ESTCP

EXECUTIVE SUMMARY

Multifaceted High-Throughput DNA Barcoding for Addressing Critical Data Gaps for At-Risk Bats on DoD Installations

Richard Lance
Eric Britzke
US Army ERDC Environmental Laboratory

Christine Edwards
Missouri Botanical Garden

June 2024

ESTCP Project RC-201417

ESTCP EXECUTIVE SUMMARY

Project: RC-201417

TABLE OF CONTENTS

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

DNA	deoxyribonucleic acid
DoD	Department of Defense
MDM	multifaceted DNA metabarcoding
NGS	next-generation sequencing
PCR	polymerase chain reaction
qPCR	quantitative polymerase chain reaction
WNS	White-nose Syndrome

ACKNOWLEDGMENTS

The project team would like to thank the following individuals who, throughout the entire project, were critical field, laboratory, or field and laboratory participants in data acquisition, analysis, and/or reporting:

- X. Guan (Bennett Aerospace), J. Swift (Missouri Botanical Garden and St. Louis University), and
- D. Lindsay (U.S. Army Engineer Research and Development Center (ERDC) Environmental Laboratory).

The project team also expresses appreciation for the many colleagues who assisted with sample acquisition and comparative sample analysis for the initial demonstration of multifaceted deoxyribonucleic acid (DNA) metabarcoding (MDM) with captive bats, well-known and monitored bat colonies, bats sampled in the wild, and archived bat specimens. The project team also expresses appreciation for those experts in molecular ecology and bat ecology that allowed access to data, provided needed feedback for or assistance with scientific manuscript preparation, or helped cross-validate project outcomes. The list of such helpful, and in many ways, critical, colleagues includes the following:

C. Bathrick (Fort Worth Zoo), R. Doege (Fort Worth Zoo), J. Elston (Fort Worth Zoo), K. Putnum (Fort Worth Zoo), A. Ward (Fort Worth Zoo), T. Nicewander (Fort Worth Zoo), P. Eiamcharoen (University of California, Davis), K. Keel (University of California, Davis), J. Nawrocki (Wildlife Ecology Institute), S. Morgan (University of Albany), J. Jung (ERDC Environmental Laboratory), K. Patton (University of Louisiana Lafayette), E. Tobin (University of Louisiana Lafayette), M. Whitby (Bat Conservation International), R. Rudd (Rabies Laboratory, New York State Department of Health), M. Jung (Bennett Aerospace), A. Piaggio (National Wildlife Research Center, U.S. Department of Agriculture), D. Bergman (Wildlife Services, U.S. Department of Agriculture), L. Van Pelt (Wildlife Services, U.S. Department of Agriculture), G. Balsamo (Infectious Disease Epidemiology Section, Louisiana Department of Health and Hospitals), B. Sasse (Arkansas Game & Fish Commission), S. Feist (ERDC Environmental Laboratory), B. Hines (Stantec Consulting Services), J. Patten (U.S. Army Fort Liberty), C. Martin, Mississippi Bat Working Group, J. Korstian (Texas Christian University and Texas Tech University), F. Walker (Northern Arizona University), and L. Patrick (Louisiana State University and University of Minnesota-Twin Cities).

The project team is grateful to the following for the assistance with scoping, permitting, and execution of Field Demonstrations 1 and 2, and/or with the preparation of the associated scientific papers: D. Brewer (Fort Huachuca, U.S. Army), C. Dobony (Fort Drum, U.S. Army), R. Rainbolt (Fort Drum, U.S. Army), Antoinette Piaggio (National Wildlife Research Center, U.S. Department of Agriculture - Animal and Plant Health Inspection Service).

Finally, the project team wishes to acknowledge R. Niver, who served as project team liaison with the U.S. Fish and Wildlife Service.

1.0 INTRODUCTION

There are over 1,400 recognized species of bat in the world, with over 40 species found in the U.S. (not including territories or freely associated states). With the exception of bases in the Arctic, near-Arctic, and some remote islands, every Department of Defense (DoD) installation will be inhabited by one, and usually numerous, bat species. Several bat species are currently federally listed as endangered, threatened, or at-risk—all of which occur on one or more DoD installations. White-nose Syndrome (WNS), an epidemic disease in North American bats caused by a fungal pathogen, has drastically reduced bat populations, leading to additional potential Endangered Species Act listings. Lack of data hampers effective bat conservation efforts and may lead to a reduction in the overall ecological resilience that supports effective and realistic training on installations, along with unnecessary regulatory impediments to military activities. Adherence to key legislative mandates, like the Sikes Act and Endangered Species Act, require reliable, robust data on plant and animal populations. In many cases, these data may be very challenging or expensive to obtain. As small, highly mobile, nocturnal taxa that often roost in inaccessible locations, bats are one such challenging group.

Innovative DNA methods that take advantage of noninvasive samples, for which animals do not need to be captured, handled, or undergo physical sampling (or for which plants do not have to be collected or physically sampled), are often ideal for difficult-to-sample species. DNA can be the key to uncovering data on many different aspects of a species' natural history or about an ecological community. DNA found in readily obtained noninvasive samples, like bat guano, can, for example, readily provide data on bat species, sex, diet, health, and population numbers. Multifaceted DNA metabarcoding (MDM) is an example of a noninvasive approach that provides a cost-effective means for accessing data that is often otherwise very difficult or costly to obtain. However, approaches like MDM require additional, broader validation in order to lay the groundwork for wider use.

2.0 OBJECTIVES

The goal of this project was to validate MDM as a rapid, user-friendly, cost-effective means for obtaining essential data for bats on DoD lands. In particular, the project sought to demonstrate how MDM would provide robust data on the following:

- Presence of different bat species in roosts
- Bat population sizes and trends
- Presence or prevalence of WNS
- General health and parasite loads of bat populations
- Sex ratios of bat populations
- Dietary resource or habitat requirements for bat populations

Phase **1** of the project focused on tests with known populations and controlled systems in order to validate the reliability of MDM-generated data. Phase **1** efforts to estimate population size via MDM were unsuccessful and were dropped from Phase 2 plans due to limitations in scheduling and budget. Phase 2 focused on applying MDM in field demonstrations with two bat communities inhabiting different military installations.

3.0 TECHNOLOGY DESCRIPTION

MDM is based on three fundamental technological and methodological pillars: 1) next-generation DNA sequencing (NGS), 2) DNA barcoding, and 3) molecular scatology. NGS platforms can produce hundreds of thousands or even millions of DNA sequence reads from a single sample or multiple samples in a single "run." Several commercial NGS platforms have been developed and are variously utilized by laboratories across the world. In this study, NGS was performed using Illumina DNA sequencing technology and chemistries, which is the most widely utilized NGS platform across all fields of biotechnology and environmental genomics. In particular, the project team utilized an Illumina MiSeq platform. For this study, the MiSeq was selected in large part because it generates paired-end DNA sequences up to 600 base pairs in length. This allowed a wider range of intact, entire loci (e.g., DNA barcode sequences, sex identification sequences) to be sequenced in a single DNA sequencing read, eliminating the need for assembly of amplicon fragments (and any associated risks of assembly errors). The MiSeq is also among the most affordable and widely available NGS platforms. The research team maintained a MiSeq in their noninvasive genetics lab, allowing for more flexible and efficient optimization for challenging samples. A key feature of NGS in this project was the ability to incorporate many individual samples into single sequencing runs, which were sorted, post-run, using multiplex identifier indexes. These indexes were short, unique, oligonucleotide sequences that were added to the ends of target DNA amplicons during the preparation of sample DNA for sequencing (i.e., library prep). Indexes allowed NGS-generated DNA sequences to be parsed by sample, barcode locus, and/or other key features of a sample amplicon set. For example, a single MDM NGS could be run for 100 guano pellet samples and one DNA locus for each of three data classes (e.g., species ID, sex determination, WNS detection). The final dataset would consist of 300 unique bins of DNA sequence data. These bins could then be analyzed individually or combined by individual, date, location, species, and/or sex, etc., as desired. Further details on NGS processing are found in Section 5.3.3 and Figures 5.11 and 5.12 of the full report.

Strictly speaking, DNA barcodes refer to DNA loci that often exhibit sequences unique to a taxonomic group, allowing the identification of samples to the level of species, genus, or family. For example, this study relied on several different barcode loci, each with sufficient levels of phylogenetic sequence diversification for the determination of bat species, or the identification of insect prey, forage plants, and parasites whose DNA occurred in the sample. Less formally, DNA loci that provided data on the sex of bats, WNS presence, and DNA genotype were included as "barcodes" in the MDM runs. The simultaneous generation of DNA related to several classes of data from a single set of samples is reflected in the term "multifaceted" component of MDM.

Molecular scatology refers to the use of DNA extracted from fecal material (or *scat)* to provide some class of genetic data of interest, including species identification, diet analysis, individual genotyping, etc. Genetic material present in scat would typically comprise DNA from the source organism (e.g., sloughed gastrointestinal wall cells), food items, endoparasites, and commensal gut microbiota. Scat samples represent an important category of "noninvasive" animal samples. Because DNA obtained from scat is often of low quality (degraded, low concentrations) and mixed with compounds that inhibit DNA processing methods, molecular scatology requires a unique set of methods and protocols in order to be effective.

While this report highlighted the successful application of MDM in detecting bat species, analyzing bat diet, identifying bat parasites, determining bat sex, and detecting WNS, *it is important to understand that MDM and like methods are not limited to bats.* The technology offers substantial advantages for ecological research and conservation efforts for *many different species,* particularly, but not exclusively, for those that are difficult to directly observe, or for which capture and handling are serious concerns. Because each species and study will have its own unique features, careful consideration and some study-specific tailoring of MDM methods will be necessary in most cases.

4.0 PERFORMANCE ASSESSMENT

MDM was demonstrated in two phases (Figure ES-1). Phase 1 included a demonstration of the comparative performance of DNA data versus conventional data for six data classes. Phase 2 included a demonstration of MDM with two different bat communities on DoD installations. Phase 1 included two steps. Step 1 focused on demonstrating that noninvasive DNA-based methods for bat species identification, sex identification, white-nose syndrome fungal pathogen detection, diet characterization, parasite characterization, and colony or population size estimation performed as well as or better than conventional or standard methods. DNA analyses included either both simple, taxonomically-narrow assays (e.g., quantitative real time polymerase chain reaction (PCR) (qPCR)) and more complex, taxonomically-broad assays (e.g., NGS metabarcoding). Step 2 focused on testing and demonstrating the capacity of NGS metabarcoding, as part of a multilocus MDM approach, to provide accurate and robust datasets for several data classes. Phase 2 focused on demonstrating MDM for two distinct bat communities of interest to DoD. One community, on Fort Drum, NY, consisted of a maternity colony of Little Brown Bats *(Myotis lucifugus).* Little Brown Bats have declined substantially in the central and eastern U.S. due to WNS and are currently under review for federal listing as a threatened or endangered species. Historically, this species was one of the most common and widespread species in the U.S. and was found on many installations. The second bat community consisted of a mixed species assemblage utilizing two roosts on Fort Huachuca, AZ. This assemblage included two nectar-feeding species, which allowed the study to include plant items in diet analyses. Additionally, one of the nectar-feeding bats, the Lesser Long-nosed Bat *(Leptonycteris yerbabuenae),* was only recently (2018) removed from the federal list of threatened and endangered species.

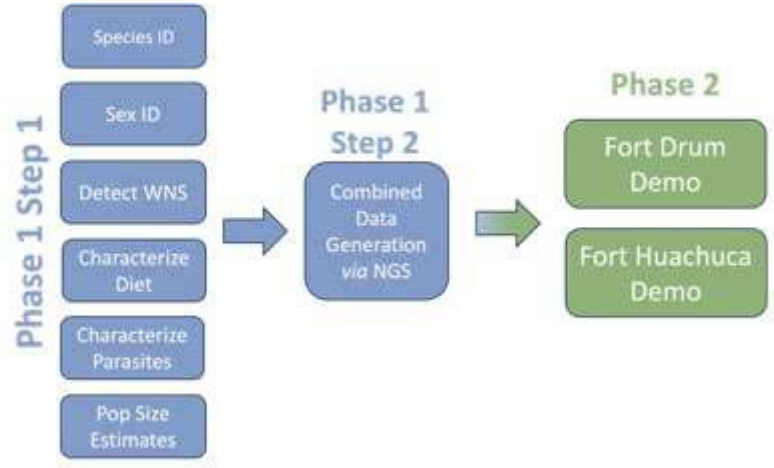

Figure ES-1. Basic Overview of MDM Bat Study Phases and Steps.

In Phase 1, the research team demonstrated that more than 50 guano samples from five species of bat (individual bats identified to species in field prior to guano collection) could be accurately assigned to species based on DNA sequences produced from different mitochondrial DNA barcode loci assayed using NGS. A 16S Ribosomal RNA gene (16S rRNA) barcode locus was particularly effective as it was amplified across all species and provided diagnostic sequence for all five species. With simple PCR-based DNA assays, 90% of samples provided "scorable" PCR results (10% failed to provide the desired PCR products), and of these 78% provided accurate sex identification. Using NGS, 72% of samples were "scored" (28% provided insufficient marker reads), and of these 84% provided accurate sex identification. Genetic assays (qPCR) for the presence of WNS pathogen DNA in eight guano samples were 100% accurate. NGS assays of parasite DNA in eight deceased Big Brown Bats *(Eptesicus fuscus)* samples produced a greater richness of parasite detections, and greater taxonomic precision than standard necroscopy methods.

For diet analysis, scat was collected from captive insect-feeding and nectar-feeding bats at the Fort Worth Zoo, TX that had been fed a controlled diet. Dietary items detected in the NGS assays closely matched bat feeding preferences observed by the zoo staff. To test the ability of NGS and MDM to provide population size estimates using DNA genotype-based capture-recapture models, the research team assessed genotyping success rates for 14 microsatellites with standard fragment analysis versus NGS for 94 Rafinesque's Big-eared Bat *(Corynorhinus rafinesquii)* guano samples obtained from a roost structure on Mammoth Cave National Park, KY. Though NGS methods required dropping two of the microsatellite loci from the test, assays of the remaining 12 loci demonstrated more than a 15% improvement on genotyping success and likely greater accuracy in microsatellite allele determination. In some cases, careful winnowing of low-quality sequences was required to eliminate spurious allele calls. An attempt to estimate the size of the Rafinesque's Big-eared Bat colony on Mammoth Cave National Park resulted in a population estimate that was approximately four times as large as the estimate using conventional techniques (video emergence monitoring). This apparent overestimation was likely the result of a combination of genotyping errors and violations of basic capture-recapture model assumptions, particularly the assumption that the population is closed to immigration and emigration.

To test the capacity of MDM to recover the same or equivalent data as single-locus PCR or NGS assays, 56 guano samples from five bat species were assayed for all six data classes using a single, multilocus NGS run. Data from the MDM run closely matched or exceeded that provided by the earlier NGS runs and other genetic assays. NGS methods also provided equivalent or better data than standard, non-genetic approaches and NGS performance did not appear to be diminished by incorporation into the multilocus MDM method. This phase of the study was reported in the scientific publication Swift et al. (2018).

During Phase 1, the research team noted the general need for additional assays for bat species identification and bat sex identification and for testing the efficacy of these and pre-existing assays across U.S. bats. The research team designed a new assay, COXl-Bat, for the COi barcode region, and a new set of sex identification assays, the XGXYC assays, that target Zinc-finger genes on the X and Y sex chromosomes of bats. Both sets of assays, along with three additional previously published species identification assays, and one previously sex identification assay, were directly tested against tissue and/or guano samples of bats of known sex for 32 species. Species that were not directly tested were reviewed *in silico* to the greatest extent possible using archived DNA sequence data.

This effort resulted in list of assays known or expected to be effective for species or species group identification of 43 species. Species groups occur among several *Myotis* species for which mitochondrial DNA lineages (all tested barcode loci are all found in the mitochondrial DNA) are not, to the extent they have been characterized, sufficiently diverged for reliable discrimination of these species as currently recognized. This effort also resulted in a list of 23 bat species for which current sex identification assays are expected to be effective. The outcomes of this effort are reported in a scientific publication (Guan et al., 2020) and short corrigendum (Guan et al., 2021).

For one component of Phase 2, in Spring 2016, the research team collected 376 guano pellets from beneath an artificial bat roost on Fort Drum, NY that was used by a Little Brown Bat maternity colony. The samples were subjected to MDM to obtain data on bat species identity, bat sex, the presence of the WNS pathogen, arthropod prey, and parasites. Bat species identification was successful for 98% (N = 368) of the samples, with 366 samples identified as coming from Little Brown Bats (as expected) and two samples identified to Big Brown Bat (known to commonly share roost structures with other bats). Sex could be determined from 85% (N = 318) of the samples, all of which were female (as expected). DNA from the WNS pathogen was detected in 62% of samples, which aligned well with past observed infection rates in the colony and annual seasonal patterns in infection levels. Arthropod prey data was recovered from 81% of samples and included both insects (Class Insecta) and spiders (Class Arachnida). The suite of prey items and relative encounter rate among samples aligned well with known diet patterns in Little Brown Bats. Parasite DNAs were detected in 35% of guano samples, with parasitic alveolates, ticks and mites, and trematodes — all from among known bat parasites— being among the mostly commonly encountered taxa. MDM performed very well in this trial, providing key data across multiple data classes for nearly all samples taken for a species with current population health, disease transmission, and future federal threatened, endangered & sensitive listing challenges. This phase of the project, including the second component described below, is reported in a scientific publication, *Multifaceted DNA metabarcoding of guano to uncover multiple classes of ecological data in two different bat communities* (Lance et al., 2022).

As a second component of Phase 2, in the late summer of 2016, the research team collected 274 guano samples from two locations on Fort Huachuca, AZ. Samples included both guano pellets produced by all bats and guano "splats" produced by nectar-feeding bats. One sampling location, where 102 samples were collected, was the outer chamber of a cave that served as a maternity colony for Lesser Long-nosed Bats and Cave Myotis *(M velifer),* as well as a night roost for Pallid Bats *(Antrozous pallidus).* The other sampling location, where 172 samples were collected, was a concrete bridge on Fort Huachuca that was known to be utilized as a night roost by several species, including among others, Lesser Long-nosed Bats, Cave Myotis, and Pallid Bats. Samples were subjected to MDM to obtain data on bat species identity, bat sex, arthropod prey, nectar plant use, and parasites (WNS does not yet occur in the Fort Huachuca region). Bat species identification was successful for 89% (N = 245) of the samples, with identification largely meeting expectations from past records of roost use. One exception was the lack of Cave Myotis samples from the outer chamber of the cave, indicating that the vast bulk of guano deposition for this species may occur deeper in the caves where this species roosts during the day. Another somewhat unexpected outcome was a lack of samples from some species known to utilize the bridge roost. In this case, it was likely that these species

were minor constituents of the bat community utilizing the bridge, that guano from these bats was comparatively rare among samples to be found under the bridge, and that additional sampling would have resulted in detections of these species. Sex could be determined from 60% (N = 163) of the samples, with sex ratios largely matching expectations based on past records. The success rate that was likely artificially low due to apparent human error or instrument malfunction in conducting these sets of assays. Arthropod prey data was recovered from 57% of guano pellet samples and included both insects, spiders, and centipedes (Class Chilopoda). Prey data aligned well with expected diet for various species from which the guano pellets originated. Nectar plant data was obtained at a comparatively lower rate (28% of splat samples), but largely agreed with known nectar plant use by Lesser Long-nosed Bats. The detection of nectar plant DNA in splat samples may have been relatively low due to a few factors, including high bat feeding rates at hummingbird feeders (little to no source DNA anticipated), low DNA concentrations in splat, and poorer cross-taxon barcode detection efficiencies for plant barcode loci in general. Plant barcode DNAs were additionally detected in 38% (N = 83) of guano pellet samples. These plant DNAs appear to have originated from occasional nectar feeding by Pallid Bats, a few pellets produced by Lesser-long nosed bats (instead of splats), and indirect plant DNA capture on the bodies and in the digestive tracts of insect prey. Parasite DNAs were detected in 51% of guano samples, with parasitic alveolates being encountered much more frequently than other parasites (including cestodes, ticks and mites, bat bugs, and kinetoplast protists). All parasites detected in these samples came from taxa known to parasitize bats.

In summary, the genetic data provided by MDM was demonstrated to be of equivalent or better to those produced through conventional approaches. The only exception was the estimation of bat colony population size. Given that other studies have effectively used guano-extracted DNA for similar capture-recapture population estimates, the poor performance of the approach in this study should be considered a function of either inadequate optimization of the technique for the particular case and/or having selected a population that violated basic model assumptions to an extent that subsequent estimates could not be accurate. In both demonstration cases, the MDM approach demonstrated a high-level of effectiveness across several data classes, particularly in comparison to the effort, cost (see below), and stress to bats that would be required to obtain the same data via conventional approaches.

For only three of the data classes included in this study can conventional methods be argued to be more accurate. Bat species identification should be extremely accurate in the hands of experienced bat biologists, though some species may still prove challenging to discriminate (e.g., *Myotis austroriparius/M septentrionalis, M californicus/M yumanensis).* Additionally, some species of *Myotis* bats cannot currently be discriminated genetically using existing mitochondrial DNA barcode loci (e.g., *M velifer/M yumanensis, M thysanodes/M evotis/M lucifugus/M occultus).* Sex identification might also be substantially more accurately provided by the conventional method (in-hand physical examination), but discriminating between sexes outside of the mating season can be a challenge. DNA-based identification of bat sex can also be challenging, in part because no one assay works for all species, and because the sex identification loci are much less abundant in each cell than the mitochondrial species barcode loci and often cannot be recovered from small, degraded samples like guano. Determination of WNS in bats, and characterization of infection intensity, is also more accurately determined by in-hand inspection of bat surfaces for the fungus. Presence of the fungal pathogen DNA in a guano sample may not correspond

to an actual infection (e.g., incidental uptake of the fungal spores) or provide evidence as to the extent of infection. *However, it must be noted that all of the above conventional or standard methods require capture and handling of bats. Capture and handling of bats may not always be feasible, especially for large numbers of samples, and carries risk of bat injuries and, for many areas, transmission of WNS pathogen spores. Additionally, capture and handling require official state and/or federal permits, whereas guano collection typically will not. Thus, MDM and related approaches will often be attractive alternative methods not only because some data classes are best determined by genetic assays, but also because the aforementioned issues with capture and handling are avoided.*

5.0 COST ASSESSMENT

Based on fairly simple, rough estimates by the research team, an MDM effort to obtain data on species, sex, WNS pathogen presence, diet, parasites, and to genotype individual samples (for population estimates, parentage or relatedness analyses, etc.) for a colony of a little less than 100 bats would require ⅓ – ½ the cost of conventional studies to obtain the same data (e.g. $35,000 – $40,000 for conventional studies versus $12,000 – $20,000 for MDM). These difference in cost may be greater with larger sample sizes and colonies, as per sample cost efficiencies typically improve with additional samples (especially if the cost of obtaining samples is low). Lab labor costs will often be the largest contributor to overall cost of an MDM effort, and efforts relying on well-described, previously established assays and methods should be considerably less costly than efforts that require to development of new assays and extensive method development and optimization. The comparative costs estimated by the research team assumed an MDM that capitalizes on well-established assays and methods.

6.0 IMPLEMENTATION ISSUES

MDM is best employed where guano or other animal scat can be readily found and collected. Colonial, cave- or edifice-roosting bats are an excellent example of taxa for which sites where guano is deposited and accessible are often known. However, MDM may not be a practical for non-colonial species that roost in tree canopies, rock crevices, etc. Guano samples from these bats would often be extremely difficult to locate or access and sample collection for any population-level effort would be very challenging and demanding. MDM could however be applied to captured bats, which often leave scat in holding bags or containers. MDM also benefits, and suffers, from the rapid advances in genomics (including environmental genomics). New assays and methods frequently arise, and trade-offs between the costs, time required to optimize, and the added power of new versus established MDM protocols should be carefully considered. Finally, much like environmental DNA, ancient DNA, and forensic DNA, MDM methods are highly susceptible to DNA contamination and require careful attention to sterile protocols on the part of those personnel procuring, processing, and assaying samples. End-users should carefully vet MDM service providers to verify that appropriate facilities (dedicated rooms for different stages of noninvasive sample processing) and methods are in place to minimize risk of contamination of or cross-contamination among sample DNA solutions and reactions.

7.0 REFERENCES

Guan, X., E. R. Britzke, A. J. Piaggio, D. L. Bergman, L. Van Pelt, and R. F. Lance. 2020.

Genetic assays for guano-based identification of species and sex in bats of the United States and Canada. Journal of Mammalogy 101: 970-978.

Guan, X., E. R. Britzke, A. J. Piaggio, D. L. Bergman, L. Van Pelt, and R. F. Lance. 2021.

Corrigendum to: Genetic assays for guano-based identification of species and sex in bats of the United States and Canada. Journal of Mammalogy 102:1444-1444.

Lance, R. F., Guan, X., Swift, J. F., Edwards, C. E., Lindsay, D. L., and E. R. Britzke. 2022.

Multifaceted DNA metabarcoding of guano to uncover multiple classes of ecological data in two different bat communities. Evolutionary Applications 15:1189-1200.

Swift, J. F., Lance, R. F., Guan, X., Britzke, E. R., Lindsay, D. L., and C. E. Edwards. 2018.

Multifaceted DNA metabarcoding: Validation of a noninvasive, next-generation approach to studying bat populations. Evolutionary Applications 11:1120-1138.