Duquesne University
Duquesne Scholarship Collection

Electronic Theses and Dissertations

2013

Evaluation of DNA Barcoding as a Technique for Elucidating the Diet of Louisiana Waterthrush Nestlings

Brian K. Trevelline

Follow this and additional works at: https://dsc.duq.edu/etd

Recommended Citation

EVALUATION OF DNA BARCODING AS A TECHNIQUE FOR ELUCIDATING

THE DIET OF LOUISIANA WATERTHRUSH NESTLINGS

A Thesis

Submitted to the Bayer School of Natural and Environmental Sciences

Duquesne University

In partial fulfillment of the requirements for

the degree of Master of Science in Environmental Science and Management

By

Brian K. Trevelline

December 2013

EVALUATION OF DNA BARCODING AS A TECHNIQUE FOR ELUCIDATING

THE DIET OF LOUISIANA WATERTHRUSH NESTLINGS

By

Brian K. Trevelline

Approved June 13, 2013

_____
Brady A. Porter, Ph.D.
Associate Professor of Biological
Sciences
(Committee Chair)

_____
Steven C. Latta, Ph.D.
Director, Conservation and Field
Research
National Aviary
(Committee Member)

_____
Kyle W. Selcer, Ph.D.
Associate Professor of Biological
Sciences
(Committee Member)

_____
Philip R. Reeder, Ph.D.
Dean, Bayer School of Natural and
Environmental Sciences

_____
John F. Stolz, Ph.D.
Director, Center for Environmental
Science and Education
Professor of Biological Sciences

ABSTRACT

EVALUATION OF DNA BARCODING AS A TECHNIQUE FOR ELUCIDATING

THE DIET OF LOUISIANA WATERTHRUSH NESTLINGS

By

Brian K. Trevelline

December 2013

Thesis supervised by Brady A. Porter, Ph.D.

The Louisiana Waterthrush (*Parkesia motacilla*) (LOWA) is a Neotropical-Nearctic migrant songbird that annually breeds in the Eastern United States. As an obligate riparian species, it preys upon a diverse community of benthic macroinvertebrates along headwater streams. Fecal material potentially contains residual DNA that can be used to molecularly identify prey species. The objective of this study was to develop a non-invasive technique capable of elucidating the diet of Waterthrush nestlings from residual DNA present in fecal sacs via DNA barcoding. Our limited analysis revealed that the majority of the analyzed fecal sacs were comprised of an acid-sensitive family of Ephemerotperan (Heptageniidae) in addition to Megalopterans and Dipterans. These results suggest that the technique of DNA barcoding can be utilized to accurately identify prey species from residual DNA found in avian fecal samples, which

may improve our understanding of landscape-level factors affecting riparian bird

communities and guide future conservation efforts.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMD – Abandoned Mine Drainage

BP – Nucleotide Base Pairs (DNA)

BLAST – Basic Local Alignment Search Tool (GenBank)

COI – Cytochrome *C* Oxidase Subunit I, a mitochondrial-coding gene

DNA – Deoxyribonucleic Acid

EPT – Ephemeroptera-Plecoptera-Trichoptera: three orders of high water quality

indicator insects

LOWA – Louisiana Waterthrush

mtDNA – Mitocondrial DNA

NGS – Next-Generation Sequencing

PCR – Polymerase Chain Reaction

PNR – Powdermill Nature Reserve

CHAPTER 1: INTRODUCTION

1.1 Louisiana Waterthrush

New World warblers (Family: Parulidae) are a diverse family of North American

passerine songbirds consisting of 22 genera and 89 species recognized by the American

Ornithologists' Union (Chesser *et al.* 2011). In general, warblers are small (6-9 grams),

primarily insectivorous birds that occupy diverse habitats ranging from temperate South

America to the Arctic (Lovette *et al.* 2010). Despite overlapping ranges and a similar

diet, multiple warbler species are capable of coexisting in a single habitat due to

extensive niche partitioning (MacArthur 1958). Since the 1980s, Neotropical migrants

breeding in Eastern North America have declined in abundance, presumably due to

anthropogenic impacts in both wintering and breeding grounds (Robbins *et al.* 1989). It is

believed that environmental issues such as climate change, acid deposition, deforestation

and exotic diseases are significant threats to warblers and may be responsible for their

long-term population declines (Faaborg *et al.* 2010a).

The Louisiana Waterthrush, *Parkesia motacilla* (Vieillot, 1809), is one of the two

warblers belonging to the genus *Parkesia*, the other being the Northern Waterthrush

(*Parkesia noveboracensis),* which share only a small portion of their distributions.

Despite the overlap of distributions and similar foraging behaviors, almost no

interspecific completion occurs between the two species (Craig 1984, 1987). Until

recently, both *P. motacilla* and *P. noveboracensis* shared the genus *Seiurus* with a third

species of wood-warbler, the Ovenbird (*Seiurus aurocapilla*) (Chesser *et al.* 2010).

However, the phylogenetic relationship between these species was revised in light of new

genetic data derived from nuclear and mitochondrial genomes. Molecular phylogenetic

analysis resulted in a major revision of warbler taxonomy, which included the genus *Seiurus* among 13 others that did not form monophyletic groups. Subsequently, the genus *Seiurus* is now considered monotypic and both the Louisiana Waterthrush and the Northern Waterthrush remain categorized together under the new genus *Parkesia* (Chesser *et al.* 2010; Lovette *et al.* 2010).

The Louisiana Waterthrush (hereafter LOWA) is a large (approximately 20 grams) wood-warbler characterized by its overall olive-brown appearance, cream-white underside, brown-streaked breast and a broad, white supercillium (Figure 1) (Mattsson *et al.* 2009). The plumage of the LOWA is not sexually dimorphic and is retained throughout the breeding and wintering seasons (Eaton 1958; Mattsson *et al.* 2009). In comparison to other species of warblers, the LOWA exhibits a more cryptic appearance, which is in stark opposition to its boisterous song (Mattsson *et al.* 2009). The LOWA's song is clearly audible over the perpetual sound of rushing water and resonates throughout its forest-interior habitat (Mattsson *et al.* 2009).

Figure 1: Adult Louisiana Waterthrush (*Parkesia motacilla*) (Mattsson *et al.* 2009). Photos by Jim Gilbert.

As a Neotropical-Nearctic migrant, the LOWA spends most of the year in the

tropics of Central America and the West Indies (Figure 2) (Mattsson *et al.* 2009). Similar

to other migrant songbirds, the LOWA flies northward throughout the night, resting at

various stopover habitats along the way and waits for favorable winds and weather

conditions (Faaborg *et al.* 2010b). By late March or early April, LOWA arrive in the

Northeastern United States as part of their reproductive strategy that takes advantage of

the ample territory and plentiful resources available for breeding (Faaborg *et al.* 2010b;

Mattsson *et al.* 2009). Despite the LOWA's large breeding range, the majority of the

population breeds along the Appalachian Mountain corridor (Bird Conservation Region

28), which spans from New York to Alabama (Mattsson *et al.* 2009). The LOWA

typically departs the breeding grounds in favor of its wintering habitat by late July or

early August (Mattsson *et al.* 2009).
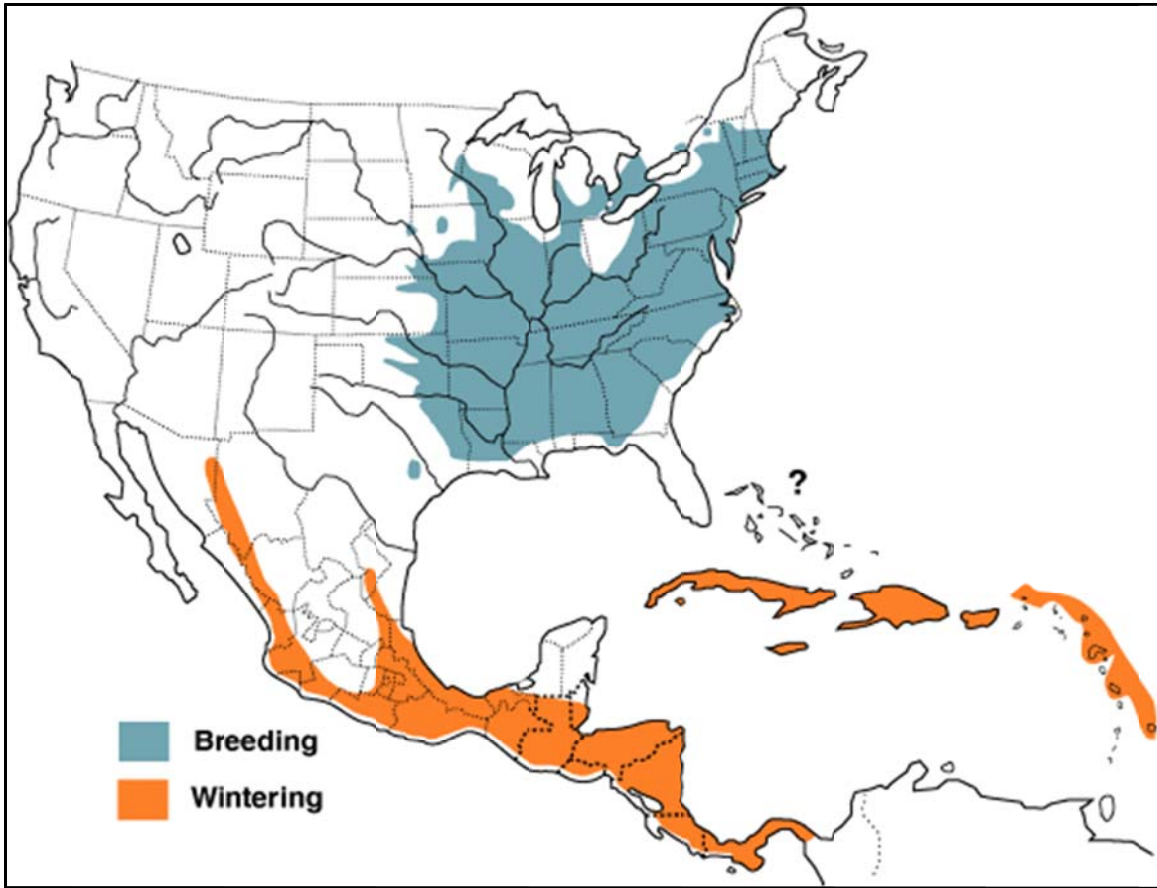
Figure 2: The breeding and wintering range of the Louisiana Waterthrush (*Parkesia motacilla*) throughout North America (Mattsson *et al.* 2009).

As an obligate riparian songbird, the LOWA breeds in closed-canopy deciduous and mixed-conifer forests surrounding a first, second or third-order gravel bottom stream (Mattsson *et al.* 2009; Mulvihill *et al.* 2008). The LOWA often utilizes exposed roots and undercut banks along streams as a safe and convenient nesting location (Figure 3) (Eaton 1958; Mattsson *et al.* 2009). Furthermore, this nest serves as an ideal location for the LOWA's constant foraging for aquatic insects, which are necessary to sustain rapidly developing nestlings (Eaton 1958). After 13 days of egg incubation, nestlings reside in the nest for a period of 10 days (Mulvihill *et al.* 2009). During this time, both adults share responsibilities for feeding the nestlings as well as maintaining the integrity of the nest (Mattsson *et al.* 2009). The feces produced by the nestlings, for example, are removed to

avoid accumulation and an increased likelihood of predation (Weatherhead 1984). In order to accomplish this, nestlings produce a protein-based mucous membrane that encompasses the feces, which is known as a fecal sac (Weatherhead 1984). The fecal sacs are produced until just before fledging, which enables the adults to easily carry the waste away from the nest in their beaks (Weatherhead 1984) (Figure 4). After fledging the nest, the brood is divided between the two adults for a 3-4 week period of extended parental care (Mulvihill *et al.* 2009).

The LOWA's dependence on a diverse community of benthic macroinvertebrates earns it the title of the only stream-dependent songbird in eastern North America (Mulvihill *et al.* 2008). The LOWA primarily forages for insects along stream banks, moss-covered rocks, and by picking though leaf litter and leaf packs found in the riffles of a headwater stream, but has also been observed capturing prey in flight (Craig 1984; Mattsson & Cooper 2006). The LOWA is known to prey upon many aquatic organisms, which are primarily comprised of immature EPT taxa (the insect orders of Ephemeroptera. Plecoptera and Tricoptera) but has also been observed consuming adult insects and other aquatic organisms such as larval salamanders and cicadas (Figure 5) (Craig 1984, 1987; Eaton 1958; Mattsson & Cooper 2006).

Figure 3: A typical nesting location of the Louisiana Waterthrush (*Parkesia motacilla*) (Mattsson *et al.* 2009). Photo by Bob Wood.

Figure 4: An adult Louisiana Waterthrush (*Parkesia Motacilla*) removing a fecal sac from the nest (Mattsson *et al.* 2009). Photo by James P. Mattsson.

Figure 5: An adult Louisiana Waterthrush (*Parkesia Motacilla*) feeding post-aquatic stage insects to nestlings (< 6 days old) (Mattsson *et al.* 2009). Photo by Bob Wood.

1.2 Benthic Macroinvertebrates

Insects are one of the most abundant and diverse forms of macroinvertebrate life found in aquatic ecosystems (Cushing & Allan 2001). In North America, there are 13 orders of insects with aquatic life stages that can be found in freshwater aquatic habitats (Cushing & Allan 2001). The aquatic life stage for a macroinvertebrate is known as a nymph or larva, depending on which type of life cycle the insect exhibits (Cushing & Allan 2001). Macroinvertebrate life cycles can be divided into two categories: hemimetabolous (nymph) and holometabolous (larva). Hemimetabolous macroinvertebrates have three distinct life stages: egg, nymph and adult. The nymph stage closely resembles the terrestrial adult form of the insect, which includes the macroinvertebrate orders of Plecoptera, Ephemeroptera and Odonata (stoneflies, mayflies and dragonflies, respectively). Holometabolous macroinvertebrates pass through an additional life cycle called a pupa, which dictates that the immature stage must be known as a larva. The orders of Trichoptera and Diptera (caddisflies and true flies, respectively) are examples of aquatic insects that undergo this life cycle. These immature forms of the terrestrial adults sometimes live in the stream for many years before fully maturing (Cushing & Allan 2001).

Each species of aquatic insect has its own unique morphology, which may or may not resemble its adult form depending on life cycle. Generally, immature macroinvertebrates can be described as having three general body regions: the head, thorax and abdomen (Cushing & Allan 2001). The head region has sensory organs such as eyes, antennae and mouthparts (Cushing & Allan 2001). The thorax is divided into three segments that each bears a set of legs, two of which will bear a set of wings,

depending on species (Cushing & Allan 2001). The abdomen consists of eleven segments, which bear the insect's gills (Cushing & Allan 2001).

Macroinvertebrate species can be divided into functional feeding groups, which is highly dependent on stream order (Vannote *et al.* 1980). Functional feeding groups such as "shredders", "collectors", "grazers" and "predators" are used to describe the role of a particular species of aquatic insect (Cushing & Allan 2001). Headwater streams with cold water and high quality riparian habitats typically foster more shredders and collectors due to high levels of detritus from the surrounding trees (Cushing & Allan 2001; Vannote *et al.* 1980). This leaf litter can be colonized by macroinvertebrates that specialize in masticating decaying plant matter, which is collected from the water column by other specialized macroinvertebrates (Cushing & Allan 2001).

Macroinvertebrate diversity within microhabitats is also highly dependent on water quality parameters, including low pH due to acidification. Some macroinvertebrate taxa are more sensitive to acidification than others, which allows for some taxa to survive and limits the availability of EPT taxa in particular (Courtney & Clements 1998; Guerold *et al.* 2000; Mulvihill *et al.* 2008). This may create problems for many higher trophic organisms that depend on a diverse macroinvertebrate community in order to survive (Graveland 1998; Schreiber & Newman 1988).

Macroinvertebrate diversity and abundance is of vital importance for a stream-dependent organism such as the Louisiana Waterthrush. Acid-tolerant macroinvertebrate taxa may be abundant in acidified drainages but are not a preferred food source for the LOWA, thus, depriving it of the calcium-rich diet necessary for egg development (Mattsson & Cooper 2006; Mulvihill *et al.* 2008; Ormerod & Rundle 1998). In the Laurel

Highlands of Southwestern Pennsylvania (our study site), macroinvertebrate communities were drastically altered due to acidification caused primarily by acid precipitation and abandoned mine drainage (AMD) (Mulvihill *et al.* 2008). Specifically, the study illustrated the importance of acid-sensitive mayflies in predicting overall LOWA breeding success (Mulvihill *et al.* 2008). The study revealed that breeding density of LOWA is highly correlated with the presence of acid-sensitive EPT taxa, which supports the conclusions of a similar study conducted in Georgia, U.S.A. (Mattsson & Cooper 2006). Furthermore, the Pennsylvania study revealed that LOWA breeding territories were nearly twice as large on acidified streams when compared to those found on circumneutral streams, presumably to compensate for insufficient macroinvertebrate prey (Mulvihill *et al.* 2008). LOWA breeding on acidified streams were also found to be less experienced and produced smaller average clutch sizes (Mulvihill *et al.* 2008). The LOWA's reliance on acid-sensitive macroinvertebrate taxa is a unique attribute among songbirds, which makes it a valuable biological indicator for water quality (Mattsson & Cooper 2006; Mulvihill *et al.* 2008).

1.3 Diet Studies

Understanding trophic relationships between insectivorous birds and their prey is a critical component of avian conservation efforts (Oehm *et al.* 2011). The LOWA's dependence on EPT taxa of macroinvertebrates has been well described but little is known regarding specific prey items that comprise its diet. However, diet studies on insectivorous passerines are notoriously difficult to perform due to intrinsic limitations associated with sampling techniques (Oehm *et al.* 2011). Popular approaches to studying the diet of warblers have included mostly antiquated techniques that primarily rely on

morphological identification of prey species. For example, stomach flushing (Moody 1970) with a saline solution or induced vomiting by the use of emetics such as copper sulfate or Ipecac (Radke & Frydendall 1974; Robinson & Holmes 1982), which enables investigators to analyze the regurgitate. In some studies (and in the case of a LOWA diet study), researchers have also resorted to sacrificing animals in order to resect stomach contents for diet analysis (Table 1) (Eaton 1958).

Table 1: Summary of common diet study techniques for warblers in order of decreasing invasiveness.

Warbler Diet Study Technique	Reference
Gut contents analysis	Eaton 1958
Stomach flushing with saline	Jenni *et al.* 1990
Use of emetics	Robinson & Holmes 1982
Use of neck ligatures	Moreby & Stoate 2000
Morphological analysis of feces	Deloria-Sheffield *et al.* 2001

Each of the above approaches offers an effective means for establishing a rudimentary understanding of dietary composition but may expose birds to unnecessary stress via invasive sampling protocols (Oehm *et al.* 2011). In the most extreme case, stomach content analysis requires the sacrifice of the birds, which presents ethical and, perhaps, legal barriers to diet studies concerning Neotropical migrant songbirds. While stomach flushing or emetics are certainly an improvement ethically, these techniques still put the animal's welfare at risk, therefore, non-invasive techniques are the preferred method of obtaining information on diet composition of warblers (Oehm *et al.* 2011). Until recently, non-invasive methods were restricted to either the morphological analysis of fecal material or direct observation of foraging behavior. DNA-based molecular techniques, however, may provide researchers with more quantitative approach to feces

analysis that also eliminates the need for time-consuming taxonomic identification of partial insect remains.

In the case of Louisiana Waterthrush, only two diet studies have been conducted. A total of 15 adults collected in Ithaca, New York, USA were analyzed via gut content analysis revealing a diet primarily composed of aquatic invertebrates, which included the orders of Ephemeroptera, Plecoptera, and Trichoptera (Eaton 1958). Another study conducted in Connecticut, USA, based entirely on observational data on three adults, described a LOWA diet primarily composed of Trichopterans (40.7%), Dipterans (20.3%) and Ephemeropterans (13.6%) (Craig 1987).

1.4 DNA Barcoding

Molecular techniques such as DNA barcoding offer a non-invasive and effective assay for performing diet studies. DNA barcoding enables researchers to identify prey items from residual DNA present in fecal material, which eliminates the need for morphological identification from insect remains. Specific regions of this DNA can be isolated and amplified using custom designed fragments of DNA known as primers, which are developed based on the target organisms. Polymerase chain reaction (PCR) is then utilized to amplify the fragment, which can then be subjected to various downstream applications that eventually provide a DNA sequence that can be identified by comparison to reference sequences.

The use of "universal" PCR primers for DNA barcoding enables researchers to compare DNA sequences between individuals by amplifying the same region of DNA across a wide variety of organisms (Folmer *et al.* 1994). The need to compare genetic information between many phyla led to development of PCR primers designed to amplify

13

a region of the Cytochrome *C* Oxidase Subunit I gene (COI), which is exhibited by nearly all organisms and is one of the most conserved protein-coding regions among animals (Folmer *et al.* 1994). The first set of PCR primers designed to amplify this region were developed for the identification of invertebrates, which was motivated by the need to conduct phylogenetic studies on the diverse communities of organisms living in deep-sea hydrothermal vents (Folmer *et al.* 1994). The COI gene is located within the mitochondrial genome and codes for proteins essential for ATP synthesis and, thus, was retained throughout evolution across most animal phyla (Folmer *et al.* 1994; Hebert *et al.* 2003). Despite being highly conserved overall, a 648 bp portion of the COI gene varies significantly between taxa, which has resulted in its frequent use for DNA barcoding applications (Table 2).

While other genes capable of taxonomic identification exist, the COI gene offers a few distinct advantages that have allowed it to become the standard DNA barcoding region. First, mitochondria are more plentiful than nuclei in cells, which provides more available template for PCR than nuclear DNA (Birky *et al.* 1983). Second, the universal primers available for this region are widely used based on their ability to amplify mtDNA from most animal phyla (Folmer *et al.* 1994). Lastly, the COI gene has a molecular evolution rate approximately three times greater than other common barcoding regions such as the 12S and 16S ribosomal DNA coding regions (Hebert *et al.* 2003).

While DNA barcoding is hardly a new technique, the inherent nature of fecal material has restricted the applicability of universal COI barcoding primers in past avian diet studies. Residual prey DNA in avian feces is highly degraded (< 300bp) due to exposure to high concentrations of uric acid (a primary component of avian feces), UV

exposure and enzymatic activity (Deagle *et al.* 2006; Oehm *et al.* 2011). Limitations

regarding the size of intact mtDNA fragments can be alleviated by amplifying a much

smaller portion of the COI barcoding region, thus, increasing the likelihood of

successfully producing the desired amplicon (Hajibabaei *et al.* 2006b; Zeale *et al.* 2011).

For our diet study, we used primers designed to amplify a 157 bp fragment of the COI

barcoding region from insect DNA (Figure 6) (Zeale *et al.* 2011). Although this region is

significantly smaller than the traditional 648 bp barcoding region (Folmer *et al.* 1994),

enough variability still exists that a high degree of taxonomic resolution can be achieved

(Figure 7) (Zeale *et al.* 2011).

Table 2: Summary of studies that used COI barcoding to molecularly identify
macroinvertebrate organisms

Target Macroinvertebrate Order	Reference
Coleoptera	Harper *et al.* 2005
	Juen & Traugott 2005
Diptera	Pfenninger *et al.* 2007
Ephemeroptera	Ball & Hebert 2005
	Hajibabaei *et al.* 2011
Hemiptera	Agustí *et al.* 2003
Lepidoptera	Hajibabaei *et al.* 2006a
	Sheppard *et al.* 2005
Trichoptera	Hajibabaei *et al.* 2011
General Arthropods	Carew *et al.* 2013
	Clare *et al.* 2009
	Folmer *et al.* 1994
	Hajibabaei *et al.* 2012
	Harper *et al.* 2007
	Zeale *et al.* 2011
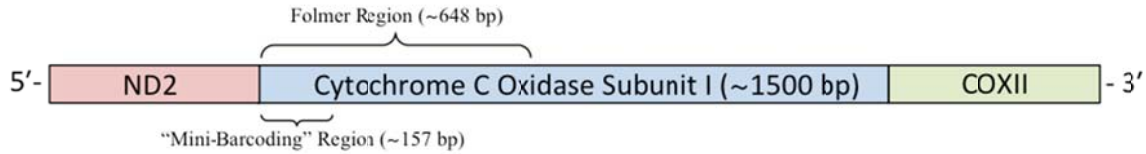
Figure 6: The COI barcoding regions and neighboring mitochondrial genes. Data obtained from GenBank.
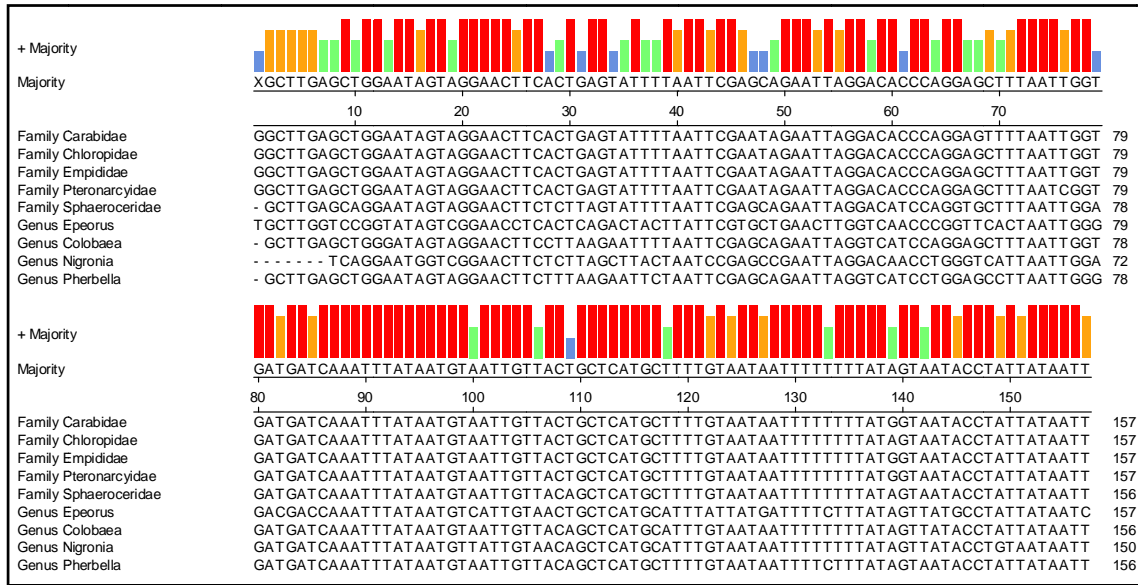
Figure 7: Alignment of COI barcode in five common families and four common genera of macroinvertebrates. Aligned sequences exhibit 49 variable sites (33%). Red bars indicate no variability at nucleotide residue; blue bars indicate the most variability nucleotide residue. Sequences obtained from GenBank and aligned via the Clustal W method in the DNASTAR MegAlign software package (DNASTAR, Madison, WI, USA).

1.5 Hypothesis and Specific Aims

In this study, we sought to develop an effective and non-invasive assay for determining dietary composition of Louisiana Waterthrush nestlings, which may improve our understanding of landscape-level factors affecting riparian bird communities and guide future conservation efforts. We hypothesize that residual prey DNA in the fecal sacs of LOWA nestlings can be amplified, sequenced and identified via DNA barcoding, which can be used to elucidate diet. Such a technique may provide ornithologists with a tool for studying the diet of many species of insectivorous songbirds in the wild, which

has never before been published in peer-reviewed literature. For this study, we aimed to: (1) develop a standardized protocol for the processing of LOWA fecal sacs for DNA barcoding analysis, (2) assess the feasibility of DNA barcoding as an approach to elucidating songbird diet and (3) to obtain preliminary data on LOWA diet composition as justification for future studies.

CHAPTER 2: MATERIALS AND METHODS

2.1 Study Site and Sample Collection

Our study was conducted within the LOWA's breeding range at the Carnegie Museum of Natural History's Powdermill Nature Reserve and the surrounding region (hereafter PNR) located in the Laurel Highlands of the Appalachian Mountains in Rector, Pennsylvania, USA (Figure 8). Historically, this region has been associated with several significant environmental issues such as aggressive timber harvesting, acid precipitation and abandoned mine drainage (AMD), which have negatively impacted the water quality of several headwater streams (Mulvihill *et al.* 2008). However, several streams in this region are in pristine condition, like Powdermill Run, which is considered one of the highest quality streams in Pennsylvania. Despite these issues, LOWA can be found nesting on both impacted and high-quality stream habitats throughout the region.

Fecal sacs were collected from LOWA nests found on both impacted (Linn Run and Loyalhanna Creek) and non-impacted streams (Camp Run and Powdermill Run) near PNR (Figure 8) (Mulvihill *et al.* 2008). Fecal sacs were opportunistically collected from both the nests (and the area surrounding them) as well as directly from the nestlings during banding (approximately 8 days post-hatching). The fecal sacs were pooled for each nest and stored in absolute (100%) ethanol at ambient temperature for a period of approximately 6 months prior to analysis.
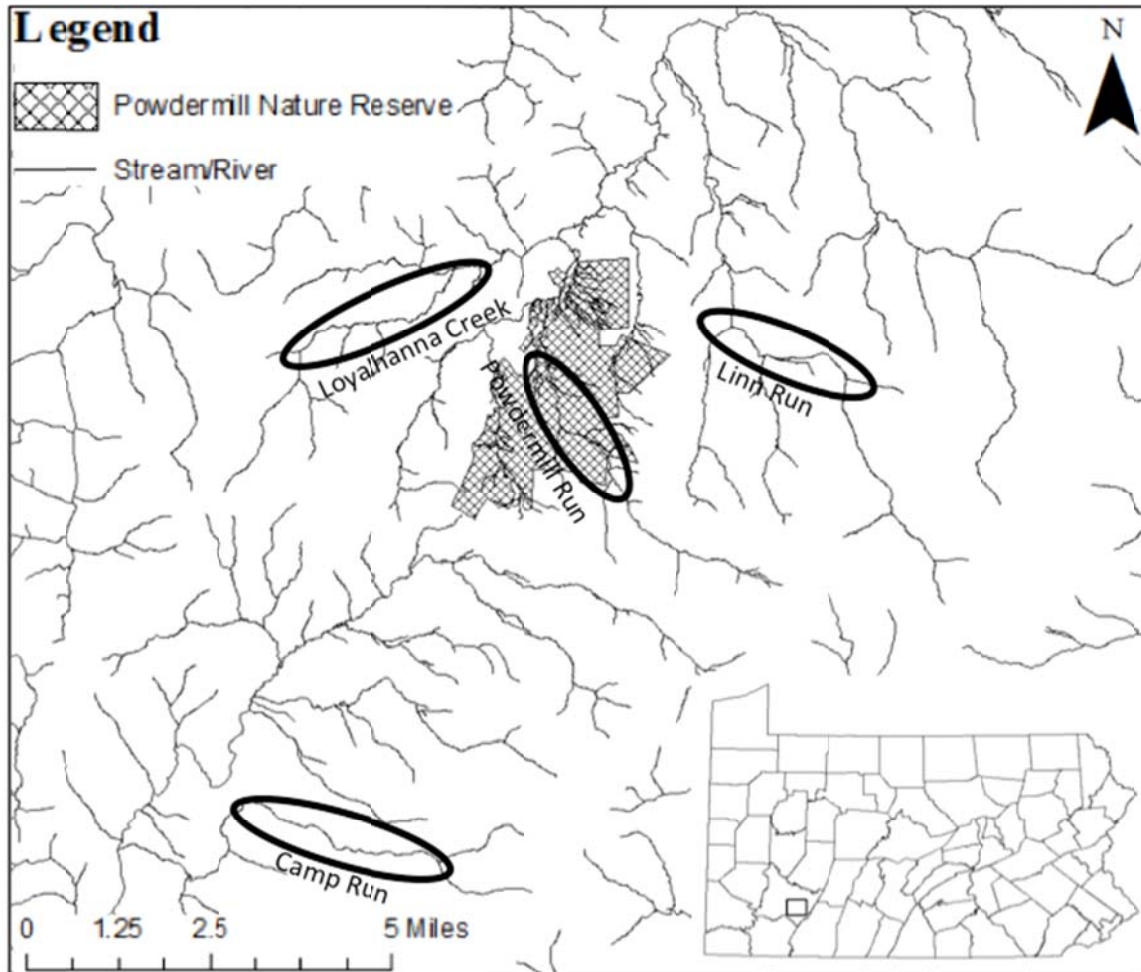
Figure 8: LOWA fecal sac collection sites on four streams found near Powdermill Nature Reserve in the Laurel Highlands of Pennsylvania (Carnegie Museum of Natural History, Pittsburgh, PA, USA; United States Geological Survey (USGS); Environmental Systems Research Institute (ESRI) ARCGIS, Redlands, CA, USA).

2.2 DNA Extraction and PCR amplification

Positive control DNA was extracted from four orders of benthic macroinvertebrates collected from Powdermill Run via phenol:chloroform:isoamyl alcohol precipitation (Maniatis *et al.* 1982) (Appendix A): Diptera (true flies), Epheroptera (maylfies), Odonata (dragonflies) and Plecoptera (stoneflies). Although fecal sacs from the same nest were pooled, DNA was extracted from each fecal sac individually. Each fecal sac was separated from its uric acid coating using sterilized instruments under magnification. The uric acid-free fecal material was dried on a slide

19

warmer for approximately 30 minutes to evaporate residual ethanol, which is believed to interfere with the DNA extraction. Approximately 20 mg (dry weight) of fecal material was subjected to the QiaAMP DNA Stool Mini Kit (Qiagen) following a modified protocol adapted from Zeale *et al.*, 2011 (Appendix B). The modifications consisted of (1) removal of uric acid coating of fecal sacs and evaporating residual ethanol for 30 minutes, (2) the addition of 40 μL of Proteinase K to buffer AL and ASL, (3) overnight incubations at 70 °C, (4) and pre-warming the buffer AE to 50 °C prior to eluting the DNA from the spin-column.

All PCR reactions using fecal sac DNA were carried out using 16 μL of the DNA template (the maximum volume of DNA template in our reactions) in 50 μL reactions using the insect-specific COI barcoding primers developed by Zeale *et al.*, 2011 and the following reagents: 5 μL 10X reaction buffer (Fisher Scientific), 10 μL 25 mM $MgCl_2$ (Fisher Scientific), 10 μL 5 mM DNTPs (Bio Express), 1.5 μL 10 μM ZBJ-ArtF1c forward primer (5'-AGATATTGGAACWTTATATTTTATTTTTGG-3') (Zeale *et al.* 2011), 1.5 μL 10 μM ZBJ-ArtR2c reverse primer (5'-WACTAATCAATTWCCAAAT CCTCC-3') (Zeale *et al.* 2011), 4.45 μL 10 mg/mL acetylated BSA (Promega), and 1.5 μL 5 U/μL DNA *Taq* Polymerase (Fisher Scientific). PCR reactions using positive control DNA were adjusted for 200 ng of total DNA template using the same buffer conditions. All reactions were prepared on ice and amplified using the following thermal cycling conditions: 3 minutes at 94 °C followed by 60 cycles of 30 s at 94 °C, 30 s at 45 °C, 30 s at 72 °C, and a final extension of 10 minutes at 72 °C. The PCR products were visualized by loading 10 μL of DNA product and 4 μL of loading dye on a 2% agarose-ethidium bromide gel. Once amplification was confirmed, the remaining volume of the

PCR reaction was loaded on a 2% TAE gel to excise the specific COI product in order to eliminate potential non-specific amplification products. The band was excised using the PureLink Quick Gel Extraction Kit (Invitrogen) via the manufacturer's instructions. The isolated DNA was stored short-term at 4 °C and at -20 °C for long-term storage.

2.3 Cloning Library Construction and Sequencing

Amplified PCR products were subjected to a tailing reaction in order to ensure that the target amplicon had not lost its poly-A tail that is necessary for TA cloning. The reaction was conducted for 30 minutes at 72 °C in a 10 μL volume using the following reagents: 1.0 μL 10X reaction buffer, 0.6 μL 25 mM $MgCl_2$, 1.0 μL 2 mM dATPs, and 1.0 μL 5 U/μL DNA *Taq* Polymerase. The tailed PCR products were cloned using the pGEM-T Easy Vector System and high-efficiency competent cells ($\geq$ 1 x 10^8 cfu μg^{-1} DNA) (Promega) using a 3:1 insert to vector ratio. Ligation reactions were carried out via the manufacturer's specifications, incubated overnight at 4 °C. Competent *E. coli* competent cells were transformed in super optimum culture (S.O.C.) (Appendix C) using 2 μL of ligation product via the manufacturer's instructions. Transformations were spread on LB agar plates and selected by X-Gal-mediated blue/white screening using 100 μg/mL ampicillin, 80 μg/mL X-Gal, and 0.5 mM IPTG. After an overnight incubation at 37 °C, white colonies were selected and placed into 1 mL LB broth containing 100 μg/mL ampicillin and incubated at 37 °C for 1 hour. An aliquot of 300 μL of LB broth containing the selected colonies was centrifuged for 3 minutes at 6000 rpm to form a bacterial pellet, which would serve as the template for colony PCR. The following reagents were added to the bacterial pellet in order to perform colony PCR: 1.25 μL 10X reaction buffer, 2.5 μL 25 mM $MgCl_2$, 2.0 μL 5 mM DNTPs, 0.5 μL 10 μM M13 forward

primer (5'-GTAAAACGACGGCCAG-3'), 0.5 µL 10 µM M13 reverse primer (5'-

CAGGAAACAGCTATGAC-3'), 0.8 µL 10 mg/mL BSA (Sigma-Aldrich), and 0.5 µL 5

U/µL DNA *Taq* Polymerase. Thermal cycling protocol for colony PCR was identical to

the protocol used for COI amplification described above.

Positive transformants were incubated overnight in 5 mL LB broth with 100

µg/mL ampicillin at 37 °C in order to generate enough plasmid DNA for sequencing.

Plasmid DNA was extracted from the bacterial colonies using the PureLink Quick

Plasmid Miniprep Kit (Invitrogen). Using 500 ng extracted plasmid DNA and the M13

forward primer, 20 µL sequencing reactions were prepared using Big Dye 3.1 Terminator

Sequencing Kit (Applied Biosystems) with the following thermal cycler protocol: 1

minute at 95 °C followed by 75 cycles of 30 s at 94 °C, 5 s at 50 °C, 4 minutes at 60 °C.

Primers were removed from the sequencing products by Sephadex G-50 (Sigma-Aldrich)

column cleaning (Appendix C). The Sephadex-cleaned sequencing reaction was run on

an ABI 310 Genetic Analyzer equipped with a 60 cm array and POP6 polymer (Applied

Biosystems). Sequences were trimmed of vector and primer sequences and entered into

GenBank via standard nucleotide BLAST (Basic Local Alignment Search Tool)

optimized for highly similar sequences.

CHAPTER 3: RESULTS

3.1 DNA Extraction

DNA was extracted from fecal samples obtained from 12 locations at PNR (Table

3). The fecal sac extraction and purification protocol outlined above typically yielded

between 400 and 5000 ng of DNA. Control DNA extracted from alcohol preserved

insects using the phenol:chloroform:isoamyl alcohol precipitation protocol (Appendix A)

had concentrations ranging between 25 and 50 μg.

Table 3: Summary of PCR amplification, cloning and sequencing results from DNA isolated from LOWA fecal sacs. * - indicates the recovery of concatenated primer sequences

Collection Site	Stream Impacted	COI Amplicon Detected	Colonies Screened	COI Positive Colonies	Colonies Sequenced	Colonies Identified in BLAST
Camp	NO	YES	25	4	1*	0
Camp	NO	YES	25	8	1*	0
Camp	NO	NO	N/A	N/A	N/A	N/A
Camp	NO	NO	N/A	N/A	N/A	N/A
Powdermill	NO	NO	N/A	N/A	N/A	N/A
Powdermill	NO	YES	195	69	3	0
Powdermill	NO	YES	61	23	13	0
Powdermill	NO	YES	24	19	17	10
Powdermill	NO	NO	N/A	N/A	N/A	N/A
Powdermill	NO	YES	25	1	1*	0
Linn	YES	NO	N/A	N/A	N/A	N/A
Loyalhanna	YES	NO	N/A	N/A	N/A	N/A

3.2 Evaluation of COI Barcoding Primers

The insect-specific, mini-barcoding COI primers designed by Zeale *et al.*, 2011 were evaluated using positive control insect DNA from four common macroinvertebrate orders: Diptera (true flies), Ephemeroptera (mayflies), Odonata (dragonflies and damselflies) and, Plecoptera (stoneflies). These orders were selected based on their published amplification by Zeale *et al.,* 2011 using COI "mini-barcoding" primers and for their abundance within the LOWA's nesting territories at PNR (Mulvihill *et al.* 2008). Using the PCR protocol outlined above and 200 ng of insect DNA as a template, the 211 bp COI barcoding amplicon (157 bp target plus 54 bp of primers) was successfully amplified (Figure 9). In order to ensure that the proposed primers would not amplify residual LOWA DNA found in fecal material, the same experiment was conducted using 200 ng of DNA isolated from the tissue four LOWA individuals. The COI amplicon was detected in our positive control DNA (order: Odonata) and was absent in the LOWA DNA from all four individuals (Figure 10).

3.3 COI Barcode Amplification from Fecal Sac DNA

Among the DNA extracted from the 12 samples in Table 3, we successfully amplified the COI barcode in 6 fecal sacs (50%) (Figure 11). Due to the degraded nature of fecal sac derived DNA and the potential presence of PCR inhibitors, reactions required a large amount of template in order to achieve successful amplification. Multiple optimization reactions were conducted in order to elucidate the required PCR conditions for COI amplification. Depending on which DNA extract was used, a minimum template concentration of 700 ng in a 50 μL reaction was required in order to achieve sufficient amplification.
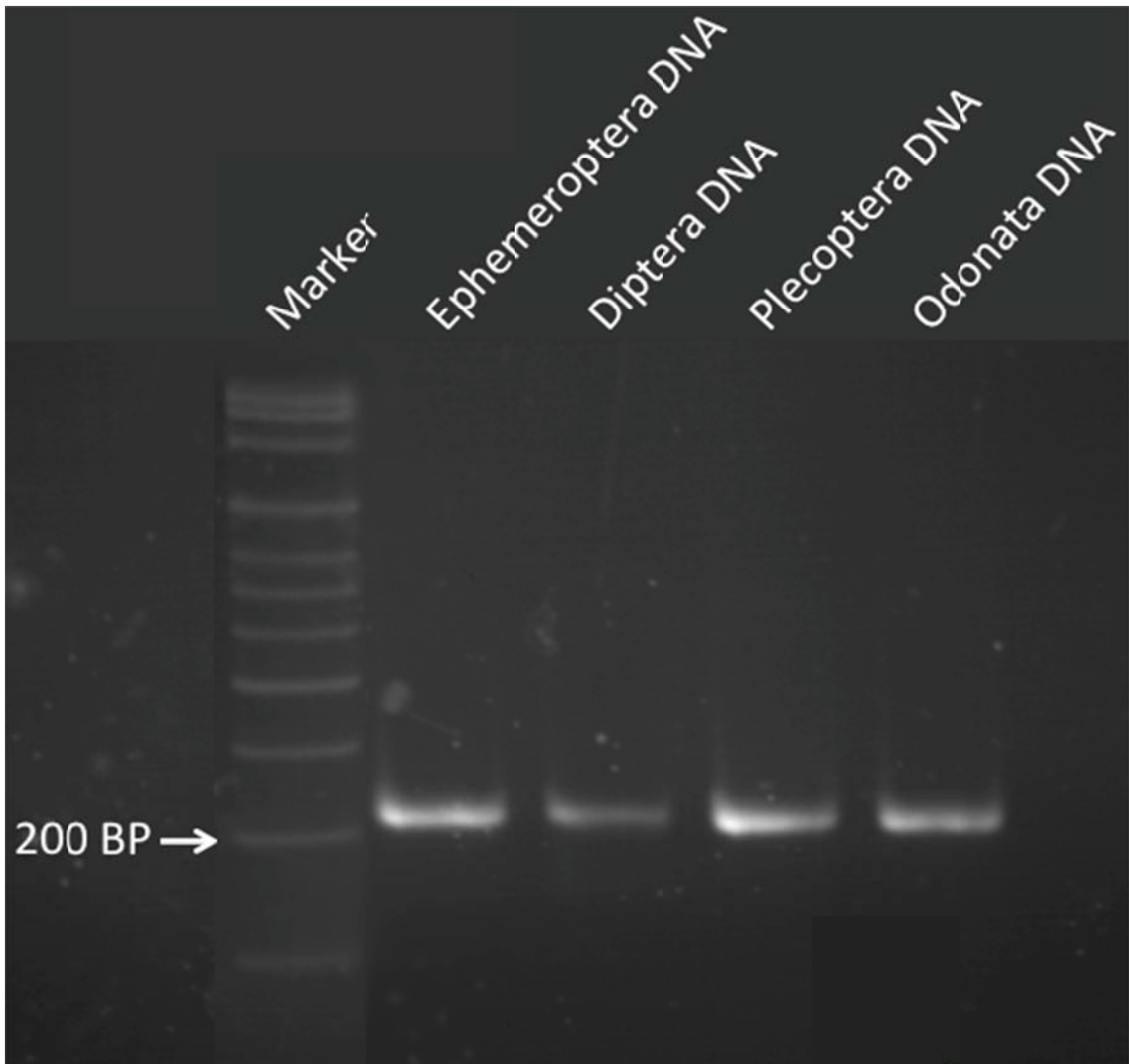
Figure 9: Amplification of the COI barcoding region (Lane 2; ~211 BP) in DNA from four macroinvertebrate orders: Ephemeroptera, Diptera, Plecoptera and Odonata.
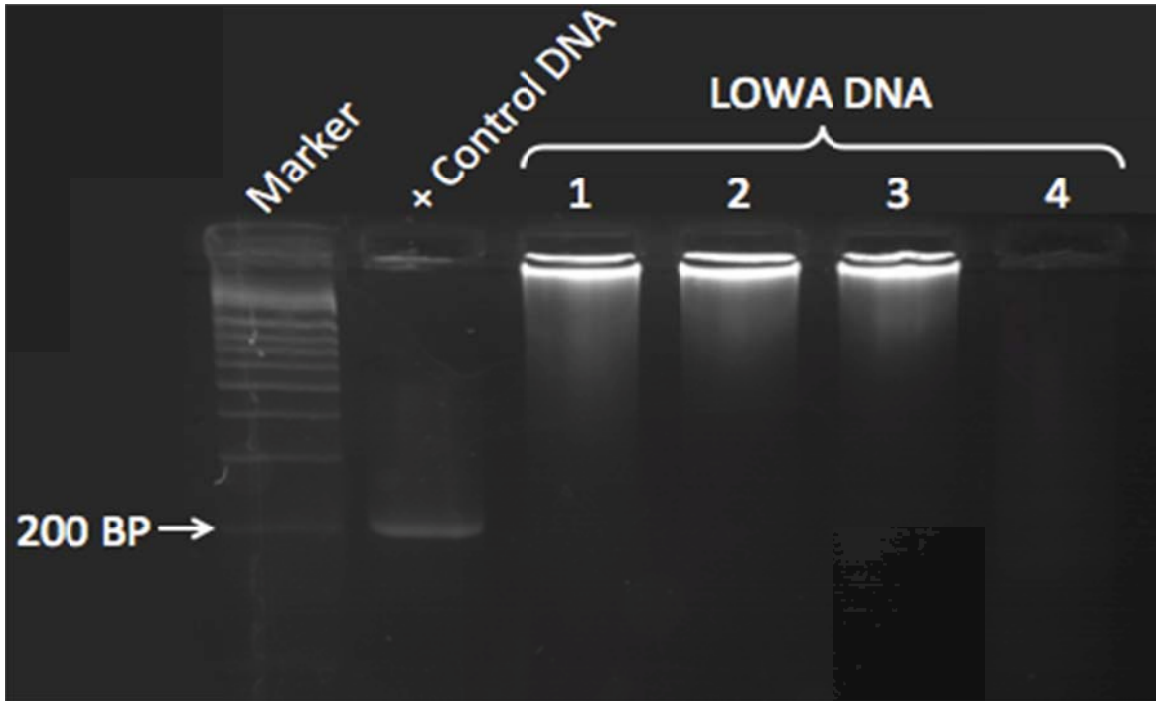
Figure 10: Detection of COI barcoding amplicon in positive control DNA (lane 2; ~211 BP) and absence in LOWA DNA from four individuals.
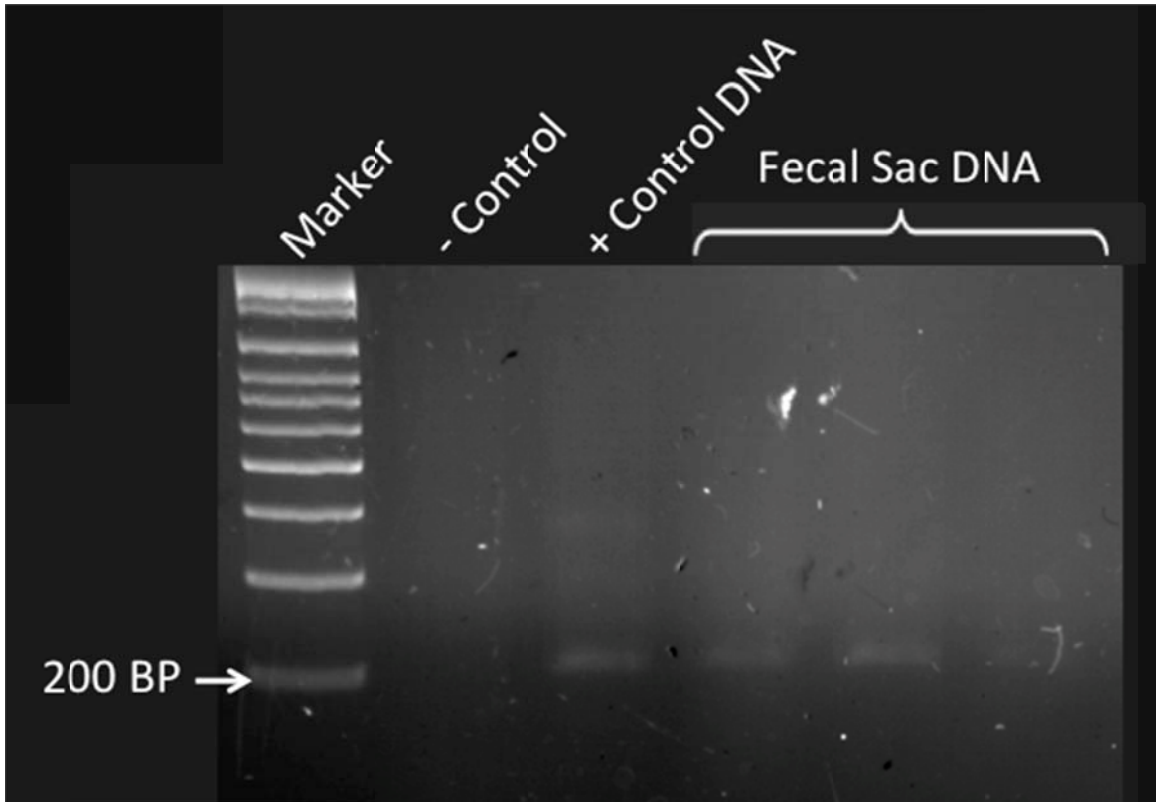
Figure 11: Detection of COI barcoding amplicon in insect DNA (lane 3; ~211BP) and fecal sacs obtained from Louisiana Waterthrush (*Parkesia motacilla*) nestlings.

3.4 Molecular Cloning and Sequencing

The ligation and transformation protocol outlined above generated between ten and fifteen selectable positive transformants per plate (approximately 5 percent of all colonies) using 200 µL of transformed bacteria in S.O.C. medium. Both white and light blue colonies (expressing β-Galactosidase) were selected and screened for the COI barcoding amplicon using the colony PCR protocol described above. M13 priming sites incorporated into the vector allowed for the confirmation of ligation by producing variable fragment lengths using PCR. Plasmids exhibiting successful and unsuccessful ligation of the COI barcode produced a 473 bp (M13 forward and reverse primers, vector backbone and COI insert) and a 262 bp (M13 forward and reverse primers and vector backbone only) amplicon, respectively (Figure 12).

A total of sixteen clones were sequenced from one fecal sac obtained from a nest found on a circumneutral stream (Powdermill Run). Among the sixteen clones sequenced, ten were positively identified to genus or species, which included one chimeric DNA insert (Table 4) (Appendix D). DNA sequencing resulted in the identification of four unique genera across three orders of insects - Diptera: *Chloropidae*, Ephemeroptera: *Epeorus* and *Heptagenia*, and Megaloptera: *Nigronia*. Amplified COI barcoding fragments were identified predominately as Ephemeroptera (54.5%) with a smaller portion belonging to the orders of Megaloptera (36.4%) and Diptera (9.1%), which was not normalized for biomass and, thus, does not necessarily reflect actual number of prey items fed to LOWA nestlings.
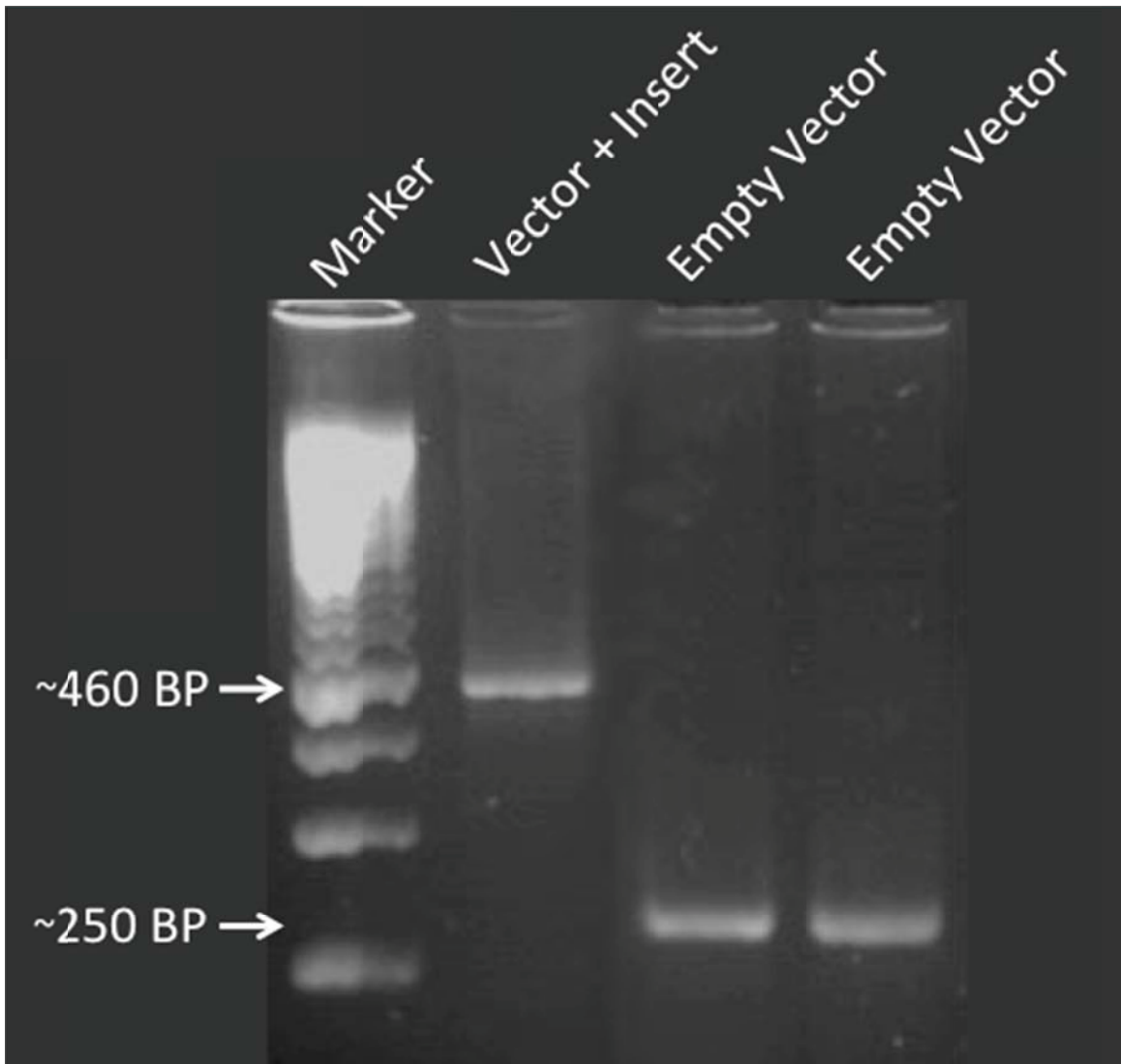
Figure 12: Amplification of pGEM Easy-T Vector with M13 forward and reverse primers to identify positive (lane 2; ~ 460 BP) and negative (lanes 2 and 3; ~250BP) transformants.

Table 4: DNA sequencing results from ten clones containing COI barcode inserts amplified from a LOWA nestling fecal sac collected from a single nest site on Powdermill Run. * - recovered DNA sequence was chimeric.

Order	Family	Genus	Species	GenBank Accession Number	Score	% Match	E-Value
Diptera	Carnoidea	*Chloropidae*	*sp.*	FR682983	241	94.0	1.00E-60
*Ephemeroptera	Heptageniidae	*Epeorus*	*vitreus*	JN200519	274	99.0	1.00E-70
Ephemeroptera	Heptageniidae	*Epeorus*	*vitreus*	JN200519	291	100.0	1.00E-75
Ephemeroptera	Heptageniidae	*Epeorus*	*vitreus*	JN200519	268	97.0	7.00E-69
Ephemeroptera	Heptageniidae	*Epeorus*	*vitreus*	JN200519	291	99.0	2.00E-75
Ephemeroptera	Heptageniidae	*Epeorus*	*vitreus*	JN200519	285	99.0	7.00E-74
Ephemeroptera	Heptageniidae	*Heptagenia*	*sp.*	JQ663258	289	100.0	5.00E-75
Megaloptera	Corydalidae	*Nigronia*	*fasciatus*	AY750518	278	100.0	1.00E-71
*Megaloptera	Corydalidae	*Nigronia*	*fasciatus*	AY750518	56	100.0	2.00E-05
Megaloptera	Corydalidae	*Nigronia*	*fasciatus*	AY750518	278	100.0	1.00E-71
Megaloptera	Corydalidae	*Nigronia*	*fasciatus*	AY750518	270	99.0	2.00E-69

CHAPTER 4: DISCUSSION

In this study, we successfully demonstrated the feasibility of a non-invasive diet study using only residual DNA found in the fecal sacs of Louisiana Waterthrush nestlings. We detected three orders and four species of aquatic macroinvertebrates from one LOWA fecal sac collected from one nest on Powdermill Run, which provides justification for its use in a large-scale investigation of LOWA diet in both impacted and non-impacted drainages. To our knowledge, no method that molecularly identified insect prey from nestling fecal sacs recovered from the wild has ever been published in peer-reviewed literature.

Despite the apparent degradation of DNA obtained from fecal material, we successfully recovered and amplified a 157 bp fragment of the COI barcoding region. We also demonstrated that this fragment could be TA cloned and sequenced with a high degree of accuracy despite its small size. The majority of our amplicons matched the GenBank reference sequences at 94% or greater, which indicates a high degree of certainty in our identification. Furthermore, the majority of the GenBank scores were relatively high while the E-values (Expect-values) were relatively low. Assessing the degree in which two sequences match each other generates these statistics and indicates the likelihood of a match being due to chance alone (Madden 2002). E-values and scores are inversely related and, thus, a low E-value and high score indicates a low probability that a match occurred by chance (Madden 2002). This is important for the assessment of our identified sequences, the majority of which reported a high degree of confidence (Table 4).

As demonstrated in this study, our method of dietary analysis from avian feces has shown to be an effective, accurate and non-invasive assay. It is possible that the method described herein could have implications for future research in ornithology; especially in studies concerning threatened or endangered species where non-invasive procedures are essential. In order for our method to be an effective technique for future avian diet studies, however, the proper collection and storage of feces is essential. As previously noted, residual prey DNA in fecal sacs is highly degraded due to environmental exposure and to the physiology of birds in general (Oehm *et al.* 2011). The uric acid coating of the fecal sacs, in conjunction with exposure to humic acid, sunlight and humidity, rapidly degrades DNA and, thus, interferes with downstream molecular analysis (Oehm *et al.* 2011). These factors emphasize the importance of appropriate collection and storage techniques, which can significantly increase the likelihood of successful DNA barcoding.

Throughout this study, we encountered a litany of issues most likely related to suboptimal collection and storage techniques by inexperienced field personnel. Multiple fecal sac samples were often stored together in one vial with only a relatively small amount of ethanol (up to 1.4 µg in 10 mL), which may have not been further diluted by the presence of water. Furthermore, samples were often stored along with environmental contaminants such as dirt, sand and detritus, which may have contributed to further DNA degradation by bacterial or DNase activity (Oehm *et al.* 2011). As noted above, LOWA also nest on acidified streams, which may add an additional level of DNA degradation to the fecal sac if recovered from the stream bank. We attempted to amplify the COI barcode from DNA derived from 12 fecal samples with only 6 successfully producing the

211 bp amplicon. Among these samples, only the DNA derived from one fecal sac was successfully cloned and sequenced. However, we were able to successfully clone our amplicon from two additional fecal sacs, which produced no match when analyzed in GenBank. The single fecal sample that could be amplified, cloned and sequenced appeared to be collected directly from the nestling during banding and was free of any organic matter, which is preferable for DNA recovery. Currently, the optimal storage conditions for avian feces intended for DNA barcoding analysis is unclear with several studies using various concentrations of ethanol (King *et al.* 2008; Oehm *et al.* 2011) and other citing no advantage to buffer-based storage techniques (Camacho-Sanchez *et al.* 2013). In the future, we plan to test the efficacy of both ethanol and buffer storage techniques to determine the optimal conditions for the preservation of DNA in LOWA fecal sacs.

Even under optimal conditions storage conditions, DNA is difficult to isolate from fecal sacs. Typically, commercial kits offer a convenient, rapid and efficient means of extracting DNA, which we found does not necessarily translate to avian fecal material. The standard protocol for the Qiagen Mini Stool Kit was not successful in extracting DNA and, thus, was significantly modified (Appendix B). The protocol has multiple modifications such as manually removing the uric acid precipitate and evaporating all ethanol used in storage, which may interfere DNA extraction buffers and the silica column's DNA binding properties. Further modification were also employed based on suggestions from the manufacturer and an extraction protocol that was originally intended for the extraction of DNA from bat feces (Zeale *et al.* 2011).

The amplification of the COI barcoding fragment was easily the most time consuming aspect of this study. Prey DNA recovered from feces is not only degraded but is also scarce in comparison to the amount of DNA from the predator itself (Symondson 2002). The most available and highest quality DNA present in fecal sacs is likely to belong to the LOWA, primarily due to the presence of epithelial cells in the feces originating from the LOWA's gastrointestinal tract (Symondson 2002). Low concentrations of intact residual prey DNA inhibits the successful amplification of the COI amplicon and, thus, cannot be identified by downstream DNA sequencing (Idaghdour *et al.* 2003). Furthermore, because the target DNA is degraded, the PCR reaction requires a higher concentration of template in order to achieve sufficient amplification. The problem with this approach, however, is that the excessive input of total DNA may actually inhibit COI amplification. There are, of course, other unknown PCR inhibitors in fecal sacs that may affect *Taq* polymerase more severely than a high concentration of nucleotides (Kohn *et al.* 1995). Even commercial kits cannot remove all PCR inhibitors, which led to the use of acetylated bovine serum albumin (BSA) to help bind inhibitors and prevent their interaction with *Taq* polymerase (King *et al.* 2008).

As indicated in Table 3, only 50% of the fecal sacs attempted successfully produced a COI amplicon. Among these, only one fecal sac from one nest produced identifiable COI fragments. Most of the unsuccessful identifications yielded no match in GenBank, which could be due to errors during PCR amplification. This may be in part due to sub-optimal thermal cycling and reaction conditions in the presence of multiple DNA templates, which may explain the presence of concatenated primer sequences (Osborne *et al.* 2005). Recombinant (chimeric) DNA sequences were also observed in

our sequencing results (Table 4), which can occur spontaneously due to the premature dissociation of polymerase during extension (Bradley & Hillis 1997). The error results in a truncated amplicon that anneals to another template and acts as a primer in subsequent PCR cycles. (Bradley & Hillis 1997). This results in an amplicon comprised of DNA derived from more than one organism, which especially common when attempting to amplify from a pool of templates derived from similarly related species (Bradley & Hillis 1997).

The recovery of concatenated primers and chimeric DNA sequences suggests that our PCR conditions may be sub-optimal. During the revision phase of this thesis, a recent publication that used the primers described by Zeale *et al*., 2011 prompted us to question the protocol being used to generate the COI barcoding amplicon. The publication that sparked our interest utilized a high-fidelity *Taq* polymerase and hot-start thermo cycling protocol to amplify the COI barcode from the feces of Gould's Wattled Bat (*Chalinolobus gouldii*, Gray 1841) (Burgar *et al.* 2013). In light of their success, conducted PCR using Phusion High-Fidelity DNA Polymerase (Thermo Scientific), which is a modified *Taq* polymerase with an additional DNA-binding domain to increase stability and decrease the likelihood of errors during PCR (Wang *et al.* 2004). Using the manufacturer's protocol, we have successfully amplified the COI barcoding amplicon in 13 of 15 fecal sacs (87%) from 3 different nest sites. While DNA sequencing on these fragments has yet to be conducted, this approach appears to be a promising alternative for future studies on LOWA nestling fecal sacs.

The COI primers discussed herein are designed to amplify a 157 bp region but also retain the attached primers throughout downstream molecular applications. This

produces a 211 bp amplicon that can be TA cloned into a plasmid vector, which is advantageous for this type of diet study. Because LOWA are believed to prey upon many species of insect, the DNA extracted from fecal sacs will contain the COI sequences of multiple organisms. This may cause multiple signal peaks during the sequencing phase of the diet analysis. Cloning the COI amplicon into a plasmid vector segregates the mixed amplicons into individual colonies, which results in unequivocal sequencing data.

The use of TA cloning, however, may have also introduced a technical issue into our analysis. Cloning with vectors designed to be selectable by blue/white screening typically provides an effective means for distinguishing clones containing successfully ligated vectors. The blue (negative) and white (positive) coloration is dependent on the expression of the *LacZ* gene, which must be significantly interrupted in order to produce white coloration. In order to disrupt the *LacZ* gene responsible for the blue pigment, the DNA fragments must either overlap the entire region (if the insert is long enough) or disrupt the reading frame. The issue arises from the length of our insert (211 bp), which is not long enough to consistently interrupt the expression of the *LacZ* gene. Furthermore, the process of TA cloning places an adenine nucleotide on each end of the fragment in order to facilitate ligation into the plasmid vector (Robles & Doers 1994). This results in an insert that is 213 bp in length and is evenly divisible by three, which will not disrupt the DNA reading frame and, thus, will not allow for white coloration. According to the manufacturer's specifications, a product less than 500 bp and is evenly divisible by three will not be able to be effectively screened using blue/white selection. A 213 bp fragment meets neither of these criteria and, thus, makes screening clones difficult. Most of the clones that were selected in this study were light blue in color but, unfortunately, the

issue was not known to us at the time and prevented us from analyzing additional clones. In our future work, we will also investigate the feasibility of Next-Generation Sequencing (NGS) in our methodology. The use of NGS would allow us to sequence DNA directly from PCR products, which would eliminate the need for molecular cloning. This approach would also be advantageous because of its ability to sequence many fragments in parallel and identify even relatively small amounts of DNA. Utilizing NGS technology would greatly increase to sensitivity of our assay as well as eliminate many technical issues and potential biases associated with TA cloning.

Despite many technical challenges, we successfully identified three macroinvertebrate orders and four species from a single LOWA nestling fecal sac originating from a single nest on Powdermill Run. Among the DNA recovered from the fecal sac, 54.5 percent of the sequenced clones were matched to two species within the order of Ephemeroptera. The detection of a high percentage of mayfly DNA in the fecal sac of a LOWA supports previous research suggesting that mayflies are an important component of Waterthrush diet (Mulvihill *et al.* 2008). Mayflies are among the most acid-sensitive macroinvertebrate orders and, thus, are virtually non-existent in acidified drainages at PNR (Mulvihill *et al.* 2008). Previous research at PNR has revealed that the presence of mayflies within macroinvertebrate communities is an effective predictor of LOWA breeding success (Mulvihill *et al.* 2008). Based on these data, we predict that mayfly DNA will be less prevalent in fecal sacs collected from acidified habitats than those collected from circumneutral habitats, which may explain variations in LOWA breeding success across streams.

Despite the apparent efficacy of our technique, potentially significant limitations may exist. The variety of invertebrates consumed by nestlings necessitates the use of primers designed to amplify the same gene region across multiple taxa, which may preferentially bind to the DNA of one species over another. While it is believed that degenerate primers will mitigate this phenomenon, the appropriate experiments are yet to be conducted. The relative biomass of prey consumed is also likely to affect dietary composition analysis. Larger prey items contain more DNA and, thus, may comprise a larger percentage of the fecal sac DNA than smaller, more preferred prey. Furthermore, DNA originating from hard-bodied prey may be better preserved than DNA from soft-bodied insects, which may result in a misrepresentation of dietary composition.

Molecular fecal sac analysis using DNA barcoding provides a convenient, non-invasive source of DNA for dietary studies. In this study, we have developed the technical framework for a much larger project that aims to describe the diet of LOWA on both acidified and circumneutral stream habitats. This technique will enable us to answer questions regarding the impact of poor water quality on LOWA breeding success by elucidating the diet of nestlings on streams at PNR. Changes in dietary composition across impacted and non-impacted aquatic ecosystems will reveal the most critical components of the LOWA's diet, which may enhance our understanding of avian food webs and guide the future conservation of riparian bird communities.

Ornithologists studying similar insectivorous avian communities may also find this method useful. While other antiquated diet analysis techniques may be harmful or even lethal, the molecular analysis of fecal material provides a quantitative and non-invasive approach to diet studies, which may also be applicable to many other

insectivorous songbirds. While this method does not offer a means of describing the diet of adults, the study of nestlings may answer important questions about diet of Neotropical songbirds during the critical period of nesting. This approach may enable researchers to study trophic dynamics in avian systems and, thus, guide conservation and management plans for important avian taxa, especially those that are threatened or endangered.

REFERENCES

Agustí N, Unruh TR and Welter SC (2003) Detecting Cacopsylla pyricola (Hemiptera: Psyllidae) in predator guts using COI mitochondrial markers. *Bulletin of Entomological Research* **93**, 179-185.

Ball SL and Hebert PD (2005) Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. *Journal of the North American Benthological Society* **24**, 508-524.

Birky WC, Maruyama T and Fuerst P (1983) An Approach to Population and Evolutionary Genetic Theory for Genes in Mitochondira and Chloroplasts, and Some Results. *Genetics* **103**, 513-527.

Bradley R, D. and Hillis DM (1997) Recombinant DNA Sequences Generated by PCR Amplification. *Molecular Biology Evolution* **14**, 592-593.

Burgar JM, Murray DC, Craig MD, *et al.* (2013) Who's for dinner? High-throughput sequencing reveals bat dietary differentiation in a biodiversity hotspot where prey taxonomy is largely undescribed. *Mol Ecol*.

Camacho-Sanchez M, Burraco P, Gomez-Mestre I and Leonard JA (2013) Preservation of RNA and DNA from mammal samples under field conditions. *Mol Ecol Resour* **13**, 663-673.

Chesser RT, Banks RC, Barker FK, *et al.* (2010) Fifty-First Supplement to the American Ornithologists' Union Check-List of North American Birds. *The Auk* **127**, 726-744.

Chesser RT, Banks RC, Barker FK, *et al.* (2011) Fifty-Second Supplement to the

    American Ornithologists' Union check-list of North American Birds. *The Auk*

    128, 600-613.

Clare EL, Fraser EE, Braid HE, Fenton MB and Hebert PD (2009) Species on the menu

    of a generalist predator, the eastern red bat (Lasiurus borealis): using a molecular

    approach to detect arthropod prey. *Mol Ecol* **18**, 2532-2542.

Courtney LA and Clements WH (1998) Effects of acidic pH on benthic

    macroinvertebrate communities in stream microcosms. *Hydrobiologia* **379**, 135-

    145.

Craig RJ (1984) Comparative Foraging Ecology of Louisiana and Northern

    Waterthrushes. *Wilson Bulletin* **96**, 173-183.

Craig RJ (1987) Divergent Prey Selection in Two Species of Waterthrushes (*Seiurus*).

    The Auk **104**, 180-187.

Cushing CE and Allan JD (2001) *Streams: Their Ecology and Life* Academic Press, San

    Diego.

Deagle BE, Eveson JP and Jarman SN (2006) Quantification of damage in DNA

    recovered from highly degraded samples--a case study on DNA in faeces. *Front*

    Zool **3**, 11.

Deloria-Sheffield CM, Millenbah KF, Bocetti CI, Sykes PW and Kepler CB (2001)

    Kirtland's Warbler Diet as Determined through Fecal Analysis. *Wilson Bulletin*

    113, 384-387.

Eaton SW (1958) A life history study of the Louisiana Waterthrush. *Wilson Bulletin* **70**,

    210-235.

Faaborg J, Holmes RT, Anders AD, *et al.* (2010a) Conserving Migratory Land Birds in

  the New World: Do We Know Enough? *Ecological Applications* **20**, 398-418.

Faaborg J, Holmes RT, Anders AD, *et al.* (2010b) Recent Advances in Understanding

  Migration Systems of New World Land Birds. *Ecological Monographs* **80**, 3-48.

Folmer O, Black M, Hoeh W, Lutz R and Vrijenhoek R (1994) DNA Primers for

  Amplification of Mitochondrial Cytochrome C Oxidase Subunit I from Diverse

  Metazoan Invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294-

  299.

Graveland J (1998) Effects of acid rain on bird populations. *Environmental Review* **6**, 41-

  54.

Guerold F, Boudot J, Jacquemin G, *et al.* (2000) Macroinvertebrate community loss as a

  result of headwater stream acidification in the Vosges mountains (N-E France).

  Biodiversity and Conservation **9**, 767-783.

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W and Hebert PD (2006a) DNA

  barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National*

  Academy of Science **103**, 968-971.

Hajibabaei M, Shokralla S, Zhou X, Singer GA and Baird DJ (2011) Environmental

  barcoding: a next-generation sequencing approach for biomonitoring applications

  using river benthos. *PLoS One* **6**, e17497.

Hajibabaei M, Smith MA, Janzen DH, *et al.* (2006b) A minimalist barcode can identify a

  specimen whose DNA is degraded. *Molecular Ecology Notes* **6**, 959-964.

Hajibabaei M, Spall JL, Shokralla S and van Konynenburg S (2012) Assessing

  biodiversity of a freshwater benthic macroinvertebrate community through non-

destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecology* **12**.

Harper GL, King RA, Dodd CS*, et al.* (2005) Rapid screening of invertebrate predators for multiple prey DNA targets. *Mol Ecol* **14**, 819-827.

Harper GL, Sheppard SK, Harwood JD*, et al.* (2007) Evaluation of temperature gradient gel electrophoresis for the analysis of prey DNA within the guts of invertebrate predators. *Bulletin of Entomological Research* **96**, 295-304.

Hebert PD, Cywinska A, Ball SL and deWaard JR (2003) Biological identifications through DNA barcodes. *Proc Biol Sci* **270**, 313-321.

Idaghdour Y, Broderick D and Korrida A (2003) Faeces as a Source of DNA for Molecular Studies in a Theatened Population of Great Bustards. *Conservation Genetics* **4**, 789-792.

Jenni L, Reutimann P and Jenni-Eiermann S (1990) Recognizability of different food types in faeces and in alimentary flushes of Sylvia warblers. *Ibis* **132**, 445-453.

Juen A and Traugott M (2005) Detecting Predation and Scavenging by DNA Gut-Content Analysis: A Case Study Using a Soil Insect Predator-Prey System. *Oecologia* **142**, 344-352.

King RA, Read DS, Traugott M and Symondson WO (2008) Molecular analysis of predation: a review of best practice for DNA-based approaches. *Mol Ecol* **17**, 947-963.

Kohn M, Knauer F, Stoffella A, Schröder W and Pääbo S (1995) Conservation Genetics of the European Brown Bear - A Study Using Excremental PCR of Nuclear and Mitochondrial Sequences. *Molecular Ecology* **4**, 95-104.

Lovette IJ, Perez-Eman JL, Sullivan JP, *et al.* (2010) A comprehensive multilocus

phylogeny for the wood-warblers and a revised classification of the Parulidae

(Aves). *Mol Phylogenet Evol* **57**, 753-770.

MacArthur R (1958) Population Ecology of Some Warblers of Northeastern Coniferous

Forests. *Ecology* **39**, 599-619.

Madden T (2002) The BLAST Sequence Analysis Tool. In: *The NCBI Handbook*

[Internet], Bethesda (MD): National Center for Biotechnology Information (US).

Maniatis T, Fritsch EF and Sambrook J (1982) *Molecular Cloning, A Laboratory Manual*

Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mattsson BJ and Cooper RJ (2006) Louisiana waterthrushes (*Seiurus motacilla*) and

habitat assessments as cost-effective indicators of instream biotic integrity.

Freshwater Biology **51**, 1941-1958.

Mattsson BJ, Master TL, Mulvihill RS and Robinson DW (2009) Louisiana Waterthrush

(*Parkesia motacilla*). In: *The Birds of North America Online* (ed. Poole A).

Cornell Lab of Ornithology, Ithaca.

Moody DT (1970) A Method for Obtaining Food Samples from Insectivorous Birds. *The*

Auk, 579-579.

Moreby SJ and Stoate C (2000) A quantitative comparison of neck-collar and faecal

analysis to determine passerine nestling diet. *Bird Study* **47**, 320-331.

Mulvihill RS, Latta SC and Newell FL (2009) Temporal Constraints on the Incidence of

Double Brooding in the Louisiana Waterthrush. *The Condor* **111**, 341-348.

Mulvihill RS, Newell FL and Latta SC (2008) Effects of acidification on the breeding

    ecology of a stream-dependent songbird, the Louisiana waterthrush (*Seiurus*

    motacilla). *Freshwater Biology* **53**, 2158–2169.

Oehm J, Juen A, Nagiller K, Neuhauser S and Traugott M (2011) Molecular scatology:

    how to improve prey DNA detection success in avian faeces? *Molecular Ecology*

    Resoures **11**, 620-628.

Ormerod SJ and Rundle SD (1998) Effects of experimental acidification and liming on

    terrestrial invertebrates: Implications for calcium availability to vertebrates.

    Environmental Pollution **103**, 183-191.

Osborne CA, Galic M, Sangwan P and Janssen PH (2005) PCR-generated artefact from

    16S rRNA gene-specific primers. *FEMS Microbiol Lett* **248**, 183-187.

Pfenninger M, Nowak C, Kley C, Steinke D and Streit B (2007) Utility of DNA

    taxonomy and barcoding for the inference of larval community structure in

    morphologically cryptic Chironomus (Diptera) species. *Mol Ecol* **16**, 1957-1968.

Radke WJ and Frydendall MJ (1974) A Survey of Emetics for Use in Stomach Contents

    Recovery in the House Sparrow. *American Midland Naturalist* **92**, 164-172.

Robbins CS, Sauer JR, Greenberg RS and Droege S (1989) Population Declines in North

    American Birds that Migrate to the Neotropics. *Proceedings of the National*

    Academy of Science **86**, 7658-7662.

Robinson S and Holmes R (1982) Foraging Behavior of Forest Birds: The Relationships

    Among Search Tactics, Diet, and Habitat Structure. *Ecology* **63**, 1918-1931.

Robles J and Doers M (1994) pGEM®-T Vector Systems troubleshooting guide.

    Promega Notes **45**, 19–20.

Schreiber KR and Newman JR (1988) Acid Precipitation Effects on Forest Habitats: Implications for Wildlife. *Conservation Biology* **2**, 249-259.

Sheppard SK, Bell J, Sunderland KD*, et al.* (2005) Detection of secondary predation by PCR analyses of the gut contents of invertebrate generalist predators. *Mol Ecol* **14**, 4461-4468.

Symondson WOC (2002) Molecular Identification of Prey in Predator Diets. *Molecular Ecology* **11**, 627-641.

Vannote RL, Minshall GW, Cummins KW, Sedell JR and Cushing CE (1980) The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences* **37**, 130-137.

Wang Y, Prosen DE, Mei L*, et al.* (2004) A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance *in vitro*. *Nucleic Acids Research* **32**, 1197-1207.

Weatherhead P (1984) Fecal Sac Removal by Tree Swallows: The Cost of Cleanliness. *The Condor* **86**, 187-191.

Zeale MR, Butlin RK, Barker GL, Lees DC and Jones G (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resoures* **11**, 236-244.

APPENDIX A: Phenol:Chloroforn:Isoamyl Alcohol (PCI) DNA Extraction from

Ethanol-Preserved Insect Tissue (adapted from Maniatis *et al*., 1982).

1. Dissect a small piece of tissue from thorax or abdomen of insect body. Blot tissue with Kimwipes until all ethanol is removed.

2. Mince tissue into 1 mm pieces with sterile scalpel and add to 1.5 mL centrifuge tube.

3. Add 500 µL ABI lysis buffer (Appendix C) and 5 µL of proteinase K (10 mg/µL). Gently flick the tube to mix and incubate overnight at 55° C or until tissue is fully digested.

4. Add 500 µL Phenol:Chloroform:Isoamyl Alcohol (equilibrated PCI) and invert 10 times. Centrifuge at full speed for 10 minutes.

5. Carefully transfer supernatant to a new 1.5 mL centrifuge tube. Repeat step 4 as many times as necessary to ensure complete removal of potential PCR inhibitors.

6. Add 500 µL Chloroform:Isoamyl Alcohol (Appendix C) and invert tube 10 times. Centrifuge at full speed for 2 minutes. Carefully transfer supernatant to a new 1.5 mL centrifuge tube.

7. Add 1000 µL pre-chilled (4° C) 95% ethanol to supernatant and invert tube 7 times. A white precipitate should form following the addition of the ethanol.

8. Centrifuge the sample for 20 minutes at 4° C to pellet DNA precipitate. Carefully remove the ethanol by decanting.

9. Rinse the DNA pellet with 200 µL of pre-chilled (4° C) 70% ethanol. Carefully remove the ethanol by decanting. With the tube open, place upside-down on a

Kimwipe overnight to dry the DNA pellet. Re-suspend the dried DNA pellet in 50

μL nuclease-free water.

APPENDIX B: Fecal Sac DNA Extraction Protocol with Qiagen DNA Stool Mini Kit (adapted from Zeale *et al.* 2011). **Modifications from original protocol indicated bold.**

1. Using sterile instruments, remove one fecal sac from ethanol storage and place into a sterile weigh boat. **Carefully remove the outermost layer (uric acid precipitate) with sterile forceps and blunt probe. Place the weigh boat containing the uric acid-free fecal sample on a slide warmer for 30 minutes to evaporate ethanol.**

2. Transfer dried fecal sample (~20 mg) to a 2 mL centrifuge tube **and add 40 µL Proteinase K (10 mg/µL)** and 1.4 mL Buffer ASL to the fecal sample. Vortex continuously for 3 minutes or until the stool sample is completely homogenized.

3. **Heat the suspension overnight in a water bath at 70º C.** Vortex occasionally.

4. Vortex for 30 seconds and centrifuge sample at 13,000 rpm for 1 minute to pellet fecal particulate. Pipet 1.2 mL of the supernatant into a new 2 mL centrifuge tube.

5. Add 1 InhibitEX tablet to the sample and vortex immediately and continuously for 1 minute or until the tablet is completely suspended. **Incubate suspension for 5 minutes at room temperature** to allow inhibitors to absorb to the InhibitEX matrix.

6. Centrifuge sample at 13,000 rpm for 3 minutes to pellet InhibitEX matrix.

7. Pipet all the supernatant into a new 1.5 mL centrifuge tube and discard the pellet. Centrifuge the sample at full speed for 3 minutes.

8. **Pipet 40 µL Proteinase K (10 mg/µL)** into a new 1.5 mL centrifuge tube. Vortex and add 400 µL of supernatant from step 7. Do not add Proteinase K directly to Buffer AL.

9. Add 400 µL Buffer AL, vortex until homogenous, and **incubate overnight in a water bath at 70 ºC.**

10. Add 400 µL of absolute ethanol to the lysate and mix by vortexing. Centrifuge briefly to remove any condensation from the lid of the centrifuge tube.

11. Apply 600 µL of the lysate to a QIAmp spin column. Centrifuge at 13,000 rpm for 1 minute. Place spin column in a new collection tube and discard the tube containing the filtrate.

12. Repeat step 11 to load the remaining aliquots of the lysate to the spin column.

13. Add 500 µL Buffer AW1 to the spin column. Centrifuge at 13,000 rpm for 1 minute. Place spin column in a new collection tube and discard the tube containing the filtrate.

14. Add 500 µL Buffer AW2 to the spin column. Centrifuge at 13,000 rpm for 3 minutes.

15. Place spin column in a new collection tube. Centrifuge at 13,000 rpm for 1 minute. Discard the tube containing filtrate.

16. Transfer the spin column into a new 1.5 mL centrifuge tube and **pipette 50 µL of pre-warmed (50° C) Buffer AE to the center of the spin column membrane.** Incubate for 5 minutes at room temperature and centrifuge at 13,000 rpm for 1 minute to elute DNA.

APPENDIX C: Laboratory Reagents

1. ABI Lysis Buffer

   - 0.1 M Tris

   - 4.0 M Urea

   - 0.2 M NaCl

   - 0.01 M EDTA

   - 0.5% n-Laurylsarcosine

2. Chloroform:Isoamyl Alcohol

   - 240 mL chloroform

   - 10 mL isoamyl alcohol

3. S.O.C. Medium (500 mL)

   - 10 g Bactotryptone

   - 2.5 g Yeast Extract

   - 1 mL NaCl (5M)

   - 186.4 mg KCl

   - 1.2 g $MgSO_4$

   - 3.6 g Dextrose

   - 5 mL $MgCl_2$ (1M)

4. Sephadex G-50 Solution (50 mL)

   - 2 g Sephadex G-50

   - 32 mL HPLC H_2O

APPENDIX D: Cytochrome *C* Oxidase Subunit I (COI) Sequences

NOTE: Underlined nucleotides represent forward and reverse insect primers. Bolded nucleotides represent the COI barcoding amplicon. Non-emphasized text represents the nucleotide sequence of the P-Gem T Easy Vector (Promega).

CLONE ID: BKT_19.1 - *Epeorus vitreus & Nigronia fasciatus (chimeric)*

GACNCCCTANGGGCGATTGGCCCGCGTCGCATGCTCCGGCCGCCATGGCGGC

CGCGGGAATTCGATT<u>TACTAATCAATTTCCAAATCC</u>**TCCGATTATAATAGGC**

ATAACTATAAAGAAAATCATAATAAATGCATGAGCAGTTACAATGACATT

ATAAATTTGGTCGTCCCCAATTAGTGAACCGGGTTGACCAAGTTCAGCA

CGAACAAGTAGTCTGAGTGAGGTTCCGACTATACCGGACCAAGCACCAG

ATCAGCTAAGAGAAGTTCCGACCATTCCTGATCATGCA<u>CCAAAAATAAAA</u>

<u>TATAAAGTTCCAATATCT</u>AATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGA

CCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGT

GTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT

TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA

AAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTC

ACTGCCCGCTTTTCCAGTCGGGAAACCTGTCGTGCCACTGCATTAATGAATCG

GCCAACGCGCGGGGGAGAGGCNGTTTGCGTATTGNGCGCTCNTCCGCTTCNT

CGCTCATGNNTCGCTGCGCTCGGTCGTTCGGCTGCCGCCANCGGTATCACCTC

CTNNAAAGCGGTAATACGGTANCCACNAAATCCGGGGGATACGCAGGAAAN

AACNTGTGGGCAAAAGGGCNGCAAAAGGCNGGGACCGTAAAAGGCCCCNTT

GCTGGCGTTTTTCATAGCTCCGCCCCTGACGAANATCNAAAATCACCCNCAA

TTNNAGNGGGGGAACNNAAGGCTNTAAANNCAGNGTTNCCNGGAACCCCCN

GNNCCCNGTNNACCGNCTTACNGACNNTCCCTTNCCTTCGGGANNGGNGNTT

TAAATNCNGGAGNACNNTCGNNAGNCNC

CLONE ID: BKT_19.2 – *Nigronia Fasciatus*

GCGCGGGAATTCGATTAGATATTGGAACTTTATATTTTATTTTTGG**TGCATGA**

TCAGGAATGGTCGGAACTTCTCTTAGCTTACTAATCCGAGCCGAATTAG

GACAACCTGGGTCATTAATTGGAGATGATCAAATTTATAATGTTATTGTA

ACAGCTCATGCATTTGTAATAATTTTTTTTATAGTTATACCTGTAATAATT

GGAGAATTTGGTAATTGATTAGTTAATCACTAGTGAATTCGCGGCCGCCTGCA

GGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATT

CTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT

GTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAACATAA

AGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTT

GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT

GAATCGGCCAACGCGCGGGGAAGAGGCGGTTTGCGTAATTGGGCGCTCTTCC

GCTTCCTCGCTCACTGGACTCNGCTGCGCTCGGTCGTTCCGGCTGCCGGCGAG

CGGTATCAACTCACTCAAANGGCGGTAATACGGTTTATCCACAGAAATCAGG

GGAAAACCCCAAGGAAAAACATGTGANCCAAAGGCCACCAAAAGGCCGGNA

ACGTTAAAAAGGCCGCGTTCCGGCGTTTTCCAAAGCNCNCCCCCTGACANCA

CNCNAAATCNNCCCCAATCAAGGNGNCAACCCAAGGCCTNTAA

CLONE ID: BKT_19.4 – *Epeorus vitreus*

GNCACCCTTGGCGATTGGGCCGACGTCGCATGCTCCCGGCCGCCATGGCGGC

CGCGGGAATTCGATT<u>AGATATTGGAACTTTATATTTTATTTTTGG</u>**TGCTTGGT**

CCGGTATAGTCGGAACCTCACTCAGACTACTTATTCGTGCTGAACTTGGT

CAACCCGGTTCACTAATTGGGGACGACCAAATTTATAATGTCATTGTAAC

TGCTCATGCATTTATTATGATTTTCTTTATAGTTATGCCTATTATAATC<u>GG</u>

<u>AGGATTTGGTAATTGATTAGTA</u>AATCACTAGTGAATTCGCGGCCGCCTGCAG

GTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTC

TATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTG

TGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAA

AGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTT

GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT

GGAATCGGCCAACGCGCCGGGGAGAGGCGGTTTGCGTATTGGGGCGCTCTTC

CGCTTCCTCGCTCATGAACTCGCTGCGCTCGGGTCGTTCGGCTNCGGCCGAGC

GGTAACAGCTCACTCCAAGGCGGTAATACGGTTATCCCANAATCAGGGGANA

ACGCCAGGAAAAACCTTGTNAACCAAAAGGCCANCAAAAGCCAGGAACCNT

TAAAAAGGCCCGTTGCTGGCGTTTTTTCCTANGCCCNCCCCCCTGNACNACNT

CNCAAAAN

CLONE ID: BKT_19.6 – *Epeorus vitreus*

TGATCNCNCTAAGGGCGACTNGGCCGACGTCGCATGCTCCCGGCCGCCATGG

CGGCCGCGGGAATTCGATT<u>ANATATTGGAACTTTATATTTTATTTTTGG</u>**TGCT**

TGGTCCGGTATAGTCNGAACCTCACTCAAACTACTTATTCGTGCTGAACT

TGGTCAACCCGGTTCACTANTTGGGGACGACCAAATTTATAATGTCANTG

TAACTGCTCATGCATTTATTATGATTTTCTTTATAGTTATGCCTGTTATAA

TC<u>GGAGGATTTGGTAATTGATTAGTA</u>AATCACTAGTGAATTCGCGGCCGCCT

GCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGT

ATTCTATACTGTCACCTAAATAACTTGGCGTAATCATGGTCATAGCTGTTTCC

TGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACCAGCCGGAAGC

ATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTG

CGTTGCGCTCACTGCCCNCTTTCCAGTCGGGAAAACTGTCGTGCCAGCTGCAT

TAATGAATCNGCCAACGCGCGGGGAAAGGCCGTTTGCGTAATGAGCGCTCTT

CCACTTCCTCGCTCAATGAATCGCTGCNCTCGGTCCTTCNGCTGCNGCAAACN

GTACNNCTCCTCAAAGCGTAATACGGTTATCCCANAACANGGGATACCNAGG

AAAAACTTGTGACAAAAGCCCNCCAAAGCCAGNACCTTAT

CLONE ID: BKT_19.7 – *Heptagenia sp.*

GTGACGCCCTANGGGCGATNCGGCCGACGTCGCATGCTCCCGGCCGCCATGG

CGGCCGCGGGAATTCGATT<u>AGATATTGGAACTTTATATTTTATTTTTGG</u>**GGCT**

TGATCTGGTATAGTAGGAACATCTCTTAGTTTACTAATTCGAGCTGAATT

GGGACAACCAGGGTCACTGATTGGAGACGACCAAATTTATAATGTAATT

GTTACAGCTCATGCTTTTATCATAATTTTCTTTATAGTTATGCCTATTATA

ATC<u>GGAGGATTTGGAAATTGATTAGTA</u>AATCACTAGTGAATTCGCGGCCGCC

TGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAG

TATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTC

CTGTGTGAAATTGTTATCCGATCACAATTCCACACAACATACGAGCCGGAAG

CATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATT

GCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCA

TTAATGAATCGGCCAACGCGCGGGGANAAGCGGTTTGCGTATTGGGGCGCTC

TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTCCGCTGCCGCGACG

GTATCACTCCTCAAAGCCGGTATACGGTTATCCACAAACAGGGGATACGCAG

GAANACTGGTGACCAA

CLONE ID: BKT_19.9 – *Epeorus vitreus*

NACNCNCTAAGGGCGATTGGCCCGCGTCGCATGCTCCCGGCCGCCATGGCGG

CCGCGGGAATTCGATTTACTAATCAATTTCCAAATCC**TCCGATTATAATAGG**

CATAACTATAAAGAAAATCATAATAAATGCATGAGCAGTTACAATGACAT

TATAAATTTGGTCGTCCCCAATTAGTGAACCGGGTTGACCAAGTTCAGCA

CGAATAAGTAGTCTGAGTGAGGNTCCGACTATACCGGACCAAGCACCAA

AAATAAAATATAAAGTTCCAATATCTAATCACTAGTGAATTCGCGGCCGCCT

GCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGT

ATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCC

TGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGC

ATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTG

CGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCAT

TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTT

CCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCNGGCGAGC

GGTATCATCTCNCTCAAAGCNGTAAACGGTTATCACCAGAACAGGGGAATAA

CGCAGGAAAGAACATGTGAGCAAAAGGCCACCAAANGGCCNGAACCNTAAA

ANGNCNCGTTGCNTGGGGTTTTTCCATAAGGCTCCGCCCCCCTGNACGAACA

TNCCAAAANTCGCGCNCAAGTCANAAGNGGNNAACCCGNCNGGATTATAAA

GAAACNNGAGTTCCCCCGGAAANCCCCTNGGGTTNCCTGTTCCACCCGCGCT

TACNGAACCNNCNCTTTCCCCTNGNATCGGGGNTTCTNANCCCNCTNNGNNT

CCANCGGNNGGC

CLONE ID: BKT_19.10 – *Nigronia fasciatus*

CNCCCTTAGGGCGATTGGCCCGCGTCGCATGCTCCCGGCCGCCATGGCGGCC

GCGGGAATTCGATTAGATATTGGAACTTTATATTTTATTTTTGG**TGCATGATC**

AGGAATGGTCGGAACTTCTCTTAGCTTACTAATCCGAGCCGAATTAGGA

CAACCTGGGTCATTAATTGGAGATGATCAAATTTATAATGTTATTGTAAC

AGCTCATGCATTTGTAATAATTTTTTTTATAGTTATACCTGTAATAATTGG

AGGATTTGGAAATTGATTAGTAAATCACTAGTGAATTCGCGGCCGCCTGCAG

GTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTC

TATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTG

TGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAA

AGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTT

GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAAT

GAATCGGCCAACGCGCGGGGGANAAGCGGTTTGCGTATTGGGCGCTCTTCCG

CTTCCTCGCTCATGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGACGGTATC

AACTCACTCAAAGGCGGTAATACCGGTATCCACCGAATCAGGGGATACGCAG

GNAAAACATGTGAACCAAAAGGCCACCAAANGGCNGGNACCGTAAAAAGGC

CCCGTTGCTGGCGTTTTTNCCNAGGCCNCCCCCCTGACGACCNCCAAAAANT

CGACCCCAAGTCAAAGTGGCNAACCCGCAGGCCTNTAAAAANCCAGGGTTTC

CCCNGGANCNCCCNNGCNCCCCGTTCCACCCNCCNTTACGGAACCTGCCCNT

TCCCCTCGGAANGGNCCTTCCACCNCTGTNAGATCCCNTTNGGGGGGGNNCCC

CNACGGGGTGGNCACCCCTTNT

CLONE ID: BKT_19.11 – *Epeorus vitreus*

CCCTAANGCGATTGGCCCGACGTCNCATGCTCCCGGCCGCCNTGGCGGCCGC

GGGAATTCGATTAGATATTGGAACTTTATATTTTATTTTTGG**TGCTTGGTCCG**

GTATAGTCGGAACCTCACTCAGACTACTTATTCGTGCTGAACTTGGTCAA

CCCGGTTCACTAATTGGGGACGACCAAATTTATAATGTCATTGTAACCGC

TCATGCATTTATTATGATTTTCTTTATAGTTATGCCTATTATAATCGGAGG

ATTTGGAAATTGATTAGTAAATCACTAGTGAATTCGCGGCCGCCTGCAGGTC

GACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTAT

AGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGA

AATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAACATAAAGT

GTAAAGCCTGNGGTGCCTAATGAGTGNAGCTAACTCACATTAATTGCGTTGC

GCTCACTGCCCGCTTTCCAGTCNGGAAACTGTCGTGCCAGCTGCATTAATGAA

TCGGCCAACGCGCGGGGAGANGCGGTTTGCGTATTGGGCGCCCATCCGCTTC

NTCGCTCCACTGAACTCGCCTNCCCTCGGTCNTTCNGCCTGCTGGCGAACNAG

TATCACTCANTTCAAAAGCCNGNAATACNNGTTATNCCACNATAAATCAGGG

GANTACCGCNNGNAAAAACATNTGATCANAAGGCCNACCANAGGCCAGNNA

NCCTTTAAAAGGCCCCTNTCCGGCNTTTTTCCNTANGCCCNCCCCTTGACNAA

NATCTAAAAACNNNCCCANTTNTAAGTGGGCNAACCCCTNGGANNNAAAAT

NCAGCGTNTCCCCNNNTAATCCCTNNNGANTCCGTNNAACCGNCCTTNACNA

AANCTNCCCTTCCCCCGGANCANNNTATTCNANACCCCGNAGNACNNATNCN

GCNCCNNCCCTNCNNAANTNN

CLONE ID: BKT_19.12 – *Chloropidae sp*.

GANCNCNCTAAGGGCGATTGGGCCGACGTCGCATGCTCCCGGCCGCCATGGC

GGCCGCGGGAATTCGATT<u>AGATATTGGAACTTTATATTTTATTTTTGG</u>**GGCTT**

GAGCTGGAATAGTTGGAACTTCCCTGAGTATTTTAATTCGAATAGAATTA

GGCCGTCCTGGAGCCTTAATTGGTGATGATCAAATTTATAATGTAATTGT

TACAGCTCATGCATTTGTAATAATTTTTTTTATAGTAATACCTATTATAAT

T<u>GGAGGATTTGGAAATTGATTAGT</u>TAATCACTAGTGAATTCGCGGCCGCCTG

CAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTA

TTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCT

GTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCA

TAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC

GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATT

AATGAATCGGCCAACGCGCGGGGAGAGGCGGGTTTGCGTATTGGGCGCTCTT

CCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGGTCGTTCGGCTGCCGGCGAG

CGGTNTCGCCCCACTCAANGCGGTNATACCGGTTACCCCAGAATCAGGGGGA

TNACGCAGGAANNACATGGTGAACAAAAGGNCACCAAANGCCAGGAANCCT

AAAAAGGCCCCGTTGCTNGCGTTTTTCCANANGCNCCCCCCCCTGACCAA

CLONE ID: BKT_19.16 – *Nigronia fasciatus*

TGATCGCCCTAAGGGCGATNCGGCCGACGTCGCATGCTCCCGGCCGCCATGG

CGGCCGCGGGAATTCGATT<u>AGATATTGGAACTTTATATTTTATTTTTGG</u>**TGCA**

TGATCAGGAATGGTCGGAACTTCTCTTAGCTTACTAATCCGAGCCGAATT

AGGACAACCTGGGTCATTAATTGGAGATGATCAAATTTATAATGTTATTG

TAACAGCTCATGCATTTGTAATAATTTTTTTATAGTTATACCTGTAATAAT

T<u>GGAGGATTTGGAAATTGATTAGT</u>TAATCACTAGTGAATTCGCGGCCGCCTG

CAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTA

TTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCT

GTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCA

TAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC

GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCACTGCATTA

ATGAATCGGCCAACGCGCGGGGAGANGCGGTTTGCGTATTGGGCGCTCTTCC

GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGCTGCCGCCNACGGTA

TCANCTCACTCNAAGGCGGTAATACGGTTACCNCANAATCNGGGATACGCAG

GAAAACTGTGANCAAAAGCCACAAANGCCGGACCGTAAAAGCCCGTTCTGG

CGTTTTCCAAGCTCNCCCCTGACACACAAAA