011; Kang & Mills, 2006). Terrestrial invertebrates comprise some of the most species-rich and diverse communities, and as such, we can anticipate that eDNA extraction steps for these taxa will need to be tailored to analyze community compositions as completely as possible. For groups undergoing decline, it becomes increasingly

urgent to determine the most efficient molecular approaches for obtaining rapid and useful data on these communities.

Arthropods denote the metazoan group most undescribed to science (Stork, 2018) and the majority of species within the group have yet to be assessed ecologically. With global declines exacerbated by factors such as habitat loss (Baillie et al., 2012), extensive efforts are being placed on conserving arthropod biodiversity. To prioritize these efforts, widespread and sensitive ecosystem monitoring is needed. For terrestrial arthropods, the majority are still monitored using traditional surveying techniques, and incorporation of novel methods such as eDNA have only recently begun. Traditionally, assessments of terrestrial arthropod community composition have been studied using pitfall traps, netting, and Malaise traps, among other methods (O'Connor et al., 2019; Campbell & Hanula, 2007). These methods can harm habitat or individuals, which is especially detrimental when monitoring for threatened or endangered groups (Tepedino & Portman, 2021; Montero-Castaño et al., 2022). Additionally, many of the methods to survey these taxa are associated with species biases, often through observers not being able to detect certain groups, or traps being biased toward capturing or excluding certain taxa (Pei et al., 2022). With such limited knowledge of existing arthropod species, there is a need for research on how molecular techniques could supplement or supplant traditional survey methods.

The need for pollinator eDNA monitoring

In addition to conservation concerns, economics also motivate the need for improved surveillance techniques of pollinators. Pollinators support numerous ecosystem services and are vital to global ecosystems and economies. In 2005, pollination was estimated at 9.5% of global agricultural production of human food (Khalifa et al., 2021). From an ecological standpoint,

reduced plant reproductive success has been connected to declines in pollinator biodiversity (Thomann et al., 2013). Further, numerous pollinator species have received a threatened or endangered listing under the United States Endangered Species Act. For these species, the use of invasive or risk-prone survey methods is an important consideration. This is particularly relevant for certain pollinator species which can only be identified under a microscope, typically requiring researchers to take lethal samples. Given this, eDNA represents a viable alternative to conducting community composition studies of pollinators.

Pollinators and many arthropods are known to interact with plants, whether it be for herbivory, florivory, or egg laying. Through these interactions, pollinators deposit genetic material on plant matter through deposition of hair and feces, for example. A study by Thomsen and Sigsgaard (2018) affirmed this and detected a variety of floral-associated arthropod species utilizing eDNA from flower samples. These detections included species of Diptera, Hymenoptera, and Lepidoptera; all of which have known pollinator species. To some extent, eDNA has been demonstrated to provide information on community composition and species occurrence for pollinators without being invasive. However, across all current pollinator eDNA studies, none have achieved detection rates that would be comparable to traditional survey methods. Overall, metabarcoding-based eDNA methods need to yield similar or superior sensitivity at a same cost basis as traditional surveys. If eDNA cannot surpass traditional techniques in this way, eDNA will likely remain applicable as a means of low-risk detection for protected species.

Chapter 2: Prospects of pollinator community surveillance using terrestrial environmental DNA metagenetics

Abstract

Current pollinator survey methods exhibit bias, require highly trained practitioners, and are difficult to scale to large sample sizes. High-throughput sequencing of terrestrial eDNA could provide a complementary tool for studying pollinator communities, but eDNA methods have not been extensively evaluated. We conducted metagenetic analysis of eDNA from 20 flower and seven honey bee-collected pollen samples and compared eDNA-derived data with traditional netting-based surveys of the pollinator communities present during sampling. We detected eight bee genera belonging to four families across COI, 16S, and 28S markers, but results varied considerably by marker and eDNA substrate. Detected bee genera were plausible for the study system and about 43 percent of total bee genera were detected with both eDNA and net-based surveys, though netting resulted in more detections across a wider diversity of genera. Data from sequenced controls suggest that eDNA identifications were unlikely to have resulted from cross-contamination. Our results demonstrate that bee communities can be documented with eDNA techniques and that the choice of marker and substrate substantially influence detection. Future improvements to our methods are required, but eDNA surveys appear well-suited to characterize diverse pollinator communities and provide novel sampling perspectives within plant-pollinator networks. Future efforts should focus on improving the selection of markers available for pollinator eDNA metagenetics, addressing taxonomic gaps within reference sequence databases and optimizing sampling and eDNA isolation protocols. We anticipate that such improvements are highly feasible and that eDNA will be a useful tool to those who study pollinators and plant-pollinator interactions.

Introduction

Aside from moral conservation concerns, changes in the abundance and distribution of insect species can influence critical ecosystem services, such as pollination (Cameron et al., 2011; Potts et al., 2010). Such changes in insect communities warrant attention from researchers, policy makers and land managers. Sensitive detection of rare pollinator species is an important goal in conservation research (Novotny et al., 2021; Strange & Tripodi, 2019; Woodard et al., 2020). Population monitoring surveys are important for tracking the status of declining species and mitigating threats. However, surveying rare pollinators is challenging since they generally exhibit low detection probability. Current methods for surveying pollinator communities have substantial limitations and pose risks to target taxa. Aerial netting requires high in-field effort and composition results vary depending on the surveyor. Passive traps, such as pan traps and blue vane traps, require less in-field effort but exhibit strong taxonomic bias for small-bodied, generalist bee taxa and are lethal (Cane et al., 2000; Portman et al., 2020; Prendergast et al., 2020). Passive trap efficacy is also influenced by environmental factors and local floral resource abundance (Kuhlman et al., 2021). Numerous, small bodied bees from passive trapping can cause bottlenecks in taxonomic identification, resulting in excess lethal take and delaying research progress (Tepedino & Portman, 2021). New pollinator survey methods that are non-lethal, sensitive, unbiased across surveyors, and scalable are needed to improve pollinator monitoring research.

High-throughput sequencing of terrestrial eDNA could provide a complimentary tool for studying pollinator communities. Within aquatic systems, eDNA-based metabarcoding and metagenetic studies have shown heightened sensitivity for species detection relative to the

traditional methods previously mentioned (Ficetola et al., 2008; Jerde et al., 2011). Less research has been conducted in terrestrial systems (Valentin et al., 2020), but recent works suggest that eDNA deposited on leaves and flowers can be used for detection of various species of Arthropoda (Allen et al., 2021; Johnson et al., 2023; Krehenwinkel et al., 2022). Thomsen & Sigsgaard (2019) provided the first proof-of-concept work documenting the application of eDNA metabarcoding to the study of floral-associated arthropod communities. However, only a single bee species (Hymenoptera; Anthophila) was detected out of 135 arthropod species across 56 flower eDNA samples, suggesting a low detection sensitivity for bee taxa. Subsequent recent studies have struggled to detect representative pollinator eDNA (Gamonal Gomez et al., 2023; Harper et al., 2023). These past works provide strong evidence that a wide diversity of invertebrates leave behind detectable DNA in the environment and also suggest that new methods and metagenetic infrastructure are needed for sensitive eDNA-based detection.

With this study, we place considerable focus on bees because they comprise some of the most important pollinators in many ecosystems and many species are of conservation interest. To aid in pollinator ecology and conservation, eDNA-based detection methods need to be accurate, sensitive, and economically scalable to large sample sizes. Compared to traditional netting, we expect eDNA methods to be more taxonomically comprehensive on a temporal scale, as plant-pollinator networks are known to be temporally dynamic (Olesen et al., 2008; Vaudo et al., 2014) and eDNA can persist on floral surfaces for days (Valentin et al., 2021). As a result, eDNA collected samples are an integrated sampling method over an extended time period, whereas net-sampling provides a ‘snapshot’ view of community composition over more limited timespans.

With regard to sampling substrate, bee-collected pollen could represent an ideal substrate for eDNA-based surveys of pollinators. Over the course of a single foraging season, honey bee colonies are capable of collecting approximately 1.3 million corbicular pollen loads, collected during foraging bouts up to 10 km from the colony (Seeley, 1995, 1997). Presumably, many pollinator species are interacting with flowers foraged by honey bees (Hung et al., 2018), potentially facilitating widespread aggregation of pollinator eDNA within a single substrate. Here we harness advances in metagenetic sequencing to test the application of eDNA-based analysis of flower samples and honey bee-collected pollen samples for surveying flower-visiting pollinator communities, with an emphasis on bees. To verify eDNA results, we compared the taxa detected using eDNA to netted flower visitors from the same time period and locations. During metagenetic analysis, we amplified the eDNA with three sets of universal arthropod primers, targeting COI, 16S, and 28S regions. We addressed three objectives; (1) assess the sensitivity of three different genetic markers for characterizing pollinator communities at the genus level (2) test whether honey bee-collected pollen eDNA provides for more sensitive pollinator detections relative to flower eDNA samples and (3) compare the results of eDNA-derived pollinator detections against traditional aerial net-derived detections.

Methods

Collection of eDNA substrates and molecular processing

Isolates of eDNA from 20 flower samples and 7 honey bee-collected pollen samples were analyzed. All samples were collected from two central Ohio sites during August and September of 2020. Flowers were collected on three dates at each site: August 20th, August 26th and September 1st at Don Scott Airport, Columbus, Ohio and August 19th, August 25th and

September 4th at the Intermediate School Land Lab, Granville, Ohio. Honey bee pollen samples were obtained using a Sundance pollen trap (Rossman Apiaries, Moultrie, Georgia) bottom mounted to two healthy and actively foraging honey bee colonies at each location. These pollen traps non-lethally collects pollen from the hind legs of honey bees as they enter the colony. Pollen samples were collected on August 20th, August 26th, and September 5th at Don Scott Airport, and on August 14th, August 24th August 31st, and September 6th at the Intermediate School Land Lab. Full details of collected samples can be found in Supporting Information Table S1. Pollen traps were active for six to ten day intervals prior to each collection event. We selected abundant plant species in full bloom that were visited by honey bees and other bee species to maximize the probability that bee eDNA would be present on the flowers. A gloved researcher used bleach-sterilized scissors to sample one inflorescence per plant, carefully avoiding contact with individual flowers, from plants chosen haphazardly throughout the area where flower visitors were observed. For each sample, flowers were collected until a one quart Ziploc bag was approximately half full of loose flower volume. Instruments were re-sterilized between flower species to avoid cross-contamination. Flower samples from each species were sealed into a plastic bag and held on ice. All samples were placed in a -20C freezer within 5 hours of collection.

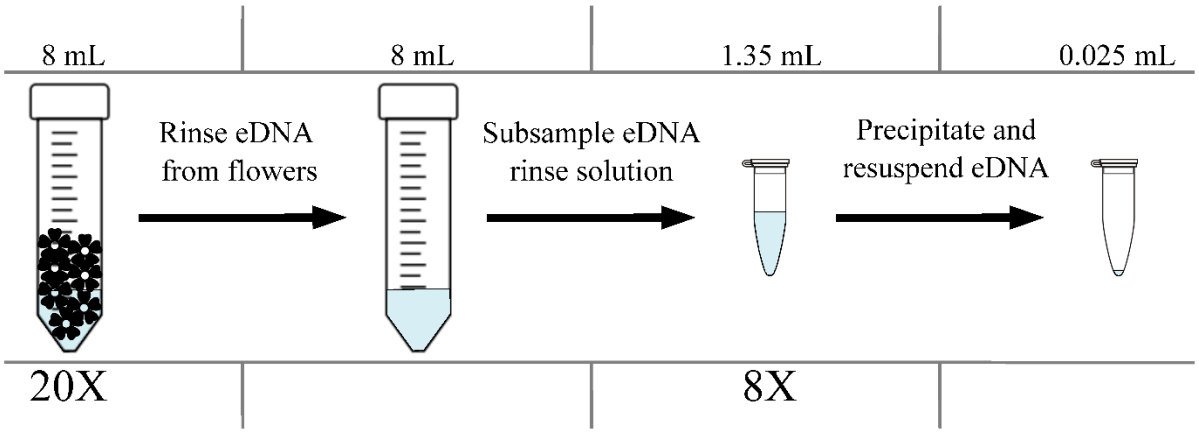


Figure 1: Illustration of sample wash procedure used to obtain surface eDNA from flower samples. Numbers at the top of the illustration show volumes used in this study. Numbers along the bottom indicate our expectations of the scalability of these methods, e.g. 20 times the volume of loose flowers could be used as input and eDNA could be isolated from all eDNA wash solution as opposed to the approximately 1/8th subsamples used here.

In the lab, flowers were carefully inspected under a dissecting microscope to remove any visible invertebrates, such as mites and thrips, with tools and working surfaces decontaminated using 10 percent bleach or flame sterilization between each flower sample. A sample wash procedure was established to obtain eDNA from the flower substrate, using an approach scalable for an increased volume of substrate as needed (Figure 1). Flower surface eDNA was obtained by gently agitating 25 mL of loose flower heads in 8 mL of 90 percent ethanol. Pollen surface eDNA was obtained by agitating 15 mL of honey bee pollen in 25 mL of 80 percent ethanol and allowing pollen to settle out of solution for 5 min. For both sample substrates, 1.35 mL of sample rinse solution was mixed with 100 µL of 3.2 M NaCl and 1 µL of 20 mg/mL glycogen. These reagents decrease the solubility of DNA and act as a DNA co-precipitant, respectively. Samples were chilled at -20°C overnight and eDNA was precipitated by centrifugation at 20,000× g for 30 min at 4°C. Supernatant was discarded and pelleted eDNA was subjected to PCR-based amplicon library preparation using the Phire Tissue Direct PCR Kit (Invitrogen, Waltham, MA).

	Primer sequence	Primer Name	Source
COI	GCHCCHGAYATRGCHTTYCC	BF2	Elbrecht & Leese, 2017
	TCDGGRTGNCCRAARAAYCA	BR2	Elbrecht & Leese, 2017
16S	TARTYCAACATCGRGGTC	Chiar16SF	Marquina et al., 2019
	CYGTRCDAAGGTAGCATA	Chiar16SR	Marquina et al., 2019
28S	GTGAAACCGTTCAGGGGTAAACC	BD0152	Darby et al., 2020
	GGTGTTTCAAGACGGGTCCTG	BD0153	Darby et al., 2020

Table 1: Primers used for amplification of each genetic marker.

Library preparation from eDNA isolates followed Richardson et al. (2019), using a 3-step PCR protocol designed to minimize taxonomic amplification biases (Berry et al., 2011; O'Donnell et al., 2016). Initial PCR reactions were performed using previously published COI, 16S, and 28S primer sets and gradient PCR-determined annealing temperatures (Table 1). Initial PCR reactions for COI were conducted with the following cycling conditions; initial denaturation step at 98°C for 5 min, then for 35 cycles: 98 °C for 5s, 52.5 °C for 5s, 72 °C for 20s; final extension step at 72 °C for 1 min. For 28S and 16S, the PCR reactions were as follows; Initial denaturation step at 98°C for 5 min, then for 35 cycles: 98 °C for 5s, 50 °C for 5s, 72 °C for 20s; final extension step at 72 °C for 1 min. Subsequent PCR reactions were conducted to append Illumina-specific oligos for dual-indexing and lane hybridization, as in Kozich et al. (2013). Between each of the three reactions, 1 µL of unpurified PCR product was transferred as template. Final library products were analyzed using gel electrophoresis to ensure the presence of appropriately sized amplicons. Libraries were then purified and normalized with the SequalPrep Normalization Plate Kit

(Invitrogen, Waltham, MA) and sequenced using a 2x250 cycle Illumina MiSeq Nano flow cell. A separate flow cell was used for each marker.

In addition to experimental eDNA sample libraries, positive controls and no-library negative controls (i.e., unused dual-index tag combinations) were included within each sequencing run to estimate mistagging and laboratory contamination rates. Sequencing of COI libraries included four positive controls of Silver-haired bat DNA template and five no-library negative controls, while 16S and 28S both used European honey bee DNA template for positive controls. With the exception of the Silver-haired bat control samples, where DNA was extracted with a proprietary kit, positive control samples underwent the same steps as field eDNA samples. 16S libraries included three positive controls and six no-library negative controls, and 28S libraries included two positive controls and seven no-library negative controls. Control samples were randomly distributed among experimental samples during processing.

Traditional netting surveys

We assessed the flower visitor community on each day that flowers were collected by netting all flower visitors to seven different focal plant species per site for which eDNA samples were also collected. These focal plant species were abundant at each site and were actively visited by bees. We conducted all insect and flower collections on fair weather days that had little or no cloud cover, temperatures between 22°C and 28°C, and little to no wind. A researcher netted continuously on a single flower species for 10 minutes. Each flower species was netted for three 10-min periods on dates when it had abundant bloom (range 1 – 4 dates per species) for a total of 690 min of netting across both sites (Supporting Information Table S2). All insects were transferred to vials to be identified. We transferred all bee specimens to a cooler, in addition to

all other insects that could not be identified to genus in the field. These remaining specimens were euthanized by freezing and identified using reference specimens and a variety of taxonomic keys and websites (e.g., Dankowicz and Dankowicz 2022, Eaton and Kaufman 2007, Discoverlife.org). Bees were identified to the lowest taxonomic level possible, usually to species. Lepidoptera and Coleoptera were identified to genus level, unless specimens were damaged or cryptic. Hemiptera, non-bee Hymenoptera, and Diptera were identified to family.

Bioinformatic and statistical analysis

After sequencing, forward and reverse reads from COI and 16S were merged and priming sites were removed using VSEARCH (v2.8.1, Rognes et al. (2016)). Since the length of 28S amplicons exceeded 500 bp, merging was skipped and forward 28S sequences were quality trimmed using a Phred score of 20 prior to priming site removal. Arthropod reference databases for each marker were produced following the MetaCurator workflow(v1.0.1; Richardson et al. (2020)) which utilizes HMMER (v3.1; Eddy (2011)), MAFFT (v7.270; Katoh et al. (2002)) and VSEARCH as dependencies. Specific search terms, used to target available arthropod sequences, were used to download reference data from NCBI Nucleotide database at the start of 2022 for each marker (Supporting Information Table S3) and a subsequent taxonomy file was created and formatted using the Taxonomizr R package (Sherrill-Mix, 2019). Semi-global VSEARCH alignment was performed to query eDNA sequences against curated reference databases trimmed to the precise amplicon regions of interest, requiring at least 80 percent query cover. Following alignment, we focused our analysis on genus-level detections with an identity match of 96 percent or greater. A sequence abundance table was then generated for each marker. Using the statistical test outlined in Richardson (2022), abundance tables were filtered to remove detections with greater than five percent chance of representing a critical mistag-associated false detection.

All computational analysis was performed on the Owens cluster of the Ohio Supercomputer Center (1987).

To assess differences among markers and substrates, we compared the genus-level richness of eDNA detections per sample across each of the three markers, as well as between pollen and flower substrates using non-parametric Kruskal-Wallis tests. These tests were performed for four taxonomic groupings, Arthropoda, Lepidoptera, Hymenoptera and Anthophila. For comparisons with significant Kruskal-Wallis results, pair-wise comparisons were conducted using *post hoc* Wilcoxon pairwise comparisons. For these comparisons by substrate and marker, we conducted an additional analysis by randomly sampling 550 sequences per sample. This standardizes the comparison at a fixed sequencing depth but is shown as supporting information since there may be systematic differences in sequencing efficiency across markers. For example, gel electrophoresis of the 28S products often revealed bands from non-target amplicons, which may explain the decreased average sequencing depth per sample for 28S libraries. Kruskal-Wallis tests were used to compare richness per sample between eDNA and traditional net surveys. Lastly, we compared the ranked experiment-wide abundance, based on total netting specimens, of genera detected using both methods to the ranked abundances of genera detected with netting alone using a Kruskal-Wallis X^2 approximation. All analyses were conducted in R v4.1 (R Core Team, 2021).

Results

Taxonomic representation of available reference data

The number and richness of reference sequences present in the curated databases varied considerably by marker (Table 2). The COI reference database had the highest total number of

reference sequences as well as highest taxonomic completeness within Arthropoda at the order and family ranks. In contrast, the 28S reference database had the lowest number of reference sequences and taxonomic completeness. Nevertheless, representation of Anthophila genera was highest within the 28S database and lowest within the 16S database. Across all three marker databases, the representation of Anthophila ranged from 76 to 385 genera and included all six North American families.

	Total References	Orders	Families	Anthophila Families	Anthophila Genera
COI	606,750	117	2,095	6	272
16S	72,939	101	1,564	6	76
28S	44,660	96	1,338	7	385

Table 2: Summary of taxonomic completeness of the MetaCurator-generated reference database for each eDNA marker.

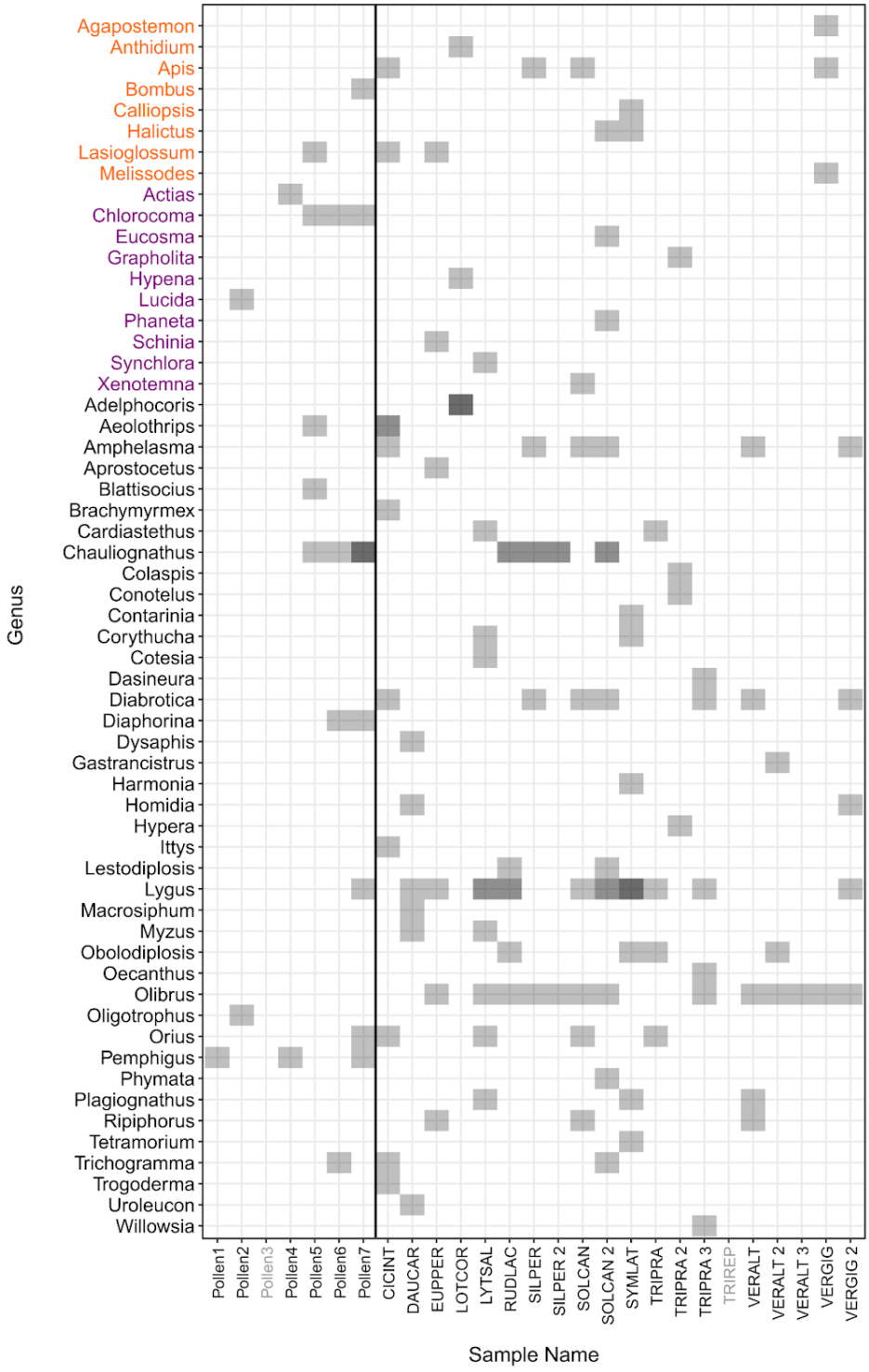


Figure 2: Tile plot of genus detections across all samples and markers, with a black line separating pollen and flower eDNA samples. Light grey tile color denotes one marker detection, medium grey tile color denotes detection by two markers, dark grey tile color denotes detection by all three markers. Genera in orange font represent Anthophila taxa while genera in purple font represent Lepidoptera. *Apis* detections were manually removed from honey bee-collected pollen

eDNA samples. X-axis labels in light grey indicate samples for which not arthropod eDNA was detected. Flower species abbreviations correspond to species listed in Supporting Information Table S1.

	Sequences Per Sample	Proportion Arthropoda	Proportion Anthophila
COI	13,972 (5,669 - 33,144)	0.4481 (1,919 - 14,698)	0.0002 (0 - 43)
16S	16,549 (562 – 42,235)	0.2053 (14 – 12,555)	0.0012 (0 - 45)
28S	7,125 (1,700 – 14,700)	0.5208 (2 – 13,485)	0.0810 (0 – 2,345)

Table 3: Average of sequences per sample and proportion of sequences belonging to Arthropoda and Anthophila with the range of sequences per sample shown in parenthesis (N = 27 for all markers). Estimates do not include control samples.

	Arthropoda Genera	Lepidoptera Genera	Hymenoptera Genera	Anthophila Genera
COI	16	6	1	1
16S	20	0	2	2
28S	33	4	13	7

Table 4: Summary of detected genera belonging to Arthropoda and Anthophila for each eDNA marker.

Quality of eDNA data

There was a limited range of reads per sample, and number of reads varied from marker to marker (Table 3). Sequencing of COI, 16S and 28S libraries yielded totals of 684,333, 812,820, and 485,919 reads, respectively. Following quality control filtering, the 16S marker had the highest average sequences per sample while 28S exhibited the lowest sequencing depth, with a per sample average of approximately half the coverage of COI and 16S (Table 3). However, 16S exhibited the lowest proportion of sequences belonging to Arthropoda, 20.53 percent. The 28S marker exhibited the highest proportion of sequences belonging to both Arthropoda and Anthophila, 52.08 percent and 8.10 percent, respectively. Across all three markers, there was a

total of 58 Arthropoda genera and 8 Anthophila genera detected (Figure 2). The 28S marker yielded the majority of both Arthropoda and Anthophila detections (Table 4). Notably, three samples contained detections belonging to two closely related Anthophila genera that appeared difficult to distinguish with 28S, *Seladonia* and *Halictus* in flower sample SOLCAN2; *Melissodes* and *Eucera* in flower sample VERGIG; and *Anthidium* and *Pseudoandthidium* in flower sample LOTCOR. To err on the side of caution, we removed the *Seladonia, Eucera* and *Pseudoandthidium* from the dataset prior to statistical analysis given that the alternate genera are relatively more common within the study system.

Traditional net survey results

Net surveys were performed on a total of 14 flower species, and each species was surveyed for an average of 49.8 person-minutes (Range 30 - 120, SD = 30, Supporting Information Table S4). Surveys yielded a total of 1,159 insect visitors, including 609 bee detections (52.5%) and 327 Lepidoptera detections (28%). Out of all detected Arthropoda, these detections encompassed four Anthophila families and 12 Anthophila genera, along with seven Lepidoptera families and 11 Lepidoptera genera. Honey bees were observed visiting all plants sampled except *Symphyotrichum lateriflorum*.

Comparison of eDNA markers

Across the eDNA dataset, significant differences in genus richness per sample were found between markers. The richness of Anthophila genera per sample was not different between markers ($H = 4.197$, df = 2, $p = 0.123$), but richness of all Arthropoda genera per sample varied significantly ($H = 6.053$, df = 2, $p = 0.048$). Post-hoc Wilcoxon pairwise comparisons revealed more Arthropoda genera detections using 28S relative to COI ($Z = 2.382$, $p = 0.017$) with a

marginal difference between 16S and 28S (Z = -1.724, p = 0.085). Hymenoptera genus richness also exhibited at least one difference by marker (H = 8.569, df = 2, p = 0.014), with 28S exhibiting significantly greater richness than COI (Z = 2.921, p = 0.003). Similarly, Lepidoptera genus richness values significantly varied by marker (H = 9.863, df = 2, p = 0.007), with COI exhibiting significantly greater richness relative to 16S (Z = -3.116, p = 0.002) and marginally greater richness relative to 28S (Z = -1.894, p = 0.058). Standardizing the sequencing depth per sample through random sampling resulted in similar interpretations (Supporting Information Figure S1).

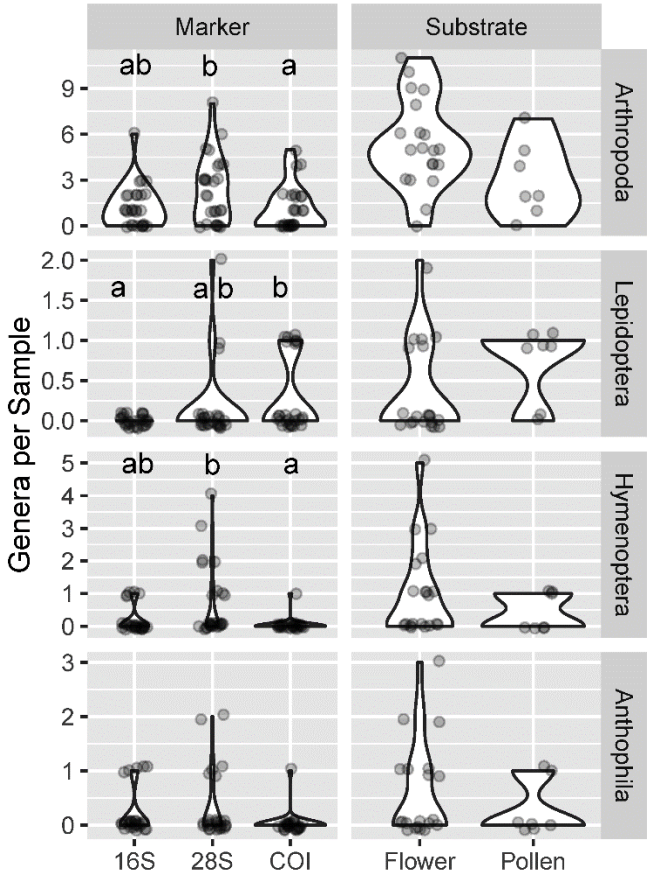


Figure 3: Violin plots depicting the richness of genera detected by marker and substrate. For comparisons exhibiting at least one significant difference, groups with different letters were found to be significantly different according to post hoc pairwise comparison. The richness of eDNA detections was summed across markers for comparison by substrate.

Comparison of eDNA substrates

In comparing eDNA-inferred genus richness per sample, evidence for variance by substrate was marginal (Figure 3). The strongest evidence for differences between pollen and flower eDNA sources were observed for the taxonomic groups Arthropoda ($H = 3.275$, $df = 1$, $p = 0.070$) and Lepidoptera ($H = 2.897$, $df = 1$, $p = 0.089$). No evidence was found when comparing results across substrates for Hymenoptera ($H = 0.573$, $df = 1$, $p = 0.449$) or Anthophila ($H = 0.505$, $df = 1$, $p = 0.477$). As with comparison across markers, standardization of sequencing depth per sample resulted in the same interpretations relative to unstandardized data (Supporting Information Figure S1).

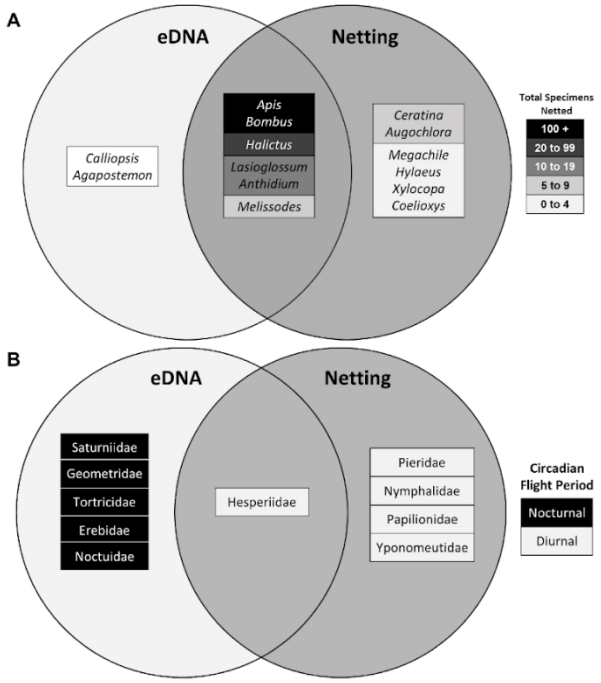


Figure 4: Bee genera (A) and Lepidopteran families (B) detected by molecular eDNA analysis, netting surveys, and both methods.

Comparison of eDNA with traditional netting

Traditional netting detected 12 Anthophila genera, in comparison to eight detected using eDNA methods, and a total of 14 with both methods combined. Accordingly, six Anthophila genera were detected with both methods (Figure 4A). The rank abundance of those detected by both methods was significantly higher than those detected solely by netting ($H = 7.05$, df = 1, $p = 0.008$; Figure 4A). However, traditional netting did detect significantly more genera per sample of Anthophila ($H = 52.55$, df = 1, $p < 0.001$), Arthropoda ($H = 22.595$, df = 1, $p < 0.001$), Hymenoptera ($H = 42.304$, df = 1, $p < 0.001$), and Lepidoptera ($H = 48.836$, df = 1, $p < 0.001$). Diptera detections did not differ significantly between the two sampling methods ($H = 0.5087$, df = 1, $p = 0.47$).

With regard to Lepidoptera, the two methods produced similar family richness, but overlapped little in taxa. Traditional netting detected five families, all with diurnally active adults. The eDNA methods detected six families, all but one of which were moths with nocturnally active adults (Supporting Information Table S5). The Lepidopteran family Hesperiidae, abundant and diurnally active skippers, was detected using both methods (Figure 4B). Despite a higher richness of Arthropoda and Anthophila from netting surveys, eDNA methods detected unique Anthophila and Lepidoptera taxa, increasing the richness and ecological diversity of the overall survey.

Discussion

Previous eDNA-based metagenetic studies have highlighted a new methodological avenue to characterize pollinator communities (Harper et al., 2023; Johnson et al., 2023; Krehenwinkel et al., 2022; Thomsen & Sigsgaard, 2019). This study adds to current knowledge by comparing

traditional netting-based surveys against eDNA analysis using three different molecular markers and two distinct sample substrates. We used eDNA methods that are feasibly scalable to considerably larger sample inputs which would be expected to yield eDNA extracts that are more representative of the broader floral-associated arthropod community (Figure 1). We also expect that these methods are cost-effectively applicable to large sample sizes (e.g. hundreds to thousands of samples). Among our data, we detected numerous arthropods, including multiple pollinating Anthophila and Lepidoptera genera. Notably, our work resulted in detection of numerous bee taxa within a small number of eDNA samples, with cases of common taxa, such as *Apis* and *Bombus,* as well as six other bee genera detected (Figure 4A). While more work is required to fully quantify the sensitivity of eDNA in detecting different taxa from flower and pollen samples, our work demonstrates the utility and complementarity of pollinator eDNA methods with more traditional netting surveys, providing a framework for future research.

Comparison of eDNA markers

Our results showed that detection of flower-associated arthropods varied by marker. We detected eight bee genera belonging to four families within flower and pollen eDNA samples. The detected genera were plausible given the study system and overlapped with those detected using net-based surveys. The majority of bee detections originated from 28S data. Among Lepidopteran detections, most were observed in COI data. Across the three markers tested, we found that 28S had the most frequent detections of Anthophila, Hymenoptera, and Arthropoda, though these trends were not always significant. To our knowledge, the 28S marker employed here has only been used for broad-scale metagenetic sequencing of pollinator communities in the Upper Midwest US (Darby et al., 2020). Our results suggest this marker represents a promising avenue toward future work with regard to Anthophila eDNA surveys, though a more thorough

quantification of species and genus-level resolution is needed in order to understand what level of accuracy can be expected across more diverse study systems.

Comparison of eDNA substrates

With the field of terrestrial eDNA research, the use of novel substrates and aggregation techniques exhibits potential for improving the sensitivity of eDNA-based surveys. Recent works have provided advances toward this goal with regard to sampling of eDNA from various sources such as spider webs (Gregorič et al., 2022), tea leaves and herbs (Krehenwinkel et al., 2022) and plant surfaces (Kirtane et al., 2022). Additionally, Valentin et al. (2020) have introduced novel methods for eDNA sample aggregation for plants and other terrestrial surfaces. Here, we provide the first implementation of eDNA surveillance of arthropods from pollen collected by honey bees. When comparing across substrates, flower and pollen samples performed similarly for detection of Arthropoda as a whole, but there was some indication of variance in efficacy across taxonomic groups. While we establish that pollen eDNA does enable detection of arthropods, including Anthophila and Lepidoptera, more work is needed to determine the full potential of this substrate. With the methods and markers employed here, eDNA results from bee-collected pollen did not exceed those from flower samples, which are easier to collect and process. Honey bees certainly interact with surfaces that are highly trafficked by other pollinators and inadvertently, accumulate some amount of eDNA during foraging. Whether this trace level of eDNA is sufficient for sensitive and broad-scale surveillance remains to be seen.

Comparison of eDNA with traditional netting

When compared to the results of our eDNA methods, traditional surveys yielded considerably greater Anthophila richness at the genus level, but detected taxa overlapped strongly with our

metagenetic results (Figure 4a). In contrast, eDNA methods yielded slightly more detections of Lepidopteran taxa (Figure 4b), and families detected were distinct from traditional surveys with the exception of Hesperiidae, common diurnally active skippers. All of the unique Lepidoptera detected with eDNA methods were families with nocturnally active adults (Supporting Information Table S5). The netting surveys, all conducted during the daytime, unsurprisingly detected only diurnal species, largely butterflies. This result suggests that eDNA surveys may be useful for monitoring nocturnal, matinal or crepuscular pollinator taxa not typically sampled during daytime flower surveys. Additionally, some of the detected genera include non-feeding adults, indicating that eDNA collected off plant components could potentially contain eggs, larval eDNA or previously airborne eDNA particles which settled onto sampled surfaces. Overall, netting visitors for 300-400 minutes per site resulted in detecting greater Anthophila richness relative to the current state of eDNA survey methods employed in this study. Future work is needed to quantify the relative cost and sensitivity of eDNA surveys as field-ready methods become available. However, given the economy of scale that can be achieved with metagenetic methods and expected increases in sensitivity, eDNA methods are likely to become more valuable over time.

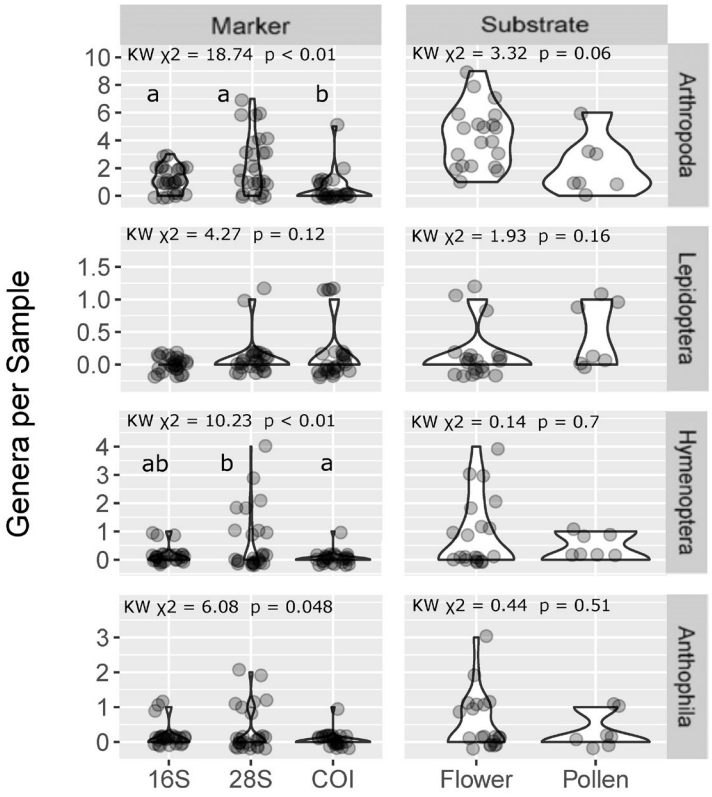
Conclusions and future directions

Through novel direct PCR-based methods, we demonstrate the application of eDNA metagenetics for broad-scale surveillance of pollinator communities. Our sample processing methods allow for greater sample inputs relative to previous works that have focused on individual florets (Thomsen & Sigsgaard, 2019; Harper et al., 2023;) or small volumes of flower wash buffer (Johnson et al., 2023). Increasing sample input mass can be accomplished by transitioning the method from a microcentrifuge to a high-volume (i.e. 50 mL conical) centrifuge

during precipitation steps. Additionally, these methods provide an alternative to laborious filtering of large-volume flower eDNA wash (Newton et al., 2023). Results indicate that the choice of metagenetic marker strongly influences the inferences that can be made and that bee-collected pollen contains traces of eDNA from the broader pollinator community. Despite these advances, our data generally suggest that future refinements will be needed before eDNA methods will be field ready as a sensitive complementary tool relative to traditional survey methods. We envision that primer sets with greater specificity to target taxa, e.g. Anthophila or Lepidoptera, will provide for some of the most impactful advances toward this goal. Further, the enhancement of reference sequence databases with new sequences spanning a greater diversity of taxa along with selection for high-resolution genetic markers such as COI and ITS1 (Hebert et al., 2003; Sheffield et al., 2009; Wang et al., 2015) will improve efforts toward species-level eDNA analysis.

Beyond serving as a new survey tool, eDNA metagenetics will provide a new and potentially more comprehensive assessment of plant-pollinator interaction networks relative to visual observations. Achievement of this goal will require mitigation of detection biases on the basis of pollinator or floral traits. Plant species traits such as floral structure, corolla length and tongue length, among others, will likely influence detectability. However, statistical methods exist for quantifying detectability and separating this process from occurrence patterns (MacKenzie et al., 2006). In the case of eDNA derived from bee-collected pollen, the unintentional amplification and sequencing of non-target taxa is the foremost obstacle apparent in our data. These obstacles are expected to be short-lived as methodological improvements arise. Whether used as a species surveillance method or applied more broadly to pollinator biology, we anticipate that eDNA metagenetics will be a valuable asset to pollinator research programs.

Appendix



Supporting Information Figure S1: Re-analysis of the dataset after standardizing for sequencing depth by randomly subsampling 550 sequences per sample.

Supporting Information Table S1: Individual data on all eDNA samples collected, along with species abbreviations.

Name	Species/Type	Abbreviation	Date Collected	Location	Landscape Type	Corresponding Net Sample?
P1_S17_L001	Pollen		08/20/2020	Don Scott	Agricultural pasture	NA
P2_S23_L001	Pollen		08/14/2020	Land Lab	Restored meadow	NA
P3_S29_L001	Pollen		08/26/2020	Don Scott	Agricultural pasture	NA
P4_S6_L001	Pollen		09/05/2020	Don Scott	Agricultural pasture	NA
P5_S12_L001	Pollen		08/24/2020	Land Lab	Restored meadow	NA
P6_S24_L001	Pollen		08/31/2020	Land Lab	Restored meadow	NA
P7_S30_L001	Pollen		09/06/2020	Land Lab	Restored meadow	NA
F1_S7_L001	*Lythrum salicaria*	LYTSAL	08/26/2020	Don Scott	Agricultural pasture	Y
F2_S13_L001	*Solidago canadensis*	SOLCAN	09/04/2020	Land Lab	Restored meadow	Y
F3_S19_L001	*Trifolium pratense*	TRIPRA	08/25/2020	Land Lab	Restored meadow	Y
F4_S2_L001	*Eupatorium perfoliatum*	EUPPER	08/26/2020	Don Scott	Agricultural pasture	Y
F5_S14_L001	*Silphium perfoliarum*	SILPER	08/25/2020	Land Lab	Restored meadow	Y
F6_S20_L001	*Verbesina alternifolia*	VERALT	09/01/2020	Don Scott	Agricultural pasture	Y
F7_S26_L001	*Verbesina alternifolia*	VERALT2	08/26/2020	Don Scott	Agricultural pasture	Y
F8_S32_L001	*Trifolium pratense*	TRIPRA2	08/18/2020	Land Lab	Restored meadow	Y
F9_S3_L001	*Daucus carota*	DAUCAR	08/20/2020	Don Scott	Agricultural pasture	Y
F10_S9_L001	*Lotus corniculatus*	LOTCOR	08/20/2020	Don Scott	Agricultural pasture	Y
F11_S15_L001	*Vernonia gigantea*	VERGIG	08/20/2020	Don Scott	Agricultural pasture	Y
F12_S21_L001	*Verbesina alternifolia*	VERALT3	08/20/2020	Don Scott	Agricultural pasture	Y
F13_S27_L001	*Silphium perfoliarum*	SILPER2	08/18/2020	Land Lab	Restored meadow	Y
F14_S33_L001	*Symphyotrichum lateriflorum*	SYMLAT	08/18/2020	Land Lab	Restored meadow	Y
F15_S4_L001	*Solidago rigida*	SOLRIG	09/04/2020	Land Lab	Restored meadow	Y
F16_S10_L001	*Vernonia gigantea*	VERGIG2	09/01/2020	Don Scott	Agricultural pasture	Y
F17_S16_L001	*Cichorium intybus*	CICINT	09/01/2020	Don Scott	Agricultural pasture	Y
F18_S28_L001	*Trifolium repens*	TRIREP	08/29/2020	Land Lab	Restored meadow	N
F19_S34_L001	*Rudbeckia laciniata*	RUDLAC	09/04/2020	Land Lab	Restored meadow	Y
F20_S11_L001	*Trifolium pratense*	TRIPRA3	09/04/2020	Land Lab	Restored meadow	Y

Supporting Information Table S2: Total netting effort, specimens collected and broad summary of taxonomic groups observed for each plant species investigated.

Plant Species	Total Effort (min)	Total visitors	Total Orders	Total Families	Total Genera	Bee Families	Bee Genera	Lepidoptera Families	Lepidoptera Genera	Honey bees
Cichorium intybus	30	25	2	2	2	1	1	1	1	Y
Daucus carota	30	90	5	12	15	3	5	2	2	Y
Eupatorium perfoliatum	30	35	2	3	4	1	1	2	3	Y
Lotus corniculatus	30	20	3	3	4	2	2	1	1	Y
Lythrum salicaria	30	25	2	2	2	1	1	1	1	Y
Verbesina alternifolia	90	280	5	9	14	3	7	3	3	Y
Vernonia gigantea	60	25	2	3	8	3	7	0	0	Y
Rudbeckia laciniata	30	96	3	4	5	1	2	2	2	Y
Silphium perfoliatum	90	288	5	11	16	3	6	5	5	Y
Solidago canadensis	30	66	3	3	4	1	2	1	1	Y
Solidago rigida	30	70	3	3	4	1	2	1	1	Y
Symphyotrichum lateriflorum	60	25	5	9	12	2	3	2	2	N
Trifolium pratense	120	83	3	6	10	2	4	4	5	Y
Trifolium repens	30	31	2	2	3	1	2	1	1	Y

Supporting Information Table S3: Search terms used to download available reference sequences from NCBI Nucleotide for each marker. Data from each search were curated using MetaCurator, as outlined in the associated Methods.

	Search Terms	Date Searched
COI	(cytochrome c oxidase subunit 1[All Fields] OR COI[All Fields]) AND (Arthropoda[All Fields] OR "Arthropoda"[Organism])	1/04/2022
16S	16S[All Fields] AND (("Arthropoda"[Organism] OR ("Arthropoda"[Organism] OR Arthropoda[All Fields])) OR "Arthropoda"[Organism]) AND animals[filter] AND (is_nuccore[filter] AND mitochondrion[filter])	1/04/2022
28S	(((28S[All Fields] AND (("Arthropoda"[Organism] OR ("Arthropoda"[Organism] OR Arthropoda[All Fields])) OR "Arthropoda"[Organism])) AND animals[filter] AND is_nuccore[filter]) AND animals[filter]) AND 28S[Title] AND animals[filter]	1/04/2022

Supporting Information Table S4: Counts of individual insect specimens netted from each of the 14 flower species investigated.

		Site	Don Scott							Granville Land Lab						
Order	Family	Genus	*Cichorium intybus*	*Daucus carota*	*Eupatorium perfoliatum*	*Lotus corniculatus*	*Lythrum salicaria*	*Verbesina alternifolia*	*Vernonia gigantea*	*Rudbeckia laciniata*	*Silphium perfoliatum*	*Solidago canadensis*	*Solidago rigida*	*Symphyotrichum lateriflorum*	*Trifolium pratense*	*Trifolium repens*
Hymenoptera (Apiformes)	Apidae	*Apis*	3	37	6	3	22	165	8	32	59	5	12		34	29
		Bombus								15	70	1	3		27	1
		Ceratina						5	1					1		
		Melissodes							4		1				1	
		Xylocopa						1								
	Colletidae	*Hylaeus*		2												
	Halictidae	*Augochlora*		3				1	2							
		Halictus		1				9	2		2			9		
		Lasioglossum		9				2	2		1			1		
	Megachilidae	*Anthidium*				13										
		Coelioxys						1								
		Megachile							1		1				1	
Coleoptera	Cantharidae	*Chauliognathus*									45					
		sp.								33	52	2	22			
	Chrysomelidae	*Systena*												1		
		sp.												1		
	Meloidae	*Epicauta*									1					
	Nitidulidae	*Brassicogethes*	22		23		3	83		15	42	58	33			
	Phalacridae	sp.						2			5			2		
Diptera	Sarcophagidae	sp.												1		
	Syrphidae	sp.		1												

	Tachinidae	sp.		1												
	Tephritidae	sp.		1												
	Unk	sp.		11		3		2	5		1			1	4	
Hemiptera	Aphididae	sp.									3					
	Cicadellidae	sp.		1				1						1		
	Membracidae	sp.												1		
	Miridae	sp.		5				2						5		
Hymenoptera	Formicidae	sp.		14												
	Ichneumonidae	sp.		1												
Lepidoptera	Hesperiidae	*Epargyreus*													3	
		sp.				1					1			1	6	
	Nymphalidae	*Danaus*								1	2				4	
		Junonia			2											
		sp.			4			5								1
	Papilionidae	*Papilio*													2	
	Pieridae	*Colias*		2											1	
		Pieris									2					
	Yponomeutidae	*Atteva*		1												
Orthoptera	Gryllidae	sp.						1								

Supporting Information Table S5: Online resources used to infer natural history traits of detected Lepidoptera.

Family	Flight Period	Source
Noctuidae	Nocturnal	https://mdc.mo.gov/discover-nature/field-guide/noctuid-moths
Saturniidae	Nocturnal/Crepuscular	https://education.mdc.mo.gov/discover-nature/field-guide/giant-silkworm-and-royal-moths
Tortrichidae	Nocturnal	https://www.butterfliesandmoths.org/taxonomy/Tortricidae
Erebidae	Nocturnal	https://www.butterfliesandmoths.org/taxonomy/Erebidae
Geometridae	Nocturnal (most)	https://andrewsforest.oregonstate.edu/sites/default/files/lter/pubs/pdf/pub3739/pub3739_09j1.pdf
Hesperiidae	Diurnal	
Pieridae	Diurnal	
Nymphalidae	Diurnal	https://www.butterfliesandmoths.org/taxonomy/Nymphalidae
Paplionidae	Diurnal	https://www.butterfliesandmoths.org/taxonomy/Papilionidae
Yponomeutidae	Diurnal	https://uwm.edu/field-station/ailanthus-webworm-moth/

References

Allen, M. C., Nielsen, A. L., Peterson, D. 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.

Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A., & Boehm, A. B. (2019). Modeling environmental DNA transport in the coastal ocean using Lagrangian particle tracking. Frontiers in Marine Science, 6, 477.

Bailiff, M. D., & Karl, D. M. (1991). Dissolved and particulate DNA dynamics during a spring bloom in the Antarctic Peninsula region, 1986–1987. Deep Sea Research Part A. Oceanographic Research Papers, 38(8-9), 1077-1095.

Baillie, J., Böhm, M., Collen, B., & Kemp, R. (2012). Spineless: status and trends of the world's invertebrates.

Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. Conservation genetics, 17(1), 1-17.

Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., BRUN, J. J., & Taberlet, P. (2012). Tracking earthworm communities from soil DNA. Molecular Ecology, 21(8), 2017-2030. Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., ... & De Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. Trends in ecology & evolution, 29(6), 358-367.

Cameron, S. A., Lozier, J. D., Strange, J. P., Koch, J. B., Cordes, N., Solter, L. F., & Griswold, T. L. (2011). Patterns of widespread decline in North American bumble bees. Proceedings of the National Academy of Sciences, 108(2), 662-667.

Campbell, J. W., & Hanula, J. L. (2007). Efficiency of Malaise traps and colored pan traps for collecting flower visiting insects from three forested ecosystems. Journal of Insect Conservation, 11(4), 399-408.

Cane, J. H., Minckley, R. L., & Kervin, L. J. (2000). Sampling bees (Hymenoptera: Apiformes) for pollinator community studies: Pitfalls of pan-trapping. Journal of the Kansas Entomological Society, 73(4), 225–231.

Coissac, E., Riaz, T., & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and animals. Molecular ecology, 21(8), 1834-1847.

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., ... & Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. Methods in Ecology and Evolution, 7(9), 1008-1018.

Cronin, M. A., Palmisciano, D. A., Vyse, E. R., & Cameron, D. G. (1991). Mitochondrial DNA in wildlife forensic science: species identification of tissues. Wildlife Society Bulletin (1973-2006), 19(1), 94-105.

Dankowicz, Z, & Dankowicz, E. (2022). Field/Photo ID for Flies. Retrieved August, 2021 from https://sites.google.com/view/flyguide.

Darby B, Bryant R, Keller A, Jochim M, Moe J, Schreiner Z, et al. (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: e0227918.

Delmont, T. O., Robe, P., Cecillon, S., Clark, I. M., Constancias, F., Simonet, P., ... & Vogel, T. M. (2011). Accessing the soil metagenome for studies of microbial diversity. Applied and environmental microbiology, 77(4), 1315-1324.

Díaz-Ferguson, E. E., & Moyer, G. R. (2014). History, applications, methodological issues and perspectives for the use environmental DNA (eDNA) in marine and freshwater environments. Revista de biologia tropical, 62(4), 1273-1284.

Eaton, E. R. and K. Kaufman. 2007. Kaufman Field Guide to Insects of North America. HMH. ISBN 0-618-15310-1

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

Fonseca, V. G., Carvalho, G. R., Nichols, B., Quince, C., Johnson, H. F., Neill, S. P., ... & Creer, S. (2014). Metagenetic analysis of patterns of distribution and diversity of marine meiobenthic eukaryotes. Global Ecology and Biogeography, 23(11), 1293-1302.

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.

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., ... & Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in ecology and evolution, 7(11), 1299-1307.

Gregorič, M., Kutnjak, D., Bačnik, K., Gostinčar, C., Pecman, A., Ravnikar, M., & Kuntner, M. (2022). Spider webs as eDNA samplers: Biodiversity assessment across the tree of life.

Molecular Ecology Resources, 22(7), 2534–2545.
Harper, L. R., Lawson Handley, L., Hahn, C., Boonham, N., Rees, H. C., Gough, K. C., ... & Hänfling, B. (2018). Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (Triturus cristatus). *Ecology and evolution*, *8*(12), 6330-6341.

Harper, L. R., Niemiller, M. L., Benito, J. B., Paddock, L. E., Knittle, E., Molano-Flores, B., & Davis, M. A. (2023). BeeDNA: Microfluidic environmental DNA metabarcoding as a tool for connecting plant and pollinator communities. Environmental DNA, 5(1), 191–211.

Hebert, P. D., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270(1512), 313-321.

Hebert, P. D., Dewaard, J. R., & Landry, J. F. (2010). DNA barcodes for 1/1000 of the animal kingdom. Biology letters, 6(3), 359-362.

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270, S96–S99.

Hering, D., Borja, A., Jones, J. I., Pont, D., Boets, P., Bouchez, A., ... & Kelly, M. (2018). Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. Water research, 138, 192-205.

Hung, K.-L. J., Kingston, J. M., Albrecht, M., Holway, D. A., & Kohn, J. R. (2018). The worldwide importance of honey bees as pollinators in natural habitats. Proceedings of the Royal Society B: Biological Sciences, 285(1870), 20172140.

Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). “Sight-unseen” detection of rare aquatic species using environmental DNA. Conservation Letters, 4(2), 150–157.

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, https://doi.org/10.1002/edn3.411

Kang, S., & Mills, A. L. (2006). The effect of sample size in studies of soil microbial community structure. Journal of microbiological methods, 66(2), 242-250.

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.

Khalifa, S. A., Elshafiey, E. H., Shetaia, A. A., El-Wahed, A. A. A., Algethami, A. F., Musharraf, S. G., ... & El-Seedi, H. R. (2021). Overview of bee pollination and its economic value for crop production. *Insects*, *12*(8), 688.

Kirtane, A., Dietschler, N. J., Bittner, T. D., Lefebvre, M. B., Celis, S., O'Connor, K., Havill, N., & Whitmore, M. C. (2022). Sensitive environmental DNA (eDNA) methods to detect hemlock woolly adelgid and its biological control predators Leucotaraxis silver flies and a Laricobius beetle. Environmental DNA, 4, 1136–1149.

Klein, A. M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *Proceedings of the royal society B: biological sciences*, *274*(1608), 303-313.

Klymus, K. E., Marshall, N. T., & Stepien, C. A. (2017). Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. PloS one, 12(5), e0177643.

Kozick, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and Environmental Microbiology, 79, 5112-5120.

Krehenwinkel, H., Weber, S., Künzel, S., & Kennedy, S. R. (2022). The bug in a teacup—Monitoring arthropod–plant associations with environmental DNA from dried plant material. Biology Letters, 18(6), 20220091.

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.

Lynggaard, C., Bertelsen, M. F., Jensen, C. V., Johnson, M. S., Frøslev, T. G., Olsen, M. T., & Bohmann, K. (2022). Airborne environmental DNA for terrestrial vertebrate community monitoring. Current Biology, 32(3), 701-707.

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. Elsevier/Academic Press; USGS Publications Warehouse.

Marquina, D., Andersson, A. F., & Ronquist, F. (2019) New mitochondrial primers for metabarcoding of insects, designed and evaluated using in silico methods. Molecular Ecology Resources, 19: 90-104.

Montero-Castaño, A., Koch, J. B. U., Lindsay, T. T. T., Love, B., Mola, J. M., Newman, K., & Sharkey, J. K. (2022). Pursuing best practices for minimizing wild bee captures to

support biological research. Conservation Science and Practice, 4(7), e12734.

Morisette, J., Burgiel, S., Brantley, K., Daniel, W. M., Darling, J., Davis, J., ... & Wilcox, T. (2021). Strategic considerations for invasive species managers in the utilization of environmental DNA (eDNA): steps for incorporating this powerful surveillance tool. Management of biological invasions: international journal of applied research on biological invasions, 12(3), 747.

Mullis KB and Faloona FA (1987) Specific synthesis of DNAinvitrovia a polymerase-catalysed chain reaction. Methods in Enzymology155: 335–351.

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.

Norrgard, K. (2008) Forensics, DNA fingerprinting, and CODIS. Nature Education 1(1):35
Novotny, J. L., Reeher, P., Varvaro, M., Lybbert, A., Smith, J., Mitchell, R. J., & Goodell, K. (2021). Bumble bee species distributions and habitat associations in the Midwestern USA, a region of declining diversity. Biodiversity and Conservation, 30(3), 865–887.

O'Connor, R. S., Kunin, W. E., Garratt, M. P., Potts, S. G., Roy, H. E., Andrews, C., ... & Carvell, C. (2019). Monitoring insect pollinators and flower visitation: The effectiveness and feasibility of different survey methods. Methods in Ecology and Evolution, 10(12), 2129-2140.

Ogden, R., Dawnay, N., & McEwing, R. (2008). Development of STR profiling systems for individual identification in wildlife: a case study of the Eurasian badger, Meles meles. Forensic Science International: Genetics Supplement Series, 1(1), 612-613.

Ohio Supercomputer Center. (1987). http://osc.edu/ark:/19495/f5s1ph73

Olesen, J. M., Bascompte, J., Elberling, H., & Jordano, P. (2008). Temporal dynamics in a pollination network. Ecology, 89(6), 1573–1582.

Othman, N., Haris, H., Fatin, Z., Najmuddin, M. F., Sariyati, N. H., Md-Zain, B. M., & Abdul-Latiff, M. A. B. (2021, April). A review on environmental DNA (eDNA) metabarcoding markers for wildlife monitoring research. In IOP Conference Series: Earth and Environmental Science (Vol. 736, No. 1, p. 012054). IOP Publishing.

Pedersen, M. W., Overballe-Petersen, S., Ermini, L., Sarkissian, C. D., Haile, J., Hellstrom, M., ... & Willerslev, E. (2015). Ancient and modern environmental DNA. Philosophical Transactions of the Royal Society B: Biological Sciences, 370(1660), 20130383.

Pei, C. K., Hovick, T. J., Duquette, C. A., Limb, R. F., Harmon, J. P., & Geaumont, B. A. (2022). Two common bee-sampling methods reflect different assemblages of the bee (Hymenoptera: Apoidea) community in mixed-grass prairie systems and are dependent on surrounding floral resource availability. *Journal of Insect Conservation*, *26*(1), 69-83.

Pereira, F., Carneiro, J., & Amorim, A. (2008). Identification of species with DNA-based technology: current progress and challenges. Recent Patents on DNA & Gene Sequences (Discontinued), 2(3), 187-200.

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.

Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: Trends, impacts and drivers. Trends in Ecology & Evolution, 25(6), 345–353.

Prendergast, K. S., Menz, M. H. M., Dixon, K. W., & Bateman, P. W. (2020). The relative performance of sampling methods for native bees: An empirical test and review of the literature. Ecosphere, 11(5), e03076.

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., 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.

Rowney, F. M., Brennan, G. L., Skjøth, C. A., Griffith, G. W., McInnes, R. N., Clewlow, Y., ... & Creer, S. (2021). Environmental DNA reveals links between abundance and composition of airborne grass pollen and respiratory health. Current Biology, 31(9), 1995-2003.

Seeley, T. D. (1995). The Wisdom of the Hive. Harvard University Press.

Seeley, T. D. (1997). Honey bee colonies are group-level adaptive units. The American Naturalist, 150(S1), s22-s41.

Sheffield, C. S., Hebert, P. D. N., Kevan, P. G., & Packer, L. (2009). DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies. Molecular Ecology Resources, 9, 196–207.

Sherrill-Mix, S. (2019). taxonomizr: Functions to Work with NCBI Accessions and Taxonomy. https://CRAN.R-project.org/package=taxonomizr

Sigsgaard, E. E., Olsen, K., Hansen, M. D., Hansen, O. L. P., Høye, T. T., Svenning, J. C., & Thomsen, P. F. (2021). Environmental DNA metabarcoding of cow dung reveals taxonomic and functional diversity of invertebrate assemblages. Molecular Ecology, 30(13), 3374-3389.

Sinniger, F., Pawlowski, J., Harii, S., Gooday, A. J., Yamamoto, H., Chevaldonné, P., ... & Creer, S. (2016). Worldwide analysis of sedimentary DNA reveals major gaps in taxonomic knowledge of deep-sea benthos. Frontiers in Marine Science, 3, 92.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., ... & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". Proceedings of the National Academy of Sciences, 103(32), 12115-12120.

Sønstebø, J. H., Gielly, L., Brysting, A. K., Elven, R., Edwards, M., Haile, J., ... & Brochmann, C. (2010). Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. Molecular Ecology Resources, 10(6), 1009-1018.

Stork, N. E. (2018). How many species of insects and other terrestrial arthropods are there on Earth?. Annual review of entomology, 63, 31-45.

Strange, J. P., & Tripodi, A. D. (2019). Characterizing bumble bee (Bombus) communities in the United States and assessing a conservation monitoring method. Ecology and Evolution, 9(3), 1061–1069.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular ecology, 21(8), 2045-2050.

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., & Kawabata, Z. I. (2012). Estimation of fish biomass using environmental DNA. PloS one, 7(4), e35868.

Takahara, T., Minamoto, T., & Doi, H. (2013). Using environmental DNA to estimate the distribution of an invasive fish species in ponds. PloS one, 8(2), e56584.

Tepedino, V. J., & Portman, Z. M. (2021). Intensive monitoring for bees in North America: Indispensable or improvident? Insect Conservation and Diversity, 14(5), 535–542.

Thomann, M., Imbert, E., Devaux, C., & Cheptou, P. O. (2013). Flowering plants under global

pollinator decline. *Trends in plant science*, *18*(7), 353-359.
Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples.

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.

Tournayre, O., Leuchtmann, M., Galan, M., Trillat, M., Piry, S., Pinaud, D., ... & Charbonnel, N. (2021). eDNA metabarcoding reveals a core and secondary diets of the greater horseshoe bat with strong spatio-temporal plasticity. Environmental DNA, 3(1), 277-296.

Townson, H., Harbach, R. E., & Callan, T. A. (1999). DNA identification of museum specimens of the Anopheles gambiae complex: an evaluation of PCR as a tool for resolving the formal taxonomy of sibling species complexes. Systematic Entomology, 24(1), 95-100.

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., 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., 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.

Vaudo, A. D., Patch, H. M., Mortensen, D. A., Grozinger, C. M., & Tooker, J. F. (2014). Bumble bees exhibit daily behavioral patterns in pollen foraging. Arthropod-Plant Interactions, 8(4), 273–283.

Wang, X.-C., Liu, C., Huang, L., Bengtsson-Palme, J., Chen, H., Zhang, J.-H., Cai, D., & Li, J.-Q. (2015). ITS1: A DNA barcode better than ITS2 in eukaryotes? Molecular Ecology Resources, 15(3), 573–586.

Wilson, C., & Wright, E. (2013). Using Environment DNA (eDNA) as a Tool in Risk-Based Decision-Making.

Woodard, S. H., Federman, S., James, R. R., Danforth, B. N., Griswold, T. L., Inouye, D., McFrederick, Q. S., Morandin, L., Paul, D. L., Sellers, E., Strange, J. P., Vaughan, M., Williams, N. M., Branstetter, M. G., Burns, C. T., Cane, J., Cariveau, A. B., Cariveau, D. P., Childers, A., … Wehling, W. (2020). Towards a U.S. national program for monitoring native bees. Biological Conservation, 252, 108821.