

UNIVERSITY OF OKLAHOMA  
GRADUATE COLLEGE

EVALUATING THE USE OF ENVIRONMENTAL DNA FOR SURVEYING SEMI-  
AQUATIC AVIAN FAUNA

A THESIS  
SUBMITTED TO THE GRADUATE FACULTY  
in partial fulfillment of the requirements for the  
Degree of  
MASTER OF SCIENCE

By  
KAYLEIGH NICOLE ANAYA  
Norman, Oklahoma  
2018

EVALUATING THE USE OF ENVIRONMENTAL DNA FOR SURVEYING SEMI-  
AQUATIC AVIAN FAUNA

A THESIS APPROVED FOR THE  
DEPARTMENT OF BIOLOGY

BY

---

Dr. Jeffrey Kelly, Chair

---

Dr. Cameron D. Siler

---

Dr. Michael Patten

© Copyright by KAYLEIGH NICOLE ANAYA 2018  
All Rights Reserved.

## **Acknowledgements**

I express my most sincere gratitude to my advisor Dr. Jeff Kelly for giving me the opportunity to pursue my graduate education under his guidance at the University of Oklahoma. I am forever grateful for his constant generosity, support, and encouragement, and I greatly admire his humility and expertise.

I thank Dr. Michael Patten, Dr. Cameron Siler, and Dr. Priscilla Crawford for serving on my thesis committee and always being available and eager to help throughout the development of my thesis project.

Special thanks to Dr. Tamaki Yuri for training me in the lab, helping me to develop lab protocols and assay designs, and for guidance in data interpretation. I could not have completed this project well without her assistance and expertise. I also thank Jessa Watters for helping me with field sampling design and equipment, training me in tissue extraction and other lab protocols, providing honest and timely feedback, and for always being willing to help. I thank Elyse Freitas for aiding me with any questions I had, training me in lab protocols, and helping with sequence trimming and alignments. Thanks to Ann Harris at the University of Oklahoma Biology Core Molecular Laboratory for her assistance with sequencing. I am also grateful for Dr. Andrea Contina, Dr. Kyle Horton, Kyle Broadfoot, Paula Cimprich for their suggestions, support, and friendship, as well as many other faculty, staff, and graduate students who aided me throughout my graduate studies at OU.

I thank the Oklahoma Ornithological Society, the University of Oklahoma Department of Biology, and the Oklahoma Biological Survey for supporting me and making this project possible.

I thank my husband, Felipe, for his unconditional love and support, for helping me with field sampling, and for sharing his knowledge and experience in writing and presenting. I thank my sister, Talisa, for helping me with field sampling, and for my parents for always supporting me and helping me to reach my dreams. Finally, I thank God for giving me the opportunity to pursue graduate education, for always loving and providing for me, and for using this time in graduate school to draw me closer to Himself; without Him I can do nothing.

# Table of Contents

Acknowledgements .....	iv
List of Tables .....	viii
List of Figures.....	ix
Abstract.....	x
Introduction .....	1
Methods and Materials .....	9
Study Species.....	9
Study Area .....	11
Landscape of Oklahoma .....	11
Ponds of Central Oklahoma.....	11
Field Sampling.....	13
Water Filtration .....	16
eDNA Filter Extraction Protocol.....	16
Tissue Extraction Protocol .....	18
Primer-Probe Assay Design .....	19
Primer-Probe Validation.....	22
eDNA Screening.....	26
Statistical Analysis .....	26
Ethics Statement .....	27
Results .....	29
Discussion.....	35
Effect of Goose abundance on eDNA Detection.....	35

Effects of Environmental Variables on eDNA Detection .....	37
Further Considerations .....	39
Conclusions and Future Research .....	40
References .....	42
Appendix I.....	50
Appendix II.....	51
Appendix III .....	52
Appendix IV .....	53
Appendix V .....	54

## List of Tables

Table 1 .....	15
Table 2 .....	22
Table 3 .....	23
Table 4 .....	24
Table 5 .....	30
Table 6 .....	34
Table 7 .....	34



## List of Figures

Figure 1.....	12
Figure 2.....	17
Figure 3.....	17
Figure 4.....	28
Figure 5.....	30
Figure 6.....	31
Figure 7.....	33

## Abstract

The identification and quantification of DNA shed into aqueous environments can be used to estimate site occupancy and population density of various taxa. It is often easier to sample for environmental DNA (eDNA) than use traditional surveying techniques, and results of eDNA analyses are often more accurate. The sensitivity and efficiency of eDNA sampling makes it a useful tool for conservation biology. Detection of eDNA is particularly useful for aquatic species that are rare or difficult to survey, though it may have promising applications for less aquatic species as well. I tested the hypothesis that eDNA can be used to detect the presence of Canada geese, a semi-aquatic bird species. I screened pond samples from central Oklahoma for the Canada goose NADH dehydrogenase subunit 2 (ND2) mitochondrial gene using a species-specific primer-probe with quantitative, real-time PCR (qPCR). Canada goose eDNA was detected at all ponds sampled, including those where Canada geese were absent during sampling. eDNA detection rate increased as goose abundance increased and was unaffected by other environmental variables (e.g. temperature, pH, water flow). These results may be due to increased eDNA production where Canada geese are more abundant. The pond environments were relatively homogenous and did not significantly affect eDNA presence and detection. Overall, eDNA can be used to detect Canada goose DNA in ponds and potentially to monitor other bird species which are rare, endangered, cryptic, or difficult to survey (e.g. black rail, *Laterallus jamaicensis*). Used in conjunction with conventional surveying techniques, eDNA can aid in the conservation of rare and threatened species across a wide range of taxa, including semi-aquatic birds.

## **Introduction**

Ongoing global decline in biodiversity is of great concern (Hull 2015).

Accelerating climate change threatens biodiversity and is detrimental to ecosystem functioning and ecological and human health (Cardinale et al. 2012; Ceballos et al. 2017). Effective conservation efforts are critical to slow or halt the loss of biodiversity. A significant priority in biodiversity conservation is accurately assessing occurrence and distribution of species, which depends on reliable species monitoring techniques (Margules & Pressey 2000; Hopkins & Freckleton 2002). For conservation biologists and policy-makers to make scientifically informed decisions to conserve, manage, and enjoy the planet's life and resources, researchers must first understand the planet's diversity and distribution patterns (Martin 2004; Wheeler et al. 2004). This knowledge is relevant for a variety of disciplines, including conservation biology, ecology, evolutionary biology, and biogeography (Ficetola et al. 2008).

A persistent decline in the number of expert taxonomists, combined with uncertainty and cost associated with field survey techniques, slows the creation of accurate species distribution maps for use in effective biodiversity conservation (Dayrat 2005). Since the 1950s there has been a decline in systematics and taxonomic researchers. This decline has resulted in similar declines in the description of discovered species, documentation of population sizes and species distributions, the discovery of previously unknown species, and the assessment of natural history and ecology (Hopkins & Freckleton 2002; Wheeler 2014). In addition, species which are locally scarce and occupy small ranges are easily overlooked by taxonomists, though these species are disproportionately more likely to be threatened with extinction than

widespread, locally-abundant species (Pimm et al. 2014). This could lead to focusing conservation efforts on less threatened and more easily identifiable taxa while neglecting species of greater conservation concern (Hopkins & Freckleton 2002; Wheeler 2014). In addition, survey methods which depend primarily on visual counting are often costly and effort-intensive, have low detection probability, are often be seasonally restricted, and may increase the risk of harming sensitive species (Wilcox et al. 2013; Turner et al. 2014). Failure to detect threatened native species or over-estimates of species density could lead to mis-management of critical conservation areas, which could be detrimental to those species' persistence and impact their functional roles within ecosystems. To improve conservation management strategies in an effort to preserve global and local biodiversity, more efficient methods for assessing biodiversity are needed. Technological advances in genetic methods and computation open new opportunities for species surveys that rely less on detection by individual observers and their skill in field identification.

One such advance in species identification involves DNA taxonomy, which identifies and differentiates species using DNA sequences. Of particular interest is the detection and identification of environmental DNA (eDNA), or DNA extracted from environmental samples (e.g. water, soil, air) without physically capturing the target organism (Wilcox et al. 2013; Thomsen & Willerslev 2014; Turner et al. 2014). The development of survey methods that detect eDNA in water samples to assess species' presence and abundance has garnered substantial interest in the last decade due to its wide availability and efficiency for species monitoring (Takahara et al. 2013). Macrofaunal eDNA typically originates from epidermal cells, saliva, urine, blood,

mucous, reproductive fluids, or feces and persists in the environment for variable amounts of time (e.g. days to thousands of years) depending on climate and an array of environmental factors, such as water flow, temperature, pH, and sediment load and type (Bohmann et al. 2014; Rees et al. 2014; Thomsen & Willerslev 2014). Used for detection of a single species, eDNA can provide information about species presence or absence, distribution, and, in some cases, abundance, density, or biomass (Ficetola et al. 2008; Thomsen et al. 2012; Takahara et al. 2013; Doi et al. 2015). A multi-species approach to eDNA using metabarcoding can assess current or historic species and community diversity at a site (Evans et al. 2016). Used in conjunction with molecular analyses (e.g. DNA barcoding or metabarcoding), eDNA makes rapid and reliable species detection possible without the need for visual confirmation or physical capture of the target organism (Thomsen & Willerslev 2014; Davy et al. 2015). The detection of eDNA in species monitoring programs is therefore of particular interest for organisms that are difficult to survey, such as invasive, endangered, or cryptic species, and species for which accurate distribution data may be lacking.

Because all living organisms shed DNA into their environment, eDNA can be a useful tool to study various taxa in an array of habitat types. To date, eDNA has been most widely used to detect microorganisms in soil, ice cores, permafrost, and in fresh and sea water (Thomsen & Willerslev 2014). Since 1991, eDNA has also been used to study macro-organisms in a variety of environments, including human forensics and agricultural transgenics. It was not until 2008 that eDNA was used to detect aquatic macrofaunal (Ficetola et al. 2008), and since then most of these studies have focused on the detection of strictly aquatic organisms (e.g. amphibians, fish), including threatened

or invasive species (Takahara et al. 2013; Wilcox et al. 2013; Rees et al. 2014; Thomsen & Willerslev 2014; Turner et al. 2014; Davy et al. 2015). Aquatic species monitoring using eDNA was first applied by Ficetola et al. (2008) to detect the presence of the invasive American bullfrog (*Lithobates catesbeianus*) in ponds in France. Since then, eDNA has been used to investigate increasingly complex biodiversity questions for aquatic species. Some studies have used eDNA to evaluate the distribution of invasive fish species in ponds (Takahara et al. 2013), estimate aquatic species' abundance and biomass (Takahara et al. 2012; Doi et al. 2015), quantify fish and amphibian species diversity using a multi-species approach (Evans et al. 2016), characterize entire faunas using next-generation sequencing (Rees et al. 2014), and investigate the effects of environmental conditions on eDNA degradation, availability, and detectability (Stoeckle et al. 2017). Because of its broad range of applications and potential generalizability across many taxa, eDNA has potential for use in many fields of study, including paleontology, ecology, disease ecology, parasitology, and conservation biology (Thomsen & Willerslev 2014; Bass et al. 2015; Huver et al. 2015). Further research will continue to illuminate the monitoring technique's full range of applications.

Species detection using eDNA involves collecting, extracting, amplifying, and sequencing DNA. The three primary approaches to species detection with eDNA are interrogation of: 1) a single known DNA sequence from a target species using species-specific primers; 2) a single unknown sequence of an unidentified species using more general primers followed by a BLAST (Basic Local Alignment Search Tool) search to identify the unknown sequence by finding regions of similarity between it and

biological sequences in a database (NCBI 2018); 3) many unknown sequences simultaneously by pooling amplicons and performing deep sequencing followed by a BLAST search (Bohmann et al. 2014). Many eDNA studies have used DNA barcoding to target a single species' DNA in environmental samples, though there is great potential for next-generation sequencing to recover sequences of thousands of specimens from a single water sample, providing a more complete measure of community biodiversity over time and space (Shokralla et al. 2012).

Although eDNA has an array of potential uses in a variety of disciplines, some studies have shown that such approaches may not be equally effective for all taxa (Davy et al. 2015). Organisms shed DNA into the environment at different rates, and factors such as the amount of time an animal spends in a particular environment (e.g. in water, on land, in the air, or underground), the organism's behavior, and habitat preferences all affect the accumulation of eDNA (Davy et al. 2015). In addition, sampling method and environmental conditions (e.g. for water samples: pH, temperature, water flow, and sediment load and type) influence the persistence, due to DNA degradation, and detectability of eDNA (Strickler et al. 2015; Stoeckle et al. 2017; Tsuji et al. 2017; Seymour et al. 2018). It is therefore vital to determine optimal sampling methods to detect eDNA for a variety of taxa, the effects of environmental conditions on eDNA degradation and persistence, and the limitations of eDNA to detect various taxa in different environments in order to accurately interpret eDNA surveillance results (Barnes et al. 2014).

Many recent studies have successfully used eDNA to detect aquatic macrofauna, such as turtles (Davy et al. 2015), frogs (Ficetola et al. 2008; Evans et al. 2016),

crayfish (Treguier et al. 2014), and fish (Takahara et al. 2013; Turner et al. 2014; Evans et al. 2016). Fewer studies have extended this surveying method to semi-aquatic species (e.g. some mammals, birds, and insects) to compare its use for taxonomic groups with different behaviors and habitat preferences than strictly aquatic species (Thomsen et al. 2012). Thomsen et al. (2012) sampled for eDNA of six animal species representing various taxa (fish, amphibians, crustaceans, mammals, and insects), including the Eurasian otter, *Lutra lutra*, in a variety of aquatic environments in Europe (i.e. ponds, lakes, streams, and temporary pools) to determine the applicability of eDNA for a variety of freshwater species. All taxa were detected successfully – even those whose life history would seemingly make them difficult to detect with these methods – though at lower rates for two species (Eurasian otter and European weather loach, *Misgurnus fossilis*) in areas where they were not observed during sampling. For the otter, a decreased detection rate could be in part due to its semiaquatic lifestyle and large territorial range. In addition, the species often inhabits streams where water retention time is low and water volume is high (Thomsen et al. 2012). Nonetheless, this study demonstrates that eDNA of semi-aquatic species can be detected successfully and at sites where species presence was unconfirmed using visual surveys at the same level of effort. As eDNA research continues to develop and sampling procedures improve and become more standardized, eDNA may prove to be a valuable complement to conventional monitoring of even semi-aquatic taxa (Thomsen et al. 2012).

Before eDNA can be used for monitoring semi-aquatic, or even terrestrial, taxa within a conservation context, further work is needed to determine the method's reliability for detecting such species in water samples. Nonetheless, it has been



demonstrated that eDNA of semi-aquatic and terrestrial taxa can be detected in water samples. In addition to using eDNA to detect six target species, Thomsen et al. (2012) recovered eDNA of four species that live in close proximity to water, including the red deer (*Cervus elaphus*), wood pigeon (*Columba palumbus*), Eurasian coot (*Fulica atra*), and marsh warbler (*Acrocephalus palustris*). Though the Eurasian coot spends much of its time in aquatic environments, the other three species have much less interaction with water. This finding demonstrates that eDNA of semi-aquatic and even non-aquatic species can be detected in water, though the interpretation of these results may differ from those of strictly aquatic species (Ficetola et al. 2008; Thomsen et al. 2012). To date, the application of eDNA techniques to monitor semi-aquatic taxa has been minimal, and though birds have been detected positively using generic primers, no research has yet targeted bird species directly as the focal taxa to test the method's generalizability. Therefore, the aim of my study was to evaluate the use of eDNA to detect a semi-aquatic bird species. Specifically, I tested the method's ability to detect the Canada goose (*Branta canadensis*), an abundant, semi-aquatic bird species, and investigated the impact of goose abundance and environmental factors on eDNA detectability.

The use of eDNA techniques for monitoring species has great potential to improve biodiversity assessments and aid in more effective conservation strategies. Though eDNA detection in water has been most widely used for aquatic species, it has potentially promising applications for monitoring semi-aquatic taxa, as well. Before applying this technique to conservation settings, it is useful to first determine whether eDNA can reliably detect the presence of a widespread and relatively abundant bird

species as a proof-of-concept. Because eDNA of three bird species was successfully detected in Thomsen et al. (2012), I predicted that Canada goose eDNA would be positively detected at ponds where Canada geese were present during sampling. In addition, I expected detection rate to increase with target species abundance and that environmental variables might affect eDNA detection rates among samples. To test these predictions, I took water samples and recorded environmental data at ponds, both with and without Canada geese present, in central Oklahoma. Ultimately, the results of my study have implications for the eDNA technique's use in monitoring other semi-aquatic bird species or taxa which are rare, cryptic, or difficult to survey, and especially for those of conservation concern.

## Methods and Materials

### Study Species

The Canada goose (*Branta canadensis*) is a common and abundant species in central Oklahoma and throughout much of the United States. Range-wide, Canada geese typically inhabit reservoirs, ponds, and rivers in temperate and tundra regions, wintering throughout most of the U.S. and in northern Mexico. They usually feed in pastures and grain fields (Baumgartner & Baumgartner 1992; Banks et al. 2004), though they have become common in urban parks and golf courses, as well (Reinking 2004), as is the case in Norman, Oklahoma, where water sampling took place.

An important aspect of eDNA study designs that use DNA barcoding to distinguish between closely related species is the inclusion of outgroup (non-target) species with which to test the sensitivity and specificity of a primer-probe assay. I selected three closely-related outgroup species in the family Anatidae (ducks, geese, swans): mallard (*Anas platyrhynchos*), gadwall (*Anas strepera*), and cackling goose (*Branta hutchinsii*). Recent studies and phylogenies have shown these species to be appropriate outgroup lineages (Gonzalez et al. 2009; Ottenburghs et al. 2016). In addition, all three species co-occur with Canada geese in Oklahoma (Sullivan et al. 2009) and are fairly common throughout the state, particularly during winter (Reinking 2004).

Canada geese, cackling geese, mallards, and gadwalls have occurred regularly in Oklahoma for at least the past century (Nice & Nice 1924; Sutton 1967; Wood & Schnell 1984). The cackling goose was not classified as a species distinct from the Canada goose until 2004 (Sibley 2004), and therefore it lacks a detailed historical

account in Oklahoma. Nonetheless, it is likely that cackling geese have long been common in the state; in 1967, George Miksch Sutton wrote of different subspecies of migratory and winter resident Canada geese (in the broad sense) in Oklahoma, with at least three subspecies of the Canada goose regularly migrating through Oklahoma. These different subspecies which Sutton hypothesized belonged to “two species rather than one,” included both the Canada and cackling goose (Sutton 1967). All four species have historically been common transients and winter residents from about mid-September to early May throughout Oklahoma (Sutton 1967; Reinking 2004). In the last few decades the populations of resident and nesting Canada geese within the United States have grown 7.9% per year (Cleary et al. 2006). Canada geese and mallards have become common permanent residents of Oklahoma and now regularly breed throughout the state (Reinking 2004; Oklahoma Bird 2014; Reinking 2017).

Historically, mallards were widespread and abundant, inhabiting reservoirs, marshes, and streams, and foraging in hay and grain fields (Sutton 1967; Baumgartner & Baumgartner 1992; Reinking 2004). Gadwalls are widespread though less common, foraging in grain fields and inhabiting ponds and the backwaters of reservoirs and rivers (Baumgartner & Baumgartner 1992). The cackling goose breeds near water on tundra (i.e. Canada, the Aleutian Islands, Alaska, etc.) and winters on inland lakes and marshes from British Columbia to California and east to western Louisiana and can be found throughout Oklahoma in winter (Banks et al. 2004; Sullivan et al. 2009).

## **Study Area**

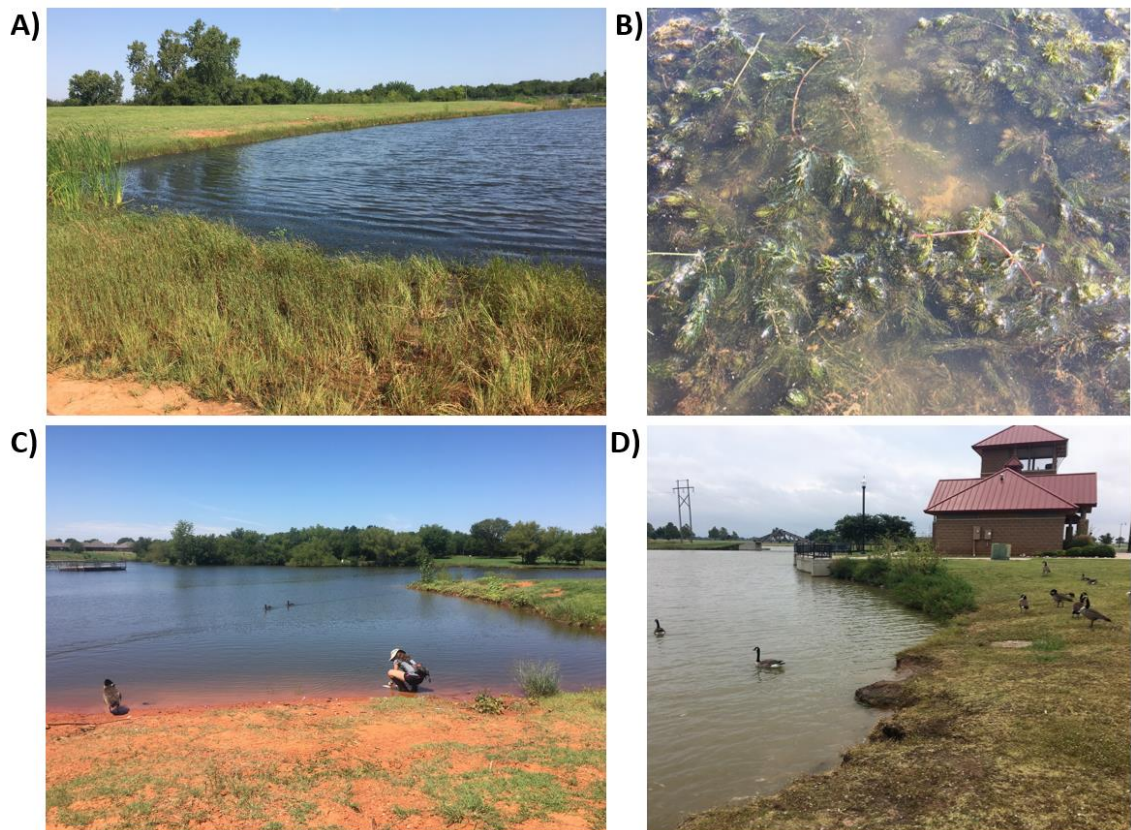
### *Landscape of Oklahoma*

Oklahoma is an ecologically diverse state consisting of flat to rolling plains, a range of soil types, six major rivers flowing from northwest to southeast, more than 200 man-made lakes, and four distinct mountainous areas (Ozarks, Ouachitas, Arbuckles, and Wichitas; Baumgartner & Baumgartner 1992; Arndt 2003). Based on a natural vegetation map from Duck and Fletcher (1943), Cleveland and surrounding counties (where environmental sampling took place) are composed of tallgrass prairie, postoak-blackjack oak forest (or cross-timbers), and bottomland forest (Hoagland 2004). Tallgrass prairie is the most extensive grassland type in Oklahoma and is composed primarily of big bluestem, little bluestem, Indian grass, and switchgrass; however, much of the state's grasslands have been hayed, grazed, or converted (Hoagland 2004). The cross-timbers, the dominant habitat type in central Oklahoma, are a mixture of forest, woodland, tallgrass, and mixed-grass prairie vegetation, with post oak and blackjack oak species contributing up to 90% of the canopy cover and 50% of the basal area in cross-timber forests (Hoagland 2004). Bottomland forests occur along major rivers throughout Oklahoma, with vegetation associations varying widely from east to west (Hoagland 2004).

### *Ponds of Central Oklahoma*

I sampled a total of nine ponds - eight in Norman, Oklahoma, and one near Lake Overholser in Oklahoma City. Two of the ponds sampled (Summit Lake and Summit 2) in Norman were residential, catch-and-release fishing ponds (Summit Lakes 2018), another was on an apartment complex-owned golf course (The Links), one pond was on

the University of Oklahoma campus (Brandt Park), three were neighborhood ponds, and one pond (Route 66) was located at a public park next to Lake Overholser, a reservoir formed by a dam on the North Canadian River (or Oklahoma River) for use as a municipal water source. Most of the ponds lacked information about their use and source of water. All ponds were in developed areas, and at least part of the landscape surrounding all ponds included mowed grass and some trees (Figure 1; Table 1).



**Figure 1. Photographs of three ponds sampled for Canada goose eDNA. A) Sampling location at the Norman Community Dog Park pond; B) Coontail (*Ceratophyllum demersum*) was abundant at the pond at the Norman Community Dog Park; C) Northeast Lions Park in Norman, OK; D) Route 66 Park in Oklahoma City, OK.**

## Field Sampling

Nine ponds were sampled for Canada goose eDNA between June and October 2017. I observed Canada geese at six of the ponds during sampling, and all species except the gadwall were observed at a minimum of one sampling location. I performed all environmental sampling myself to standardize methods. I wore latex gloves during sampling to prevent contamination and collected samples using VWR sterile round wire sample bags (catalog number 82007-726), which remained sealed (i.e. sterile) until opening at each respective sampling point. For each pond I submerged a sterile bag filled with approximately 500mL distilled water under the pond surface near the shore for 30 seconds to serve as a negative control, which would indicate whether contamination occurred during transportation (Bohmann et al. 2014). I took eight environmental samples along the shore of each pond in a 4x2 fashion, which consisted of four points, 5 meters apart, in a line transect along the shore, with two samples per point at 1 and 2 meters from the edge of the shore (see Figure 2). At two sites (Links and Summit Lake) the water was too deep, or the substrate was too unstable, to sample safely two meters from the shore without a kayak. In these cases, I used an 8x1 sampling regime instead, sampling at 8 points 5 meters apart and 1 meter in from the shore.

I took a variety of measurements at each pond site and sampling point. I used a Garmin GPSMAP 64st to record GPS coordinates and elevation of each water body and, later, at each sampling point (n=8 per pond) prior to water collection. I described relative pond size (small, medium, large) in the field based on personal observations and later calculated pond area (m<sup>2</sup>) in Google Earth Pro. I conducted a 5-minute direct

counting survey of the target species (*B. canadensis*) prior to sampling, recording the number of birds, their general location (i.e. in water, on shore), and their proximity to the initial GPS point (i.e. <25m or >25 meters), and documented presence of outgroup species. At each sampling point, prior to water collection, I measured water temperature (°C), pH, conductivity (µS), and total dissolved solids (TDS, in ppt) using an ExTech probe. The ExTech probe did not tare properly for some measurements of TDS and conductivity, so these variables were omitted from analyses. A wind flow meter was inserted just below the water surface to approximate water flow (m/s). Samples were taken by dragging the bag through the water, about 5cm below the water surface. I walked along the dry shore between sampling points (1 – 4) to avoid disturbing the water and sediment. After sampling was completed, I recorded secchi depth (cm) of the water body as an indicator of turbidity. At one pond I collected a Canada goose fecal sample and mixed it with distilled water to create a fecal positive control. I stored all water samples in a cooler on ice in the field. When possible, I transported samples to the lab immediately after sampling for filtration; however, to prevent DNA degradation when immediate filtration was not possible, I stored samples in a refrigerator for up to 25 hours or in a freezer until filtration was completed. Some eDNA degradation could have occurred in samples that were not filtered immediately (Thomsen et al. 2012; Stoeckle et al. 2017; Tsuji et al. 2017); however, correlation analyses revealed no significant effect of storage time prior to filtration on eDNA detection rates.



**Table 1. Ponds sampled for Canada goose eDNA.**

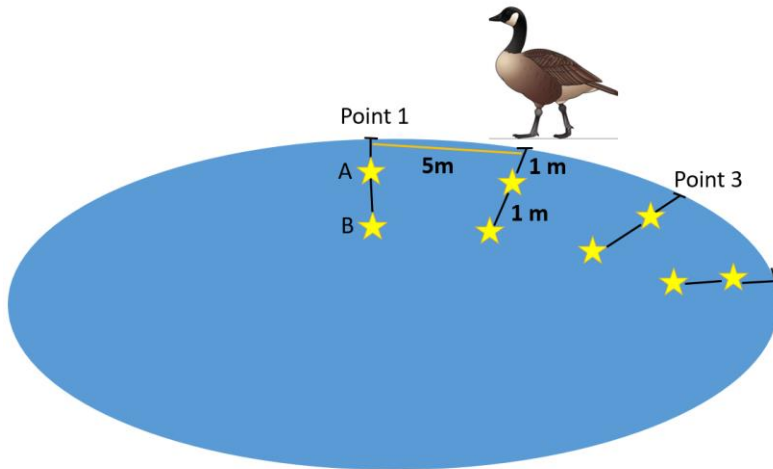
Pond	City	Latitude (N)	Longitude (W)	Pond Size (m <sup>2</sup> )	Geese	Description
Dog Park	Norman	35.23644	97.4277	17,332	Absent	Public park, mowed grass, some exposed soil, aquatic plants
Hallbrook	Norman	35.24511	97.41203	8,873	Absent	Residential reservoir, mowed grass, clay-like pond bottom
Hall Park Lake	Norman	35.23353	97.41082	28,109	Absent	Residential, mowed grass, cat tails, watermilfoil
The Links	Norman	35.16976	97.40065	27,324	Present	Golf course, soft clay-like sediment
Summit Lake	Norman	35.21664	97.39973	62,031	Present	Residential, stocked with fish, mowed grass, patches of trees
Summit 2	Norman	35.32436	97.39882	10,294	Present	*Same as Summit Lake
NE Lions Park	Norman	35.24159	97.41679	22,541	Present	Residential, short grass, some soil exposed, some trees
Brandt Park	Norman	35.20455	97.43599	11,231	Present	College campus, mowed grass, trees surrounding
Route 66 Park	Oklahoma City	35.49664	97.69112	38,259	Present	Public park pond near Lake Overholser, mowed grass, few trees

## **Water Filtration**

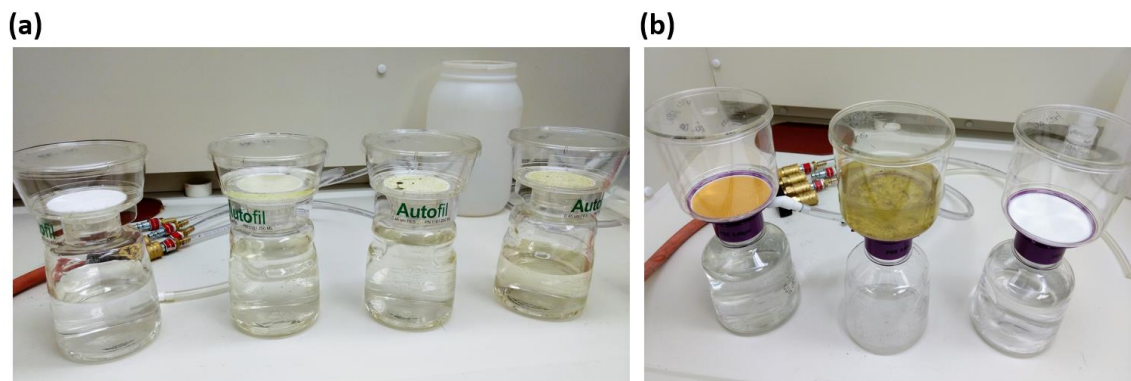
I performed water filtration in a fume hood in the pre-PCR room at the Sam Noble Museum. Most samples were filtered using 250mL Autofil 0.45 $\mu$ m High Flow PES top filters with a 66mm filter diameter, though some samples were filtered using 500mL VWR 0.45 $\mu$ m 75mm PES filters, which filtered higher water volumes more rapidly than the Autofil filters (Figure 3). Negative control samples were often filtered first. I filtered a maximum of four samples at once using a four-way air splitter connected to the hood vacuum system to increase filtration speed. When possible, the entire sample volume (500 – 1000mL) was filtered, but for many samples the filter became clogged with pond sediment and biological material, preventing complete filtration. In these cases, once filtration stopped, I discarded any remaining unfiltered water. I cut each filter out of its plastic cup using sterilized forceps (soaked in Eliminate and rinsed with doubly distilled H<sub>2</sub>O) and a disposable sterile #11 blade scalpel. I preserved each filter in its own cryotube filled with 95% ethanol at -20°C until screening.

## **eDNA Filter Extraction Protocol**

Day 1: Scissors and forceps were rinsed, sterilized with DNA Eliminate (item number DE32330102), and rinsed twice more with RODI H<sub>2</sub>O in autoclaved beakers to prevent contamination between filters. Filter pipette tips were used to prevent contamination. I created a negative lab control by removing the filter paper from a sterile filter top, rinsing it in RODI H<sub>2</sub>O in an autoclaved beaker, and finally soaking it in EtOH until further use. About one-quarter of each filter was used for eDNA extraction, and the remainder was preserved at -20°C. I cut each quarter into halves and



**Figure 2.** The 4x2 sampling regime used to collect a total of eight water samples at each pond. Yellow stars indicate sampling location at each pond. Locations within ponds were based on accessibility and were usually near the target species where possible. The distance between each point was 5m along the shore. “A” samples were taken 1m from the shore, and “B” samples were taken 2 meters from the shore. (Goose image from <https://www.vectorstock.com/royalty-free-vector/canada-goose-vector-479739>).



**Figure 3.** A) Setup of four Autofil 0.45µm High Flow PES (polyethersulfone membrane) top filters connected to the fume hood vacuum system via a four-way air splitter. The left-most filter is a negative control. Sediment and biological material collected on the filters, and filtered water was discarded. Chambers were sterilized with Eliminate and reused. B) 500mL VWR 0.45µm 75mm PES filters were used for some samples. A sample with high sediment load, the fecal positive control, and a negative control are pictured.

then into smaller pieces and placed them in two separate, sterilized microcentrifuge tubes. A SpeedVac was used to dry the tubes on low heat for about 20 minutes. I added

400µL Buffer ATL and 20µL Proteinase K (Qiagen DNeasy kit) to each microcentrifuge tube using filter pipette tips and then incubated the tubes in an Eppendorf ThermoMixer C at 500RPM and 55°C overnight.

Day 2: I transferred all contents from the Day 1 microcentrifuge tubes into Qiashreder spin columns and collected the filtrate. I added 800µL Buffer AL to each filtrate sample and incubated the samples for 10 minutes at 70°C. 800µL of 200 proof ethanol was added to the tubes and the mixture was then added to a Qiamsp Spin column (DNeasy kit). The filtrate was discarded, and the spin columns were placed into new collection tubes. 500µL of Buffer AW1 was added to each spin column, followed by 500µL Buffer AW2. In a new collection tube, 50µL Buffer AE was added to the spin column to elute the DNA. The columns were centrifuged after incubating at room temperature for 5 minutes. 50µL Buffer AE was added again to each spin column; each tube sat at room temperature for 2 hours before centrifuging. Following centrifugation, eluted DNA was transferred to Zymo Inhibition kit spin columns and centrifuged once again. The final eDNA samples were then stored at -20°C until ready for Canada goose eDNA screening using quantitative, real-time PCR.

### **Tissue Extraction Protocol**

A total of thirty-three tissue aliquots from the target and non-target species were obtained from four museums for use in primer-probe assay design and testing (Appendix I). Six tissue aliquots from Oklahoma specimens (4, *B. canadensis*; 1, *A. platyrhynchos*; 1, *A. strepera*) loaned from the Sam Noble Oklahoma Museum of Natural History represented the local population. Additional tissues were obtained from the Denver Museum of Nature and Science (n = 11), the Museum of Vertebrate

Zoology at the University of California Berkeley (n = 4), and the Field Museum of Nature and Science (n = 12) to account for nationwide genetic diversity during assay design and to meet the recommended sample sizes suggest by Pilliod et al. (2013).

I used Fujita's modified guanidine thiocyanate extraction method (Esselstyn et al. 2008) to extract genomic DNA from all 33 tissue aliquots. This protocol consisted of three phases: cell lysis, DNA precipitation, and DNA resuspension. I performed cell lysis using a solution composed of 410 $\mu$ L extraction buffer (1M Tris, 5M NaCl, 0.5M EDTA, and sterile ddH<sub>2</sub>O), 80 $\mu$ L 10% SDS, and 10 $\mu$ L Proteinase K (10mg/ml). Samples were digested overnight in an Eppendorf ThermoMixer C at 55°C and 500RPM. The following day I removed the samples from the incubator and centrifuged them to remove excess protein. I then poured the supernatant into microcentrifuge tubes containing 200 $\mu$ L 5M NaCl. I inverted the tubes 50 times and centrifuged them to remove salts. I poured the supernatant into microcentrifuge tubes containing 500 $\mu$ L cold isopropanol and cooled them at -20°C for about 10 minutes. I centrifuged the samples once again to precipitate DNA, forming DNA pellets. I washed the DNA pellets in three rounds of 500 $\mu$ L cold 80% EtOH, centrifuging and discarding supernatant in between. I air-dried the tubes overnight at room temperature until evaporation was complete. The following day I resuspended the dried DNA in 100 $\mu$ L 0.25xTE Buffer and stored them at -20°C until further use in PCR and serial dilution trials.

### **Primer-Probe Assay Design**

A key component of this study was the design of a species-specific primer-probe assay with which to amplify Canada goose eDNA extracted from water samples. To ensure that I designed an assay specific to Canada geese, I first used relatively general

primers, or short sequences of DNA that attach to a complementary strand of DNA to begin replication, amplify target and non-target species' DNA for eventual sequencing, alignment, and primer-probe assay design.

I designed thirty-five primer pair combinations for three mitochondrial genes—cytochrome c oxidase I (COI), cytochrome b (Cyt b), NADH dehydrogenase subunit 2 (ND2)—using primers published in the literature (Appendix II; Johnson & Sorenson 1998; Donne-Gousse et al. 2002; Paxinos et al. 2002; McCracken & Sorenson 2005; Kerr et al. 2007) and tested them on seven tissue DNA extractions (four Canada goose and one each of mallard, gadwall, and cackling goose) using polymerase chain reaction (PCR) to determine which gene and primer pairs best amplified all four species' DNA (PCR products and eventually sequences of all four species would be used to design a primer-probe specific to Canada geese). Samples were prepared in 10 $\mu$ L volumes consisting of 1 $\mu$ L DNA extraction, 6.5 $\mu$ L doubly distilled H<sub>2</sub>O, 1 $\mu$ L 10xPCR Buffer, 0.8 $\mu$ L dNTP mixture (2.5mM), 0.05 $\mu$ L Takara Taq HS, 0.4 $\mu$ L forward primer, and 0.4 $\mu$ L reverse primer. PCR was performed in a Bio-Rad C1000 touch Thermal Cycler at the following conditions: 2.5 min at 98°C, then 35 cycles of 10s at 98°C, 30s at 55°C, and 60s at 55°C, and followed by 5min at 72°C and infinite hold at 12°C.

I used gel electrophoresis to visualize the PCR products. I created a 1% agarose gel submerged in 1xTBE buffer to wet-load 2 $\mu$ L of each PCR product and a DNA ladder (100mg/mL 1kb plus from Thermo Fisher) stained with GelRed (Biotium). I ran the gel for 50 minutes at 110mV using a Labnet Enduro Power Supply Mini and performed imaging in a Bio-Rad Molecular Imager Gel Doc XR+ with Image Lab Software 5.1. Following imaging, I selected two primer pairs for each gene (COI, Cyt b,

ND2) to optimize band strength, fragment length, and amplification across all species and localities. I tested the resulting six primer pairs on fifteen DNA extractions (eight Canada goose, three cackling goose, two mallards, and two gadwall) from different localities to confirm their ability to amplify DNA of all four study species. I visualized the PCR products using the same gel electrophoresis protocol as before.

I selected one primer pair for each gene (Table 2), maximizing the number of samples successfully amplified, especially for the target species. I submitted the PCR products (18 $\mu$ L) from each of these three primer pairs to the University of Oklahoma Biology Core Molecular Laboratory for PCR cleanup, Sanger sequencing using ABI Prism BigDye Terminator, version 3.1 (Applied Biosystems, Foster City, CA), and sequencing cleanup. Novel sequence data were deposited in GenBank (Table 3).

I trimmed and aligned DNA sequences for all three genes in Geneious 9.0.5. Only one cackling goose amplified well using Cyt b, while multiple cackling goose sequences amplified successfully for COI and ND2. Because it was important to incorporate sufficient outgroup species diversity (especially of the cackling goose) to ensure a robust assay design, I proceeded to design primer-probe assays for COI and ND2 only (Table 4). I tried to maximize base pair mismatches between the target and non-target species in both the primer- and probe-binding regions to reduce or prevent amplification and fluorescence of non-target species during quantitative PCR reactions (qPCR; Wilcox et al. 2013). One probe and two primers (forward and reverse) were designed for each gene in Primer Express 3.0.1 using Taqman MGB Allelic Discrimination. The ND2 probe and primers targeted an 82-base pair region of the Canada goose ND2 gene, while the COI probe and primers targeted a 71-base pair

region of the Canada goose mitochondrial COI gene (Table 4). Primers were ordered from Thermo Fisher.

**Table 1. Three primer pairs (one each for COI, Cyt b, and ND2) were selected for optimizing DNA amplification in PCR, visualized with gel electrophoresis. The forward primer is listed above the reverse primer for each gene.**

<b>Gene</b>	<b>Primer</b>	<b>Primer Sequence (5'–3')</b>
COI	FalcoFA	TCAACAAACCACAAAGACATCGGCAC
	BirdR2	ACTACATGTGAGATGATTCCGAATCCAG
Cyt b	L15191	ATCTGCATCTACCTACACATCGG
	H16064	CTTCGATTTTTGGTTTACAAGACC
ND2	L5216	GGCCCATACCCCGRAAATG
	H6313	CTCTTATTTAAGGCTTTGAAGGC

### **Primer-Probe Validation**

I performed single-tube DNA quantitation (ng/μL) on all tissue DNA extractions using the QuantiFluor dsDNA System in a Promega Quantus Fluorometer (product number E6150). I prepared each tissue extraction sample, a blank sample, and a 200ng DNA standard in 0.5mL thin-wall PCR tubes containing 200μL QuantiFluor dsDNA dye working solution (1:400 dilution in 1xTE buffer).

I ran a serial dilutions experiment on fifteen tissue extractions that were not used for assay design to test the specificity and sensitivity of the COI and ND2 primer-probes. I used the DNA quantitation data to create 50μL solutions (20ng/μL) of each of the fifteen tissue extractions in 0.25xTE Buffer. For each sample I created 1:10, 1:100, 1:1,000, and 1:10,000 dilutions in series by combining 2μL of the 20ng/μL solution (or subsequent dilution) with 18μL 0.25xTE Buffer. I resuspended the COI and ND2 primers in Buffer AE and tested them on the serial dilutions of Canada goose (1:10,000 dilution) and nontarget species DNA (1:1000 dilution) and a negative control in



**Table 2. Museum tissues were sequenced for mitochondrial genes COI, ND2, and/or Cyt b and can be found on Genbank under the accession numbers listed.**

<b>Species</b>	<b>Tissue Specimen</b>	<b>Locality</b>	<b>COI</b>	<b>ND2</b>	<b>Cytb</b>
A. platyrhynchos	OCCR 8821/OMNH 22601	Oklahoma	MH676083	MH383306	MH676102
A. s. strepera	FMNH 492427	Wisconsin	-	MH383307	MH676103
B. c. maxima	FMNH 438243	Illinois	MH676090	MH383303	-
B. canadensis	FMNH 449116	Minnesota	MH676091	MH383302	MH676100
B. canadensis	FMNH 488508	Illinois	MH676084	MH383299	MH676101
B. canadensis	MVZ 184923	California	MH676087	-	-
B. canadensis	OCCR 10645/OMNH 22918	Oklahoma	MH676088	MH383300	MH676097
B. canadensis	OCCR 11246/OMNH 23166	Oklahoma	MH676086	MH383301	MH676099
B. canadensis	DMNS 35091	Colorado	MH676085	-	MH676096
B. canadensis	DMNS 46079	Colorado	MH676089	-	MH676098
B. hutchinsii	FMNH 480459	Illinois	MH676092	MH383305	-
B. hutchinsii	FMNH 486719	Wisconsin	MH676094	-	-
B. hutchinsii	DMNS 34864	Colorado	MH676093	MH383304	MH676095

**Table 3. Primers and probes (5'–3') were designed to target the mitochondrial genes ND2 and COI for Canada goose (*Branta canadensis*). Melting temperature (T<sub>m</sub>) of each is included. Both assays were tested in qPCR against serial dilutions of target and non-target species DNA extracted from tissue samples. Though both amplified target species DNA, ND2 was more consistent and was therefore used for eDNA screening.**

<b>Gene</b>		<b>Sequence</b>	<b>T<sub>m</sub></b>
<b>ND2</b>	Forward Primer	CCGCCCTGGTCCTATTCTC	58
	Reverse Primer	GAGGTTGGGTGGTTTATTTGTGTAA	59
	Probe	CATAACTAACGCCTGAG	67
<b>COI</b>	Forward Primer	CCGCGCAGAACTAGGACAAC	59
	Reverse Primer	GGGCGGTGACGATTACATTG	60
	Probe	CTCTCCTAGGCGACG	66

triplicate using quantitative, real-time PCR (qPCR). These dilutions were used to simulate relatively low DNA concentrations observable in pond samples, as well as more concentrated non-target species DNA than target DNA. qPCR was preferred over PCR because it amplifies a target region of DNA specific to a target organism in real-time and uses both primers and a probe to provide additional specificity and potentially quantitative data. I prepared PCR plates in a separate room from qPCR equipment to avoid contamination. Each tissue extraction dilution was screened in triplicate using 3µL of DNA sample (or doubly distilled H<sub>2</sub>O for the negative control) and 7µL of reagent consisting of 0.75µL ddH<sub>2</sub>O, 5µL MasterMix, 0.5µL of each primer (10µM), 0.25µL probe (10µM). An Applied Biosystems QuantStudio 3 Real-Time PCR System and QuantStudio Design & Analysis Software v1.4 were used for screening and analysis. The qPCR conditions were as follows: 20 seconds at 95°C, 50 cycles of 1 second at 95°C and 20 seconds at 60°C, and a post-read stage of 30 seconds at 60°C. A sample is considered positive for target DNA if its normalized reporter intensity (R<sub>n</sub>, the ratio of the fluorescence emission intensity of the reporter dye divided by that of the

passive reference dye) is above the fluorescence intensity threshold value, which is calculated by the software based on the negative controls. A high  $R_n$  value indicates stronger fluorescence. With a threshold value of 3.230, all screening reactions containing non-target species DNA for both gene assays resulted in no positive detection. For the ND2 assay, all Canada goose samples were positive for Canada goose DNA, and all non-target species and the negative controls were negative for Canada goose DNA (Appendix IV). For the COI assay, all but two of twenty-one Canada goose replicates confirmed the presence of target DNA. I found that these two negative replicates had  $\Delta R_n$  values just below threshold (3.188 and 3.244) but considerably higher than any non-target species (mean  $\Delta R_n=1.284$  for non-target samples; the next-highest  $\Delta R_n$  was an outlier of 2.980 for one gadwall replicate). Because of the potential for Type II error when using the COI assay, and because the ND2 assay more reliably called presence/absence without needing to adjust the threshold, I continued assay testing for ND2 only.

Using the same qPCR protocol as before, I tested the sensitivity and specificity of the ND2 assay on fifteen different mixtures of target and non-target species DNA, each at 1:10,000 dilution, in an attempt to mimic environmental samples, which can contain DNA fragments from thousands of specimens of hundreds of species in a single sample (Shokralla et al. 2012). I created various DNA combination samples in which target DNA was at a lower relative concentration than, and mixed with, outgroup species' DNA. Amplification occurred in all samples containing Canada goose DNA (100% detection rate). All samples lacking target DNA, including the negative control, had negative detection, confirming assay specificity (Appendix V). Interestingly,

Welch's t-test revealed that the mean  $\Delta R_n$  for combination samples containing Canada goose DNA but lacking cackling goose was significantly higher ( $=3.65$ ) than samples with both Canada and cackling goose DNA present (mean= $2.69$ ;  $t=6.45$ ,  $df=6.74$ ,  $p=0.00041$ ). In other words, when cackling goose DNA was present along with Canada goose DNA,  $\Delta R_n$  was significantly lower, though still above the threshold value.

### **eDNA Screening**

I screened 81 eDNA extractions ( $n=8$  per pond + 1 negative field control per pond), one negative lab control (doubly distilled  $H_2O$ ) from each eDNA extraction event ( $n=8$ ), and one positive control fecal sample for Canada goose eDNA using the ND2 assay and following the same qPCR protocol as before. If at least two out of three replications for a pond sample yielded a positive result, I considered the sample positive for the presence of Canada goose eDNA. Samples in which only one out of three replicates were positive were re-screened; if at least one replicate in the second run yielded a positive result, Canada goose eDNA was considered present.

### **Statistical Analysis**

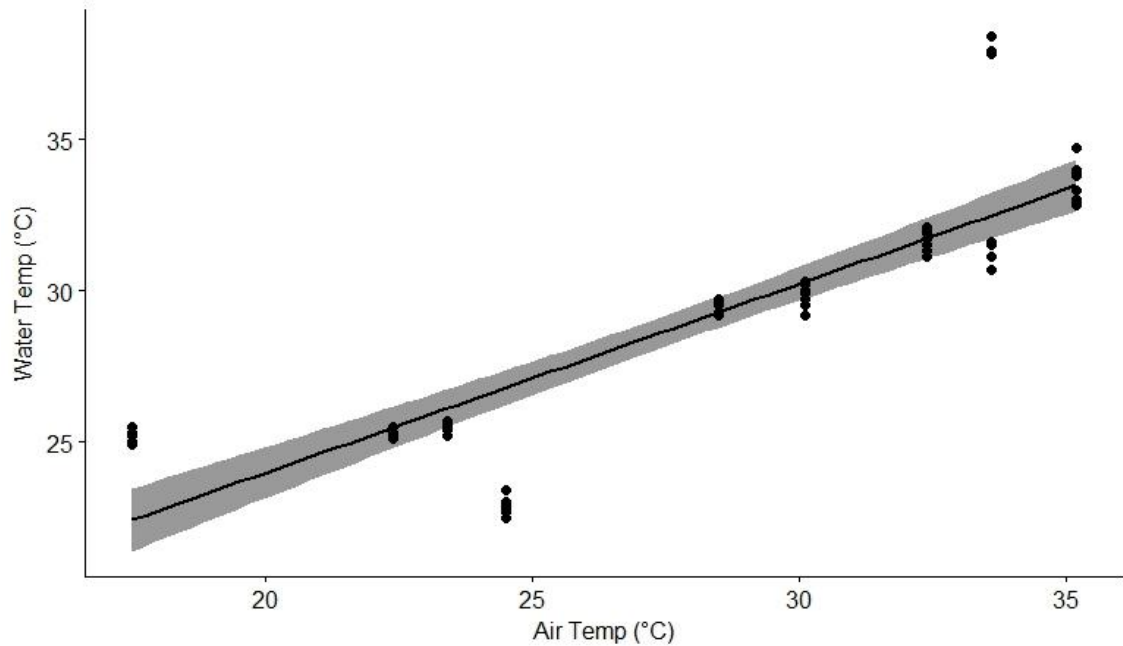
All statistical analyses were performed in R version 3.3.1. I used the `ggplot2` (Wickham 2009) and `ggpubr` (Kassambara 2017) packages to create scatterplots. Spearman's rank correlation was used to determine the strength of the relationships between environmental variables and eDNA presence or detection. I ran classification and regression tree (CART) analyses (`rpart` package; Therneau & Atkinson 2018) to measure the effects of environmental variables (goose abundance, pond area [ $m^2$ ], goose density, water pH and temperature [ $^{\circ}C$ ] of each sample, secchi depth of the pond [cm], and sampling distance [m] from the shore) on the presence of Canada goose

eDNA. CART is designed to explore and model ecological data that is complex, unbalanced, and may contain missing values and can be used for both description and prediction (De'ath & Fabricius 2000). Classification and regression trees are represented graphically and explain variation in a dependent variable by one or more explanatory variables (numeric or categorical) by splitting the data into two mutually exclusive, relatively homogenous groups repeatedly (De'ath & Fabricius 2000). In classification tree analyses, eDNA presence/absence was used as the response variable. In regression tree analyses, eDNA detection rate (% of positive qPCR replicates) was used as the response variable.

Because sampling date and time varied widely (between June to October and between 7am and 4pm), and because water temperature was highly correlated with air temperature (Figure 4), I regressed water temperature against air temperature to create unstandardized residuals (“residual temperature”; Figure 4). I used these residual temperatures in further analyses to investigate the effects of temperature while controlling for sampling bias.

### **Ethics Statement**

No permits were required for sampling at any of the pond sites, although permission was granted to sample at a few privately-owned lakes. Field sampling did not involve any endangered or protected species, and no live vertebrate animals were used in this study. Lab and field protocols (protocol R17-029) were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Oklahoma. No live animals were used or harmed in this study.



**Figure 4. Air temperature (°C) at time of sampling was highly correlated with water temperature (°C) of samples ( $r=0.87$ ;  $y=0.62x+11.49$ ). Because water temperature of each sample varied with sampling date (June to October) and time (early morning to late afternoon), residual temperature values (distance of each point from the regression line) were used in analyses.**

## Results

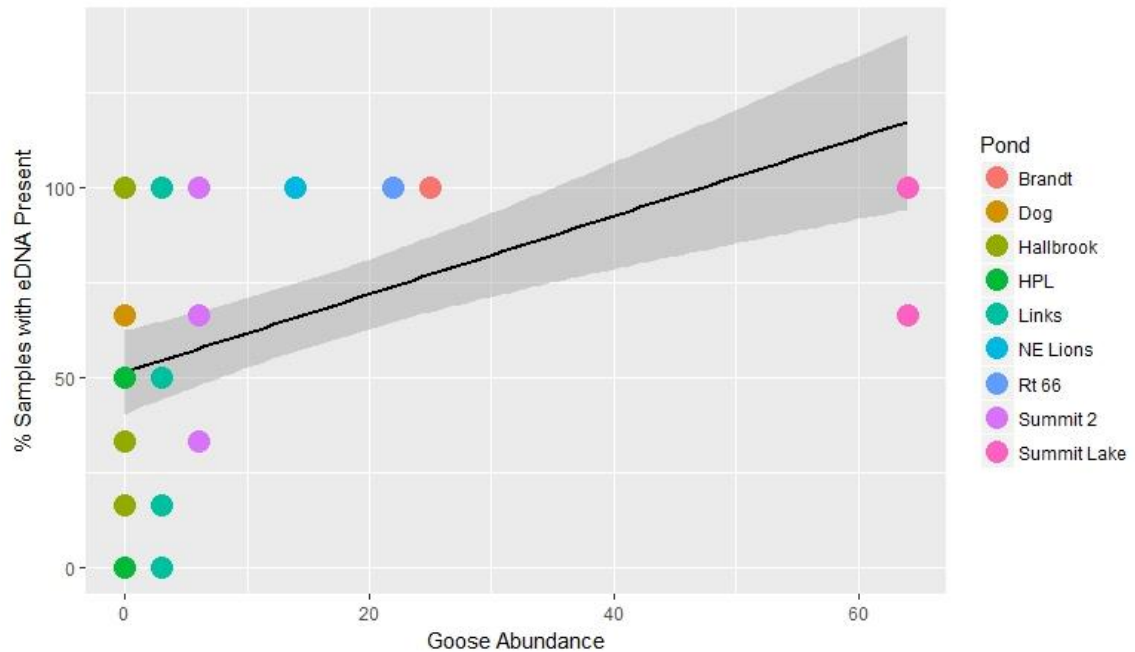
Canada goose eDNA was detected at all nine ponds sampled, including the three locations where Canada geese were absent (Table 5). Both the percentage of pond samples and the percentage of qPCR replicates (each pond sample was run in triplicate) with eDNA present for each pond were analyzed. All field and lab negative controls were negative for Canada goose eDNA, with one exception each due to contamination during one round of filter extractions. Any samples extracted simultaneously with these negative controls (except for those which still resulted in total negative detection and indicated no contamination) were excluded from analyses.

In general, the percentage of both pond and qPCR samples with eDNA present increased with increasing goose abundance ( $r=0.5527$  and  $r=0.7097$ , respectively; Table 5, Figure 5). In addition, goose abundance was positively correlated with pond area ( $r=0.4267$ , Figure 6), though pond area did not have a significant effect on eDNA detection rates. Initially, the water temperature ( $^{\circ}\text{C}$ ) of each pond sample was positively correlated with both eDNA detection rate (% samples with eDNA present,  $r=0.4060$ ) and goose abundance ( $r=0.5626$ ); however, these correlations were insignificant (spearman's rank correlation) when residual temperature was used.

Both the classification and regression trees agree that goose abundance accounted for the greatest variation in eDNA detection, occupying the first split in each tree (Figure 7), and was the best predictor for eDNA presence or absence with the highest improvement scores of all variables considered (Tables 6, 7). The classification tree ( $R^2=0.4$ , Figure 7a) indicates that temperature explained the detection data best when goose abundance was low (improvement score=2.41), and eDNA detection was

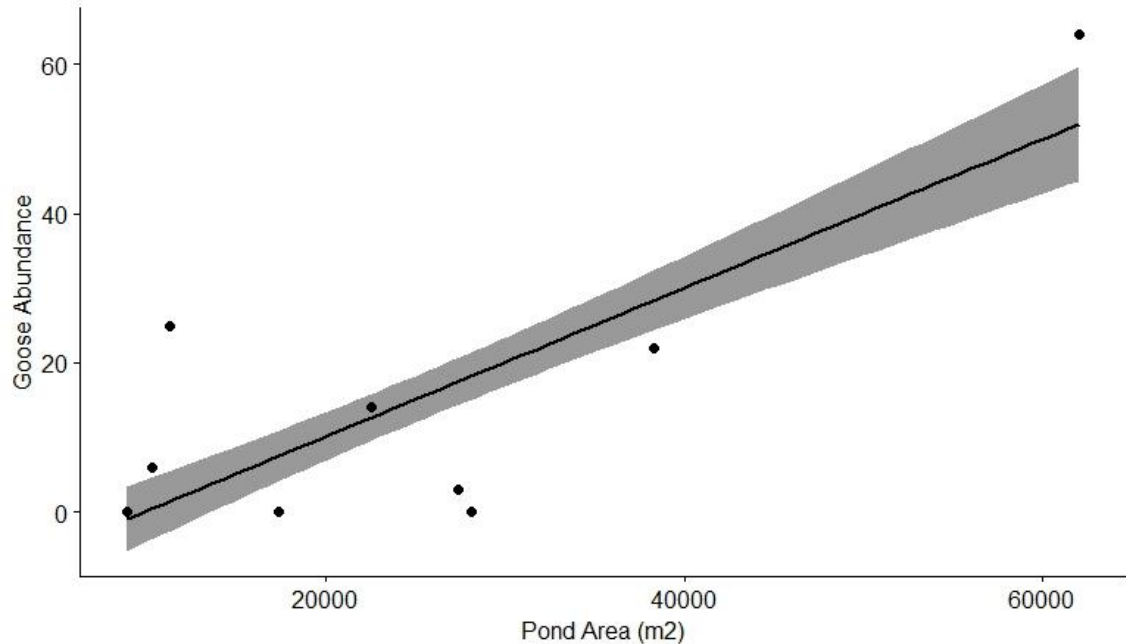
**Table 4. Canada goose (*Branta canadensis*) eDNA detection results of pond water samples taken from nine ponds around Norman, Oklahoma and Oklahoma City from June to October 2017. eDNA was confirmed present at a pond if at least one field sample had two or more positive qPCR replicates.**

Locality	# Positive		%		eDNA Present	# Geese Observed
	Pond Samples	# Total Samples	Samples Positive	% qPCR Positive		
Dog Park	4	8	50	33.33	Y	0
Hallbrook	4	8	50	33.33	Y	0
Hall Park Lake	2	6	33.33	22.22	Y	0
The Links	5	8	62.5	41.03	Y	3
Summit 2	7	7	100	77.78	Y	6
NE Lions Park	7	7	100	100	Y	14
Route 66 Park	5	5	100	100	Y	22
Brandt Park	7	7	100	100	Y	25
Summit Lake	8	8	100	95.83	Y	64



**Figure 5. Goose abundance (# individuals present during sampling) was positively correlated with eDNA detection rate (% qPCR replicates with eDNA present from 64 screened pond samples) using Spearman’s rank correlation ( $r=0.71$ ).**





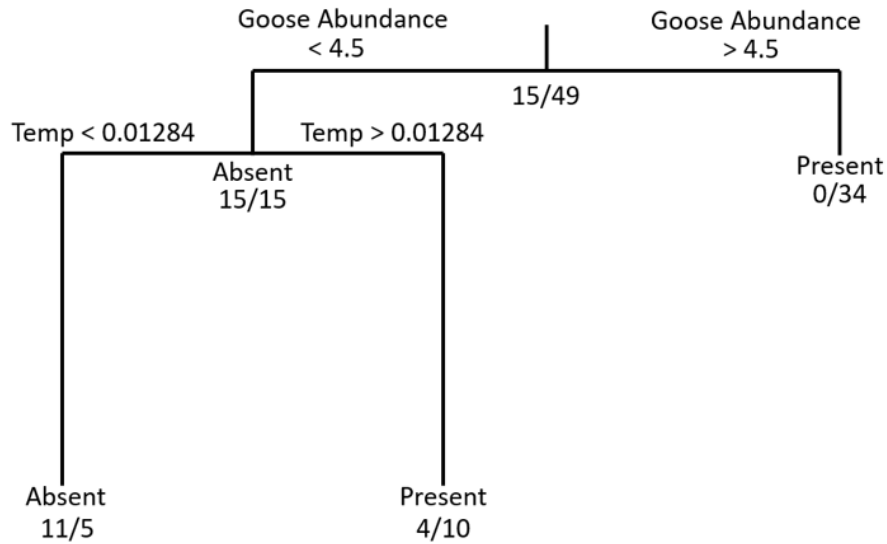
**Figure 6. Pond area (m<sup>2</sup>) was positively correlated with goose abundance (number of individuals present during sampling) at each of nine ponds in central Oklahoma according to Spearman’s rank correlation ( $r=0.43$ ).**

predominantly positive when residual temperature was higher than expected ( $>0.013$  °C; Table 6). According to this model, when goose abundance was low ( $<4.5$  individuals), ponds with higher-than-normal water temperatures had higher eDNA detection rates than cooler ponds. However, the regression tree indicates that no variables explained the data well when goose abundance was low, resulting in a less complex tree with only one split (Figure 7b). Overall, the regression tree, which used continuous rather than categorical eDNA detection data, had an improved fit ( $R^2=0.576$ , Figure 7b) over the classification tree, indicating that eDNA detection is best explained solely by goose abundance, and when more than 4.5 geese are present, eDNA detection is assured.

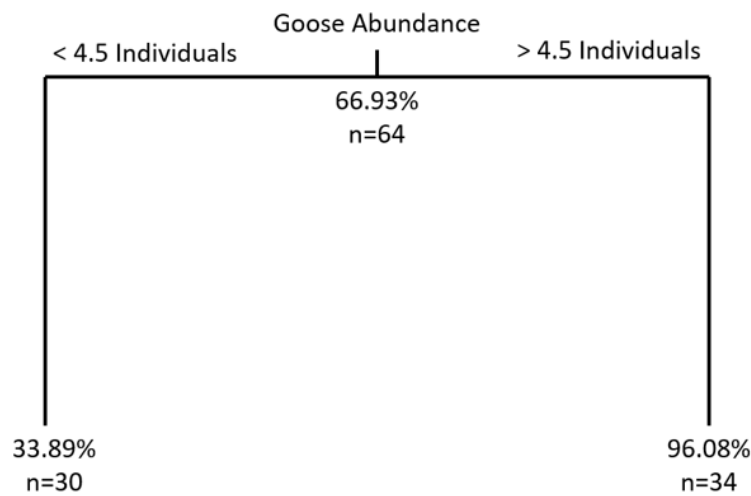
The correlation between eDNA detection and goose density ( $\# \text{ geese/m}^2 \cdot 1000$ ;  $r=0.704$ ,  $p=8.507e^{-11}$ ) was slightly less significant than the correlation between

detection and goose abundance ( $r=0.710$ ,  $p=5.209e^{-11}$ ), and the cross-validation error of the regression tree was slightly higher when using goose density ( $xerror=0.458$ ) compared to goose abundance ( $xerror=0.448$ ). In addition, in the regression tree analysis, goose density had a variable importance of 48, whereas when goose abundance and pond area ( $m^2$ ) were analyzed separately, goose abundance had an importance value of 42 and area dropped to 13. Because goose abundance was slightly more correlated with eDNA detection, resulted in a regression tree with slightly higher fit, and had a high variable importance score, analyses using goose abundance instead of density were used.

A)



B)



**Figure 7. CART analyses of the detection of *Branta canadensis* eDNA in 64 water samples taken from nine ponds in central Oklahoma. Explanatory variables included goose abundance (number of individuals present during sampling), pond area (m<sup>2</sup>), pH, residual water temperature (°C), secchi depth (cm) of the water body, and sampling distance (m) from the pond's shore. Terminal nodes indicate whether eDNA was absent or present among samples as well as the ratio of negative to positive samples (0/1). a) Tree based on binary presence/absence data ( $R^2=0.4$ ,  $xerror=0.67$ ,  $xstd=0.19$ , misclassification rate =  $15/64 = 23.44\%$ ). b) Tree based on % detection of eDNA ( $R^2=0.58$ ,  $xerror=0.46$ ,  $xstd=0.10$ ).**

**Table 5. Explanatory variables considered for classification tree splitting. The primary split for a particular node includes the improvement score for that split. Each node was also analyzed for surrogate splits. Explanatory variables are ranked according to variable importance (sum=100) in splitting or predicting the dependent variable (eDNA presence/absence).**

Variables	Node 1 (N=64)		Node 2 (N=30)		Importance
	Primary Splits	Surrogate Splits	Primary Splits	Surrogate Splits	
	Improvement	Agree	Improvement	Agree	
Geese	7.9688	-	0.3409	0.800	44
pH	3.9142	0.688	1.1111	0.733	17
Area	1.7923	0.672	-	0.800	18
Res. Temp	1.4593	0.641	2.4107	-	20
Secchi	1.4083	-	-	-	-
Distance	-	-	0.0667	0.567	1

**Table 6. Explanatory variables considered for regression tree splitting. This tree includes only one node and split. The primary split includes the improvement score of each variable for that split. The node was also analyzed for surrogate splits. Explanatory variables are ranked according to variable importance (sum=100) in splitting or predicting the dependent variable (eDNA detection rate).**

Variables	Node 1 (N=64)		Importance
	Primary Splits	Surrogate Splits	
	Improvement	Agree	
Geese	0.5764	-	48
Area	0.1420	0.672	15
pH	0.1251	0.688	18
Residual Temp	0.1121	0.641	18
Secchi	0.0877	-	1

## Discussion

The first step in evaluating a new survey method is determining whether the method positively detects the target species when present. My results show that eDNA effectively detects the presence of Canada geese (*B. canadensis*) in ponds. Using a protocol I developed for determining the presence of the Canada goose at ponds based on the detection of eDNA in water samples (Takahara et al. 2013), I detected Canada goose eDNA, using species-specific primers and probes targeting the ND2 mitochondrial gene, in all six ponds where the species was visually observed, as well as in three ponds where geese were not observed prior to or during sampling (Table 5). These results indicate that eDNA is a promising approach for surveying Canada geese and possibly other semi-aquatic bird species, as well.

### *Effect of Goose abundance on eDNA Detection*

Correlation and CART analyses revealed that goose abundance is the best indicator of and accounts for the greatest variation in eDNA presence and detection rate. eDNA detection rate (% samples with eDNA present) was significantly positively correlated with goose abundance, so that as goose abundance increased, eDNA detection rate also increased (Figure 4). Both classification and regression trees found that goose abundance was the most important variable for explaining eDNA detection, while the effects of other environmental variables were generally insignificant. CART analyses revealed that where goose abundance was greater than 4.5 individuals, eDNA detection rate was very high (96.08%), and where goose abundance was less than 4.5 individuals, eDNA detection rate was low (33.89%). These results suggest that eDNA presence and detection is assured when goose abundance is high, though detection rates

drop significantly and are more variable when goose abundance is low. According to Ellison et al. (2006), inconsistency in eDNA detection between qPCR replicates is expected when target DNA concentration is extremely low (<100 copies/reaction). In addition, a model by Barnes et al. (2014) estimated that qPCR became more likely to fail to detect common carp (*C. carpio*) eDNA than to detect it after 47.2 hours (~2 days), and eDNA detection failure became 95% probable after 101.1 hrs (4.2 days) following carp removal. Based on these findings, it is possible that: 1) eDNA concentration was low in pond samples which had low detection rates, and 2) at ponds where geese were not observed, geese could have been present at some point in the 4 days previous to sampling. Because ponds with low goose abundance had more variable detection rates between samples than ponds with high goose abundance, it is likely that Canada goose eDNA concentration was low at ponds with low goose abundance. This would support my hypothesis that increased goose abundance results in higher eDNA concentration and is likely due to increased eDNA production (the rate at which DNA is released into the environment; Strickler et al. 2015).

Some studies have shown that target species density or biomass might be a more reliable predictor of eDNA concentration and detectability than target species abundance (Stoeckle et al. 2017; Takahara et al. 2012; Thomsen et al. 2012; Ficetola et al. 2008). However, my CART and correlation analyses revealed that goose abundance explained the data slightly better than goose density, a finding consistent with that of other studies (Lacoursiere-Roussel et al. 2016; Doi et al. 2015). Pond area (m<sup>2</sup>) had relatively low variable importance in all CART models (Tables 6, 7), and the relationship between pond area (m<sup>2</sup>) and goose abundance, though positive, was

relatively weak (Figure 6). Similarly, Doi et al. (2015) found that common carp (*Cyprinus carpio* L.) abundance more accurately estimated eDNA concentration than did biomass. Though there is some uncertainty as to which parameter best explains eDNA concentration and detection, my results align with current findings which state that target species abundance, density, and biomass have a significant, positive effect on eDNA concentration and amplification rates (Doi et al. 2015; Stoeckle et al. 2017; Thomsen et al. 2012; Takahara et al. 2012; Ficetola et al. 2008; Lacoursiere-Roussel et al. 2016). Further work is needed to determine which variable best explains eDNA detection, as well as how reliably eDNA concentration can estimate target species abundance (Stoeckle et al. 2017).

#### *Effects of Environmental Variables on eDNA Detection*

Spearman's correlation and regression tree analysis revealed that goose abundance was the only variable that had a significant effect on eDNA detection rate. Studies have shown that, in general, increased temperatures cause increased DNA degradation rates directly via denaturation (though only at temperatures > 50°C) and indirectly through increased enzyme kinetics (i.e. exonuclease activity) and microbial metabolism (Barnes et al. 2014, Strickler et al. 2015; Tsuji et al. 2017). It therefore might be expected that samples with higher temperatures would result in lower or more variable eDNA detection rates. However, the range of temperatures recorded in this study was between 22.5 and 38.4°C (or -4.3°C and 5.9°C residual temperature), which may have been too small to detect an effect of temperature on eDNA detection rate (Seymour et al. 2018). In addition, the maximum temperature (38.4°C) was likely too low to cause direct denaturation. Strickler et al. (2015) found that, though temperature

(5°C, 25°C, 35°C treatments) had a significant effect on bullfrog tadpole eDNA concentration over time, eDNA concentration at the outset of their experiment was similar in all microcosms, irrespective of temperature treatment. It is possible that the eDNA in my samples was fresh, especially at ponds with geese present during sampling, and had not yet undergone significant degradation. If this were the case, no effect of temperature would be expected (Strickler et al. 2015). However, eDNA concentration and degradation rate were not quantified in this study, so no definite conclusions relating detection rate and eDNA concentration can be made. Overall, though, it is unlikely that the moderate water temperatures recorded in this study had a significant effect on eDNA detection rates, as demonstrated by correlation and regression tree analyses (Figure 7b).

Because acidic conditions are known to catalyze hydrolytic processes that degrade DNA (Strickler et al. 2015), one might expect pH to have a negative effect on eDNA detection rates. However, studies have suggested that pH alone may be insufficient to accelerate eDNA degradation (Strickler et al. 2015). In a study investigating the effects of temperature, pH, and UV-B on bullfrog tadpole eDNA degradation rate, pH (=4) was only significant when interacting with other variables (i.e. when temperatures were high; Strickler et al. 2015). In my study, variation in pH (7.05 to 9.46) was low, and all samples were alkaline. This, combined with the lack of an effect of temperature on eDNA detection, and thus a lack of any interaction effects, suggests that an effect of pH on eDNA detection was unlikely (Barnes et al. 2014). Though evidence suggests that pH and temperature, along with other variables, affect eDNA persistence and detection (Seymour et al. 2018; Tsuji et al. 2017; Stoeckle et al.



2017; Strickler et al. 2015), greater variation and larger samples sizes are needed to investigate the effects more fully. In this study, where samples sizes were small and variation in pH and temperature were low, goose abundance alone best explained the data.

#### *Further Considerations*

eDNA availability and persistence depends on several factors, including: (1) the eDNA production rate of a species or individual, (2) the effects of abiotic environmental characteristics on eDNA degradation, (3) transport and removal of eDNA such as through volatilization or water flow downstream, and (4) eDNA molecule characteristics (length, sequence, conformation) which affect how DNA interacts with the environment (Barnes et al. 2014; Strickler et al. 2015). Though eDNA may become undetectable in less than a day after removal of the target species from a microcosm, studies have shown that eDNA can persist in water at detectable concentrations for up to 58 days after species removal (Strickler et al. 2015). In addition, eDNA can persist in sediment for hundreds of days to hundreds of thousands of years (Turner et al. 2015; Bohmann et al. 2014). It is therefore critical to understand how the above factors influence eDNA degradation and concentration to accurately interpret eDNA surveillance results, minimizing false negative and false positives, and to improve sampling strategies (Barnes et al. 2014). In addition, researchers must take care when analyzing eDNA data (e.g. eDNA presence-absence vs. concentration over time), as different approaches dramatically affect how eDNA surveillance results are interpreted (Barnes et al. 2014).

### *Conclusions and Future Research*

eDNA is an effective method for detecting Canada goose eDNA in ponds, and where goose abundance is high, eDNA detection is assured. When used for species monitoring, eDNA detection rate may provide some indication about goose abundance. However, further work is needed to understand Canada goose eDNA production rates and how environmental variables affect eDNA detection, concentration, and persistence to better interpret surveillance results and apply this surveying technique to other species and taxa, particularly for conservation applications. It is of interest to apply the eDNA technique to semi-aquatic or marsh birds of conservation concern which are rare or cryptic or which have minimal vocalization, such as rails, bitterns, or woodcocks. However, because eDNA is so sensitive, care must be taken not to over- or underestimate population sizes and ranges to avoid forming misguided conservation plans. eDNA could also be tested using soil samples to survey for more terrestrial bird species, such as Henslow's sparrow, or other terrestrial taxa, which could have applications in tracking temporal and spatial migratory movements.

Although initially time-intensive during the design and testing stages, once markers are developed eDNA could serve as a rapid, cost-effective, and highly sensitive surveying technique to aid conventional surveying in identifying high-priority monitoring areas for species of conservation concern and to enhance species detection, population monitoring, distribution mapping, and conservation planning (Davy et al. 2015). Improvements in these areas could have far-reaching consequences for a variety of disciplines that depend on accurate surveying and species distribution assessments, including conservation biology, molecular biology, biogeography, ecology,

paleontology, and environmental sciences (Ficetola et al. 2008; Thomsen & Willerslev 2014).

Many eDNA studies have used DNA barcoding to detect a single target species. This approach has high detection sensitivity, specificity, and quantification ability but is limited in scope (Thomsen & Willerslev 2014). As technologies improve, techniques such as next-generation sequencing will allow researchers to exploit eDNA's potential to recover the DNA of hundreds of specimens of multiple species at once from a single sample, which will allow efficient processing of complex environmental samples (Shokralla et al. 2012). Nonetheless, to better understand the implications of eDNA surveying for birds, the DNA barcoding approach demonstrated here should be tested on species of conservation concern in the hope of aiding in future conservation planning and management.

## References

- Arndt, D. (2003). The climate of Oklahoma. Oklahoma Climatological Survey. at <  
<http://cig.mesonet.org/climateatlas/doc60.html>>. Accessed 18 June 2018.
- Banks, R.C., C. Cicero, J.L. Dunn, A.W. Kratter, P.C. Rasmussen, J.V. Remsen, Jr.,  
J.D. Rising, and D.F. Stotz. (2004) Forty-fifth supplement to the American  
Ornithologists' Union *check-list of North American birds*. Auk 121(3), 985-  
95.
- Barnes, M.A., C.R. Turner, C.L. Jerde, M.A. Renshaw, W.L. Chadderton, and D.M.  
Lodge. (2014) Environmental conditions influence eDNA persistence in aquatic  
systems. *Environmental Science and Technology* 48, 1819-27.
- Bass, D., G.D. Stentiford, D.T.J. Littlewood, and H. Hartikainen. (2015) Diverse  
applications of environmental DNA methods in parasitology. *Trends in  
Parasitology* 31(10), 499-513.
- Baumgartner, F.M., and A.M. Baumgartner. (1992). *Oklahoma Bird Life*. University of  
Oklahoma Press: Norman, Oklahoma. 443 p.
- Bohmann, K., A. Evans, M.T.P. Gilbert, G.R. Carvalho, S. Creer, M. Knapp, D.W.  
Yu, and M. de Bruyn. (2014) Environmental DNA for wildlife biology and  
biodiversity monitoring. *Trends in Ecology & Evolution* 29(6), 358-367.
- Cardinale, B.J. et al. (2012) Biodiversity loss and its impact on humanity. *Nature* 486,  
59–67.
- Ceballos, G., P.R. Ehrlich, and R. Dirzo. (2017) Biological annihilation via the ongoing  
sixth mass extinction signaled by vertebrate population losses and

- declines. *Proceedings of the National Academy of Sciences* 114(30), E6089-E6096.
- Cleary, E.C., R.A. Dolbeer, and S.E. Wright. (2006) *Wildlife strikes to civil aircraft in the United States 1990–2005*. U.S. Department of Transportation, Federal Aviation Administration, and Wildlife Services. *National Wildlife Strike Database Serial Report 12*, Washington, D.C., USA.
- Davy, C.M., A.G. Kidd, and C.C. Wilson. (2015) Development and Validation of Environmental DNA (eDNA) Markers for Detection of Freshwater Turtles. *PLoS ONE* 10(7), e0130965.
- Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society* 85(3), 407-15.
- De'ath, G., and K.E. Fabricius. (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81(11), 3178-92.
- Doi, H., K. Uchii, T. Takahara, S. Matsushashi, H. Yamanaka, and T. Minamoto. (2015) Use of droplet digital PCR for estimation of fish abundance and biomass in environmental DNA surveys. *PLoS ONE* 10(3), e0122763.
- Donne-Goussé, C., V. Laudet, and C. Hänni. (2002) A molecular phylogeny of Anseriformes based on mitochondrial DNA analysis. *Molecular Phylogenetics and Evolutions* 23, 339-56.
- Duck, L.G., and J.B. Fletcher. (1943) *A game type map of Oklahoma. A Survey of the Game and Fur-bearing Animals of Oklahoma*. Oklahoma Department of Wildlife Conservation: Oklahoma City, Oklahoma.

- Ellison, S.L.R., C.A. English, M.J. Burns, and J.T. Keer. (2006) Routes to improving the reliability of low level DNA analysis using real-time PCR. *BMC Biotechnology* 6, 33.
- Esselstyn, J.A., H.J.D. Garcia, M.G. Saulog, and L.R. Heaney. (2008) A new species of *Desmalopex* (Pteropodidae) from the Philippines, with a phylogenetic analysis of the Pteropodini. *Journal of Mammalogy* 89, 815–825.
- Evans, N.T., B.P. Olds, M.A. Renshaw, C.R. Turner, Y. Li, C.L. Jerde, A.R. Mahon, M.E. Pfrender, G.A. Lamberti, and D.M. Lodge. (2016) Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources* 16, 29-41.
- Ficetola, G.F., C. Miaud, F. Pompanon, and P. Taberlet. (2008) Species detection using environmental DNA from water samples. *Biology Letters* 4 (4), 423-5.
- Gonzalez, J., H. Düttmann, and M. Wink. (2009) Phylogenetic relationships based on two mitochondrial genes and hybridization patterns in Anatidae. *Journal of Zoology* 279 (3), 310-318.
- Hoagland, B.W. (2004). *Vegetation of Oklahoma*. In D. Reinking (Ed.), *Oklahoma Breeding Bird Atlas*. George M. Sutton Avian Research Center. University of Oklahoma Press: Norman, Oklahoma.
- Hopkins, G.W., and R.P. Freckleton. (2002) Declines in the numbers of amateur and professional taxonomists: implications for conservation. *Animal Conservation* 5(3), 245-49.
- Hull, P. (2015). Life in the aftermath of mass extinctions. *Current Biology* 25(19), 941-52.

- Huver, J.R., J. Koprivnikar, P.T.J. Johnson, and S. Whyard. (2015) Development and application of an eDNA method to detect and quantify a parasite in aquatic ecosystems. *Ecological Applications* 25(4), 991-1002.
- Johnson, K.P., and M.D. Sorenson. (1998) Comparing molecular evolution in two mitochondrial protein coding genes (cytochrome b and ND2) in the dabbling ducks (tribe: Anatini). *Molecular Phylogenetics and Evolution* 10(1), 82-94.
- Kassambara, A. (2017). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.1.6. <https://CRAN.R-project.org/package=ggpubr>.
- Kerr, K.C.R., M.Y. Stoeckle, C.J. Dove, L.A. Weigt, C.M. Francis, and P.D.N. Hebert. (2007) Comprehensive DNA barcode coverage of North American birds. *Molecular Ecology Notes* 7(4), 535-543.
- Lacoursière-Roussel, A., Y. Dubois, E. Normandeau, and L. Bernatchez. (2016) Improving herpetological surveys in eastern North America using the environmental DNA method. *Genome* 59, 991-1007.
- Margules, C.R., and R.L. Pressey. (2000) Systematic conservation planning. *Nature* 405, 243-53.
- Martin, J.T. (2004) Taxonomists and Conservation. *Science* 305, 1104-05.
- McCracken, K.G., and M.D. Sorenson. (2005) Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (*Nomonyx-Oxyura*)? *Systematic Biology* 54(1), 35-55.
- National Center for Biotechnology Information (NCBI). Basic Local Alignment Search Tool (BLAST). <<https://blast.ncbi.nlm.nih.gov/Blast.cgi>> Accessed 29 June 2018.

- Nice, M.M., and L.B. Nice. (1924) The Birds of Oklahoma. University of Oklahoma Bulletin. Norman, OK. 122 p.
- Oklahoma Bird Records Committee. (2014) Date guide to the occurrences of birds in Oklahoma. Sixth edition. Oklahoma Ornithological Society: Edmond, Oklahoma. 43 p.
- Ottenburghs, J., H.J. Megens, R.H.S. Kraus, O. Madsen, P. van Hooft, S.E. van Wieren, R.P.M.A Crooijmans, R.C. Ydenberg, M.A.M. Groenen, and H.H.T. Prins. (2016) A tree of geese: A phylogenomic perspective on the evolutionary history of True Geese. *Molecular Phylogenetics and Evolution* 101 (Supplement C), 303-313.
- Paxinos, E.E., H.F. James, S.L. Olson, M.D. Sorenson, J. Jackson, and R.C. Fleischer. (2002) mtDNA from fossils reveals a radiation of Hawaiian geese recently derived from the Canada goose (*Branta canadensis*). *Proceedings of the National Academy of Sciences* 99 (3), 1399-1404.
- Pilliod, D.S., C.S. Goldberg, M.B. Laramie, and L.P. Waits. (2013) Application of environmental DNA for inventory and monitoring of aquatic species: U.S. Geological Survey Fact Sheet 2012-3146, 4 p.
- Pimm, S.L., C.N. Jenkins, R. Abell, T.M. Brooks, J.L. Gittleman, L.N. Joppa, P.H. Raven, C.M. Roberts, and J.O. Sexton. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science* 344(6187), 1246752.



- Rees, H.C., B.C. Maddison, D.J. Middleditch, J.R.M. Patmore, and K.C. Gough. (2014) The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology* 51, 1450-59.
- Reinking, D.L. (2004) *Oklahoma Breeding Bird Atlas*. George M. Sutton Avian Research Center. University of Oklahoma Press: Norman, Oklahoma. 528 p.
- Reinking, D.L. (2017) *Oklahoma Winter Bird Atlas*. George M. Sutton Avian Research Center. University of Oklahoma Press: Norman, Oklahoma. 552 p.
- Seymour, M. et al. (2018) Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. *Communications Biology* 1, 1-8.
- Shokralla, S., J.L. Spall, J.F. Gibson, and M. Hajibabaei. (2012) Next-generation sequencing technologies for environmental DNA research. *Molecular Ecology* 21(8), 1794-1805.
- Sibley, D. (2004). Distinguishing cackling and Canada goose. at <http://www.sibleyguides.com/2007/07/identification-of-cackling-and-canada-goose/>. Accessed 18 June 2018.
- Stoeckle, B.C., S. Beggel, A.F. Cerwenka, E. Motivans, R. Kuehn, and J. Geist. (2017) A systematic approach to evaluate the influence of environmental conditions on eDNA detection success in aquatic ecosystems. *PLoS ONE* 12(12), e0189119.
- Strickler, K.M., A.K. Fremier, and C.S. Goldberg. (2015) Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation* 183, 85-92.

- Sullivan, B.L., C.L. Wood, M.J. Iliff, R.E. Bonney, D. Fink, and S. Kelling. (2009) eBird: a citizen-based bird observation network in the biological sciences. *Biological Conservation* 142, 2282-2292.
- Summit Lakes. (2018) Homeowners Association Website. About Us. at <http://www.summitlakespoa.com/info.php?pnum=5>>. Accessed 3 July 2018.
- Sutton, G.M. (1967). Oklahoma birds: Their ecology and distribution, with comments on the avifauna of the Southern Great Plains. University of Oklahoma Press: Norman, Oklahoma. 674 p.
- Takahara, T., T. Minamoto, H. Yamanaka, H. Doi, and Z. Kawabata. (2012) Estimation of fish biomass using environmental DNA. *PLoS ONE* 7(4), e35868.
- Takahara, T., T. Minamoto, and H. Doi. (2013) Using environmental DNA to estimate the distribution of an invasive fish species in ponds. *PLoS ONE* 8(2), e56584.
- Therneau, T., and B. Atkinson. (2018). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-13. <https://CRAN.R-project.org/package=rpart>.
- Thomsen, P.F., J.O.S. Kielgast, L.L. Iversen, C. Wiuf, M. Rasmussen, M.T.P Gilbert, L. Orlando, and E. Willerslev. (2012) Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology* 21 (11), 2565-73.
- Thomsen, P.F. and E. Willerslev. (2014) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation* 183 (Supplement C), 4-18.
- Tréguier, A., J. Paillisson, T. Dejean, A. Valentini, M.A. Schlaepfer, and J. Roussel. (2014) Environmental DNA surveillance for invertebrate species: advantages

- and technical limitations to detect invasive crayfish *Procambarus clarkia* in freshwater ponds. *Journal of Applied Ecology* 51, 871-79.
- Tsuji, S., M. Ushio, S. Sakurai, T. Minamoto, and H. Yamanaka. (2017) Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. *PLoS ONE* 12(4), e0176608.
- Turner, C.R., D.J. Miller, K.J. Coyne, and J. Corush. (2014) Improved methods for capture, extraction, and quantitative assay of environmental DNA from Asian bigheaded carp (*Hypophthalmichthys spp.*). *PLoS One* 9(12), e114329.
- Turner, C.R., K.L. Uy, and R.C. Everhart. (2015) Fish environmental DNA is more concentration in aquatic sediments than surface water. *Biological Conservation* 183, 93-102.
- Wheeler, Q. (2014) Are reports of the death of taxonomy an exaggeration? *New Phytologist*, 201(2), 370-371.
- Wheeler, Q., P. Raven, and E. Wilson. (2004) Taxonomy: Impediment or expedient? *Science* 303(5656), 285.
- Wickham, H. (2009) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wilcox, T.M., K.S. McKelvey, M.K. Young, S.F. Jane, W.H. Lowe, A.R. Whiteley, and M.K. Schwartz. (2013) Robust detection of rare species using environmental DNA: the importance of primer specificity. *PLoS One* 8 (3), e59520.
- Wood, D.S., and G.D. Schnell. (1984) *Distributions of Oklahoma Birds*. University of Oklahoma Press: Norman, Oklahoma. 209 p.

## Appendix I

**Thirty-three museum tissue aliquots were used in this study. OCGR specimens were from the Sam Noble Oklahoma Museum of Natural History; DMNS - Denver Museum of Nature and Science; FMNH - Field Museum of Natural History; MVZ - Museum of Vertebrate Zoology.**

Sample ID	Species	State
OCGR 8821	<i>Anas platyrhynchos</i>	Oklahoma
OCGR 8837	<i>Anas strepera</i>	Oklahoma
OCGR 9930	<i>Branta canadensis</i>	Oklahoma
OCGR 10645	<i>Branta canadensis</i>	Oklahoma
OCGR 11246	<i>Branta canadensis</i>	Oklahoma
OCGR 11835	<i>Branta canadensis</i>	Oklahoma
ZB 34858	<i>Branta canadensis</i>	Colorado
ZB 34864	<i>Branta hutchinsii</i>	Colorado
ZB 35091	<i>Branta canadensis</i>	Colorado
ZB 35224	<i>Anas platyrhynchos</i>	Kansas
ZB 35225	<i>Anas strepera</i>	North Dakota
ZB 44017	<i>Branta canadensis</i>	Colorado
ZB 45763	<i>Anas platyrhynchos</i>	Colorado
ZB 46079	<i>Branta canadensis</i>	Colorado
ZB 47055	<i>Anas platyrhynchos</i>	Colorado
ZB 47686	<i>Branta hutchinsii</i>	Colorado
ZB 47705	<i>Branta hutchinsii</i>	Colorado
FMNH 363337	<i>Branta canadensis minima</i>	California
FMNH 438243	<i>Branta canadensis maxima</i>	Illinois
FMNH 438244	<i>Branta canadensis canadensis</i>	Illinois
FMNH 440351	<i>Branta canadensis maxima</i>	Wisconsin
FMNH 449116	<i>Branta canadensis</i>	Minnesota
FMNH 480459	<i>Branta hutchinsii</i>	Illinois
FMNH 486719	<i>Branta hutchinsii</i>	Wisconsin
FMNH 488416	<i>Anas platyrhynchos platyrhynchos</i>	Illinois
FMNH 488508	<i>Branta canadensis</i>	Illinois
FMNH 492427	<i>Anas strepera strepera</i>	Wisconsin
FMNH 496455	<i>Anas platyrhynchos platyrhynchos</i>	Minnesota
FMNH 500383	<i>Anas strepera strepera</i>	Illinois
MVZ 181825	<i>Anas platyrhynchos platyrhynchos</i>	California
MVZ 183934	<i>Anas platyrhynchos platyrhynchos</i>	California
MVZ 182091	<i>Anas strepera</i>	Oregon
MVZ 184923	<i>Branta canadensis</i>	California

## Appendix II

**Forward and reverse primers from the literature were used to create thirty-five primer-pair combinations, which were tested on seven tissue DNA extractions.**

Primer Pair	Gene	Forward primer	Reverse primer
1	COI	BirdF1	BirdR1
2	COI	BirdF1	BirdR2
3	COI	BirdF1	VertebrateR1
4	COI	FalcoFA	BirdR1
5	COI	FalcoFA	BirdR2
6	COI	FalcoFA	VertebrateR1
7	Cytb	L14770	H15021
8	Cytb	L14770	H15545
9	Cytb	L14770	H15646
10	Cytb	L14770	H15742
11	Cytb	L14770	H16064
12	Cytb	L14990	H15021
13	Cytb	L14990	H15545
14	Cytb	L14990	H15646
15	Cytb	L14990	H15742
16	Cytb	L14990	H16064
17	Cytb	L14996	H15021
18	Cytb	L14996	H15545
19	Cytb	L14996	H15646
20	Cytb	L14996	H15742
21	Cytb	L14996	H16064
22	Cytb	L15191	H15021
23	Cytb	L15191	H15545
24	Cytb	L15191	H15646
25	Cytb	L15191	H15742
26	Cytb	L15191	H16064
27	ND2	L5216	H5766
28	ND2	L5216	H6031
29	ND2	L5216	H6313
30	ND2	L5219	H5766
31	ND2	L5219	H6031
32	ND2	L5219	H6313
33	ND2	L5524	H5766
34	ND2	L5524	H6031
35	ND2	L5524	H6313

### Appendix III

**Twenty primers that were reported to successfully amplify DNA of a variety of taxa were ordered and used for this study. Mitochondrial genes COI, Cyt b, and ND2 were targeted for Canada goose DNA amplification.**

Gene	Primer Sequence	Reference
COI	BirdF1: 5'-TTCTCCAACCAAGACATTTGGCAC-3'	Kerr et al. 2007
COI	BirdR1: 5'-ACGTGGGAGATAATTCCAATCCTG-3'	Kerr et al. 2007
COI	FalcoFA: TCAACAAACCACAAAGACATCGGCAC	Kerr et al. 2007
COI	BirdR2: 5'-ACTACATGTGAGATGATCCGAAATCCAG-3'	Kerr et al. 2007
COI	VertebrateR1: TAGACTTCTGGGTGGCCAAAAGAAATCA	Kerr et al. 2007
Cyt b	H15021: 5'-GTATGGGTGAAATGGAAATTTTGTC-3'	Paxinos et al. 2002
Cyt b	L14770: 5'-AATAGGMCCMGAAGMCTNGC-3'	McCracken & Sorenson (2005)
Cyt b	L14990: 5'-AACATCTCCGCATGATGAAA-3'	Johnson & Sorenson (1998)
Cyt b	L14996: 5'-AAYATYTCWGYHTGATGAAAYTTYGG-3'	McCracken & Sorenson (2005)
Cyt b	L15191: 5'-ATCTGCATCTACCTACACATCGG-3'	Johnson & Sorenson (1998)
Cyt b	H15545: 5'-GTATGGGTGAAATGGAAATTT-3'	Johnson & Sorenson (1998)
Cyt b	H15646: 5'-GGNGTRAAGTTTCTGGGTCNCC-3'	McCracken & Sorenson (2005)
Cyt b	H15742: 5'-TGCTAGTACGCCCTCCTAGTTTGGGATTGA-3'	Johnson & Sorenson (1998)
Cyt b	H16064: 5'-CTTCGATTTTGGTTTACAAGACC-3'	Donne-Goussé et al. (2002)
ND2	L5216: 5'-GGCCCATACCCCGRAAATG-3'	McCracken & Sorenson (2005)
ND2	L5219: 5'-CCCATACCCCGAAAATGATG-3'	Johnson & Sorenson (1998)
ND2	L5524: 5'-AGGCCTGGTCCCATTTCACT-3'	Donne-Goussé et al. (2002)
ND2	H5766: 5'-GGATGAGAAGGCTAGGATTTKCG-3'	Johnson & Sorenson (1998)
ND2	H6031: 5'-CACTTGGTATAAACCCCTGT-3'	Donne-Goussé et al. (2002)
ND2	H6313: 5'-CTCTATTAAAGGCTTTGAAGGC-3'	Johnson & Sorenson (1998)

## Appendix IV

Fifteen DNA extractions from museum tissues collected from various states in the U.S. were screened for the presence of Canada goose DNA using two primer-probe sets targeting mitochondrial genes COI and ND2. Each tissue extraction was run in triplicate in qPCR, and the percentage of positive replicates for each sample is reported. Target species DNA was diluted to 1:10,000, with non-target species at 1:1,000 dilution. The average change in normalized reporter intensity ( $\Delta Rn$ ) for each sample run in triplicate is indicated. Presence of the target DNA sequence was determined by a  $\Delta Rn$  value higher than threshold (=3.230).

Tissue ID	Species	Locality	COI % Amp	COI $\Delta Rn$	ND2 % Amp	ND2 $\Delta Rn$
OCGR 9930	<i>B. canadensis</i>	OK	100	3.638	100	3.325
OCGR 11835	<i>B. canadensis</i>	OK	66.7	3.395	100	3.306
ZB 34858	<i>B. canadensis</i>	CO	66.7	3.435	100	3.266
ZB 44017	<i>B. canadensis</i>	CO	100	3.428	100	3.332
MVZ 184923	<i>B. canadensis</i>	CA	100	3.522	100	3.533
FMNH 438244	<i>B. c. canadensis</i>	IL	100	3.536	100	3.417
FMNH 440351	<i>B. c. maxima</i>	WI	100	3.576	100	3.472
ZB 35224	<i>A. platyrhynchos</i>	KS	0	0.734	0	1.578
ZB 47705	<i>A. platyrhynchos</i>	CO	0	1.294	0	1.988
OCGR 8821	<i>A. platyrhynchos</i>	OK	0	0.844	0	1.625
OCGR 8837	<i>A. strepera</i>	OK	0	2.161	0	1.709
FMNH 480459	<i>B. hutchinsii</i>	IL	0	1.318	0	2.069
FMNH 486719	<i>B. hutchinsii</i>	WI	0	1.326	0	2.13
FMNH 363337	<i>B.c. minima</i>	CA	0	1.277	0	2.023
ZB 34864	<i>B. hutchinsii</i>	CO	0	1.31	0	2.045
Neg. control	-	-	0	0.774	0	1.678

## Appendix V

**An ND2 primer-probe assay was designed and tested on fifteen different combinations of DNA extracts (1:1000 dilution) from museum tissue aliquots of four closely related species (mallard, gadwall, cackling goose, Canada goose) vouchered around the United States. “Presence” indicates whether DNA was positively amplified (Y=Yes, N=No) using the ND2 assay targeting Canada goose DNA.  $\Delta Rn$  mean is the average change in normalized reporter intensity for each sample run in triplicate in qPCR. Presence of the target DNA sequence is determined by a  $\Delta Rn$  value higher than the threshold (=2.285).**

Sample Name	Presence	$\Delta Rn$ mean	Species Combination
Combo 1	Y	3.817	Canada, mallard, gadwall
Combo 2	Y	2.536	Canada, cackling
Combo 3	N	2.033	Cackling
Combo 4	Y	2.809	Mallard, gadwall, Canada, cackling
Combo 5	N	2.037	Mallard, gadwall, cackling
Combo 6	Y	2.836	Canada, gadwall, mallard, cackling
Combo 7	N	1.654	Mallard, gadwall
Combo 8	Y	3.853	Canada, mallard, gadwall
Combo 9	Y	3.461	Canada, mallard, gadwall
Combo 10	N	1.587	Mallard, gadwall
Combo 11	Y	2.912	Mallard, cackling, gadwall, Canada
Combo 12	N	1.896	Mallard, gadwall, cackling
Combo 13	Y	2.373	Mallard, gadwall, cackling, Canada
Combo 14	Y	3.466	Canada, mallard, gadwall
Combo 15	N	1.960	Mallard, gadwall, cackling
Neg. Ctrl	N	1.621	None