EXECUTIVE SUMMARY

Assessing Pollinator Communities via Environmental DNA (eDNA) Metacommunity Assay

SERDP Project RC19-1102

APRIL 2020

Dr. Mark Davis
Dr. Lynsey Harper
Dr. Brenda Molano-Flores
University of Illinois Urbana-Champaign

Mr. Joseph Benito
Dr. Matthew Niemiller
The University of Alabama in Huntsville

Distribution Statement A
This document has been cleared for public release

SERDP
DOD ▪ EPA ▪ DOE

This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.

EXECUTIVE SUMMARY

Project: RC19-1102

TABLE OF CONTENTS

Page

1.0 INTRODUCTION 1
2.0 OBJECTIVES 2
3.0 TECHNOLOGY DESCRIPTION 2
3.1 GREENHOUSE EXPERIMENT 2
3.2 FIELD EXPERIMENT 3
3.3 DNA EXTRACTION, QUANTIFICATION, AND PLATING 4
3.4 PRIMER SELECTION 5
3.5 IN SILICO PRIMER VALIDATION 5
3.6 IN VITRO PRIMER VALIDATION 5
3.7 MICROFLUIDIC METABARCODING 6
3.8 METABEAT PIPELINE 6
3.9 ANACAPA TOOLKIT PIPELINE 7
4.0 PERFORMANCE ASSESSMENT 8
4.1 PRIMER DEVELOPMENT AND VALIDATION 8
4.2 CONTAMINATION 9
4.3 GREENHOUSE RESULTS 12
4.4 FIELD RESULTS 12
4.5 DIVERSITY ACROSS FLOWER SPECIES 16
4.6 DIVERSITY ACROSS SAMPLE TYPES 19
4.7 DIVERSITY ACROSS PRESERVATION/EXTRACTION METHOD 22
4.8 DIVERSITY ACROSS PRIMER SETS 25
4.9 DIVERSITY ACROSS BIOINFORMATICS PIPELINES 29
5.0 IMPLEMENTATION ISSUES 32
5.1 BENEFITS 33
6.0 REFERENCES 33
APPENDIX A FLOWER SPECIES THAT COMPRISED THE GREENHOUSE COMMUNITY A-1
APPENDIX B DETAILS OF THE PRIMER PANEL SELECTED FOR POLLINATOR EDNA MICROFLUIDIC METABARCODING. B-1
APPENDIX C RESULTS OF *IN SILICO* PRIMER TESTING FOR EACH PRIMER PAIR CONSIDERED C-1

LIST OF FIGURES

Page

Figure 1. The Team Established a Mixed Species Colony of Plants to Ensure a Constant Supply of Pollen and Nectar (A). A Subset of Plants Was then Moved to an Isolated Greenhouse Room Where a Colony of Common Eastern Bumblebees Was Established (B). 3
Figure 2. Heat Map Showing Frequency of Contaminants Detected in Metabarcoding Process Controls. 11
Figure 3. Heat Map Showing the Frequency of Taxa Detected in Baseline Samples Collected from Flower Species Before They Were Introduced to the Greenhouse Containing Common Eastern Bumblebee (B. impatiens). 13
Figure 4. Heat Map Showing the Frequency of Taxa Detected in Greenhouse and Field Samples Collected from the Four Focal Flower Species. 14
Figure 5. Bipartite Network Showing Interactions Between Invertebrate Families and the Four Focal Flower Species. 15
Figure 6. Boxplot Showing the Number of Taxa Detected in eDNA Samples from Different Flower Species. 16
Figure 7. NMDS Plots of Communities (Jaccard dissimilarity) from Different Flower Species (Colored Points/Ellipses). 18
Figure 8. Boxplot Showing the Number of Taxa Detected in eDNA Samples Taken from Different Sources of Plant Material. 19
Figure 9. NMDS Plots of Communities (Jaccard dissimilarity) Produced by Different Sample Types (Colored Points/Ellipses). 21
Figure 10. Boxplot Showing the Number of Taxa Detected in eDNA Samples that Were Preserved and Extracted Using Different Methods. 22
Figure 11. NMDS Plots of Communities (Jaccard Dissimilarity) Produced by Different Preservation/Extraction Methods (Colored Points/Ellipse). 24
Figure 12. Boxplot Showing the Number of Taxa Detected in eDNA Samples Using Different Primer Sets. 25
Figure 13. NMDS Plots of Communities (Jaccard Dissimilarity) Produced by Different Primer Sets (Colored Points/Ellipse). 28
Figure 14. Boxplot Showing the Number of Taxa Detected in eDNA Samples Processed using Different Bioinformatics Pipelines. 29
Figure 15. NMDS Plots of Communities (Jaccard Dissimilarity) Produced by Different Bioinformatics Pipelines (Colored Points/Ellipse). 31

LIST OF TABLES

Page

Table 1. The Subset of Primers with Approximately Appropriate Annealing Temperature and Fragment Length Were Tested with 11 Locally Occurring Bee Species Using PCR and Verified via Gel Electrophoresis. 10

Table 2. Summary of Analyses Statistically Comparing Homogeneity of MVDISP Between the Communities of Different Flower Species (ANOVA) and Variation in Community Composition of eDNA Samples from Different Flower Species (PERMANOVA). 17
Table 3. Summary of Analyses Statistically Comparing Homogeneity of MVDISP between the Communities Produced by Different Sample Types (ANOVA) and Variation in Community Composition of eDNA Samples Sourced from Different Material (PERMANOVA). 20
Table 4. Summary of Analyses Statistically Comparing Homogeneity of MVDISP Between the Communities Produced by Different Preservation/Extraction Methods (ANOVA) and Variation in Community Composition of eDNA Samples Preserved and Extracted Using Different Protocols (PERMANOVA). 23
Table 5. Summary of Analyses Statistically Comparing Homogeneity of MVDISP Between the Communities Produced by Different Primer Sets (ANOVA) and Variation in Community Composition of eDNA Samples When Amplified with Different Primer Sets (PERMANOVA). 26
Table 6. Summary of Analyses Statistically Comparing Homogeneity of MVDISP Between the Communities Produced by Different Bioinformatics Pipelines (ANOVA) and Variation in Community Composition of eDNA Samples Processed Using Anacapa or MetaBEAT (PERMANOVA). 30
Table 7. Greenhouse Community Flower Species A-1
Table 8. Primer Panel Details B-1
Table 9. Results of in Silico Primer Testing C-1

ACRONYMS AND ABBREVIATIONS

16s	mitochondrial 16s ribosomal RNA locus
18s	nuclear ribosomal RNA locus
28S	nuclear ribosomal RNA locus
AE	Qiagen elution buffer
AL	Qiagen lysis buffer
ANOVA	analysis of variance
ASV	amplicon sequence variant
ATL	Qiagen tissue lysis buffer
AW1	Qiagen wash buffer
AW2	Qiagen wash buffer
BLAST	Basic Local Alignment Search Tool
BLCA	Bayesian Least Common Ancestor
BOLD	Barcode of Life Data System
bp	base pairs
ºC	degrees Celsius
COI	mitochondrial cytochrome c oxidase subunit 1 locus
CRUX	Creating Reference libraries Using eXisting tools
CTAB	cetyl trimethyl ammonium bromide
df	degrees of freedom
DNA	deoxyribonucleic acid
DoD	U.S. Department of Defense
eDNA	environmental DNA
EDTA	ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
EtOH	ethanol (ethyl alcohol)
FLASH	fast length adjustment of short reads
FR	Federal Register
g	gram or gravitation constant
GLM	generalized linear model
hr	hour
IBCP	Illinois Bat Conservation Program
INHS	Illinois Natural History Survey
LCA	least common ancestor

M	million
metaBEAT	Metabarcoding and Environmental DNA Analysis Tool
min	minute
mL	milliliter
mtDNA	mitochondrial DNA
MVDISP	multivariate dispersion
NaCl	sodium chloride
NCBI	National Center for Biotechnology Information
nDNA	Nuclear DNA
ng	nanogram
NGS	next-generation sequencing
NMDS	non-metric multidimensional scaling
nt	nucleotide
OTU	operational taxonomic unit
P	probability value (p-value)
PCI	phenol-chloroform-isoamyl
PCR	polymerase chain reaction
PERMANOVA	permutational multivariate analysis of variance
PRI	Prairie Research Institute
r	correlation coefficient
R^2	coefficient of determination
RNA	ribonucleic acid
RTE	rare, threatened, and endangered
SE	standard error
sec	second
SERDP	Strategic Environmental Research and Development Program
TAE	Tris-acetate-EDTA
TE	Tris-EDTA
T&E	threatened and endangered
UIUC	University of Illinois at Urbana-Champaign
µL	microliter
µM	micromolar
UV	ultraviolet
V	volt
v	version
x	times (multiple of)
χ^2	chi-square

ACKNOWLEDGEMENTS

We are grateful for the Strategic Environmental Research and Development Program for funding this research. We thank Patricia Dickerson, Evelyn Knittle, and Lauren Paddock for technical assistance. We are deeply grateful to Rosalie Metallo, Montgomery Flack, and Clinton Shipley at the University of Illinois at Urbana-Champaign Plant Care (UIUC) facility for technical, logistical, and moral support of our research. We appreciate Dr. Mark R. Band, Dr. Alvaro G. Hernandez, Dr. Christy L. Wright and the staff at the Roy J. Carver Biotechnology Center, Functional Genomics Unit, for technical support and conducting microfluidic metabarcoding. We thank Emily Curd for assistance with implementing Anacapa. We recognize Danielle Ruffatto, Tiffany Jolley, and Joe Spencer for imaging and design support. We thank Catherine Dana, Tara Hohoff, and Tommy McElrath for providing samples, as well as encouragement and creative criticism to improve this research. We thank Angella Moorehouse at the Illinois Nature Preserves Commission for supplying species lists used to develop the reference DNA sequence database.

1.0 INTRODUCTION

Biodiversity is being lost at an alarming rate. Current extinction rates exceed those of the last five mass extinction events, and biodiversity loss cuts across habitats, ecosystems, and geopolitical boundaries (Pimm and Brooks, 2000; Wake and Vredenburg, 2008; Barnosky et al., 2011; Kolbert, 2014). This modern mass extinction event has been precipitated by anthropogenic activities (Krutzen, 2006), and the impacts have been particularly hard on pollinator species (Potts, 2010; Cameron et al., 2011; Lever et al., 2014). Consequently, numerous pollinators have been added to the Federal Register (81 FR 67786, 82 FR 10285) as threatened or endangered, signaling the vital importance of pollinators to North American ecosystems and economies (Southwick and Southwick, 1992; Losey and Vaughn, 2006). The loss of pollinator biodiversity is of substantial concern, as it reduces plant reproductive success (Thomann et al., 2013), erodes ecosystem services provided by pollination (Allsopp et al., 2008; Winfree et al., 2011; Vanbergen, 2013), and ultimately drives staggering economic losses (Kevan and Phillips, 2001; Gallai et al., 2009). Given the above, rapid, efficient, and accurate assessment of pollinator communities is a conservation imperative to inform adaptive management strategies to stanch the loss of this critical component of biodiversity.

Such assessments are both technically difficult and financially costly (Chacoff et al., 2012; Bartomeus, 2013; Plein et al., 2017). Traditional monitoring often involves comparison of a scant number of communities, requires exorbitant time commitments to document pollinator species presence/absence, and is reliant on extensive time and ever dwindling taxonomic expertise adjudicating pollinators observed (Sheffield et al., 2009; Tur et al., 2013; Weiner et al., 2014). Exacerbating the above is that monitoring is simply unlikely to be performed for the vast majority of pollinator species (Young et al., 2017). It is likely that these problems are further amplified for threatened and endangered species, whose populations tend to be low in numbers and patchily distributed (Goodman, 1987; Menges et al., 1991; Matthies et al., 2004).

Yet, pollinator biomonitoring is critical to both monitor the status and trend of rare, threatened, and endangered (RTE) pollinators and evaluate the impacts of biodiversity loss on ecosystem services (Vamos et al., 2017). Approaches that can assay pollinator biodiversity in a time and cost-effective manner will vastly enhance the ability to develop effective interventions to mitigate further loss.

Environmental DNA (eDNA) analysis has emerged as a viable candidate for such biodiversity assessments (Taberlet et al., 2012; Goldberg et al., 2015; Deiner et al., 2017). Organisms are continually shedding DNA, which persists, albeit temporarily, in their habitats and can be harnessed to identify single species (e.g., RTE species) or entire communities (Dejean et al., 2012; Thomsen et al., 2012). Single-species eDNA assays have been increasingly leveraged as a reliable means of surveying rare or invasive species (Thomsen and Willerslev, 2015), and the United Sates Department of Defense (DoD) has been at the leading edge of the eDNA revolution as a tool to aid in conservation. However, given the costs associated with single-species assays in terms of research and development investments, coupled with the increasing number of imperiled species that require monitoring (particularly in the case of pollinators), methods that can screen for multiple taxa simultaneously are essential to increase efficiency and ostensibly reduce economic burden.

Coupling pollinator eDNA with microfluidic metabarcoding may in fact provide the mechanism for faster, more accurate, and less costly assessments of invertebrate communities (Taberlet et al., 2012;

Ficetola et al., 2015; Valentini et al., 2016; Deiner et al., 2017). eDNA metabarcoding utilizes a very short genetic sequence from a standard part of the genome (typically mitochondrial genome) that possesses conserved primer binding sites across a range of taxa yet has enough sequence variability within intervening regions to distinguish among species in an environmental sample, such as soil or water (Hebert et al., 2003). Since the turn of the 21st century, the scientific community has largely adopted a small (~600 bp) subsection of the mitochondrial DNA (mtDNA) cytochrome c oxidase subunit I (COI) region as the standard for DNA barcoding, although shorter fragments are required for Next Generation Sequencing (NGS) platforms. Online genetic repositories like the Barcode of Life Data System (BOLD) and the National Center for Biotechnology Information's (NCBI) GenBank have been populated with vast amounts of DNA barcodes (approximately 1.5 M in BOLD and over 1.9 M in GenBank) cutting across wide swaths of biodiversity.

Moreover, the emergence of NGS has elevated the resolution with which molecular genetic approaches can capture snapshots of biodiversity (Yu et al., 2012). In particular, DNA and eDNA metabarcoding have become established as innovative and cost-effective means of rapidly assaying biodiversity (Ji et al., 2013; Aylagas et al., 2016; Elbrecht et al., 2017). Finally, microfluidic metagenomics has emerged as an additional advancement that could elucidate broader communities revealed via eDNA metabarcoding due to simultaneous application of multiple primer sets potentially targeting different taxa, different genes, or different regions of the same gene (Hauck et al., 2019). Prior to the submission of the proposal, eDNA metabarcoding assays for pollinators had yet to be developed and deployed. Just prior to the initiation of the SERDP Limited Scope Project, the first paper to assess the potential of eDNA metabarcoding to document pollinator communities was published (Thomsen and Sigsgaard, 2019). However, this study merely sought to determine whether pollinator eDNA could be detected. The project builds upon this work investigating the detection of pollinator biodiversity by combining eDNA sampling, microfluidic metabarcoding, NGS, and comparative bioinformatics to assess pollinator networks associated with host plants. Ultimately, the work advances the state-of-the art and expands the knowledge base to facilitate widespread adoption of pollinator microfluidic eDNA metabarcoding across DoD installations.

2.0 OBJECTIVES

The technical objectives were: 1) develop and validate (both *in silico* and *in vitro*) a panel of broad-spectrum pollinator eDNA metabarcoding primers, 2) test the efficacy of multiple capture methods to obtain NGS-ready pollinator eDNA from flowers, and 3) validate eDNA sampling and metabarcoding methods in controlled and natural field experiments with mixed plant/pollinator communities.

3.0 TECHNOLOGY DESCRIPTION

3.1 GREENHOUSE EXPERIMENT

A mixed species flowering plant colony was established in the University of Illinois Urbana-Champaign (UIUC) Plant Care Facility (Appendix 1). Plantings were staggered so that a constant supply of pollen and nectar was available (Fig. 1). For the experiment, a subset of flowering plants was moved to a secure room where a Natupol colony (Koppert Biological Systems) of common eastern bumblebees (*Bombus impatiens*) was released (Fig. 1). The team then systematically

introduced four focal flower species (*Penstemon*, *Monarda*, *Solanum*, and *Cynoglossum*) into the greenhouse. Upon introducing a focal flower species, the team observed the flowers for bumblebee visits. When a visit by a bumblebee was observed, and upon the bee leaving the flower, the flower was sampled by one of three methods: 1) whole flower heads were harvested and placed in preservative (*Penstemon*, *Monarda*, and *Solanum*), 2) the flower was swabbed with a flocked cotton swab, and the swab was placed in preservative (*Penstemon* and *Monarda* only), or 3) nectar was extracted from the flower via microcapillary tube and placed in preservative (*Cynoglossum* only). Preservatives included Qiagen a tissue lysis (ATL) buffer or cetyl trimethyl ammonium bromide (CTAB), and in volumes of 300, 600, 900, or 1200 µL depending on size/type of sample (i.e., large *Monarda* flowerheads preserved in 1200 µL, nectar draws preserved in 300 µL). Twenty-five samples were collected per flower type, sampling method, and preservative. In addition to sampling flowers where bee visits were directly observed, flowers were allowed to remain in the greenhouse for 24 hours after visits and were then randomly sampled to assess the ability to detect pollinator visits in absence of direct observation. Twenty-five samples from each flower type, using each sampling method, and both preservation methods were acquired.

Figure 1. The Team Established a Mixed Species Colony of Plants to Ensure a Constant Supply of Pollen and Nectar (A). A Subset of Plants Was then Moved to an Isolated Greenhouse Room Where a Colony of Common Eastern Bumblebees Was Established (B).

3.2 FIELD EXPERIMENT

Two of the four focal flower species (*Penstemon* and *Monarda*) also occur on the UIUC campus, and so a pilot field component to the greenhouse experiment was conducted. Here, the team employed a similar approach as in the greenhouse experiment. On days when the weather was suitable (9am, temperatures > 13 ºC and sunny), the team observed flowers. When a bee species was observed visiting a flower, the bee was identified, a photo was taken, and after the bee left the flower, the flower was sampled by one of two methods (whole flower harvest or swabbing). The flower was preserved in either ATL buffer or CTAB in volumes of 300, 600, 900, or 1200 µL, depending on size of the flowerhead, until 25 flower samples via each sampling method, for each preservation type, and for each flower species were obtained. In addition, 25 random flower samples were acquired via each sampling method, for each preservation type, and for each flower species.

3.3 DNA EXTRACTION, QUANTIFICATION, AND PLATING

DNA extractions were performed in laboratories at the Collaborative Ecological Genetics Laboratory, Illinois Natural History Survey, University of Illinois Urbana-Champaign, where dedicated labs and established protocols for working with samples of low DNA concentration are in place. Decontamination procedures include regular bleaching of work surfaces, UV-light irradiation of consumables and workspaces, isolated PCR hoods, and physical separation of pre- and post-PCR work (i.e., a dedicated eDNA "clean room," where no high-copy DNA has been introduced).

DNA was extracted using two methods. First, a modified Qiagen DNeasy® Blood & Tissue approach (per Thomsen and Sigsgaard, 2019) was used for whole flower, swab, and nectar samples preserved in ATL buffer. Lysis was performed in the plastic tubes containing the flowers by adding 60, 100, 200 or µL proteinase K respectively depending on the volume of preservative the sample was originally stored in. Samples were lysed at 56 °C with agitation in a rotor for 3 hr. After lysis, samples were mixed on a vortexer for 10 sec and a total of 500, 800, or 1,500 µL lysis mixture was retrieved, respectively. Equal amounts of AL buffer and absolute ethanol (EtOH), corresponding to the volume of retrieved lysis mixture, were added to the tubes and vortexed thoroughly before the samples were added to spin columns and spun through the membrane filters over several rounds (700 µL per round). Columns were washed by first adding 600 µL AW1 and then 600 µL AW2 buffers. Finally, DNA was eluted in 2 × 60 µL AE buffer, each time with a 15 min incubation step at 37 °C before spinning. All spinning steps were performed at 10,000 g. Each extraction batch (n = 23) also included an all-buffers-only negative extraction control. DNA extracts were stored at −20 °C until quantification and plating.

Second, a modified Phenol-Chloroform-Isoamyl (PCI) extraction and ethanol precipitation method was employed (following Renshaw et al., 2015) for whole flower, swab, and nectar samples preserved in CTAB. From original preservation tubes, 450 µL of CTAB was removed and transferred to a sterile, labeled 2 mL microcentrifuge tube, and 450 µL Chloroform- Isoamyl alcohol was added, then vortexed, and centrifuged at 15,000 g for 5 mins. After centrifugation, 400 µL of the aqueous layer was transferred to a new, sterile 2 mL microcentrifuge tube, and 350 µL of ice-cold isopropanol and 175 µL of room temperature 5 M NaCl were added. Samples were then allowed to precipitate at -20 °C overnight. Following overnight precipitation, samples were centrifuged at 15,000 g for 10 mins to pellet the DNA. Liquid was decanted off, then 150 µL of room temperature 70% EtOH was added to wash the pellet, and samples were again centrifuged at 15,000 g for 5 mins. After EtOH was decanted off, 150 µL of room temperature 70% EtOH was again added to wash pellets, and samples were again centrifuged at 15,000 g for 5 mins. After EtOH was decanted off, tubes were inverted and placed on a paper towel for a minimum of 10 mins to remove any excess liquid. Samples were then dried in a vacufuge at 45 °C for 15 mins, followed by air drying until no visible EtOH remained. DNA was then rehydrated with 100 µL of 1x TE buffer.

Extracted DNA samples from both methods were quantified using a Qubit® 3.0 Fluorometer with the Qubit™ dsDNA High Sensitivity Assay (Thermo Fisher Scientific). Samples were quantified in triplicate and the average concentration (ng/µL) was used as the final value. Finally, 30 µL aliquots of each sample were plated on a 96-well 0.2 mL optical reaction plate. A total of 42 samples, two PCR positive controls, two field blanks (negative controls), and two extraction blanks (negative controls) were included on each plate. Plates were then stored at -20 °C until submission (typically less than 24 hours) for microfluidic metabarcoding (see below).

3.4 PRIMER SELECTION

The team conducted a comprehensive literature review to identify potential primer sets for further testing (Appendix 2), resulting in a total of 60 candidate primer sets across two mtDNA regions: cytochrome oxidase subunit 1 (COI), 16s ribosomal RNA (16S), and two nuclear genes: 18s rRNA (18S), and 28s rRNA (28S). Based on the parameters of the microfluidic metabarcoding platform, the team sought primers with published annealing temperatures of approximately 55 °C and with a mean fragment size of approximately 150-400 bp in length.

Primer sets within these parameters were subjected to *in silico* (i.e., simulating PCR via the ecoPCR (Ficetola et al., 2010) software package) and *in vitro* (i.e., subjecting extracted DNA from field collected bumblebees to PCR and confirmatory gel electrophoresis) validation. This included a total of 15 primer sets designed to amplify regions of the COI and 16S loci (Appendix 2). These primer sets have been used in studies investigating diversity of terrestrial arthropods (Clarke et al., 2014a), freshwater macroinvertebrates (Elbrecht & Leese, 2017; Vamos et al., 2017), freshwater mollusks (Klymus et al., 2017), and marine metazoans (Günther et al., 2018; Leray et al., 2013; Meusnier et al., 2008) as well as diet of various predators (Corse et al., 2019; Zeale et al., 2011). Several of the selected primers were recently vetted by Elbrecht et al. (2019) for terrestrial arthropod biodiversity assessment using DNA metabarcoding, where seven published primers outperformed 35 others that were tested: BF3/BR2 (Elbrecht et al., 2019), BF1/BR2 (Elbrecht & Leese, 2017), fwhF2/fwhR2n (Vamos et al., 2017), and mlCOIintF (Leray et al., 2013). By employing 15 primer sets that amplify different markers for microfluidic metabarcoding, the team cast a much wider net of species detection than a single primer set with standard metabarcoding (Hauck et al., 2019).

3.5 IN SILICO PRIMER VALIDATION

The team constructed a reference sequence database for invertebrate species that occur in Illinois using a list of binomial names based on observations by the Illinois Natural History Survey (INHS) greenhouse staff, light-trapping data from the Illinois Bat Conservation Program (IBCP), Arthropods and Mollusks catalogued in the INHS Insect Collection (Illinois Natural History Survey Insect Collection, 2015), cicadas recorded by Dr. Catherine Dana (INHS), and threatened and endangered (T&E) priorities provided by Angella Moorehouse (Illinois Nature Preserves Commission). With this list, all available COI and 16S rRNA sequences for Illinois invertebrate species were downloaded from GenBank using ReproPhylo v1.3 (Szitenberg et al., 2015) in July 2019. Reference sequences in GenBank format were then converted to ecoPCR database format using obiconvert (Boyer et al., 2016). ecoPCR parameters were set to allow a 50–500 bp fragment and 0, 1, 2, or 3 mismatches between each primer and each sequence in the reference database. ecoPCR results were summarized using Microsoft Excel (Appendix 3).

3.6 IN VITRO PRIMER VALIDATION

The subset of 15 primer sets was subjected to standard polymerase chain reaction (PCR) and gel electrophoresis was conducted. Eleven bee pollinator species known to occur in Illinois were selected. Reactions were conducted in 25 µL reactions using the following volumes: 12.5 µL of sterile water, 11.1 µL of GoTaq (Thermo Fisher) colorless Mastermix, 0.2 µL 10 µM forward primer, 0.2 µL 10 µM reverse primer, and 1 µL of template DNA. The PCR program consisted of a 5 min 94 °C denaturation step, followed by 40 cycles consisting of a 45 sec 94 °C denaturation

step, a 1 min 55 °C annealing step, and a 1 min 30 sec 72 °C extension step. Once these cycles were complete, a final 5 min 72 °C extension step was completed, followed by a hold at 4 °C. The team confirmed PCR amplification via gel electrophoresis using 2.0% agarose gels in Tris-acetate-EDTA (TAE) buffer. Electrophoresis settings were 120 V for 90 minutes. Gels were visualized in a LabNet Systems EnduroGel Gel Documentation System. To account for PCR stochasticity, three replicates were conducted for all combinations. Those comparisons in which 3 of 3 replicates showed amplification were deemed successful. Comparisons in which 1 or 2 of 3 replicates showed amplification were deemed moderately successful, and those in which 0 of 3 replicates showed amplification were deemed unsuccessful.

3.7 MICROFLUIDIC METABARCODING

The team used the Fluidigm 48.48 Access Array™ (Fluidigm, San Francisco, California, USA; Fluidigm Corporation, 2016) to employ generic primers for pollinator amplification. The Access Array™ uses integrated fluidic circuits and a 4-primer amplicon tagging scheme in which target-specific primer pairs amplify up to 48 different targets, allowing for the simultaneous amplification of barcoded targets in each of 2,304 reaction chambers. The team selected primer sets for amplification based on taxonomic coverage (see Appendix 3), that had Fluidigm-compatible annealing temperatures of approximately 55 °C, and target amplicon lengths ranging from 150–400 bp (lower bound to meet post-amplification removal of primer-adapted dimers; upper bound to limit amplicon length for paired-end sequencing). Forward and reverse amplification primers were modified by the addition of 5' common sequence tags (CS1, CS2, www.fluidigm.com) which serve as the binding site for the addition of the P5 and P7 Illumina sequences and dual-index multiplex barcodes. Microfluidic metabarcoding, including Access Array™ amplification, Illumina sequencing, and demultiplexing, was conducted by the Roy J. Carver Biotechnology Center Functional Genomic Unit at UIUC.

3.8 METABEAT PIPELINE

The demultiplexed FASTQ files for each primer set were processed using metaBEAT v0.97.11 (https://github.com/HullUnibioinformatics/metaBEAT), which incorporates open-source software for quality filtering, trimming, merging, chimera removal, clustering, and taxonomic assignment. After quality filtering (phred score Q30) and trimming using Trimmomatic v0.32 (Bolger et al., 2014), merging (10 bp overlap minimum and 10% mismatch maximum) using fast length adjustment of short reads (FLASH) v1.2.11 (Magoč and Salzberg, 2011), chimera detection using the UCHIME algorithm (Edgar et al., 2011) in vsearch v1.1 (Rognes et al., 2016), and clustering (97% identity with three sequences minimum per cluster) with vsearch v1.1 (Rognes et al., 2016), non-redundant query sequences were compared against the reference database using Basic Local Alignment Search Tool (BLAST) (Zhang et al., 2000). Putative taxonomic identity was assigned using a lowest common ancestor (LCA) approach based on the top 10% BLAST matches for any query matching with at least 90% identity to a reference sequence across more than 80% of its length. Unassigned sequences were subjected to a separate BLAST against the complete NCBI nucleotide (nt) database at 90% identity to determine the source via LCA as described above.

3.9 ANACAPA TOOLKIT PIPELINE

The demultiplexed FASTQ files for each primer set were processed using the Anacapa Toolkit v1 (archived version doi:10.5281/zenodo.3064152 at github.com/limey-bean/Anacapa; Curd et al., 2019). The Anacapa pipeline consists of four modules to build comprehensive reference sequence databases and assign taxonomy to raw multilocus metabarcode sequence datasets. The first module, called Creating Reference libraries Using eXisting tools (CRUX), generates a custom reference sequence database by running *in silico* PCR for a primer set using ecoPCR (Ficetola et al., 2010) on the European Molecular Biology Laboratory (EMBL) standard nucleotide database, then uses *blastn* (Camacho et al., 2009) to query the seed database against the NCBI non-redundant nucleotide database, retaining only the longest version of each sequence, then retrieving taxonomy using *entrez-qiime* v2.0 (Baker, 2016). CRUX generates an unfiltered database with all accessions and taxonomic path information as well as a paired filtered database that excludes accessions with ambiguous taxonomic paths. A *Bowtie2*-formated index library (Langmead and Salzberg, 2012) is also generated. The team generated a reference 16s rRNA database using CRUX (NCBI nr/nt databases queried on 16 November 2019).

The second module conducts DNA sequence quality control and amplicon sequence variant (ASV) parsing. ASVs are unique sequence reads that differ by as little as 1–2 base pairs and are resolved without the arbitrary similarity thresholds that define operational taxonomic units (OTUs) (Edgar, 2016; Amir et al., 2017; Callahan et al., 2017). The ASV methods infer biological sequences in a sample from sequences with amplification and sequencing errors using a de novo process and the expectation that biological sequences are more likely to be repeatedly observed in a sample than erroneous samples (Callahan et al., 2016, 2017; Edgar, 2016). ASV methods have been shown to outperform OTU methods in the sensitivity and accuracy of correcting erroneous sequences, improving taxonomic resolution and increasing observed diversity in a sample (Callahan et al., 2016; Edgar, 2016; Needham et al., 2017).

First, *cutadapt* (Martin, 2011) and *FastX-toolkit* (Gordon and Hannon, 2010) were used to trim primers, Illumina adapters, and low-quality bases from raw FASTQ files for each sample using the sample quality filtering and trimming settings as in the metaBEAT pipeline. A custom Python script sorted sequence reads into three sets – paired-end, forward only, and reverse only – then processed each set separately through *dada2* v1.14 (Callahan et al., 2016) to denoise, dereplicate, merge paired reads, and remove chimeric sequences. This module generated ASV FASTA files and ASV count summary tables for four read types: merged paired-end reads, unmerged paired-end reads, forward only reads, and reverse only reads.

The resulting ASV FASTA files and count summary tables were input into the Anacapa Classifier module, which assigns taxonomy to ASVs using *Bowtie2* and a modified version of the Bayesian Least Common Ancestor (BLCA) algorithm (Gao et al., 2017). *Bowtie2* was used to query ASVs against the CRUX-generated reference databases using the very-sensitive option to determine up to 100 reference matches. *Bowtie2-BCLA* then was used to process the output, using multiple sequence alignment via *muscle* v3.8.31 (Edgar, 2004), with 100 bootstraps to probabilistically determine taxonomic identity by selecting the lowest common ancestor from the multiple weighted *Bowtie2* hits for each ASV (Curd et al., 2019).

4.0 PERFORMANCE ASSESSMENT

All data manipulation and downstream analysis was conducted in R v3.6.3 (R Core Team, 2020). Datasets for each primer set from both pipelines were combined into one master dataset and contamination in controls assessed (Fig. 1). A false positive sequence threshold was calculated based on the maximum frequency of exogenous DNA from metazoan taxa found in the negative controls (1.17%), i.e., field blanks, extraction blanks, and PCR negative controls. The false positive sequence threshold was applied to the metabarcoding data, and the data filtered to remove non-metazoan taxa before combining with greenhouse and field metadata. The team then examined several biotic and technical grouping variables that may influence pollinator detection: 1) flower species sampled, 2) sample type, 3) preservation/extraction method, 4) bioinformatics pipeline, and 5) primer set.

The read count data were converted to site x taxonomy presence-absence matrices using the *decostand* function in the R package vegan v2.5-6 (Oksanen et al., 2019). Presence-absence matrices were used, as potential bias introduced by PCR amplification may prevent reliable abundance or biomass estimation from sequence reads produced by DNA or eDNA metabarcoding (Elbrecht et al., 2017). The team compared taxon richness of eDNA samples according to each grouping variable. Alpha diversity was obtained using the *specnumber* function in vegan (Oksanen et al., 2019). The data were not normally distributed, and the assumptions of a one-way Analysis of Variance (ANOVA) and Generalized Linear Models (GLMs) with different error families and link-functions were violated. Therefore, the team compared taxon richness according to each grouping variable using the non-parametric Kruskal-Wallis test and performed multiple pairwise comparisons using the non-parametric Dunn's test.

The R package betapart v1.5.1 (Baselga and Orme, 2012) was used to estimate total beta diversity, partitioned by nestedness (i.e., community dissimilarity due to taxon subsets) and turnover (i.e., community dissimilarity due to taxon replacement), across all samples with the *beta multi-function*. These three components of beta diversity (Jaccard dissimilarity) were then estimated for eDNA samples according to each grouping variable (e.g., flower species) using the *beta.pair* function. For each component of beta diversity, the team compared the variance in each group of samples by calculating homogeneity of multivariate dispersions (MVDISP) using the *betadisper* function from vegan (Oksanen et al., 2019). Differences in MVDISP between groups of eDNA samples was then statistically tested using an ANOVA. Community dissimilarity for each component of beta diversity was visualized using Non-metric Multidimensional Scaling (NMDS) with the *metaMDS* function and tested statistically using permutational multivariate analysis of variance (PERMANOVA) with the function *adonis* in vegan (Oksanen et al., 2019). Pre-defined cut-off values were used for effect size, where PERMANOVA results were interpreted as moderate and strong effects if $R^2 > 0.09$ and $R^2 > 0.25$ respectively. These values are broadly equivalent to correlation coefficients of $r = 0.3$ and 0.5 which represent moderate and strong effects accordingly (Nakagawa and Cuthill, 2007; Macher et al., 2018).

4.1 PRIMER DEVELOPMENT AND VALIDATION

Literature review revealed almost 120 individual arthropod primers that yielded well over 60 primer combinations across two mtDNA regions and two nuclear genes. A coarse assessment of primer performance for microfluidic metabarcoding (based upon an approximately 55 °C annealing temperature and a 150–400 bp fragment length) yielded a subset of 15 primers for extensive *in silico*

and *in vitro* validation. *In silico* and *in vitro* testing revealed that the selected primer sets were largely successful in amplifying bee pollinator DNA, specifically 11 species that occur in Illinois (Table 1). Given these results, the team chose to retain the full primer panel for pollinator microfluidic eDNA metabarcoding. More critically, these results reveal that the first objective was successful, i.e., to develop and validate, both *in silico* and *in vitro*, a panel of broad-spectrum pollinator eDNA metabarcoding primers.

4.2 CONTAMINATION

Before false positive threshold application, 712 taxa were detected from 345 eDNA samples. Microfluidic eDNA metabarcoding correctly identified seven out of ten bee species used as PCR positive controls to species-level, with three species identified to genus or family-level. However, the PCR positive controls contained a number of contaminants (Fig. 2) due to the manner in which bee specimens were collected, stored, and preserved by museum staff. In contrast, the negative process controls (field, extraction, and PCR) possessed very little contamination (Fig. 2). A total of 517 taxa were consistently detected below the sequence threshold (1.17%) calculated using negative process controls. Therefore, the final dataset after threshold application contained 195 taxa for downstream analyses. Among the taxa retained were floral visitors, floral residents, and key pollinator species, such as common eastern bumblebee (*Bombus impatiens*), western flower thrip (*Frankliniella occidentalis*), silverleaf whitefly (*Bemisia tabaci*), aphids (*Aphis* spp.), banded garden spider (*Argiope trifasciata*), and white-lined sphinx (*Hyles lineata*).

Table 1. The Subset of Primers with Approximately Appropriate Annealing Temperature and Fragment Length Were Tested with 11 Locally Occurring Bee Species Using PCR and Verified via Gel Electrophoresis.

Blue cells represent comparisons in which all three replicates were positive, gold cells represent comparisons in which one or two replicates were positive, and grey cells represent comparisons in which all three replicates failed

Species	16s rRNA		COI												
	Ins16S-1F & Ins16S-1Rshort	MOL16S_F & MOL16S_R	BF1 & BR1	BF1 & BR2	BF2 & BR2	BF3 & BR2	nsCOIFo & mlCOIintK	mlCOIintF & jgHCO2198	fwhF1 & fwhR1	fwhF2 & fwhR2n	Uni-MinibarF1 & Uni-MinibarR1	ZBJ-ArtF1c & ZBJ-ArtR2c	LepLCO & McoiR1	LepLCO & McoiR2	LepLCO & MLepF1rev
Bombus pennsylvanicus															
Melissodes bimaculata															
Eucera hamata															
Svastra petuka															
Bombus perplexus															
Xylocopa virginica															
Bombus auricomis															
Bombus vagans															
Bombus bimaculatus															
Bombus impatiens															
Bombus grisecollis															

Figure 2. Heat Map Showing Frequency of Contaminants Detected in Metabarcoding Process Controls.

Assignments that were not detected in a control are colored white with no border

4.3 GREENHOUSE RESULTS

The target bee species, the common eastern bumblebee (*B. impatiens*) was not detected in any of the baseline eDNA samples collected from flower species before they were introduced to the greenhouse containing the target pollinator species (Fig. 3). Microfluidic metabarcoding successfully detected the common eastern bumblebee (*B. impatiens*) from eDNA samples upon introduction. Both whole flowers (n = 1) and swabs (n = 3) from *Monarda* preserved in CTAB yielded positive detections for the target species, but not samples preserved in ATL buffer. All sample types from *Penstemon, Solanum*, and *Cyngolossum amabile* failed to detect the target species. In addition to the target species, beneficial insects that had been released in the greenhouse as well as other species known to occur in the greenhouse facilities were detected via the microfluidic eDNA metabarcoding approach (Figure 4). Specifically, the insidious flower bug (*Orius insidiosus*), dark-winged fungus gnats (*Bradysia impatiens*), shore flies (*Scatella stagnalis*), western flower thrips (*Frankliniella occidentalis*), and silverleaf whitefly (*Bemisia tabaci*) were detected. However, several beneficial arthropod species as well as species known to occur in the greenhouse were not detected, including *Hippodamia convergens, Amblyseius californicus, Encarsia formosa, Phidippus audax, Brasysia coprophilia, Ceuthophilus* spp*., Tetranychidae* spp*., Trialeurodes abutiloneus,* and *T. vaporariorum*. Finally, several species not previously known to occur in the greenhouse were detected, including aphids (*Aphis* spp.) and banded garden spiders (*Argiope trifasciata*).

4.4 FIELD RESULTS

The team observed seven bee pollinator species in the field (B. auricomus, B. bimaculatus, B. griseocollis, B. impatiens, B. pensylvanicus, Apis mellifera, and Xylocopa virginica). Prior to false positive sequence threshold application, B. bimaculatus, B. griseocollis, and B. vagans were detected at low frequencies from Penstemon whole flowers, but not Penstemon swabs or any Mondarda samples. None of the observed bee species were detected using microfluidic eDNA metabarcoding after application of the false positive sequence threshold. This included the carpenter bee (X. virginica), despite widespread observations and extensive evidence of nectar robbing by this species. Pollinators that were not directly observed but were recovered via microfluidic eDNA metabarcoding included white-lined sphinx moth (Hyles lineata), cigarette beetle (Lasioderma serricorne), Melanophthalma inermis, Eastern calligrapher (Toxomerus geminatus), common paragus (Paragus haemorrhous), house fly (Musca domestica), Paraliburnia kilmani, Psychoda alternata, insidious flower bug (Orius insidiosus), western flower thrips (Frankliniella occidentalis), and eastern flower thrips (Frankliniella tritici) (Figure 5).

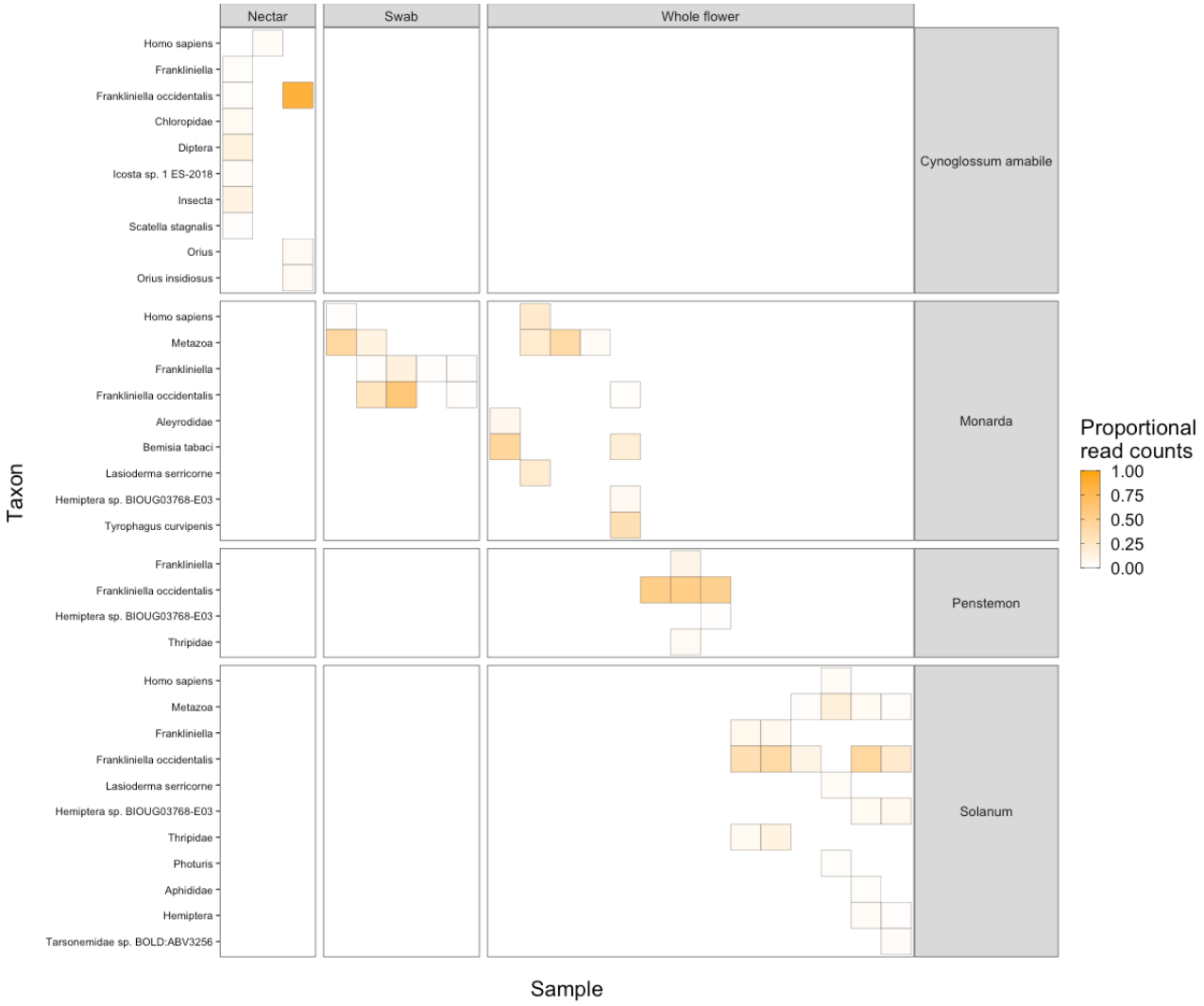

Figure 3. Heat Map Showing the Frequency of Taxa Detected in Baseline Samples Collected from Flower Species Before They Were Introduced to the Greenhouse Containing Common Eastern Bumblebee (B. impatiens).

Assignments that were not detected in a control are colored white with no border.

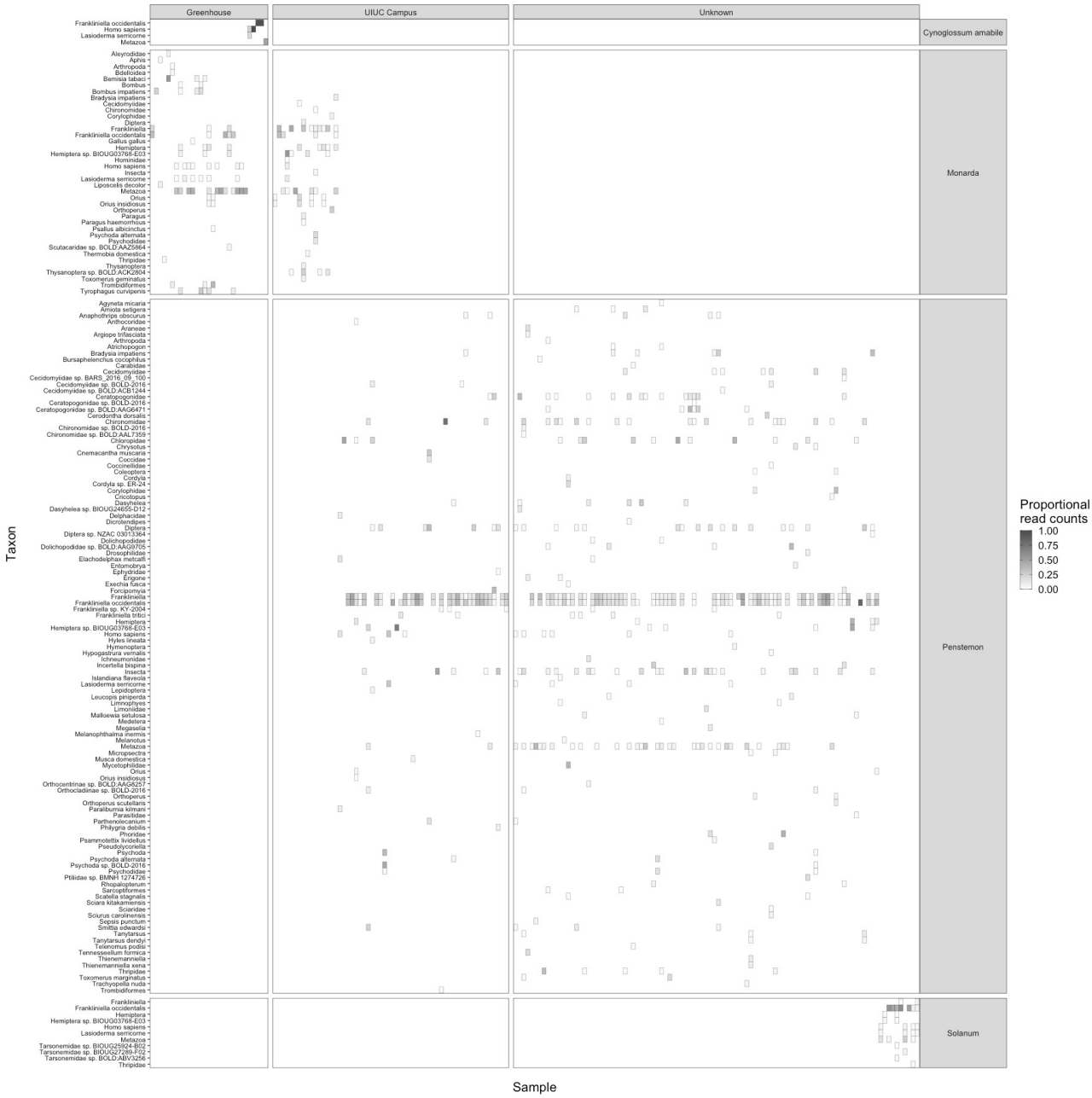

Figure 4. Heat Map Showing the Frequency of Taxa Detected in Greenhouse and Field Samples Collected from the Four Focal Flower Species.

Assignments that were not detected in a control are colored white with no border.

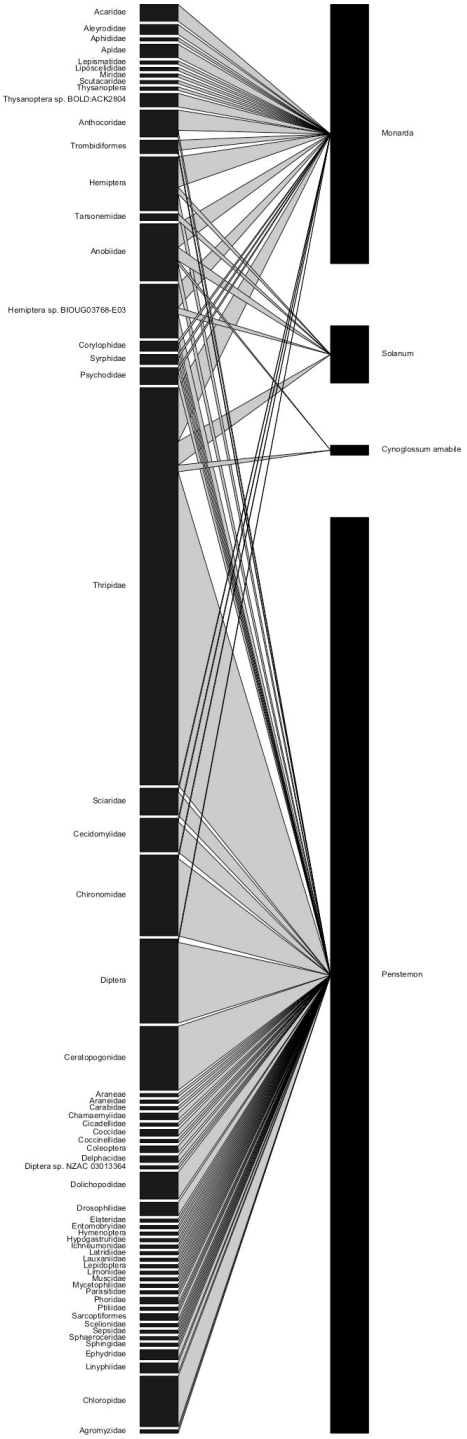

Figure 5. Bipartite Network Showing Interactions Between Invertebrate Families and the Four Focal Flower Species.

Monarda and Penstemon possessed the most interactions suggesting that these flower species support more pollinator diversity or retain more pollinator eDNA.

4.5 DIVERSITY ACROSS FLOWER SPECIES

Flower species had a significant effect on alpha diversity of eDNA samples ($\chi^2_3 = 16.328$, $P < 0.001$). Taxon richness for *Cynoglossum amabile* (nectar draws only) was significantly lower than all other flower species (*Monarda* Z = -2.569, $P = 0.010$; *Penstemon* Z = -3.490, $P = 0.001$; *Solanum* Z = -2.044; $P = 0.041$), and greater for *Penstemon* than *Monarda* (-2.058, $P = 0.040$) (Fig. 6). Beta diversity among eDNA samples was largely driven by turnover (99.65%) as opposed to nestedness (0.35%). MVDISP did not significantly differ between flower species for turnover, nestedness, and total beta diversity (Table 2). Flower species had a weak positive effect on turnover and total beta diversity of pollinator communities (Table 2; Fig. 7).

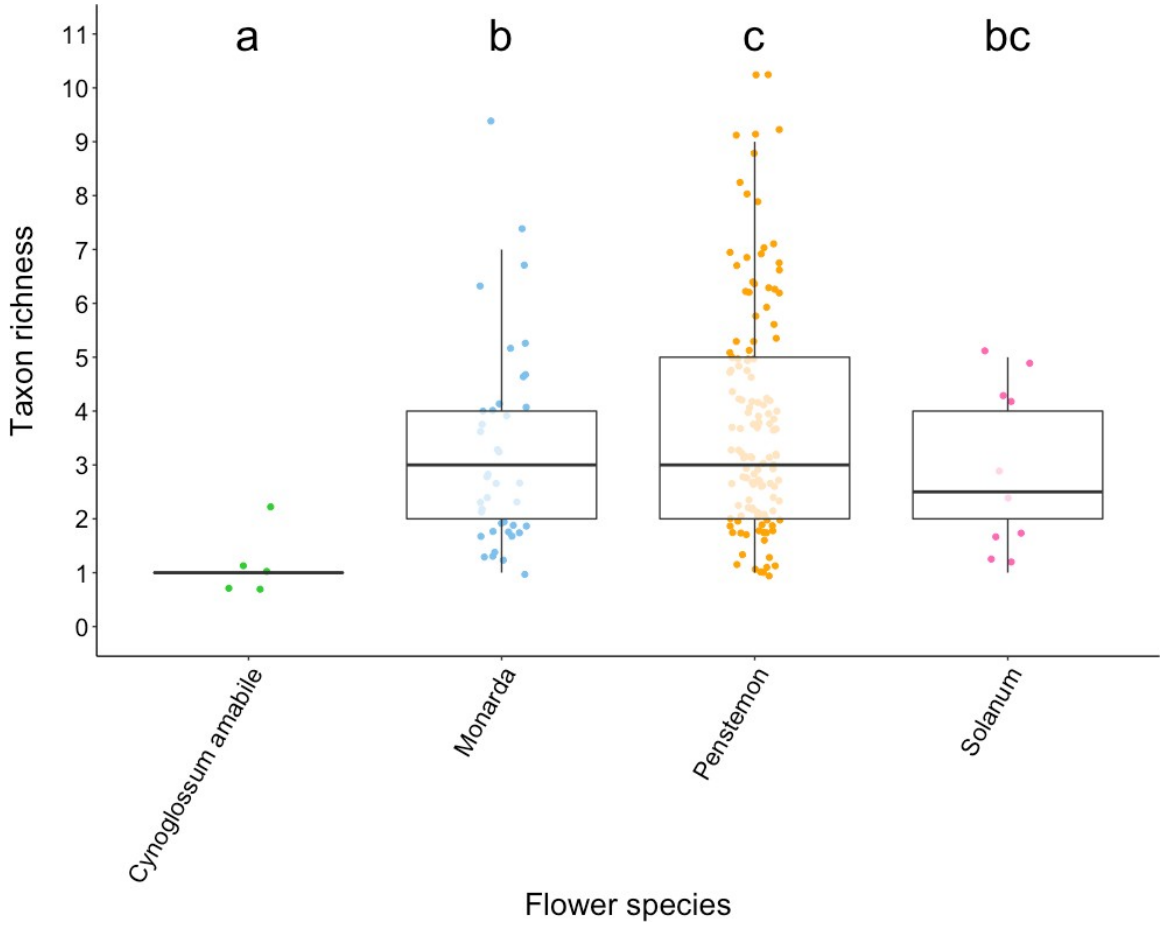

Figure 6. Boxplot Showing the Number of Taxa Detected in eDNA Samples from Different Flower Species.

Taxon richness was significantly different in relation to flower species. Boxes show 25th, 50th, and 75th percentiles, and whiskers show 5th and 95th percentiles.

Table 2. Summary of Analyses Statistically Comparing Homogeneity of MVDISP Between the Communities of Different Flower Species (ANOVA) and Variation in Community Composition of eDNA Samples from Different Flower Species (PERMANOVA).

	Homogeneity of multivariate dispersions (ANOVA)				Community similarity (PERMANOVA)			
Mean distance to centroid	**± SE**	**df**	***F***	***P***	**df**	***F***	R^2	***P***
Turnover		3	2.587	0.055	3	5.582	0.084	**0.001**
Cynoglossum amabile	0.559 ± 0.010							
Monarda	0.582 ± 0.028							
Penstemon	0.449 ± 0.099							
Solanum	0.408 ± 0.060							
Nestedness		3	1.403	0.243	3	-1.512	-0.025	0.958
Cynoglossum amabile	0.138 ± 0.021							
Monarda	0.124 ± 0.016							
Penstemon	0.169 ± 0.024							
Solanum	0.214 ± 0.031							
Total beta diversity		3	1.956	0.122	3	4.394	0.067	**0.001**
Cynoglossum amabile	0.569 ± 0.018							
Monarda	0.629 ± 0.008							
Penstemon	0.550 ± 0.045							
Solanum	0.529 ± 0.014							

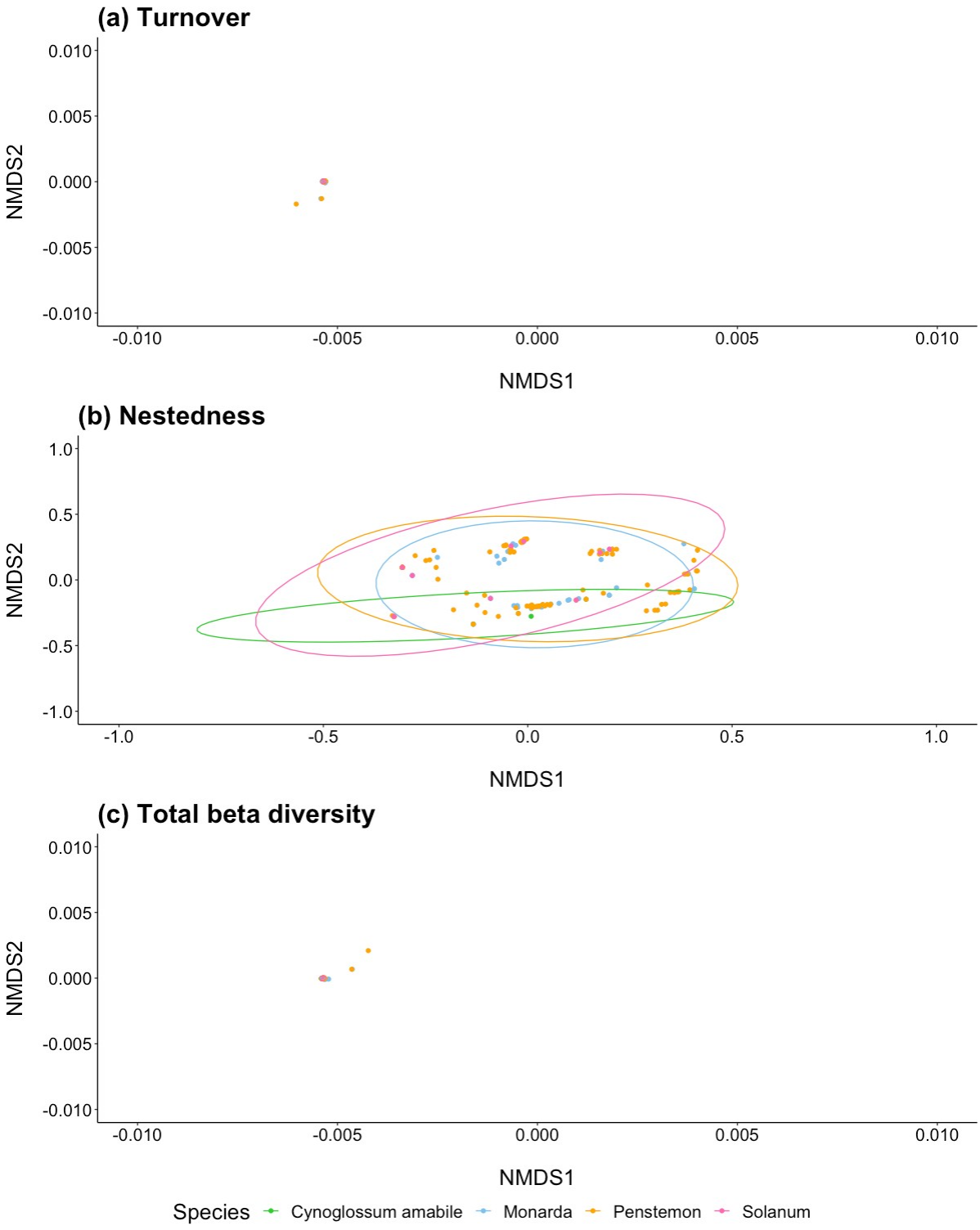

Figure 7. NMDS Plots of Communities (Jaccard dissimilarity) from Different Flower Species (Colored Points/Ellipses).

*The turnover **(a)** and nestedness **(b)** partitions of total beta diversity **(c)** are shown.*

4.6 DIVERSITY ACROSS SAMPLE TYPES

Sample type (i.e., swab, nectar, or whole flower) significantly influenced alpha diversity of eDNA samples ($\chi^2_2 = 14.47$, $P < 0.001$). Taxon richness from nectar draws (*Cynoglossum amabile* only) was lower than swabs (Z = -2.841, $P = 0.005$) and whole flowers (Z = -3.493, $P < 0.001$). Swabs also produced lower taxon richness than whole flowers (Z = -2.001, $P = 0.045$) (Fig. 8). MVDISP did not differ between sample types for turnover, nestedness, or total beta diversity (Table 3). Sample type had a weak positive effect on nestedness of pollinator communities, but not turnover or total beta diversity (Table 3; Fig. 9).

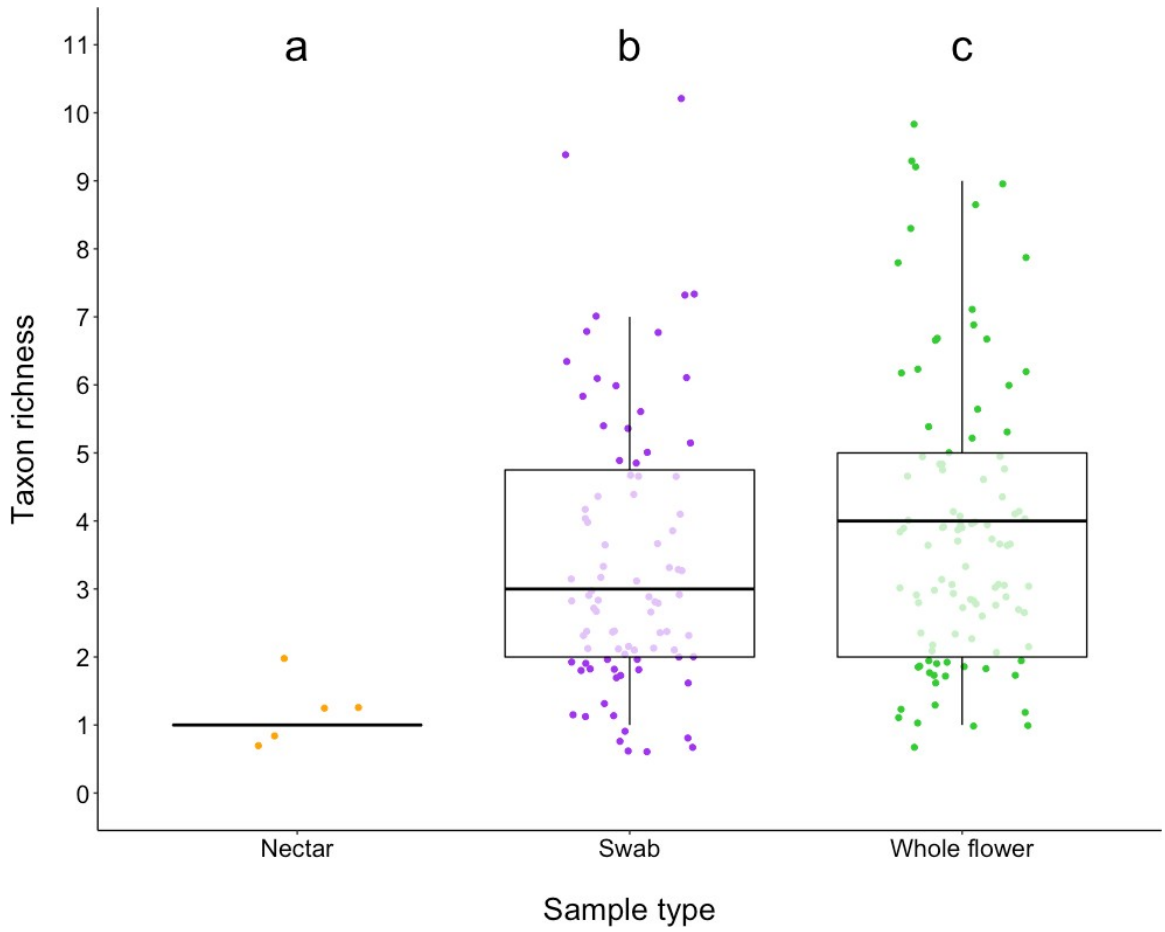

Figure 8. Boxplot Showing the Number of Taxa Detected in eDNA Samples Taken from Different Sources of Plant Material.

Taxon richness was significantly different in relation to sample type. Boxes show 25th, 50th, and 75th percentiles, and whiskers show 5th and 95th percentiles.

Table 3. Summary of Analyses Statistically Comparing Homogeneity of MVDISP between the Communities Produced by Different Sample Types (ANOVA) and Variation in Community Composition of eDNA Samples Sourced from Different Material (PERMANOVA).

	Homogeneity of multivariate dispersions (ANOVA)				Community similarity (PERMANOVA)			
	Mean distance to centroid ± SE	**df**	***F***	***P***	**df**	***F***	R^2	***P***
Turnover		2	0.398	0.673	2	0.730	0.008	0.707
Nectar	0.560 ± 0.010							
Swab	0.541 ± 0.069							
Whole flower	0.511 ± 0.055							
Nestedness		2	0.842	0.432	2	8.035	0.078	**0.003**
Nectar	0.141 ± 0.022							
Swab	0.131 ± 0.017							
Whole flower	0.158 ± 0.024							
Total beta diversity		2	0.073	0.929	2	1.310	0.014	0.135
Nectar	0.569 ± 0.018							
Swab	0.596 ± 0.037							
Whole flower	0.590 ± 0.022							

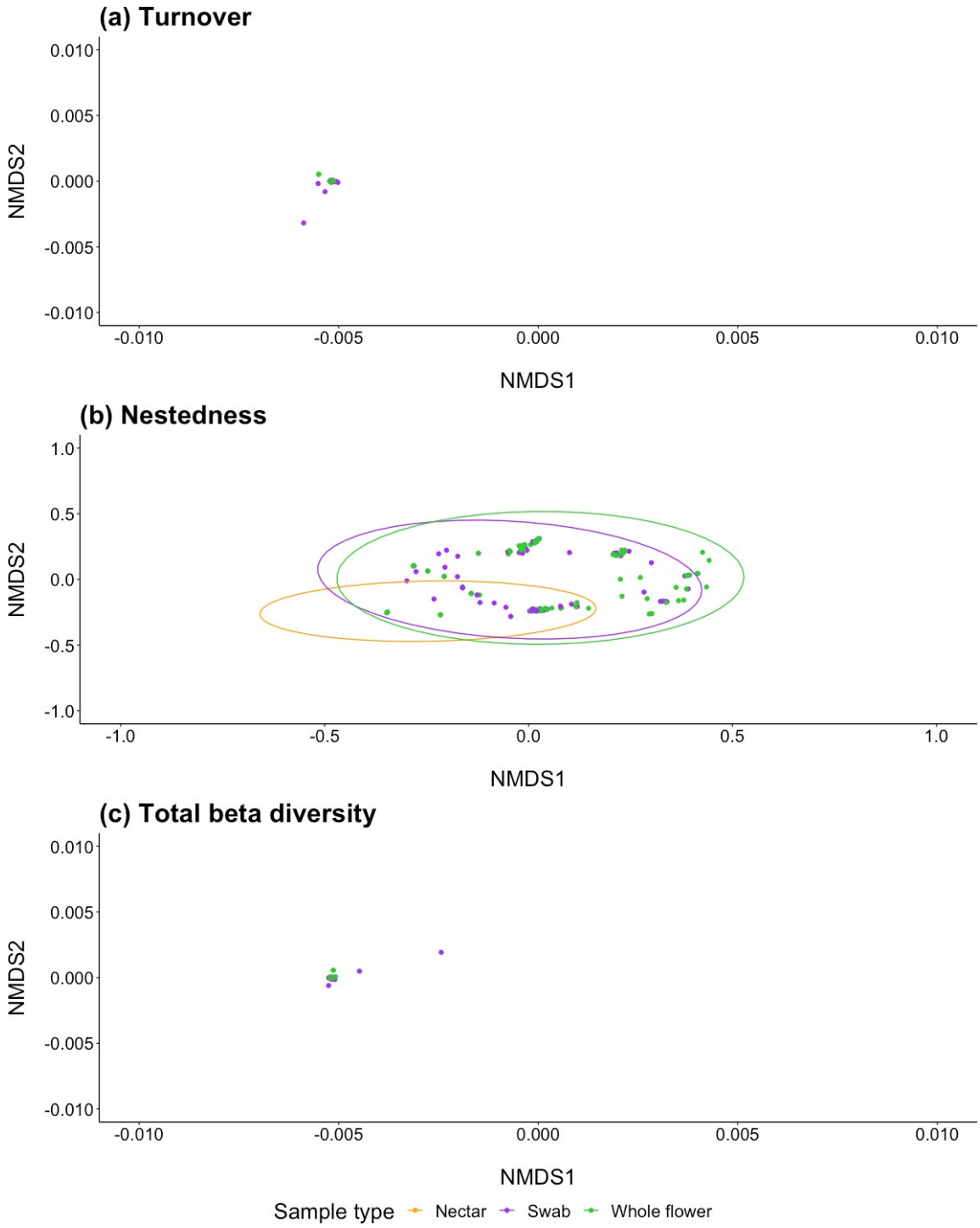

Figure 9. NMDS Plots of Communities (Jaccard dissimilarity) Produced by Different Sample Types (Colored Points/Ellipses).

The turnover (a) and nestedness (b) partitions of total beta diversity (c) are shown.

4.7 DIVERSITY ACROSS PRESERVATION/EXTRACTION METHOD

Preservation/extraction method did not have an effect on alpha diversity of eDNA samples (GLM: $\chi^2_1 = 10.346$, $P = 0.001$). Taxon richness was greater when samples were preserved in ATL buffer and extraction with a commercial Qiagen kit ($Z = 3.217$, $P = 0.001$) as opposed to CTAB and extraction with a PCI protocol (Fig. 10).

MVDISP was not different between preservation/extraction methods for turnover, nestedness, or total beta diversity (Table 4). Preservation/extraction method had no effect on turnover of pollinator communities but did weakly and moderately influence nestedness and total beta diversity respectively (Table 4; Fig. 11).

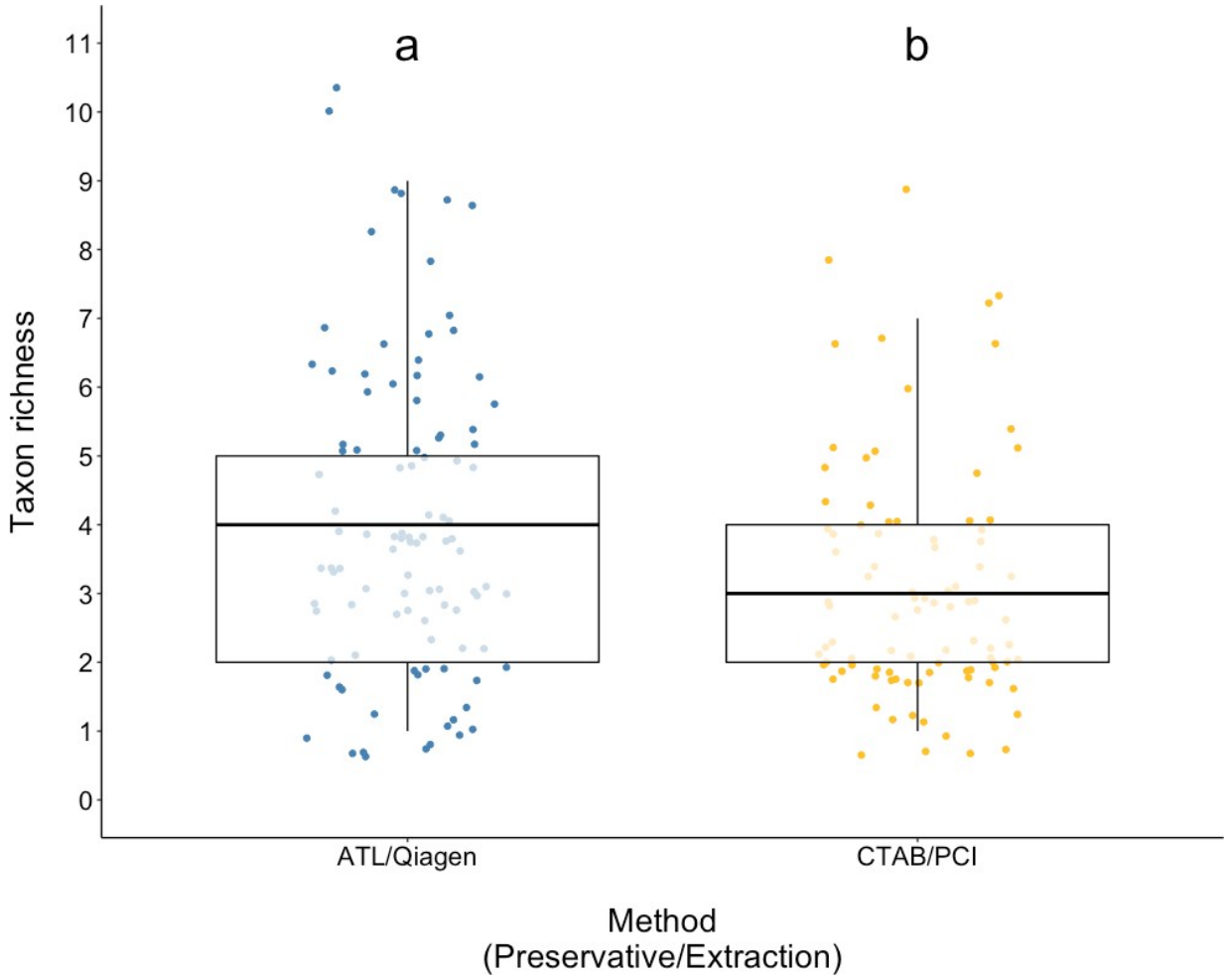

Figure 10. Boxplot Showing the Number of Taxa Detected in eDNA Samples that Were Preserved and Extracted Using Different Methods.

Taxon richness was significantly different in relation to preservation/extraction method. Boxes show 25th, 50th, and 75th percentiles, and whiskers show 5th and 95th percentiles.

Table 4. Summary of Analyses Statistically Comparing Homogeneity of MVDISP Between the Communities Produced by Different Preservation/Extraction Methods (ANOVA) and Variation in Community Composition of eDNA Samples Preserved and Extracted Using Different Protocols (PERMANOVA).

		Homogeneity of multivariate dispersions (ANOVA)			Community similarity (PERMANOVA)			
Mean distance to	**centroid ± SE**	**df**	***F***	***P***	**df**	***F***	R^2	***P***
Turnover		1	0.0001	0.991	1	0.334	0.005	0.473
ATL/Qiagen	0.522 ± 0.045							
CTAB/PCI	0.521 ± 0.098							
Nestedness		1	1.118	0.292	1	21.094	0.099	**0.001**
ATL/Qiagen	0.149 ± 0.023							
CTAB/PCI	0.127 ± 0.019							
Total beta diversity		1	0.192	0.662	1	2.116	0.011	**0.015**
ATL/Qiagen	0.596 ± 0.017							
CTAB/PCI	0.584 ± 0.053							

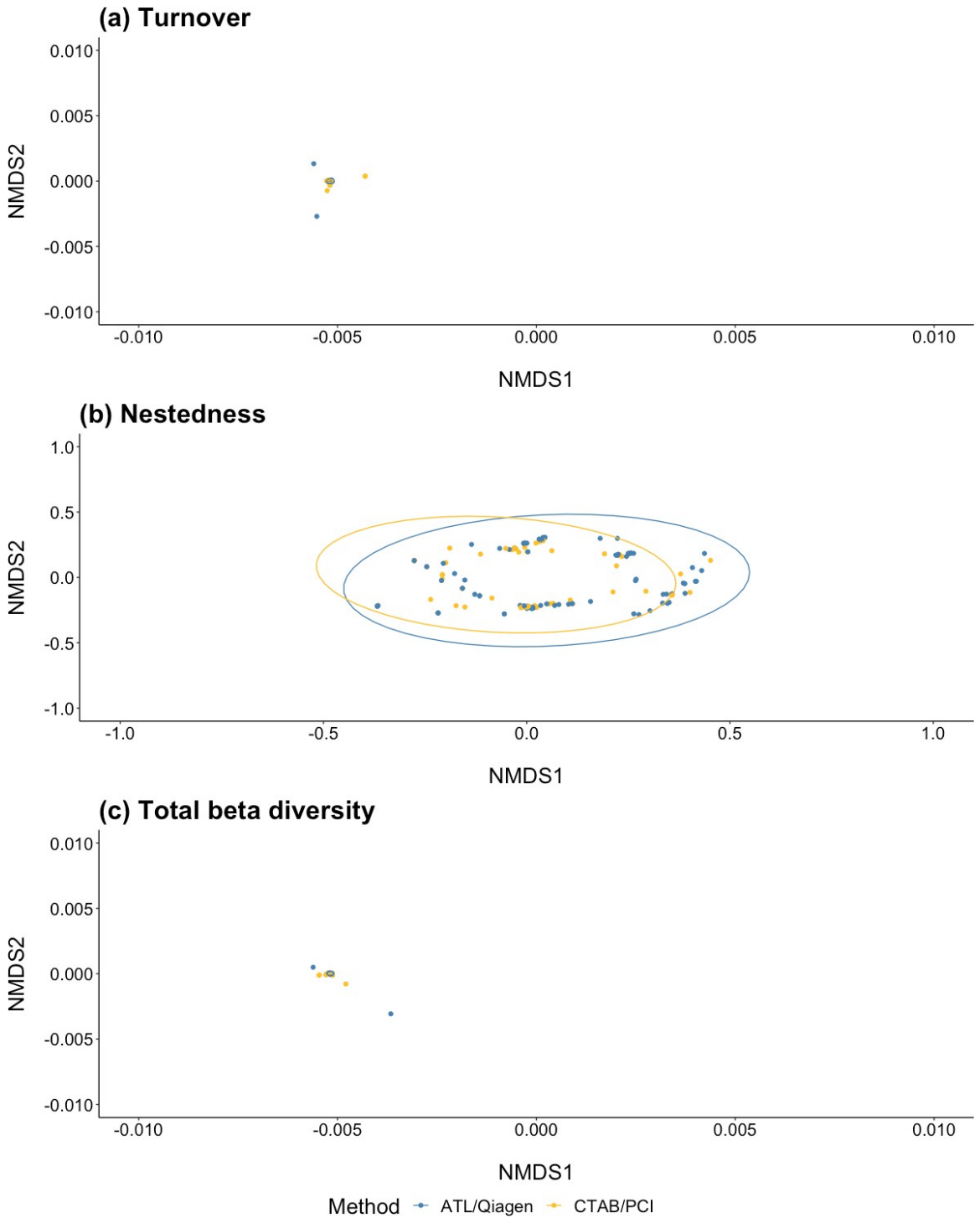

Figure 11. NMDS Plots of Communities (Jaccard Dissimilarity) Produced by Different Preservation/Extraction Methods (Colored Points/Ellipse).

*The turnover **(a)** and nestedness **(b)** partitions of total beta diversity **(c)** are shown.*

4.8 DIVERSITY ACROSS PRIMER SETS

Primer set used for microfluidic eDNA metabarcoding had an effect on alpha diversity of eDNA samples (GLM: $\chi^2_{14} = 352.86$, $P < 0.001$). Overall, taxon richness was higher using BF1/BR2 and mlCOIintF/jgHCO2198, but taxon richness significantly differed between most primer sets (Fig. 12). MVDISP differed between primer sets for turnover, nestedness, and total beta diversity (Table 5). Primer set exerted a moderate influence on turnover and total beta diversity of communities but had no effect on nestedness (Table 5; Fig.13).

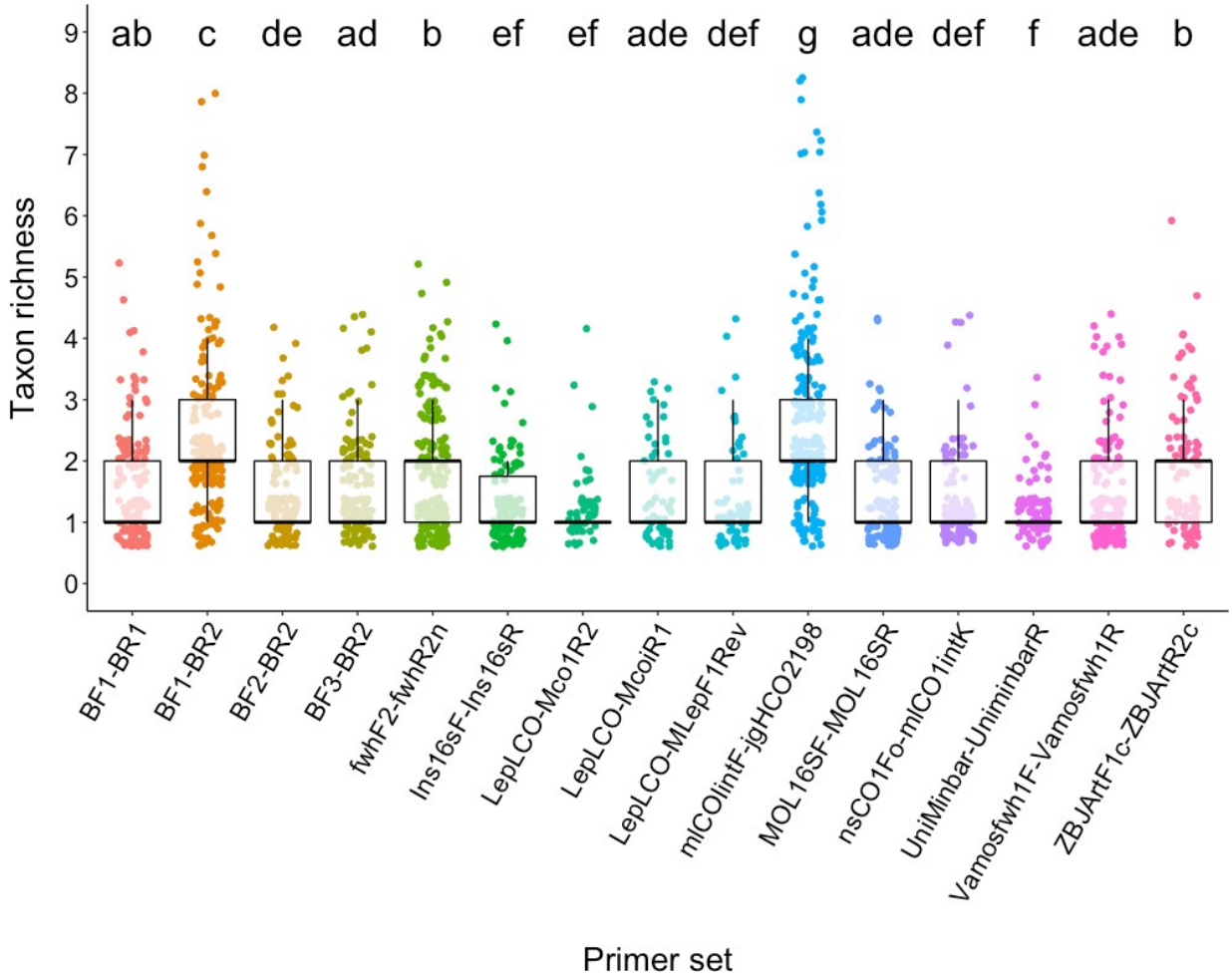

Figure 12. Boxplot Showing the Number of Taxa Detected in eDNA Samples Using Different Primer Sets.

Taxon richness was significantly different in relation to primer set used. Boxes show 25th, 50th, and 75th percentiles, and whiskers show 5th and 95th percentiles.

Table 5. Summary of Analyses Statistically Comparing Homogeneity of MVDISP Between the Communities Produced by Different Primer Sets (ANOVA) and Variation in Community Composition of eDNA Samples When Amplified with Different Primer Sets (PERMANOVA).

	Homogeneity of multivariate dispersions (ANOVA)				Community similarity (PERMANOVA)			
Mean distance to centroid ± SE		**df**	***F***	***P***	**df**	***F***	R^2	***P***
Turnover		14	10.400	**0.001**	14	24.639	0.162	**0.001**
BF1/BR1	0.375 ± 0.169							
BF1/BR2	0.369 ± 0.195							
BF2/BR2	0.666 ± 0.129							
BF3/BR2	0.404 ± 0.184							
fwhF1/fwhR1	0.499 ± 0.131							
fwhF2/fwhR2n	0.404 ± 0.184							
LepLCO/McoiR1	0.629 ± 0.011							
LepLCO/McoiR2	0.593 ± 0.052							
LepLCO/MLepF1Rev	0.625 ± 0.011							
mlCOIintF/jgHCO2198	0.456 ± 0.143							
nsCO1Fo/mlCO1intK	0.657 ± 0.007							
UniMinbar/UniminbarR	0.549 ± 0.120							
ZBJArtF1c/ZBJArtR2c	0.613 ± 0.019							
Ins16sF/Ins16sR	0.512 ± 0.135							
MOL16SF/MOL16SR	0.477 ± 0.137							
Nestedness		14	9.331	**<0.001**	14	-28.71	-0.291	1.000
BF1/BR1	0.204 ± 0.027							
BF1/BR2	0.147 ± 0.044							
BF2/BR2	0.064 ± 0.006							
BF3/BR2	0.136 ± 0.022							
fwhF1/fwhR1	0.144 ± 0.025							
fwhF2/fwhR2n	0.194 ± 0.029							
LepLCO/McoiR1	0.090 ± 0.011							
LepLCO/McoiR2	0.098 ± 0.017							
LepLCO/MLepF1Rev	0.110 ± 0.013							
mlCOIintF/jgHCO2198	0.158 ± 0.031							
nsCO1Fo/mlCO1intK	0.064 ± 0.005							
UniMinbar/UniminbarR	0.096 ± 0.016							
ZBJArtF1c/ZBJArtR2c	0.128 ± 0.014							
Ins16sF/Ins16sR	0.098 ± 0.022							
MOL16SF/MOL16SR	0.151 ± 0.023							
Total beta diversity		14	6.235	**<0.001**	14	20.123	0.136	**0.001**
BF1/BR1	0.468 ± 0.166							
BF1/BR2	0.482 ± 0.163							
BF2/BR2	0.673 ± 0.005							
BF3/BR2	0.586 ± 0.077							
fwhF1/fwhR1	0.569 ± 0.082							
fwhF2/fwhR2n	0.496 ± 0.178							

LepLCO/McoiR1	0.647 ± 0.005
LepLCO/McoiR2	0.619 ± 0.033
LepLCO/MLepF1Rev	0.647 ± 0.005
mlCOIintF/jgHCO2198	0.548 ± 0.076
nsCO1Fo/mlCO1intK	0.664 ± 0.006
UniMinbar/UniminbarR	0.584 ± 0.091
ZBJArtF1c/ZBJArtR2c	0.646 ± 0.008
Ins16sF/Ins16sR	0.562 ± 0.086
MOL16SF/MOL16SR	0.555 ± 0.073

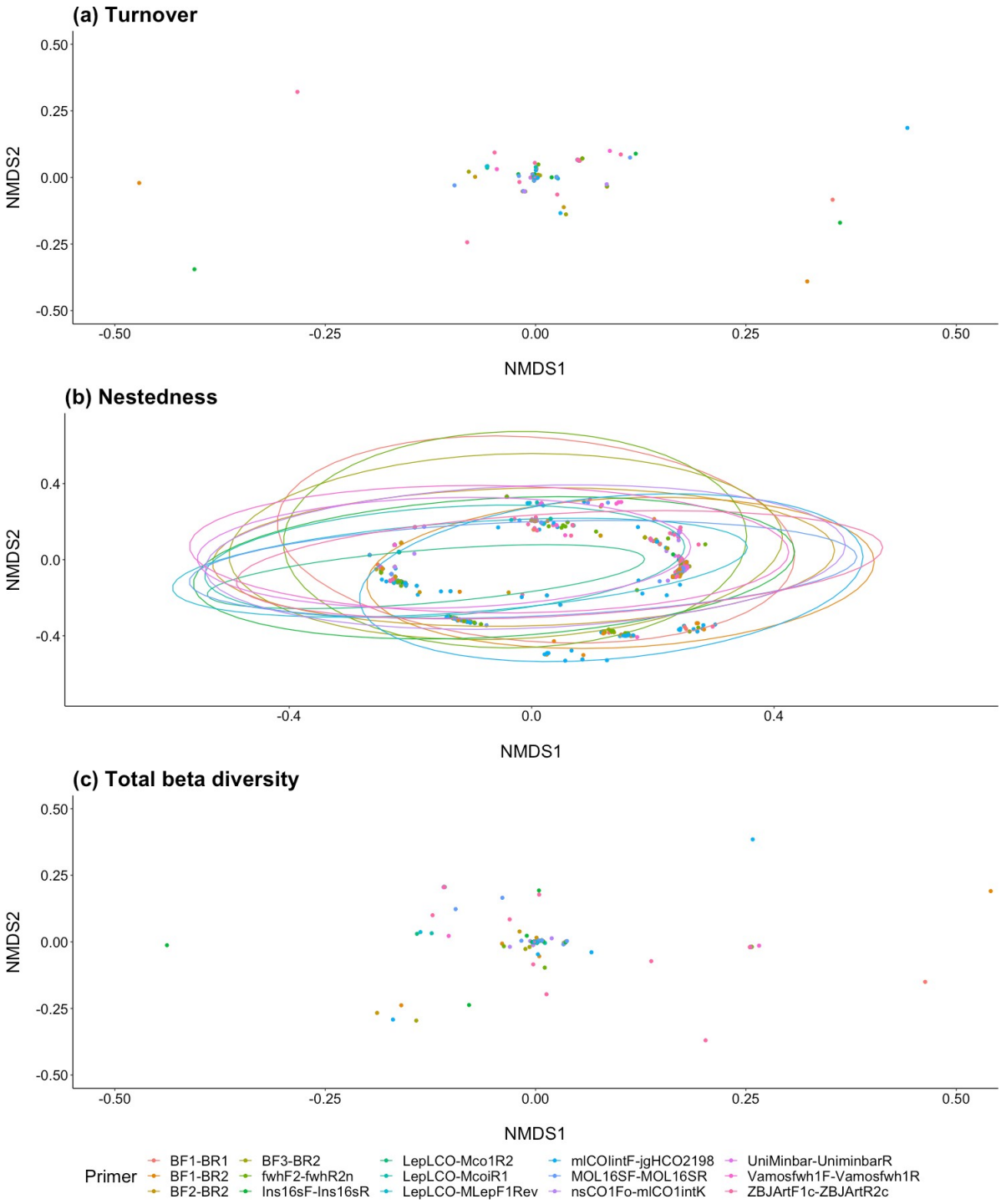

Figure 13. NMDS Plots of Communities (Jaccard Dissimilarity) Produced by Different Primer Sets (Colored Points/Ellipse).

*The turnover **(a)** and nestedness **(b)** partitions of total beta diversity **(c)** are shown.*

4.9 DIVERSITY ACROSS BIOINFORMATICS PIPELINES

Bioinformatics pipeline used to process the metabarcoding data influenced alpha diversity of eDNA samples (GLM: $\chi^2_1 = 33.741$, $P < 0.001$). Taxon richness generated by Anacapa was lower than taxon richness generated by metaBEAT (Z = -5.809, $P < 0.001$) (Fig. 14). MVDISP was not different between bioinformatics pipelines for turnover, nestedness, or total beta diversity (Table 6). Bioinformatics pipeline exerted a weak influence on turnover and total beta diversity of pollinator communities, but not nestedness (Table 6; Fig. 15).

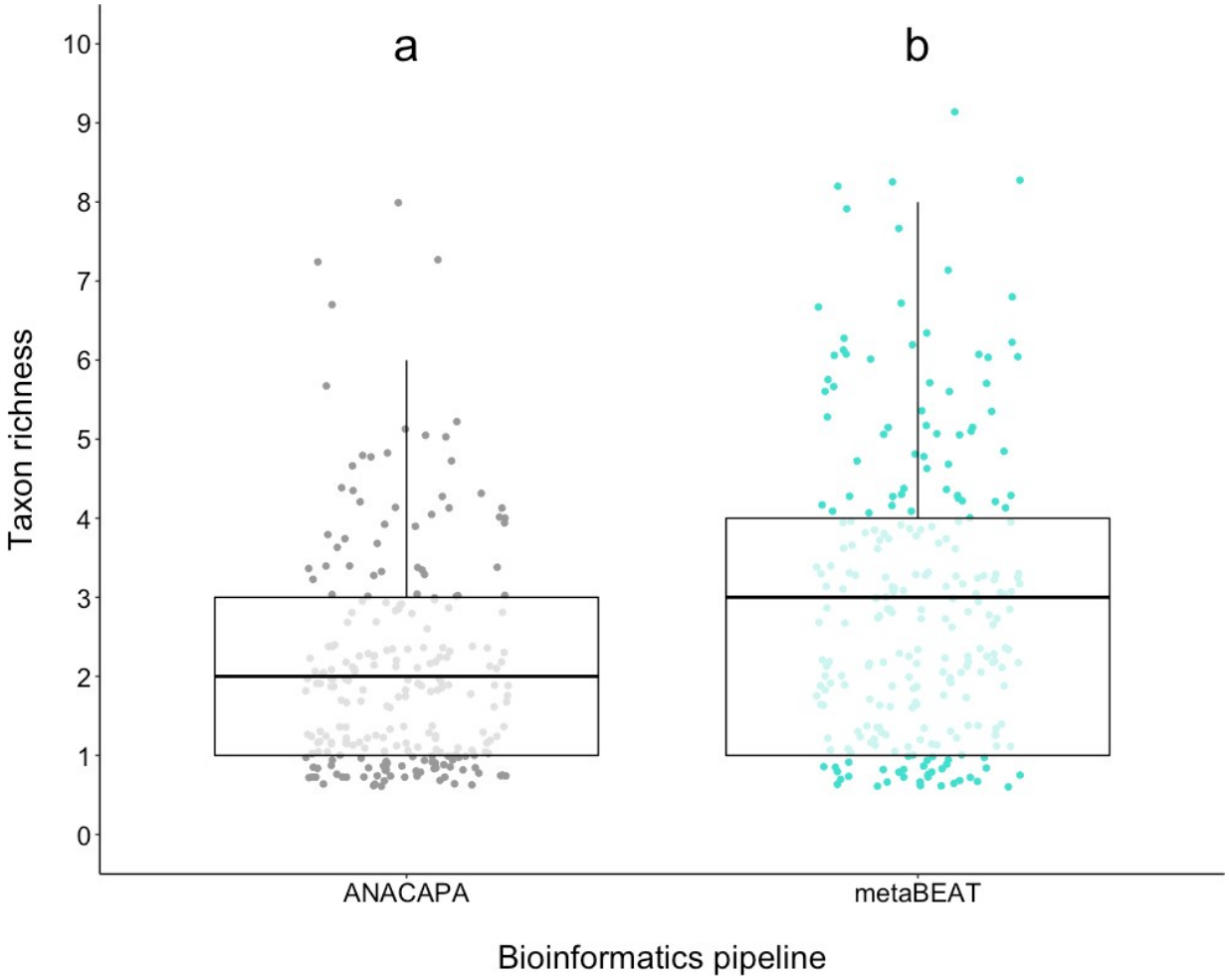

Figure 14. Boxplot Showing the Number of Taxa Detected in eDNA Samples Processed using Different Bioinformatics Pipelines.

Taxon richness was not significantly different in relation to bioinformatics pipeline. Boxes show 25th, 50th, and 75th percentiles, and whiskers show 5th and 95th percentiles.

Table 6. Summary of Analyses Statistically Comparing Homogeneity of MVDISP Between the Communities Produced by Different Bioinformatics Pipelines (ANOVA) and Variation in Community Composition of eDNA Samples Processed Using Anacapa or MetaBEAT (PERMANOVA).

		Homogeneity of multivariate dispersions (ANOVA)			Community similarity (PERMANOVA)			
Mean distance to	**centroid ± SE**	**df**	***F***	***P***	**df**	***F***	R^2	***P***
Turnover		1	0.312	0.577	1	39.573	0.076	**0.001**
ANACAPA	0.548 ± 0.073							
metaBEAT	0.536 ± 0.052							
Nestedness		1	3.306	0.070	1	-56.459	-0.132	1.000
ANACAPA	0.143 ± 0.021							
metaBEAT	0.169 ± 0.025							
Total beta diversity		1	0.045	0.832	1	29.059	0.057	**0.001**
ANACAPA	0.612 ± 0.031							
metaBEAT	0.610 ± 0.017							

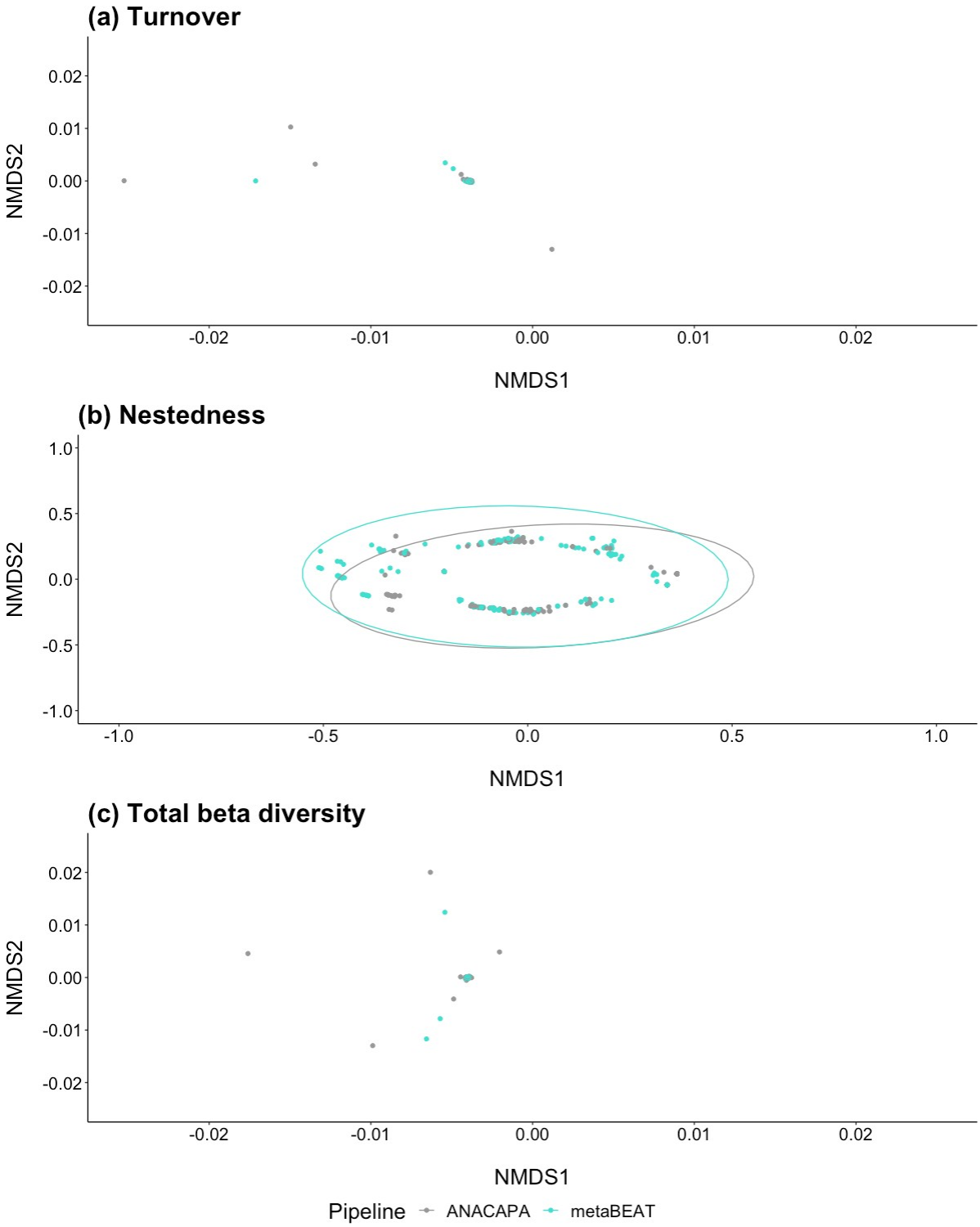

Figure 15. NMDS Plots of Communities (Jaccard Dissimilarity) Produced by Different Bioinformatics Pipelines (Colored Points/Ellipse).

*The turnover **(a)** and nestedness **(b)** partitions of total beta diversity **(c)** are shown*

5.0 IMPLEMENTATION ISSUES

The ability to leverage eDNA metabarcoding to document pollinator communities (and other plant-associated arthropods) is an emerging field which can revolutionize understanding of pollinator biodiversity (Van Zandt et al., 2020), and opens avenues for research at DoD installations in numerous contexts. The authors believe two subsequent studies are needed to further assess the effectiveness of this technology to benefit installations.

First, while the team has demonstrated that pollinator microfluidic eDNA metabarcoding is effective in documenting arthropod pollinator communities, there are also many vertebrates (particularly bats and birds) that are key pollinators and are increasingly afforded protections under the U.S. Endangered Species Act. Yet, controlled lab studies using these pollinator species are difficult. As such, the team proposes research at Fort Huachuca, Arizona, to assess whether or not this approach is able to detect the federally listed lesser long-nosed bat (*Leptonycteris yerbabuena*), and the broader pollinator community from cacti flowers and an assessment of bird pollinator networks (in collaboration with Dr. Jinelle Sperry) on DoD land holdings on Oahu, Hawaii. These studies will validate this approach for non-arthropod and yet critically important pollinator species.

Second, the team proposes a study to compare this approach with conventional methods to document pollinator networks in natural settings. Specifically, the team proposes a study of grassland pollinator communities at Fort McCoy in Wisconsin, home to the karner blue butterfly (*Lycaeides melissa samuelis*), federally listed as endangered. In this study, a robust sampling design that samples diverse flower species at multiple temporal and spatial scales will be used to document pollinator communities via microfluidic eDNA metabarcoding. This will be paired with a conventional observational assessment of pollinator communities. Here, the team will compare and contrast both the economic cost of each approach as well as the time investment required to document pollinator communities. This will allow an understanding of which approach yields a greater return on investment, and which approach minimizes range closure times on installations.

In addition to the critical studies identified above, the team proposes deploying this methodology in real-world applications to provide actionable information to DoD installations, and examples are provided below. First, as concerns over the loss of pollinator biodiversity have grown (e.g., Potts et al., 2016), species are increasingly being afforded protection under the U.S. Endangered Species Act. As such, pollinator eDNA metabarcoding can be rapidly deployed on large spatial scales to provide a preliminary screening for these species. This will allow for high intensity conventional sampling to be strategically targeted to locations with a high likelihood of success and maximizing return-on-investment.

Secondly, biodiversity is essential for resilient ecosystems (Oliver et al., 2015a,b), and pollinators deliver critical ecosystem services (Hein, 2009; Vanbergen, 2013; Winfree, 2013). In the interest of promoting installation resilience, assessing the health of plant and pollinator networks is essential. By coupling this approach with microfluidic metabarcoding of pollen sacs, critical information of plant-pollinator interactions can be provided, ultimately serving as a proxy for resilience.

5.1 BENEFITS

The team believes the benefits of this approach are three-fold. First, conventional measuring and monitoring of pollinator communities is both costly and time consuming. Researchers with expert taxonomic knowledge (an increasingly rare commodity) must spend large amounts of time observing pollinator/plant interactions. Reducing the time and effort needed to acquire the same information reduces the financial burden of surveying. This has an added benefit in that once species of conservation concern are identified, intensive monitoring and adaptive management can be strategically targeted. Thus, in both phases (data and information acquisition and adaptive management) costs are reduced, thereby maximizing return-on-investment.

Second, the team believes that eDNA sampling requires much less time in the field, resulting in reduced range closure time, increasing mission readiness, and ultimately lethality. To provide real-world context, PI Mark Davis led an eDNA study at the U.S. Army Joint Readiness Training Center at Fort Polk assessing imperiled freshwater crayfish and mussels (Davis et al., 2019). Using conventional sampling (timed searches), the team was able to survey a total of 9 sites over three days. Using eDNA sampling, the team was able to sample a total of 29 sites over an identical three-day period. The rapidity with which eDNA sampling proceeds often provides a distinct advantage to conventional sampling.

Finally, integrated pollinator microfluidic eDNA metabarcoding approaches (as above) can reveal complex networks that undergird critical ecosystem functions and services, and ultimately allow for an assessment of ecosystem resiliency. By combining data from pollinator microfluidic eDNA metabarcoding with, for example, standard metabarcoding of pollen sacs using plant genetic markers (e.g., rbcL, ITS2, trnL), plant/pollinator networks can be formally assessed. Moreover, by assessing a gradient of sites (from degraded to restored/enhanced/protected to protected/pristine) researchers can develop indices of biotic integrity to use as a proxy for ecosystem functions and services, ultimately providing insights into the resilience of ecological systems on DoD land holdings.

6.0 REFERENCES

Allsop, M.H., de Lange, W.J., & Veldtman, R. (2008) Valuing insect pollination services with cost of replacement. *PLoS ONE*. 10.1371/journal.pone.0003128

Amir, A., McDonald, D., Navas-Molina, J.A., Kopylova, E., Morton, J.T., Xu, Z.Z., Kightley, E.P., Thompson, L.R., Hyde, E.R., Gonzalez, A., & Knight, R. (2017) Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems*, **2**, e00191–16. https://doi.org/10.1128/mSystems.00191-16

Aylagas, E., Borja, A., Irigoien, Z., & Rodriguez-Ezpeleta, N. (2016) Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. *Frontiers in Marine Science.* DOI: 10.3389/fmars.2016.00096

Baker, C. (2016) bakerccm/entrez_qiime: entrez_qiime v2.0. https://doi.org/10.5281/zenodo.159607

Barnosky, A.D., Matzke, N., Momiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C, Mcguiter, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B., & Ferrer, E.A. (2011) Has the Earth's sixth mass extinction already arrived? *Nature*, **471**, 51–57.

Bartomeus, I. (2013) Understanding linkage rules in plant-pollinator networks by using hierarchical models that incorporate pollinator detectability and plant traits. *PLoS ONE* DOI: 10.1371/journal.pone.0069200

Baselga A., & Orme C.D.L. (2012) betapart: an R package for the study of beta diversity: Betapart package. *Methods in Ecology and Evolution,* **3**, 808–812.

Bolger, A.M., Lohse, M., & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016) obitools: a unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, **16**, 176–182. https://doi.org/10.1111/1755-0998.12428

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., & Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, **13**, 581–583. https://doi.org/10.1038/nmeth.3869

Callahan, B.J., McMurdie, P.J., & Holmes, S.P. (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, **11**, 2639–2643. https://doi.org/10.1038/ismej.2017.119

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T.L. (2009) BLAST+: Architecture and applications. *BMC Bioinformatics*, **10**, 421. https://doi.org/10.1186/1471-2105-10-421

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

Chacoff, N.P., Vazquez, D.P., Lomascolo, S.B., Stevani, E.K., Dorado, J., & Padron, B. (2012) Evaluating sampling completeness in a desert plant-pollinator network. *Journal of Animal Ecology,* **81**, 190–200.

Clarke, L.J., Czechowski, P., Soubrier, J., Stevens, M.I., & Cooper, A. (2014a) Modular tagging of amplicons using a single PCR for high-throughput sequencing. *Molecular Ecology Resources*, **14**, 117–121. https://doi.org/10.1111/1755-0998.12162

Corse, E., Tougard, C., Archambaud-Suard, G., Agnèse, J.-F., Messu Mandeng, F.D., Bilong, C.F., Duneau, D., Zinger, L., Chappaz, R., Xu, C.C.Y., Meglécz, E., & Dubut, V. (2019) One-locus-several-primers: A strategy to improve the taxonomic and haplotypic coverage in diet metabarcoding studies. *Ecology and Evolution*, **9**, 4603–4620. https://doi.org/10.1002/ece3.5063

Curd, E.E., Gold, Z., Kandlikar, G.S., Gomer, J., Ogden, M., O'Connell, T., Pipes, L., Schweizer, T.M., Rabichow, L., Lin, M., Shi, B., Barber, P.H., Kraft, N., Wayne, R., & Meyer, R.S. (2019) Anacapa toolkit: an environmental DNA toolkit for processing multilocus metabarcode datasets. *Methods in Ecology and Evolution*, **10**, 1469–1475. https://doi.org/10.1111/2041-210X.13214

Davis, M.A., Larson, E.R., Sperry, J., Tiemann, J., Harper, L.R., Tetzlaff, S., Sternhagen, E.C., Holtswarth, J., & Snyder, E.D. (2019). Assessing the status and distribution of at-risk aquatic species in the New Lands region of the Fort Polk Army Installation. Technical Report submitted to the Louisiana Department of Wildlife and Fisheries, and the United States Department of Defense, Joint Readiness Training Center - Fort Polk.

Deiner, K., Bik, H.M., Machler, E., Seymor, M., Lacoursiere-Roussel, A., Altermatt, F., Creer, F., Bista, I., Lodge, D.M., de Vere, N., Pfrener, M.E., & Bernatchez, L. (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology,* **26**, 5872–5895.

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., & Miaud, C. (2012) Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates cathesbeianus. Journal of Applied Ecology,* **49**, 953–959.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acid Research*, **32**, 1792–1797. https://doi.org/10.1093/nar/gkh340

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., & Knight, R. (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, **27**, 2194–2200. https://doi.org/10.1093/bioinformatics/btr381

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

Elbrecht, V., & Leese, F. (2017) Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, **5**, 1–11. https://doi.org/10.3389/fenvs.2017.00011

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

Elbrecht, V., Braukmann, T.W.A., Ivanova, N.V., Prosser, S.W.J., Hajibabaei, M., Wright, M., Zakharov, E.V., Hebert, P.D.N., & Steinke, D. (2019) Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, **7**, e7745. https://doi.org/10.7717/peerj.7745

Ficetola, G.F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessiere, J., Taberlet, P., & Pompanon, F. (2010) An *in silico* approach for the evaluation of DNA barcodes. *BMC Genomics*, **11**, 434. https://doi.org/10.1186/1471-2164-11-434

Ficetola, G.F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., DeBarba, M., Gielly, L., Lopes, C.M., Boyer, F., Pompanon, F., Raye, G., & Taberlet, P. (2015) Replication levels, false presences and the estimation of presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, **15**, 543–555. https://doi.org/10.1111/1755-0998.12338

Gallai, N., Salles, J.M., Settele, J., & Vaissiere, B.E. (2009) Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. *Ecological Economics,* **68**, 810–821.

Gao, X., Lin, H., Revanna, K., & Dong, Q. (2017) A Bayesian taxonomic classification method for 16S rRNA gene sequences with improved species-level accuracy. *BMC Bioinformatics*, **18**, 247. https://doi.org/10.1186/s12859-017-1670-4

Goldberg, C.S., Strickler, K.M., & Pilliod, D.S. (2015) Moving environmental DNA methods from concept to practice for monitoring aquatic macroorganisms. *Biological Conservation,* **183**, 1–3.

Goodman, D. (1987) The demography of chance extinction. In: Soule´, M. E. (ed.), Viable populations for conservation. Cambridge Univ. Press: 11–34.

Gordon, A., & Hannon, G.J. (2010) Fastx-toolkit. FASTQ/A Short-Reads Preprocessing Tools (Unpublished). http://hannonlab.cshl.edu/ Fastx_toolkit, 5.

Günther, B., Knebelsberger, T., Neumann, H., Laakmann, S., & Martínez Arbizu, P. (2018) Metabarcoding of marine environmental DNA based on mitochondrial and nuclear genes. *Scientific Reports*, **8**, 14822. https://doi.org/10.1038/s41598-018-32917-x

Hauck, L.L., Weitemier, K.A., Penaluna, B.E., Garcia, T.S., & Cronn, R. (2019) Casting a broader net: Using microfluidic metagenomics to capture aquatic biodiversity data from diverse taxonomic targets. *Environmental DNA*, **1**, 2008. https://doi.org/10.1002/edn3.26

Hebert, P.D.N., Cywinska, A., Ball, S.L., & Dewaard, J.R. (2003). Biological indentifications through DNA barcodes. *Proceedings of the Royal Society of London*, **270**, 313-321.

Hein, L. (2009) The economic value of the pollination service, a review across scales. *The Open Ecology Journal*, **2**, 74–82.

Illinois Natural History Survey Insect Collection. (2015) Illinois Natural History Survey. Occurrence dataset. https://doi.org/10.15468/eol0pe accessed via GBIF.org on 2019-06-18.

Ji, Y., Ashton, L., Pedley, S.M., Edwards, D.P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P.M., Woodcock, P., Edwards, F.A., Larsen, T.H., Hsu, W.W., Benedick, S., Hamer, K.C., Wilcove, D.S., Bruce, C., Wang, X., Levi, T., Lott, M., Emerson, G.C., & Yu, D.W. (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters,* **16**, 1245–1257.

Kevan, P.G., & Phillips, T.P. (2001) The economic impacts of pollinator declines: an approach to assessing the consequences. *Conservation Ecology,* **5**, 8.

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**, e0177643. https://doi.org/10.1371/journal.pone.0177643

Kolbert, E. 2014. The sixth extinction: an unnatural history. Henry Holt and Co., New York, NY. Krehenwinkel, H., Wolf, M., Lim, J.Y., Rominger, A.J., Simison, W.B., & Gillespie, R.G.

(2017) Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, **7**, 17668. https://doi.org/10.1038/s41598-017-17333-x

Krutzen, P.J. 2006. The "Anthropocene." In: Ehlers, E., and T. Krafft (eds). Earth System Science in the Anthropocene. Springer, Berlin, Heidelberg.

Langmead, B., & Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357–359. https://doi.org/10.1038/nmeth.1923

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T. & Machida, R.J. (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, **10**, 34. https://doi.org/10.1186/1742-9994-10-34

Lever, J.J., van Nes, E.H., Scheffer, M., & Bascompte, J. (2014) The sudden collapse of pollinator communities. *Ecology Letters,* **17**, 350–359.

Losey, J.E., and Vaughn, M. (2006) The economic value of ecosystem services provided by insects. *BioScience,* **56**, 311–323.

Macher, J-N., Vivancos, A., Piggott, J.J., Centeno, F.C., Matthaei, C.D., & Leese, F. (2018) Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate COI primers. *Molecular Ecology Resources*, **18**, 1456–1468.

Magoč, T., & Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, **27**, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507

Matthies, D., Brauer, I., Maibom, W., & Tscharntke, T. (2004) Population size and risk of local extinction: empirical evidence from rare plants. *Oikos,* **105**, 481–488.

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

Menges, E. S. (1991) The application of minimum viable population theory to plants. *In*: Falk, D. A. and Holsinger, K. E. (eds), Genetics and conservation of rare plants. Oxford Univ. Press: 47–61.

Meusnier, I., Singer, G.A.C., Landry, J.-F., Hickey, D.A., Hebert, P.D.N., & Hajibabaei, M. (2008) A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*, **9**, 214. https://doi.org/10.1186/1471-2164-9-214

Nakagawa S., & Cuthill I.C. (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological reviews of the Cambridge Philosophical Society* **82**, 591–605.

Needham, D.M., Sachdeva, R., & Fuhrman, J.A. (2017) Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. *The ISME Journal*, **11**, 1614–1629. https://doi.org/10.1038/ismej.2017.29

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., & Wagner, H. (2019) vegan: Community Ecology Package. R package version 2.5-4. https://CRAN.R-project.org/package=vegan

Oliver, T.H., Heard, M.S., Isaac, N.J.B., Roy, D.B., Procter, D., Eigenbrod, F., Freckleton, R., Hector, A., Orme, C.D.L., Petchey, O.L., Proenca, V., Raffaelli, D., Suttle, K.B., Mace, G.M., Martin-Lopez, B., Woodcock, B.A., & Bullock, J.M. (2015a). Biodiversity and resilience of ecosystem functions. *Trends in Ecology & Evolution*, **30**, 673–684.

Oliver, T.H., Isaac, N.J.B., August, T.A., Woodcock, B.A., Roy, D.B., & Bullock, J.M. (2015b) Declining resilience of ecosystem functions under biodiversity loss. *Nature Communications*, **6**, 10122. https://doi.org/10.1038/ncomms10122

Pimm S.L., & Brooks, T.M. (2000) The sixth extinction: how large, where, and when. Nature and human society: the quest for a sustainable world. National Academy Press, Washington DC.

Plein, M., Morris, W.K., Moir, M.L., & Vesk, P.A. (2017) Identifying species at coextinction risk when detection is imperfect: Model evaluation and case study. *PLoS ONE* DOI: 10.1371/journal.pone.0183351

Potts, S.G., Beismeijer, J.C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W.E. (2010) Global pollinator declines: trends, impacts, and drivers. *Trends in Ecology and Evolution,* **25**, 3450-3453.

Potts, S.G., Imperatriz-Fonseca, V., Ngo, H.T., Aizen, M.A., Biesmeijer, J.C., Breeze, T.D., Dicks, L.V., Garibaldi, L.A., Hill, R., Settele, J., & Vanbergen, A.J. (2016) Safeguarding pollinators and their values to human well-being. *Nature*, **540**, 220–229. https://doi.org/10.1038/nature20588

R Core Team. (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Renshaw, M.A., Olds, B.P., Jerde, C.L., McVeigh, M.M., & Loge, D.M. (2015). The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol-chloroform-isoamyl alcohol DNA extraction. *Molecular Ecology Resources*, **15**, 168– 176. https://doi.org/10.1111/1755-0998.12281

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, **4**, e2584. https://doi.org/10.7717/peerj.2584

Sheffield, C.S., Herbet, 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.

Southwick, E.E., & Southwick, J. (1996) Estimating the economic value of honeybees (Hymenoptera: Apidae) as agricultural pollinators in the United States. *Journal of Economic Entomology,* **85**, 621–633.

Szitenberg, A., John, M., Blaxter, M.L., & Lunt, D.H. (2015) ReproPhylo: An Environment for Reproducible Phylogenomics. *PLoS Computational Biology*, **11**, e1004447. https://doi.org/10.1371/journal.pcbi.1004447

Taberlet, P., Cojssac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology,* **21**. 2045–2050.

Thomann, M., Imbert, E., Devaux, C., & Cheptou, P-O. (2013) Flowering plants under global pollinator decline. *Trends in Plant Science,* **18**, 353–359.

Thomsen, P.F., Kielgast, J., Iversen, L.L., Wiuf, C., Rasmussen, M., Gilbert, M.T.P., Orlando, L., & Willerslev, E. (2012) Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology,* **21**, 2565–2573.

Thomsen, P.F., & Willerslev, E. (2015) Environmental DNA – an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation.* **183**, 4–18.

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

Tur, C., Castro-Urgal, R., & Traveset, A. (2013) Linking plant specialization to dependence in interactions for seed set in pollination networks. *PLoS ONE* 8: DOI: e78294.

Valentini, A., Taberlet, P., Miaud, C., Ciade, R., Herder, J., Thomsen, P.F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N, Roset, N., Copp, G.H., Geniez, P., Pont, D., Argillier, C., Baudoin, J-M., Peroux, T., Crivelli, A.J., Olivier, A., Acqueberge, M., Le Brun, M., Moller, P.R., Willerslev, E., & Dejean, T. (2016)

Next- generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology,* **25**, 929–942.

Vamos, E.E., Elbrecht, V., & Leese, F. (2017) Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics*, **1**, e14625. https://doi.org/10.3897/mbmg.1.14625

Vanbergen, A.J. (2013) Threats to an ecosystem service: pressures on pollinators. *Frontiers in Ecology and the Environment*, **11**, 251–259. https://doi.org/10.1890/120126

Van Zandt, P.A. Johnson, D.D., Hartley, C., LeCroy, K.A., Shew, H.W., Davis, B.T., & Lehnert, M.S. (2020) Which moths might be pollinators? Approaches in the search for flower-visiting needles in the Lepidopteran haystack. *Ecological Entomology,* **45**, 13–25. https://doi.org/10.1111/een.12782

Wake, D.B., & Vredenburg, V.T. (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of the Sciences USA,* **105**, 11466–11473.

Weiner, C.N., Werner, M., Linsenmair, K.E., & Blutgen, N. (2014) Land-use impacts on plant-pollinator networks: interaction strength and specialization predict pollinator declines. *Ecology,* **95***,* 466–474. https://doi.org/10.1890/13-0436.1

Winfree, R., Gross, B.J., & Kremen, C. (2011) Valuing pollination services to agriculture. *Ecological Economics,* **71**, 80–88.

Winfree, R. (2013) Global change, biodiversity, and ecosystem services: what can we learn from studies of pollination? *Basic and Applied Ecology*, **14**, 453–460. https://doi.org/10.1016/j.baae.2013.07.004

Young, B.E., Auer, S., Ormes, M., Rapacciuolo, G., Schweitzer, D., & Sears, N. (2017) Are pollinating hawk moths declining in the Northern United States? An analysis of collection records. *PLoS ONE* DOI: 10.1371/journal.pone.0185683

Yu, D.W., Ji, Y., Emerson, B.C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution,* **3**, 613–623.

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

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000) A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, **7**, 203–214. https://doi.org/10.1089/10665270050081478

APPENDIX A FLOWER SPECIES THAT COMPRISED THE GREENHOUSE COMMUNITY

These flowering plants provided a constant supply of pollen and nectar to the introduced common eastern bumblebee (*Bombus impatiens*) established in the greenhouse

Table 7. Greenhouse Community Flower Species

Botanical Name	Common Name
Nemophila menziesii	Baby Blue Eyes
Gypsophila elegans	Baby's Breath
Linaria maroccana	Baby Snapdragon, Moroccan Toadflax
Eschscholzia californica	California Poppy
Cynoglossum amabile	Chinese Forget-Me-Not
Centaurea cyannus	Cornflower
Helianthus annuus	Dwarf Sunflower Sunspot
Clarkia amoena	Farwell-To-Spring, Godetia
Mirabilis jalapa	Four O-Clock
Gilia capitata	Globe Gillia, Blue-Thimble-Flower
Clarkia unguiculata	Elegent Clarkia, Mountain Garland
Gaillardia pulchella	Indian Blanket
Lupinus	Lupine
Silene armeria	None-So-Pretty, Sweet William Catchfly
Coreopsis tinctoria	Plains Coreopsis
Calendula officinalis	Pot Marigold
Ammi majus	Queen Anne's Lace
Papaver rhoeas	Red Poppy, Common Poppy
Delphinium ajacis	Rocket Larkspur
Lavatera trimestris	Rose Mallow
Linum grandiflorum "Rubrum"	Scarlet Flax
Cosmos sulphureus	Sulphur Cosmos

APPENDIX B DETAILS OF THE PRIMER PANEL SELECTED FOR POLLINATOR EDNA MICROFLUIDIC METABARCODING.

Table 8. Primer Panel Details

Primer set			Details			% species amplified *in silico* (N = 5313) mismatches			
Name	Forward primer (5' - 3')	Reverse primer (5' - 3')	Marker	Fragment size (bp)	Target taxa	1	2	3	4
Ins16S-1F & Ins16S-1Rshort	TRRGACGAGAAGACCCTATA	ACGCTGTTATCCCTAARGTA	16S	156	Invertebrates	1.45	8.02	10.20	10.90
MOL16S_F & MOL16S_R	RRWRGACRAGAAGACCCT	ARTCCAACATCGAGGT	16S	183-310	Molluscs	3.65	6.61	9.71	10.75
BF1 & BR1	ACWGGWTGRACWGTNTAYCC	ARYATDGTRATDGCHCCDGC	COI	217	Invertebrates	29.68	43.93	46.19	46.45
BF1 & BR2	ACWGGWTGRACWGTNTAYCC	TCDGGRTGNCCRAARAAYCA	COI	316	Invertebrates	7.06	11.67	12.35	12.67
BF2 & BR2	GCHCCHGAYATRGCHTTYCC	TCDGGRTGNCCRAARAAYCA	COI	421	Invertebrates	6.08	7.91	8.85	9.94
BF3 & BR2	CCHGAYATRGCHTTYCCHCG	TCDGGRTGNCCRAARAAYCA	COI	418	Arthropods	7.27	9.62	9.82	9.98
nsCOIFo & mlCOIintK	THATRATNGGNGGNTTYGGNAAHTG	GGRGGRTAWACWGTTCAWCCWGTWCC	COI	124	Invertebrates	16.83	33.99	42.10	44.02
mlCOIintF & jgHCO2198	GGWACWGGWTGAACWGTWTAYCCYCC	TAIACYTCIGGRTGICCRAARAAYCA	COI	313	Metazoans	0.00	0.00	0.00	11.71
Vamos_fwh1	YTCHACWAAYCAYAARGAYATYGG	ARTCARTTWCCRAAHCCHCC	COI	178	Invertebrates	2.88	4.20	4.69	4.74
Vamos_fwh2n	GGDACWGGWTGAACWGTWTAYCCHCC	GTRATWGCHCCDGCTARWACWGG	COI	205	Invertebrates	19.73	39.02	44.91	46.23
Uni-MinibarF1 & Uni-MinibarR1	TCCACTAATCACAARGATATTGGTAC	GAAAATCATAATGAAGGCATGAGC	COI	127	Metazoans	0.00	0.00	0.00	0.17
ZBJ-ArtF1c & ZBJ-ArtR2c	AGATATTGGAACWTTATATTTTATTTTTGG	WACTAATCAATTWCCAAATCCTCC	COI	157	Arthropods	0.26	1.32	2.88	4.40
LepLCO & McoiR1	RKTCAACMAATCATAAAGATATTGG	AATCCBCCRATTAWAATKGGTAT	COI	~150	Invertebrates	0.23	1.05	2.09	3.33
LepLCO & McoiR2	RKTCAACMAATCATAAAGATATTGG	CCBCCRATTAWAATKGGTATHAC	COI	~150	Invertebrates	0.24	1.13	2.03	3.20
LepLCO & MLepF1rev	RKTCAACMAATCATAAAGATATTGG	CGTGGAAAWGCTATATCWGGTG	COI	~150	Invertebrates	0.02	0.26	1.07	2.99

APPENDIX C RESULTS OF *IN SILICO* PRIMER TESTING FOR EACH PRIMER PAIR CONSIDERED

The total number of species and the proportion of species that amplified out of a 5313 Illinois native pollinator species database are depicted.

Table 9. Results of in Silico Primer Testing

		Total Amplified (mismatches)				Proportion Amplified (mismatches)			
Primer Set	**Region**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Corse_LepLCO-McoiR1	COI	12	56	111	177	0	1	2	3
Corse_LepLCO-McoiR2	COI	13	60	108	170	0	1	2	3
Corse_LepLCO-MLepF1rev	COI	1	14	57	159	0	0	1	3
Corse_MFZR	COI	0	0	7	63	0	0	0	1
Elbrecht_BF1-BR1	COI	1577	2334	2454	2468	30	44	46	46
Elbrecht_BF1-BR2	COI	375	620	656	673	7	12	12	13
Elbrecht_BF2-BR2	COI	323	420	470	528	6	8	9	10
Elbrecht_BF3-BR2	COI	386	511	522	530	7	10	10	10
Galan_MG	COI	48	140	232	244	1	3	4	5
Gunther_nsCOIFo-mlCOIintK	COI	894	1806	2237	2339	17	34	42	44
Leray_Geller	COI	0	0	0	622	0	0	0	12
Meusnier_Uni-Minibar	COI	0	0	0	9	0	0	0	0
Shokralla_Ill-B-F-HCO2198	COI	0	0	0	353	0	0	0	7
Shokralla_LCO1490-Ill-C-R	COI	0	0	0	0	0	0	0	0
Vamos_fwh1	COI	153	223	249	252	3	4	5	5
Vamos_fwh2	COI	71	1795	2336	2449	1	34	44	46
Vamos_fwh2n	COI	1048	2073	2386	2456	20	39	45	46
Zeale_ZBJ-Art	COI	14	70	153	234	0	1	3	4
Clarke_Ins16S	16s	77	426	542	579	1	8	10	11
DeBarba_16SMAV	16s	0	1	1	1	0	0	0	0
Klymus_MOL16S	16s	194	351	516	571	4	7	10	11
Klymus_SPH16S	16s	0	0	0	3	0	0	0	0