Using eDNA and Habitat Suitability Modeling to Better Understand the Range and Habitat Requirements of the Eastern Black Rail

by

Amberly A. Neice

August 2020

Director of Thesis: Dr. Susan B. McRae

Major Department: Biology

Abstract:

The Black Rail (*Laterallus jamaicensis*) is a marsh bird that is globally listed as Near Threatened and is being considered for listing under the federal Endangered Species Act. This species has experienced concerning population declines throughout its range. Black Rails are difficult to detect due to their small size, concealing habitat, and cryptic behavior. The most common survey method for rails uses audio callback but does not detect unresponsive individuals, is constrained seasonally as well as temporally, and requires significant personnel effort. New methods are needed to provide information on the distribution and habitat requirements for this threatened species. Here, I describe a novel detection method for Black Rail using environmental DNA (eDNA) and an ecological niche model identifying areas and characteristics of suitable habitat for this species.

To detect Black Rail eDNA I developed a qPCR assay that targets a 219-bp region of the cytochrome c oxidase subunit 1 gene (COI) and uses a fluorescent reporter probe to increase specificity. The assay reliably produces a signal when sufficient copies of Black Rail template are present, and does not produce signal when tested for cross-species amplification using genomic DNA from sympatric rail species. The assay successfully amplified Black Rail eDNA

from environmental samples taken from locations with positive detections. I tested statistically whether various environmental factors, as well as sampling and handling variables, affected eDNA detectability. Among the factors tested for their influence on amplification success (time between collection and DNA extraction, storage temperature before filtering, field detection method (audio, visual, camera trap, none), time between detection and sample collection, water salinity, and air temperature), only water depth was found to have a significant effect.

I also created a habitat suitability model for the Eastern Black Rail focusing on the Atlantic coastal plain using eBird data contributed by citizen scientists and environmental variable data from the Esri databank using a maximum entropy model framework. The map generated by the *MaxEnt* model indicated habitat suitability in areas known for Black Rail occupation. The environmental factors that best predicted Black Rail presence were flooded areas of shrub and herbaceous vegetation, proximity to water, and flat plains. These environmental variable associations were congruent with other habitat association studies conducted in other parts of the species' range that focused on smaller areas and used presence data collected through surveys. My habitat suitability model had comparable statistical parameters to other *MaxEnt* models created for birds. Correlation with known areas of Black Rail occupation and previous habitat associations confirms the validity of the model and importance of high marsh habitat for the species.

The uses of eDNA adds a novel tool to the avian conservation toolbox that can be improved and adapted for other species of concern. The habitat suability model provides a starting point for land management and habitat restoration efforts for Black Rail now and in the future. The information gained using these two techniques can add much needed insight into the range and ecological needs of this imperiled species.

Using eDNA and Habitat Suitability Modeling to Better Understand the Range and Habitat Requirements of the Eastern Black Rail

A Thesis

Presented to the Faculty of the Department of Biology

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Biology

by

Amberly A. Neice

August 2020

Title: Using eDNA and Habitat Suitability Modeling to Better Understand the Range and Habitat Requirements of the Eastern Black Rail

By: Amberly A. Neice

Approved by:

Director of Thesis:

_____

Susan B. McRae Ph.D

Committee Members:

_____

Christopher K. Balakrishnan Ph.D

_____

Erin K. Field Ph.D

_____

Ariane L. Peralta Ph.D

Department Chairperson:

_____

David R. Chalcraft, Ph.D.

Dean of the Graduate school:

_____

Paul J. Gemperline, Ph.D.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Susan McRae for giving me the opportunity to work on this challenging and rewarding project with the most amazing birds. Thank you so much for your patience, guidance, and support that allowed this project to be a success. I would also like to acknowledge the support and generosity of my committee Dr. Ariane Peralta, Dr. Erin Field, and Dr. Chris Balakrishnan. Thank you all for always being willing to answer my questions and for all the advice. Thanks also goes out to my field assistants Cody McGregor and Ryan Pless who braved the heat and mud to help me find these elusive birds. Last but not least I would like to thank my family, both near and far, for their unwavering support and confidence.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One: Using eDNA to Detect Secretive Marsh Birds

Summary

The Black rail is a small, globally-threatened marsh bird that has experienced population declines throughout its range. The most common method of detection, audio call back survey, does not detect unresponsive individuals, is constrained seasonally as well as temporally, and requires significant personnel effort. New, minimally invasive detection methods are needed to provide information on distribution and habitat requirements for this threatened species. Here, I describe a novel method to detect a secretive marsh bird, the Black Rail (*Laterallus jamaicensis*) using environmental DNA (eDNA). To obtain positive environmental controls for validation of eDNA diagnostics, audio callback surveys targeting Black Rail were performed at sites along the North Carolina coast and Virginia tidewater region during the 2018 and 2019 breeding seasons. Only 12 out of 620 callback surveys (2%) produced an auditory detection, and eDNA samples were collected from the 6 sites in North Carolina where Black Rails were detected. Additional samples were obtained from collaborators working on Black Rails in South Carolina and Florida. I developed a qPCR assay to detect small amounts of Black Rail DNA. The assay targets a 219-bp region of the cytochrome c oxidase subunit 1 gene (COI) and uses a fluorescent reporter probe to increase specificity. The assay successfully amplified Black Rail eDNA from 30-47% (depending on the level of test stringency) of the environmental samples taken from locations with positive detections. When tested for cross-species amplification using genomic DNA from sympatric rail species, the assay only produced a signal for Black Rail exemplifying the assay's high degree of species specificity. I tested statistically whether various environmental factors, as well as sampling and handling variables, affected eDNA detectability. Only water depth met the statistical criteria for importance, possibly suggesting the role of water in protecting and

dispersing DNA thus improving eDNA detectability. This eDNA diagnostic test for Black Rails offers a new way to ascertain Black Rail occupancy without temporal constraints, and will contribute to informing the species conservation plan.

Introduction

The Black Rail is a sparrow-sized rail found in either high marsh (saltmarsh, brackish or freshwater) or infrequently flooded upland wetlands generally dominated by shorter grasses (Flores & Eddleman, 1995). Comparatively little is known of its ecology and behavior because its concealing habitat and secretive nature make detection difficult. Black Rails are currently listed as globally Near Threatened by Birdlife International (BirdLife International, 2019). An alarming loss of inland marsh habitat has driven Black Rails, among other wetland species, to fragmented, marginal habitats and to coastal marshes that are impacted by human development and sea level rise (Conway & Sulzman, 2007; Stedman & Dahl, 2013). Having a reliable and time-efficient detection method to assess the range and current occupancy of these species is becoming more critical.

The U.S. Fish and Wildlife Service recently proposed that the Black Rail should be uplisted to federally Threatened, and it is a candidate for federal listing under the Endangered Species Act (U.S. Fish and Wildlife Service, 2018). Of concern is the paucity of detections of Black Rails in much of the northeastern region of its historical range, with possible extirpation in some states (Watts, 2016). They are notoriously difficult to detect due to their rarity, small size, and secretive nature, in addition to the fact they are found in habitat that is difficult to traverse. They prefer marshes with relatively shallow water dominated by *Spartina* marsh grass species and have been found to occupy managed impoundments (Roach & Barrett, 2015). Rails in general are weak flyers and tend to opt for a speedy foot retreat. The fact that they flush only when necessary makes them even less likely to be detected visually (Davidson, 1992; Stuart, 1920).

Currently, the most effective method for detecting secretive marsh birds is the Standardized North American Marsh Bird Survey Protocol (Conway, 2011). These surveys are carried out in suitable habitat at multiple points separated by at least 400 m. After a period of silent listening, a recording of a series of different calls of each target species are broadcast, each followed by phases of listening for a response. These surveys are time consuming, minimally taking 6 minutes to complete a single point. The protocol states that surveys should be completed in prescribed time windows either during the morning (thirty minutes before sunrise until two hours after) or evening (two hours before sunset until thirty minutes after). Ideally, surveys should be conducted at the same point location three times during the breeding season with at least 10 days between visits. Studies have shown that the peak in vocalizations for Black Rail can be heard anywhere from March to late June depending on the region (Flores & Eddleman, 1995; Kerlinger & Wiedner, 1990; Spear et al., 1999). The recommended two-month callback survey window for rails breeding in North Carolina is between April 1st and May 31st, annually.

Callback surveys have limitations including the fact that they are most effective during the breeding season. They rely on the bird making its presence known through a vocal response, and they are likely to only elicit a response from breeding males (Legare et al., 1999). Their effectiveness relies on the trained surveyor hearing and recognizing a responding individual, and they require the effort of multiple technicians over an extended period of time. A lab-based molecular detection method could be a way to circumvent some of these limitations.

Animals release DNA into the environment during everyday activities resulting in shedding skin cells or depositing intestinal epithelial cells along with waste (Valiere & Taberlet, 2000). This ‘environmental’ DNA (hereafter, eDNA) can be extracted from a substrate sample to confirm the target organism’s presence in the area based on a unique sequence signature. This

technique has been used successfully to detect the presence of rare species such as the great crested newt (*Triturus cristatus*) in the United Kingdom (Biggs et al., 2015), and also in the early detection of a few individuals of an alien species, such as at the invasion frontiers of two species of Asian carp (*Hypophthalmichthys molitrix* and *H. nobilis*) in tributaries of the Great Lakes (Jerde et al., 2011). Environmental DNA can be especially useful for organisms that are difficult to detect visually or by sound. It holds great promise for the detection of elusive animals such as a secretive marsh birds.

Several aquatic and semi-aquatic species have been detected successfully using eDNA. For example, environmental DNA has been used to detect invasive Asian carp in the Great Lakes at sites where traditional survey methods had failed to detect it (Jerde et al., 2011). Positive detections in areas lacking boat traffic or where boats were present only downstream, and multiple detections from the same place at different time points enabled the authors to reasonably rule out other possible sources of DNA such as ballast water from ships and excrement from fish-eating birds or mammals.

Small numbers of aquatic organisms can be detected from eDNA and based on occupancy of limited duration. In an experiment conducted in the lab, aquaria with specified densities of the invasive American Bullfrog (*Lithobates catesbeianus*) were compared to natural ponds where the frogs had or had not been detected (Ficetola et al., 2008). Using eDNA, they were able to detect their presence in ~ 40% of ponds where only one or two non-breeding individuals were found, and in ~ 80% of ponds where frogs were observed breeding. In these aquarium experiments, they were able to detect the presence of one tadpole in three liters of water after the tadpole had been present for only 24 hours.

The use of eDNA can help better understand circannual and spatial patterns of activity in species that are difficult to follow or where data collection could result in disrupting those activities. A study of hibernacula of northern map turtles (*Graptemys geographica*) in Canada used eDNA collected from under lake ice to provide valuable insights into where the turtles overwintered (Feng et al., 2020). The experiment was unique in that it focused on a rather inactive time for the species.

Environmental DNA can also be used to assess general biodiversity through targeting a region highly conserved among species and sequencing the resulting fragments to determine species in a process called meta-barcoding. In a study testing samples taken off the coast of Denmark, researchers were able to ascertain the biodiversity of the coastal area and also detected a rare vagrant fish, the European Pilchard (*Sardina pilchardus*) (Thomsen et al., 2012) . In addition to fish species, this study also detected several common seabirds. This included the unexpected detection of a Red-Throated Loon (*Gavia stellata*) which was later validated by a bird watch database confirming the species presence during the time of sampling. Thus, birds leave detectable amounts of their DNA in the environment and there is a good chance rare species could be detected using eDNA. Sea bird eDNA (from cormorants *Phalacrocorax spp.*) was also identified in another study of biodiversity, in near shore water samples of California kelp forests (Port et al., 2016). A recent metabarcoding study of U.K. pondscapes not only found DNA from aquatic avian species, such as ducks and coots, but also terrestrial species such as jays, owls, and even mammals (Harper, L. et al., 2020).

More recent studies have used metabarcoding using environmental DNA to detect bird species from water. An avian metagenomic study was completed by designing universal bird primers that targeted the 12S RNA and using MISeq technology to recover the sequences (Ushio

et al., 2018). The main validation used water samples collected from the Yokohama zoo, and confirmed the presence of reads corresponding to the non-native species known to be in the enclosures and not from local wildlife. All the zoo samples returned results corresponding with the species present in the respective enclosures. The presence of a smaller number of reads from species from other enclosures was attributed to the zoo staff transporting DNA and shared husbandry equipment. The authors also analyzed samples taken from a local pond. These samples detected DNA sequences from expected known local bird species, though species presence was not physically validated by other survey methods before or after.

Environmental DNA provides particular promise in being a tool for detecting species of conservation concern. A recent study made use of samples from drinking water sources from both captive and wild populations of the Gouldian Finch (*Erythrura gouldiae*), an endangered desert species from Australia. In this environment, finch flocks are known to congregate daily at watering holes, and the authors collected samples from these water sources to examine the efficacy of using eDNA to detect this species in the wild (Day et al., 2019). They were able to detect DNA from watering holes where wild finch flocks had been observed and from water dishes used by captive birds. Detectable eDNA was found to persist in the captive finches' water dishes up to 144 hours after the dishes had been removed from the enclosure and exposed to sunlight.

Innovative sampling methods have been employed to collect DNA non-invasively from birds. One such source of eDNA is saliva from food remains. In Costa Rica, Scarlet Macaw (*Ara macao*) DNA was recovered from partially eaten almond fruits (Monge et al., 2020). The researchers were able to collect enough quality DNA from the food remains to do population genetic tests such as sexing and amplifying several microsatellite loci. These studies demonstrate

that DNA deposited into the environment by rare birds can be recovered and analyzed using PCR methods.

An important step in developing an eDNA protocol is determining what informative part of the genome to target. For detecting diploid species, mitochondrial genes have the advantage that individuals carry only one haplotype inherited through the maternal line. In addition, mtDNA is known to be highly variable among species, with higher rates of mutation accumulation than nuclear genomes, probably due to small effective population size and haploid inheritance (Neiman & Taylor, 2009). Many meta-barcoding studies employ mitochondrial sequences that can be amplified with universal primers, are short, and unique to species. Cytochrome c oxidase subunit I (COI), the "bar-coding" gene, works well because of its low within-species and high between-species sequence variability (Hebert et al., 2004). First identified while looking for a short region of DNA to compare between species to help inform lepidopteran phylogenies, its general usefulness in distinguishing among species more widely was recognized (Hebert et al., 2003). A subsequent study showed that 260 bird species could be distinguished using one 648-base pair region of the COI gene (Hebert et al., 2004).

A significant methodological consideration for eDNA is the type and amount of environmental sample that would be optimal to reliably detect an organism. A study investigating water sample volume and eDNA detectability concluded that for three macroinvertebrate species, *Ancylus fluviatilis*, *Baetis buceratus* and *Gammarus pulex*, there was little difference in detectability between a 0.25L sample and a 2L sample (Mächler et al., 2016). The authors suggested using a minimum of 1L sample and analyzing at least 14µL of DNA extract (eDNA concentrations of 0.306 to 7.46 ng DNA per 15µL reaction) to provide repeatable results.

Concentrating eDNA from a larger volume of water requires a filtering step. Filters made of different materials and pore sizes can be used to concentrate eDNA prior to DNA extraction. Two separate studies looked at the effect of filter type and extraction method on DNA yield. Both studies tested cellulose nitrate, polyether sulfone, and polycarbonate track-etched filters with two different DNA extraction kits, the DNeasy tissue kit (Qiagen) and the MoBio Power water kit (now sold as Powersoil Kit by Qiagen). Each study concluded that the highest DNA yields were obtained when a DNeasy tissue kit (Qiagen) was paired with a cellulose nitrate filter. A close second was the combination of the MoBio Power Water kit paired with a polyester sulfone filter (Djurhuus et al., 2017; Hinlo et al., 2017).

Compared to conducting callback surveys, eDNA samples are less time consuming to collect. An additional advantage of this method is that the individual collecting the sample does not need specialized training or be able to reliably identify the target species. As part of a study in the United Kingdom, citizen scientists collected water samples from wetlands in the known range of the protected great crested newt (Biggs et al., 2015).The results showed that, given only a written description of the sampling procedure, citizen scientists were able to produce samples that had a 91% positive detection rate. This study also looked at the effect of environmental variables on newt detectability. The variables ranged from water quality and pond size to habitat suitability index score and the presence of waterfowl. The only environmental factor that was significant was habitat suitability. Many of the studies that compared traditional survey methods with eDNA methods found that eDNA was nearly as accurate as, and was more time and cost effective, than the survey method (Ficetola et al., 2008; Thomsen et al., 2012).

Environmental DNA is a relatively new method for detecting macro organisms and the methodology is still in development. Many questions remain about its effectiveness. For

example, it is unclear how long eDNA persists in the environment at a level that is detectable. In an attempt to determine how long DNA might persist in the environment, experiments have been conducted in tandem to test this in a freshwater ecosystem and in a controlled mesocosm (Thomsen et al., 2012). The mesocosm experiment consisted of introducing larval amphibians at different densities to purpose-built outdoor artificial ponds outfitted with "clean" water, typical pond plants, and algae. The water was then tested for the presence of the target species DNA before, after, and during the occupancy of the larvae. Target DNA was only detectable for up to two weeks after the removal of the larvae. This suggested that detectable amounts of DNA under these general environmental conditions are likely to represent only organisms currently occupying the area.

The objective of my study was to develop and validate a molecular diagnostic test using eDNA to detect the presence of a secretive marsh bird, the Black Rail. To my knowledge, this was the first study to investigate whether DNA from rare secretive marsh birds can be detected from environmental samples. Previous eDNA studies guided basic sample collection and preparation techniques that I was able to adapt. My specific aims were (1) to develop a diagnostic PCR-based test specific to Black Rail that is reliable in amplifying a signal from a small amount of template DNA, (2) to demonstrate the efficacy of the test using environmental samples collected from locations where occupancy was confirmed, and (3) to identify factors affecting detectability to be able to make recommendations about sample collection and handling for future eDNA studies.

Methods

Field sample collection

To validate the diagnostic eDNA assay, I needed positive controls from sites where Black Rail were detected and likely deposited DNA into the environment. During the 2018 and 2019 breeding seasons, I conducted auditory surveys using the Standardized Marsh Bird Callback Protocol (Conway, 2011) in appropriate emergent vegetation wetlands along the North Carolina coastal plain. Sites were chosen based on accessibility and suitability of the habitat for rails. I considered wetland habitats to be suitable for Black Rail if they were at least half a hectare in size, dominated by native marsh grasses particularly *Spartina* species, and irregularly flooded to a water level at or below 15 cm. These included oligohaline estuarine marshes and freshwater landlocked marshes on public and private lands. Each site was visited 3 times, with at least 9 days between visits, per the callback protocol (Conway, 2011).

The recording used for surveys consisted of a three-minute passive listening phase then thirty seconds of Black Rail *ki-ki-kur*, a minute of silence followed by another *ki-ki-kur*, a minute of silence then a Black Rail *growl*, two minutes of silence then thirty seconds of a King Rail *kek*, thirty seconds of silence then a King Rail *grunt*, then a final two minutes of silence. The King Rail calls were added for two reasons. King Rails are another candidate species for developing an eDNA diagnostic test, and these surveys supplemented a resurveying project documenting regional King Rail occupancy. Moreover, broadcasting calls of other conspecific rails along with a Black Rail call increases likelihood of eliciting a vocal response from Black Rails (Nadeau et al., 2013). The calls were broadcast at maximum volume (~80 dB) using a game caller (FoxPro NX4 or FoxPro Spitfire). Surveys were performed from half an hour before to three hours after sunrise or within three hours before and a half hour after sunset.

Between April 17th and July 8th, 2018, I conducted 186 call back surveys for Black Rails at 66 individual survey points along the North Carolina coast. I made three visits to each site with at least nine days between visits (Figure 1-1). Between March 22nd and July 14th, 2019, I conducted 434 callback surveys for Black Rails at 155 sites, with three visits to each site at least 17 days between visits (Figure 1-2). Although each survey site was visited three times, each individual survey point was not always visited three times (see Table 1-6 for more detail). Occasionally, individual points were skipped during a subsequent visit if the point was inaccessible due to tidal water level changes. Individual points were sometimes added when new suitable habitat was discovered. The first survey visit to a site was often exploratory and typically took longer to complete the first survey round. A survey point or two would occasionally be added during the second visit when there was extra time available during the survey time window; these individual points received only two visits. During the 2019 season, 14 points at 3 sites near Wilmington, NC were surveyed once then abandoned due to logistical constraints. These sites had sparse monotypic vegetation structure and their high, tidally-influenced water level made the habitat less optimal for Black Rail. Out of 115 points surveyed during 2019, 26 were points that I had surveyed in 2018.

In addition to the samples collected during my own surveys I also received samples from collaborators working on Black Rail populations in South Carolina and Florida. Collaborators on this project were Christine Hand, Wildlife Biologist, South Carolina Department of Natural Resources and Amanda Bessler, Zone Ecologist, U.S. Fish and Wildlife Service, Region 4 Inventory and Monitoring Network (Florida) (see Table 1-6 for more detail). Motion sensor camera traps were used at the Bear Island NWR site in South Carolina to collect visual data on Black Rails during both the 2018 and 2019 seasons(Hand et al., 2019). Hand and S. McRae

collected environmental samples where Black Rails were detected on camera. The USFWS had technicians conducting audio surveys for Black Rails and other marsh birds during the 2018 season at three different refuges in Florida: St. Marks NWR, St. John NWR, and St. Vincent NWR. The samples contributed by the USFWS were at sites of positive detections from these surveys.

Water samples. When a Black Rail was detected, eDNA samples were collected as close as possible to its perceived location, targeting areas of higher probability for rail traffic such as footprints and tunnels. Samples were also collected from locations where Black Rail had been detected by myself or others earlier in the season, regardless if there was a detection at the time of sampling. Where possible, I collected a 1-2L water sample in a 2-L sterilizable PC square media bottle (TriForest) or a 2-L single-use Whirl-Pak (Nasco). A metal loop with a 0.5m handle holding a single-use plastic disposable cup (Hefty) was used as a ladle to extend the reach of the observer when collecting some water samples. The loop was sanitized with bleach and dried between uses. Other samples were collected directly into the storage vessel. Water samples were either refrigerated at 4°C and then vacuum-filtered within 24 hours, or frozen at -20°C as soon after collection as possible and always within 8 hours. Frozen water samples were thawed and vacuum filtered in the lab at a later date (mean = 162 days after collection, range = 6 to 640).

Soil samples. Soil samples were collected by filling 50mL falcon tubes at most sites of Black Rail detection. Care was taken to not sample directly where the observer had walked to avoid possible contamination from outerwear. Areas of exposed moist soil next to gaps or tunnels in the vegetation were targeted. Soil samples were stored at -20ºC until DNA extraction.

Sample processing

Each water sample was filtered to concentrate the eDNA. The water was vacuum filtered using a vacuum pump (KNF, Trenton, NJ) and filtering apparatus (Nalgene ThermoFisher Scientific, Waltham, MA) through a 47 mm diameter cellulose nitrate filter with a pore size of 0.45 µm (Whatman, ThermoFisher Scientific, Waltham, MA). The filtering apparatus was dissembled and soaked in a bleach solution for a minimum of 10 minutes, thoroughly rinsed with deionized water and dried before each use. I discovered that allowing the sediment in the samples to settle and filtering the top of the water that had less particulate first before the more turbid water at the bottom helped reduce filter time. Typically, two filtering units would be run at the same time to process one 2-L sample resulting in two cellulose nitrate filters for the sample. Occasionally, if the filtering took longer that 24 hours the filters would be removed and replaced. The first set of filters were preferentially chosen for extraction and any later filters were kept as back-ups which were ultimately not used.

DNA extraction

Due to the challenging nature of environmental samples containing potential inhibitory compounds, several extractions methods were tested to determine which would be the most effective at removing inhibitors from these samples. The methods tested were an ammonium acetate and ethanol precipitation ('Salting-out method', protocol based on Bruford et al., 1998), DNeasy tissue kit (Qiagen), PowerWater kit (Qiagen), and PowerSoil kit (Qiagen). All extractions using kits were performed following manufacturer's instructions. For each extraction, either half a cellulose nitrate filter disk with deposited precipitate or 0.4-0.7 g of soil was used. Filters were cut into small pieces with a clean razor blade to facilitate fit in the extraction tube and increase surface area exposed to lysis buffer.

To compare the performance of the DNA extraction methods, environmental samples were tested with each method and extracts tested for the presence of inhibitors by spiking a standardized PCR reaction using the BLRA COI2 primers with genomic DNA template extracted from Black Rail blood. If the sample contained compounds inhibitory to PCR, this would be evident from the product (or lack thereof) in comparison to unspiked positive controls in the same run.

Species-specific assay development

Rails (Family Rallidae) are an ancient lineage of birds (Taylor & Van Perlo, 1998). This has the advantage that due to sequence divergence they are relatively different from other birds genetically reducing the possibility of false positives based on other avian DNA during the amplification of eDNA. However, published universal bird primers for the COI gene (Hebert et al., 2004) would not amplify mtDNA sequence in the development of the diagnostic test. Thus, the primer design process required two stages. First, rail-specific primers within the COI gene were (Table 1-1, BLRA COI primer set 1) designed based on published Rallidae sequences in Genbank (Clark et al., 2016; alignment sequence list in Table 1-2). I acquired two Black Rail blood samples (eastern subspecies) from Louisiana (contributed by collaborator Erik Johnson, Louisiana Audubon), and extracted DNA from them to amplify and sequence 591bp of the COI gene (the most complete Black Rail COI published sequence has 652bp: Genbank GI #: 116832367). Two King Rail (*Rallus elegans*) blood samples collected from Mackay Island National Wildlife Refuge in North Carolina (during the course of the McRae lab's long-term monitoring project of the population) were amplified using the Rail COI primer set (Primers in Table 1-1, King Rail sequence in Appendix Figure A2) and sequenced. No previously published COI sequence existed for the Clapper Rail (*Rallus crepitans)* so I sequenced muscle samples of

two individuals (contributed by local hunter, G. Huntsman) using the Rail COI primer set (Primers in Table 1-1, Clapper Rail COI sequence in Appendix Figure A3). Sanger sequencing was performed using a standard Big Dye protocol on an ABI 3130 Genetic Analyzer. The reactions used to create the template were 25µL reactions with 11.5µL nuclease free water 2.5 APEX 10x Mg free PCR buffer, 0.75µL 5µM $MgCl_2$ (final concentration: 0.15µM), 2.5µL 10µM dNTPs (final concentration: 1µM), 1.25µL of each BLRA COI2 primer (final concentration: 0.5µM each), 0.25µL Taq Polymerase, and 2µL template (DNA extracted from a single Black Rail blood sample). The thermocycling conditions were 95°C for 2 min, followed by 40 cycles of 95°C for 30s, 60°C for 30s, and 72°C for 1min, followed by a 72°C hold for 5 min and an infinite hold at 12°C. Approximately 8µL of the PCR product was visualized on a 2% agarose gel to check that the reaction was successful. The remainder of the PCR product was cleaned and purified using a ExoSAP-it product clean up kit per the manufacturer's instructions (ThermoFisher). The sequencer output was then trimmed and the forward and reverse reads were aligned to create a consensus sequence using the program Geneious.

My aim was to develop a PCR-based diagnostic test that would uniquely amplify Black Rail DNA, so I compared the COI sequence of other North American rail species to identify sites of divergence and identify nucleotide substitutions unique to the Black Rail. Using the program TCOFFEE (Notredame et al., 2000), I aligned the COI sequences I generated for Black Rail, King rail and Clapper Rail to published COI sequences downloaded from National Center for Biotechnology Information Genbank database from a selection of rail species sympatric with the Black Rail over at least part of its eastern range: Yellow Rail (*Coturnicops noveboracensis*), Virginia Rail (*Rallus limicola*), Common Gallinule (*Gallinula galeata*), Sora (*Porzana carolina*), and American Coot (*Fulica americana*) (Table 1-2).

I used the online program Primer3 (Koressaar & Remm, 2007) to design primers targeting portions of the Black Rail COI sequence with the most mismatches compared to other rails. The resulting primer pair (Table 1-1, BLRA COI2) amplified a 219-bp segment of the Black Rail COI gene. I tested for cross-species amplification with genomic DNA from blood samples of sympatric rail species (King Rail, Clapper Rail, Yellow Rail, Common Gallinule, Virginia Rail, Sora) using traditional PCR (BioRad T100 thermocycler). The reactions were comprised of 6.1µL nuclease free water, 1µL APEX 10x Mg free PCR buffer, 0.3µL 5µM $MgCl_2$ (final concentration: 0.15µM), 1µL 10µM dNTPs (final concentration: 1µM), 0.5µL of each BLRA COI2 primer (final concentration: 0.5µM each), 0.1µL Taq Polymerase, and 2µL template. Thermal cycling conditions were set at 95°C for 2 min, followed by 40 cycles of 95°C for 30s, 60°C for 30s, and 72°C for 1min, followed by a 72°C hold for 5 min and an infinite hold at 12°C. PCR products were visualized by electrophoresis on a 2% agarose gel. The primers were also tested for their detection limits using the same PCR reaction conditions and a dilution series of Black Rail PCR genomic DNA.

SYBR Green qPCR

Considering the low amounts of target DNA expected in the environmental samples, I switched to a more sensitive detection method. This change was supported by the results of a test run of qPCR with the same primer set and SYBR green showing eDNA having detectable amplification at lower concentrations of target DNA than the conventional PCR limit of detection. The BLRA COI2 primer pair was then tested using real-time or quantitative PCR (qPCR) (Bio-Rad CFX Connect Real-Time System) with SYBR green reporter (BIO-RAD SsoAdvanced Universal IT SYBR® Green Supermix). The reactions consisted of 4µL nuclease free Fisher water, 5µL SsoAdvanced Universal IT SYBR® Green Supermix, and 0.5µL of each

BLRA COI2 primer (final concentration: 0.05µM each) added to 2µL template. Thermal cycling conditions were set at 95°C for 2 min, followed by 40 cycles of 95°C for 30s, 60°C for 30s, and 72°C for 1min followed by the pre-set melt curve. A purified product was used as a standard for quantifying eDNA using the new assay. This was made by running a traditional PCR in 25µL reactions with 11.5µL nuclease free water 2.5 APEX 10x Mg free PCR buffer, 0.75µL 5µM $MgCl_2$ (final concentration: 0.15µM), 2.5µL 10µM dNTPs (final concentration: 1µM), 1.25µL of each BLRA COI2 primer (final concentration: 0.5µM each), 0.25µL Taq Polymerase, and 5µL template (DNA extracted from a single Black Rail blood sample). Approximately 8µL of the PCR product was visualized on a 2% agarose gel to check that the reaction was successful. The remainder of the PCR product was cleaned and purified using a MoBio Ultraclean Gelspin DNA Purification Kit per the manufacturer's instructions. The end concentration was 1.18 ng/µL (5x 10^9 copies), quantified with a Qubit 3 fluorometer (Invitrogen) using Qubit high sensitivity dsDNA assay reagents (ThermoFisher). A standard curve from $2.34x10^{-4}$ ng target DNA to $2.34x10^{-11}$ ng target DNA in 1/10 dilutions was made to help quantify unknown samples and determine a detection limit for the assay. This primer set with SYBR green assay was tested for cross-species amplification using genomic DNA extracted from the blood of non-target species (Clapper Rail, King Rail, Yellow Rail, Common Gallinule, Virginia Rail, Sora).

TaqMan Dual MGB Probe Assay

To resolve the problem of non-specificity, I turned to TaqMan probe chemistry to develop a new qPCR assay. SYBR green produces a signal by binding with any double stranded DNA present, thus relying on the two primers for specificity. The TaqMan assays require greater sequence specificity to produce an amplification signal due to requiring complementarity with both primers and an additional probe. A probe is a short oligonucleotide that is designed to

specifically match an internal section of a primer product. The probe contains a reporter molecule that fluoresces when activated indicating the presence of the target sequence (Livak et al., 1995). This probe chemistry works by having the fluorescent reporter molecule at the 5' end of the probe and a quencher molecule at the 3' end that suppresses the reporter molecule fluorescence while they are in close proximity to each other. If the probe target sequence is present the probe sits between the two primers in the assay during the aneling phase. During extension *taq* DNA polymerase removes the probe thus detaching the reporter molecule which then fluoresces in the absence of the quencher molecule's immediate proximity(Livak et al., 1995). This fluorescence is recorded by the machine after each PCR cycle. The fluorescence is cumulative thus samples with a higher starting quantity of target DNA require fewer cycles to produce a detectable fluorescence. The TaqMan MGB probe also has a minor grove binder (MGB) molecule that help stabilize the probe/template complex and increase the melting temperature of the complex allowing for the use of shorter probe sequences (Kutyavin et al., 2000). I designed this assay with the ThermoFisher TaqMan genotyping assay design tool that was recommended by the manufacturer but is normally used for distinguishing SNPs for genotyping purposes. The species-specific primer pair targeted a 65-bp variable region of the COI gene bracketing a central region where two alternative probes labeled with different fluorophores were designed that differed by a couple of base pairs between the Black Rail and the other rails (Table 1-3, TaqMan dual probe assay). There were at least 7 SNPs in the primers and probes combined between the Black Rail and the other sympatric rail species (Clapper Rail, Yellow Rail, Common Gallinule, Virginia Rail, Sora). The probe labeled with VIC (Table 1-1, BLRA_TD_Probe1) was an exact match to Black Rail only and the probe labeled with FAM

(Table 1-1, BLRA_TD_Probe2) matched the consensus sequence for the other sympatric rail species.

Real-time PCR was performed in duplicate 10µL reaction volumes containing 4µL nuclease free water (Fisher Scientific), 1µL10X assay mix (final reaction concentration: 0.2µM each probe, 0.9µM each primer), 5 µL 2X TaqMan Environmental Master Mix 2.0 (Applied Biosystems) and 2µL template. Following Applied Biosystems TaqMan protocol, thermal cycling conditions were set at 95°C for 10 min, followed by 50 cycles of 95°C for 15s and 60°C for 1 min. A standard curve of genomic DNA in 1/10 dilutions was made to help compare unknown samples and determine a detection limit for the assay. Each set of reactions contained two no-template controls (equal volume of nuclease free water) to detect contamination. This assay was tested with other non-target species (King Rail, Clapper Rail, Yellow Rail, Common Gallinule, Virginia Rail, Sora). Since the primers and probes were sold as a mixture, no additional optimization of this assay was possible to reduce cross-species amplification.

TaqMan Single MGB Probe Assay

Following previous eDNA studies (Dunn et al., 2017; Takahara et al., 2013), I used the TaqMan gene expression assay design tool for designing a new assay using a fluorescent probe and quencher, providing a signal for only one target sequence. I identified a downstream segment of the COI gene in order to include two SNPs distinguishing Black Rail from the other sympatric species in the probe sequence and designed an assay in the identified region with an 89-bp product. The assay mix consisted of a primer pair and a fluorescently labeled probe with a quencher (Table 1-1, TaqMan single probe assay). Real-time PCRs were performed in triplicate 10µL reaction volumes containing 4µL nuclease free water (Fisher), 1µL 10X assay mix (final reaction concentration: 0.25µM probe, 0.9µM each primer), 5µL 2X TaqMan Environmental

Master Mix 2.0 (Applied Biosystems) and 2µL template. Following Applied Biosystems' TaqMan protocol, thermal cycling conditions were set at 95°C for 10 min, followed by 50 cycles of 95°C for 15s and 62°C for 1 min. A standard curve from genomic DNA in 1/10 dilutions series was made to help compare unknown samples and determine a detection limit for the assay. Each set of reactions contained three no-template controls (equal amounts nuclease free water) to detect contamination. This assay was tested with a panel of non-target species (King Rail, Clapper Rail, Yellow Rail, Common Gallinule, Virginia Rail, Sora).

Affinity Plus Probe Assay

A new assay was designed in the same variable region of the COI gene as the TaqMan dual probe assay but using an Affinity Plus probe with an additional SNP in the probe sequence and different primers targeting a 65-bp region (Table 1-1, Affinity Plus assay). Primers from Bioneer were ordered separately from the probe (Integrated DNA Technologies). The Affinity Plus probe's Iowa Black dark quencher preforms the same function as the non-fluorescent quencher (NFQ) that is attached to the TaqMan probe, they are simply different proprietary versions of the same type of molecule. The Affinity Plus probe does not have an MGB, which is a proprietary molecule owned by ThermoFisher, but does have Affinity Plus monomers which are proprietary molecule owned by Integrated Technologies. These monomers are locked nucleotides that are placed strategically throughout the probe that, like the MGB, help stabilize and increase the melting temperature of the probe/template complex (Owczarzy et al., 2011). Real-time PCR was performed in triplicate 10µL reaction volumes containing 2.75 µL nuclease free water (Fisher), 0.25 µL of 10µM probe (final reaction concentration: 0.25µM), 1µL each 10µM primer (final reaction concentration: 1µM each), 5µL 2X TaqMan Environmental Master Mix 2.0 (Applied Biosystems), and 2µL template DNA. Following Applied Biosystems TaqMan

protocol, thermal cycling conditions were set at 95°C for 10 min, followed by 50 cycles of 95°C for 15s and 60°C for 1 min. A standard curve from $2.36x10^{-6}$ng target DNA to $2.36x10^{-9}$ng target DNA in 1/10 dilutions was made to quantify unknown samples and determine a detection limit for the assay. Each set of reactions contained three no-template controls to detect contamination.

Modified Affinity Plus Probe Assay

A modified version of the Affinity Plus probe assay used newly designed primers (Bioneer) in combination with the same Integrated DNA Technologies quencher probe (Table 1-1, modified Affinity Plus assay). Primers and probe were ordered separately enabling me to adjust their concentrations independently to better optimize the reaction. Real-time PCR was performed in triplicate 10µL reaction volumes containing 2.75 µL nuclease free water (Fisher), 0.25 µL of 10µM probe (final reaction concentration: 0.25µM), 1µL each 10µM primer (final reaction concentration: 1µM each), 5µL 2X TaqMan Environmental Master Mix 2.0 (Applied Biosystems), and 2µL template DNA. Following the Applied Biosystems TaqMan protocol, thermal cycling conditions were set at 95°C for 10 min, followed by 50 cycles of 95°C for 15s and 60°C for 1 min. A standard curve from $2.36x10^{-6}$ng purified target DNA to $2.36x10^{-9}$ng target DNA in 1/10 dilutions was made to help quantify unknown samples and determine a detection limit for the assay. Each set of reactions contained three no-template controls.

BLRA COI2 Primers with Affinity Plus Probe: A combination assay as a diagnostic test for Black Rail eDNA

Each of the assays had problems of cross-species amplification with related non-target species, and additionally, sporadic signal in the no template water controls (NTC). I suspected that small product size of the TaqMan assays could contribute to these inconsistent results (Table

1-1, TaqMan dual probe assay: 65bp, TaqMan single probe assay: 89bp, Affinity Plus assay: 65bp, modified Affinity Plus assay: 63bp). To address this, I developed an assay amplifying a larger fragment while retaining the quencher probe: I combined the original Black Rail COI2 primers which has a 219-bp product size (Table 1-1: BLRA COI primer set 2, 219bp product) with the Affinity Plus probe to add specificity.

A purified product was used as a standard for quantifying eDNA using the new assay. This was made by running a traditional PCR in 25µL reactions with 11.5µL nuclease free water 2.5 APEX 10x Mg free PCR buffer, 0.75µL 5µM $MgCl_2$ (final concentration: 0.15µM), 2.5µL 10µM dNTPs (final concentration: 1µM), 1.25µL of each BLRA COI2 primer (final concentration: 0.5µM each), 0.25µL Taq Polymerase, and 5µL template (DNA extracted from a single Black Rail blood sample). Approximately 8µL of the PCR product was visualized on a 3% agarose gel to check that the reaction was successful. The remainder of the PCR product was cleaned and purified using a MoBio Ultraclean Gelspin DNA Purification Kit per the manufacturer's instructions. The end concentration was 1.26 ng/µL (5.46 x 10^9 copies), quantified with a Qubit 3 fluorometer (Invitrogen) using Qubit high sensitivity dsDNA assay reagents (ThermoFisher).

Separate primer and probe aliquots allowed me to vary their concentrations to optimize the assay. A test of this combination assay consisted of a plate with five replicates of a dilution series ranging from 109,145 target copies to 1 copy, in 1/10 dilution increments, as quantification standards. On the same plate, a test of non-target species was done in duplicate using two individuals from each non-target species (total of 4 reaction for each species, King Rail, Clapper Rail, Virginia Rail, Sora, Yellow Rail, and Common Gallinule) to test for cross amplification.

Real-time PCR was performed with reactions containing 2.75µL nuclease free water (Fisher), 1 µL of each 10µM BLRA COI2 primers (final concentration: 1µM each primer), 0.25µL of 10µM Affinity Plus probe (final concentration: 0.25µM), 5µL 2X TaqMan Environmental Master Mix 2.0 (Applied Biosystems) and 2µL template. Based on the Applied Biosystems TaqMan protocol, thermal cycling conditions were set at 95°C for 10 min, followed by 50 cycles of 95°C for 15s and 60°C for 1 min.

Following publication standards for quantitative real-time PCR experiments, I set the limit of quantification (LOQ) at 95% amplification calculated from the five replicates of a dilution series of purified product (Bustin et al., 2009). The 95% was calculated by plotting the percent of the replicates that had positive amplification and the concentration and using the equation of that line to determine where 95% positive would be if the dilution series had more resolution. Following other eDNA the limit of detection (LOD) was set at the point where only one out of five replicates produced a signal (Dunn et al., 2017; Harper, K. et al., 2020). The same dilutions were run using traditional PCR with the same primers and the sensitivity of the two methods compared. The traditional PCR dilution series reactions consisted of 5.6µL nuclease free water, 1µL APEX 10x Mg free PCR buffer, 0.3µL 5µM $MgCl_2$ (final concentration: 0.15µM), 1µL 10µM dNTPs (final concentration: 1µM), 1µL of each BLRA COI2 primer (final concentration: 1µM each), 0.1µL Taq Polymerase, and 2µL template. The dilutions ranged from 10,914,478,000 target copies to 10 copies, in 1/10 dilution increments. The information from the five replicate qPCR standard curve dilution series was used to reduce the number of standards needed on each eDNA analysis plate. Duplicates of two standard dilutions above the LOQ and two dilutions below the LOQ including the LOD were run on each subsequent eDNA analysis plate to ensure plate results were comparable.

The combination assay was used to validate all of the Black Rail eDNA samples I extracted. For an eDNA sample to be considered positive for Black Rail DNA, it had to produce a qPCR signal at a cycle number (Cq) lower than or equivalent to the LOD; to be reliably quantified, it had to score a Cq that preceded or equaled the LOQ. Any non-target species that produced a qPCR signal at a Cq that exceeded the LOD was considered a negative result.

A diluted master mix was tested because eDNA samples are notorious for having inhibitors and other eDNA studies have reduced their effects by dilution with water (Biggs et al., 2015; Dunn et al., 2017). The diluted master mix contained 4.35µL nuclease free water (Fisher Scientific), 0.64 µL of each 10µM BLRA COI2 primer (final concentration 0.64µM each primer), 0.16 µL of 10µM Affinity Plus probe (final concentration: 0.16µM), 4.35µL 2X TaqMan Environmental Master Mix 2.0 (Applied Biosystems) and was added to 2µL template. Following Applied Biosystems TaqMan protocol, thermal cycling conditions were set at 95°C for 10 min, followed by 50 cycles of 95°C for 15s and 60°C for 1 min. This dilution was used for processing the eDNA samples.

Each eDNA sample was tested in triplicate on two separate plates, to reveal any batch plate effects, for 6 replicates per sample. Each eDNA plate also included the small standard dilution series (four dilutions from 10,900 copies to 10 copies, in duplicate), as well as a water NTC (Fisher Scientific). Among the eDNA samples collected from sites of positive Black Rail detection were two environmental samples collect from a marsh location where no Black Rails had been detected. This site was unlikely to support this species due to unsuitable vegetation and water depth and these samples were considered environmental negative samples.

Table 1-1. List of primer and probe sequences used in this study. For ease of reference the primers and probes are organized into primer sets and assay sets. All primers amplify portions of the COI gene.

Primer or Probe name	Sequence
Rail COI Primer set	Product size: 610bp
RailCOI1F	ACCTAATCTTTGGGGCCTGA
RailCOI1R	GGGTCGAAGAATGTGGTGTT
BLRA COI primer set 1	Product size: 591bp
LaterallusCOIF	AGCCGGCATAATTGGTACTG
RailCOI1R	GGGTCGAAGAATGTGGTGTT
BLRA COI primer set 2	Product size: 219bp
BLRA COI2F	CTTCCTCCCTCTTTCCTGCT
BLRA COI2R	GGATAGTGCGGGTGGTTTTA
TaqMan dual probe assay	Product size: 65bp
BLRA_TD_F	TCTGACTCCTTCCTCCCTCTTTC
BLRA_TD_Probe1	VIC-CTGCTACTACTAGCTTCATC-NFQ
BLRA_TD_Probe2	FAM-TGCTACTACTAGCTTCCTC-NFQ
BLRA_TD_R	TCCTGCTCCTGCTTCTACTGT
TaqMan single probe assay	Product size: 89bp
BLRA_TS_F	CCGCACTATCCCAATACCAAACC
BLRA_TS_Probe	FAM-ATGAGGACGGACCATACAA-NFQ
BLRA_TS_R	GGCAAGTACAGGTAGGGATAGTAGT
Affinity Plus probe assay	Product size: 65bp
BLRA_AP_F	TCTGACTCCTTCCTCCC
BLRA_AP_Probe	FAM-CTA+C+TA+GCTT+C+A+TCA-1ABkFQ
BLRA_AP_R	TCCTGCTCCTGCTTCTA
Modified Affinity Plus probe assay	Product size: 63bp
BLRA_MAP_F	TCCTTCCTCCCTCTTTCCTG
BLRA_AP_Probe	FAM-CTA+C+TA+GCTT+C+A+TCA-1ABkFQ
BLRA_MAP_R	CTGTTCCTGCTCCTGCTTCT

Table 1-2. List of sequences and their source used in primer and assay development alignments.

Species	Sequence source	Sequence length
Black Rail	consensus sequence from BLRA COI primer set 1 product	591bp
King Rail	consensus sequence from Rail COI primer set product	610bp
Clapper Rail	consensus sequence from Rail COI primer set product	610bp
Virginia Rail	Genbank GI#: 262179842	732bp
Yellow Rail	Genbank GI#: 116876767	633bp
Common Gallinule	Genbank GI#: 327494068	694bp
Sora	Genbank GI#: 294514074	694bp
American Coot	Genbank GI#: 117372219	697bp

Table 1-3. Non-target species mismatch alignment. Black Rail COI sequences are aligned with the COI sequence from sympatric rail species. Primers are underlined and probes are highlighted in dark grey with white lettering to distinguish them from the rest of the sequence. Mismatches are highlighted in black and have white lettering. The number of mismatches for each assay for each species is summarized to the right of the alignment.

BLRA COI2 primers

		Number of mismatches		
		Total	Forward	Reverse
Black Rail	CTTCCTCCCTCTTTCCTGCT, TAAAACCACCCGCACTATCC			
King Rail	CTCCCTCCTTCCTTCCTACT, TGAAACCCCCAGCCCTCTCC	9	4	5
Clapper Rail	CTCCCTCCTTCCTTCCTACT, TGAAACCCCCAGCCCTCTCC	9	4	5
Virginia Rail	CTCCCTCCCTCCTTCCTACT, TAAAACCCCCAGCCCTCTCT	8	3	5
Yellow Rail	CTCCCCACCCTCCTTCCTAAT, TAAAACCGCCCGCCCTGTCC	8	5	3
Common Gallinule	CTTCCCCCCCTCCTTCCTGCT, TAAAACCACCAGCCCTATCC	4	2	2
Sora	CTCCCCCCCTTCCTTCCTTCT, TAAAACCACCCGCCCTATCT	7	5	2
American Coot	CTCCCCCCCCTCCTTCCTACT, TAAAACCACCCGCCCTATCC	5	4	1

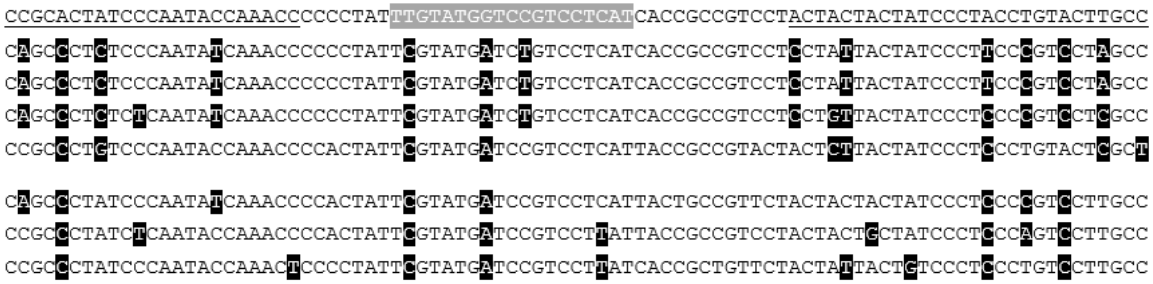
TaqMan dual probe assay

		Number of mismatches			
		Total	Forward	Probe	Reverse
Black Rail	TCTGACTCCTTCCTCCCTCTTTCCTGCTACTACTAGCTTCATCAACAGTAGAAGCAGGAGCAGGA				
King Rail	TCTGACTCCTCCCTCCTTCCTTCCTACTCCTCCTAGCATCCTCCACAGTCGAAGCAGGGGCAGGA	10	3	5	2
Clapper Rail	TCTGACTCCTCCCTCCTTCCTTCCTACTCCTCCTAGCATCCTCCACAGTCGAAGCAGGGGCAGGA	10	3	5	2
Virginia Rail	TCTGACTTCTCCCTCCCTCCTTCCTACTCCTTCTAGCCTCCTCCACAGTAGAAGCAGGGGCAGGA	9	3	5	1
Yellow Rail	TCTGGCTACTCCCACCCTCCTTCCTAATACTACTAGCTTCCTCCACAGTCGAAGCAGGAGTAGGC	11	5	3	3
Common Gallinule	TCTGACTCCTTCCCCCCTCCTTCCTGCTCCTCCTAGCATCCTCCATAGTAGAGGCAGGAGCAGGT	9	2	4	3
Sora	TCTGACTTCTCCCCCCCTTCCTTCCTTCTACTCCTAGCATCATCCACAGTAGAAGCAGGAGCAGGC	9	5	3	1
American Coot	TCTGACTCCTCCCCCCCTCCTTCCTACTACTCCTAGCATCATCCACAGTAGAAGCAGGAGCAGGC	7	3	3	1

Affinity Plus assay

	Number of mismatches		
Total	Forward	Probe	Reverse
8	2	4	2
8	2	4	2
7	2	4	1
9	4	2	3
7	1	4	2
8	4	3	1
6	2	3	1

TaqMan single probe assay



	Number of mismatches		
Total	Forward	Probe	Reverse
13	4	3	6
13	4	3	6
15	5	3	7
9	2	2	5
8	3	2	3
9	2	3	4
9	2	3	4

Modified Affinity Plus assay

	Number of mismatches		
Total	Forward	Probe	Reverse
10	4	4	2
10	4	4	2
9	4	4	1
10	5	2	3
8	2	4	2
10	6	3	1
8	4	3	1

Table 1-4. Alignment of the 219-bp fragment of the Black Rail COI gene amplified using the combination assay aligned with the COI sequences of sympatric rails. The BLRA COI2 primers are underlined and the Affinity Plus probe is highlighted in dark grey with white lettering. Mismatches are highlighted in black and have white lettering. The number of mismatches for each species is indicated in the table to right of the alignment.

		Number of mismatches			
		Total	Forward	Probe	Reverse
Black Rail	CTTCCTCCCTCTTTCCTGCTACTACTAGCTTCATCAACAGTAGAAGCAGGAGCAGGAACAGGC				
King Rail	CTCCCTCCATCCTTCCTACTCCTCCTAGCATCCTCCACAGTCGAAGCAGGGGCAGGAACAGGA	13	4	4	5
Clapper Rail	CTCCCTCCATCCTTCCTACTCCTCCTAGCATCCTCCACAGTCGAAGCAGGGGCAGGAACAGGA	13	4	4	5
Virginia Rail	CTCCCTCCCTCCTTCCTACTCCTTCTAGCCTCCTCCACAGTAGAAGCAGGGGCAGGAACAGGA	12	3	4	5
Yellow Rail	CTCCCACCCTCCTTCCTAATACTACTAGCTTCCTCCACAGTGGAAGCAGGAGTAGGCACAGGT	10	5	2	3
Common Gallinule	CTTCCCCCCTCCTTCCTGCTCCTCCTAGCATCCTCCATAGTAGAGGCAGGAGCAGGTACAGGT	8	2	4	2
Sora	CTCCCCCCATTCCTTCCTTCTACTCCTAGCATCATCCACAGTAGAAGCAGGAGCAGGCACAGGC	10	5	3	2
American Coot	CTCCCCCCCTCCTTCCTACTACTCCTAGCATCATCCACAGTAGAAGCAGGAGCAGGCACAGGA	8	4	3	1

Black Rail	TGAACCGTCTACCCTCCACTTGCCGGCAACCTAGCCCACGCAGGAGCCTCAGTAGACCTAGCC
King Rail	TGAACAGTATACCCCCCACTAGCTGGCAACCTAGCCCACGCAGGAGCTTCAGTAGACCTAGCC
Clapper Rail	TGAACAGTATACCCCCCACTAGCTGGCAACCTAGYCCACGCAGGAGCTTCAGTAGACCTAGCC
Virginia Rail	TGAACAGTCTACCCCCCATTAGCCGGCAACCTAGCCCACGCAGGAGCCTCGGTAGACCTAGCC
Yellow Rail	TGAACCGTATACCCTCCGCTTGCCGGCAACCTAGCCCACGCAGGGGCCTCAGTAGACCTAGCC
Common Gallinule	TGAACAGTCTATCCCCCACTAGCCGGCAACCTAGCACACGCAGGTGCTTCAGTAGACCTAGCT
Sora	TGAACAGTCTACCCCCCACTAGCTGGCAACCTAGCACATGCAGGAGCTTCAGTGGACTTAGCC
American Coot	TGAACAGTTTACCCCCCACTAGCCGGCAACCTAGCACATGCAGGCGCTTCAGTTGACCTAGCC

Black Rail	ATCTTCTCCCTGCACTTAGCAGGAGTTTCATCCATCCTGGGCGCCATCAATTTTATTACAACT
King Rail	ATTTTCTCACTCCACCTGGCAGGTGTATCATCNATCCTGGGCGCAATCAACTTTATCACGACC
Clapper Rail	ATTTTCTCACTCCACCTGGCAGGTGTATCATCTATCCTGGGCGCAATCAACTTCATCACGACC
Virginia Rail	ATTTTCTCACTTCACCTAGCAGGAGTATCATCCATCCTAGGTGCAATCAACTTCATTACAACC
Yellow Rail	ATCTTCTCCCTACACCTGGCAGGAGTATCATCCATCCTAGGCGCTATCAACTTTATCACAACA
Common Gallinule	ATCTTCTCCCTCCACTTAGCAGGAGTCTCATCTATCCTAGGTGCCATCAATTTCATCACAACT
Sora	ATCTTCTCCCTCCATCTAGCAGGTGTCTCATCCATTCTAGGCGCCATCAACTTCATCACAACC
American Coot	ATCTTCTCCCTCCACTTAGCGGGCGTCTCATCTATCCTAGGCGCCATCAATTTTATTACAACT

Black Rail	GCCATTAACATAAAACCACCCGCACTATCC
King Rail	GCCATTAACATCAAACCCCCAGCCCTCTCC
Clapper Rail	GCCATTAACATCAAACCCCCAGCCCTCTCC
Virginia Rail	GCCATTAACATAAAACCCCCAGCCCTCTCT
Yellow Rail	GCAATCAACATAAAACCCCCGCCCTCTCC
Common Gallinule	GCCATCAACATAAAACCACCAGCCCTATCC
Sora	GCTATCAACATAAAACCACCCGCCCTATCT
American Coot	GCCATCAACATAAAACCACCCGCCCTATCC

Statistical analysis

After the eDNA samples were assayed, the number of replicates that scored as positive (Cq before the LOD) was divided by the total number of replicates to produce an amplification success rate for each sample. Amplification success rate was related to variables pertaining to the sample's origin, collection and handling methods.

Statistical analyses were performed in R (R core team, 2019). To determine whether sample type, collection and storage methods and environmental conditions affected amplification success rate of eDNA samples taken from sites of Black rail detection, I conducted Generalized Linear Mixed Models with a binomial distribution. The binomial distribution was selected because the response variable, amplification success, was a rate. Mixed models allowed the inclusion of random effects of location and sample group in cases where locations had been sampled multiple times and where water and soil samples were taken from the same place at the same time. (I assumed that samples taken from the same area were more likely to have similar amplification success, and samples taken at the same time would similarly be non-independent).

All of the samples with camera trap detections came from Bear Island Wildlife Management area so more detail was needed when it came to sample location and groups. For analysis looking at the solely the camera data a separate sample group notation was used to more specifically look at samples taken from the same camera trap on the same day. When appropriate, subsets of the data were used to look at conditions specific to certain variables. For example, I used only water sample data to look at filtering effects, and only the camera detections to look at the number of days since a rail was recorded at the sampling site. These models were assessed using the Wald test (Pr (>|z|)) and bootstrap confidence intervals were set at 95%. The variables that were tested for an effect on detectability were the number of days

between collection and DNA extraction, the storage temperature of the water environmental sample before filtering, sample location, field detection method (audio, visual, camera trap, none), number of days since last camera detection before sample collection, water sample salinity, air temperature, and estimated water depth at sampling point. To look at the effect of sample type I used the data from the few instances where paired water and soil samples were taken at the same place at the same time and did a Wilcoxon rank test to see if the there was a significant difference between the paired water sample average and the paired soil average.

Results

Field survey detections

Out of the 620 callback surveys conducted at 195 different survey points (66 in 2018, 155 in 2019, 26 points were surveyed both years) by myself and my field assistants during both years of the study, Black Rail were detected via audio detection during 12 (2%) of these surveys: 13 Black Rails were detected at 7 (4%) of the sites over two years. A total of 19 eDNA samples were collected from these sites at the time of detection (Table 1-5). 16 samples were collected outside of official surveys (i.e. without audio callback or adhering to survey protocol) in North Carolina at locations where Black Rail were previously detected earlier in the breeding season either by myself or others, of these samples 5 had detections at the time of sampling and 11 did not.

I received a total of 39 additional samples (12 water and 27 soil) over the two breeding seasons from collaborators working on Black Rail populations in South Carolina and Florida. Of these, 28 samples were taken from camera trap sites where a Black Rail had been detected visually as little as 12 hours before sampling, and in a few cases more than one bird was seen multiple times in the 5 days prior to sampling. The USFWS contributed 5 eDNA soil samples from positive detection sites on federal refuge lands in Florida (Table 1-6).

Many of the samples were collected as pairs from the same point at the same time: a water and a soil sample: 6 pairs (12 samples) from South Carolina and 11 pairs (22 samples) from North Carolina. Two water samples from North Carolina were unable to be extracted due to thawing accidents and were not processed further making the final eDNA sample count 74.

Table 1-5. Summary of the call back surveys done in North Carolina and Virginia and eDNA samples collected as a direct result of the surveys. The "total survey points", "survey points with detections", and "locations sampled" (with grey background) indicate the number of points/locations that were surveyed or sampled both years, all other numbers in the "both years" column are totals for both years.

Survey Summary	2018	2019	Both years
Total surveys	186	434	620
Total number of survey points	66	155	26
Survey points with only 2 visits	10	23	N/A
Points abandoned after first survey round	1	14	N/A
Surveys with detections	5	7	12
Survey points with detections	3	6	2
Individual BLRA detected	6	7	13
eDNA samples collected	9 (6 water, 3 soil)	10 (3 water, 7 soil)	19
Locations sampled	2	5	1

Table 1-6. Summary of the sample type, detection type, sample state, and water storage condition of eDNA samples collected during each year of the study. For clarification, "samples taken from areas of previous detection" means that Black Rails had been detected in the area earlier in the breeding season but were not detected during the time of sampling.

		2018	2019
Detection type	Samples with audio detection	19 (8 water, 11 soil)	17 (7 water, 10 soil)
	Samples with visual detection	0	2 (2 soil)
	Samples with camera detection	12 (5 water, 7 soil)	16 (5 water, 11 soil)
	No detection: Samples from areas of previous detection	0	11 (4 water, 6 soil)
State of origin	Samples from North Carolina	9 (6 water, 3 soil)	28 (11 water, 17)
	Samples from South Carolina	17 (7 water, 10 soil)	17 (5 water, 12 soil)
	Samples from Florida	5 (5 soil)	0
Water sample storage temperature	Number of water samples frozen (-20°C)	10	13
	Number of water samples refrigerated (4°C)	3	3
	Total locations sampled	7	7
	Total samples	31 (13 water, 18 soil)	45 (16 water, 29 soil)

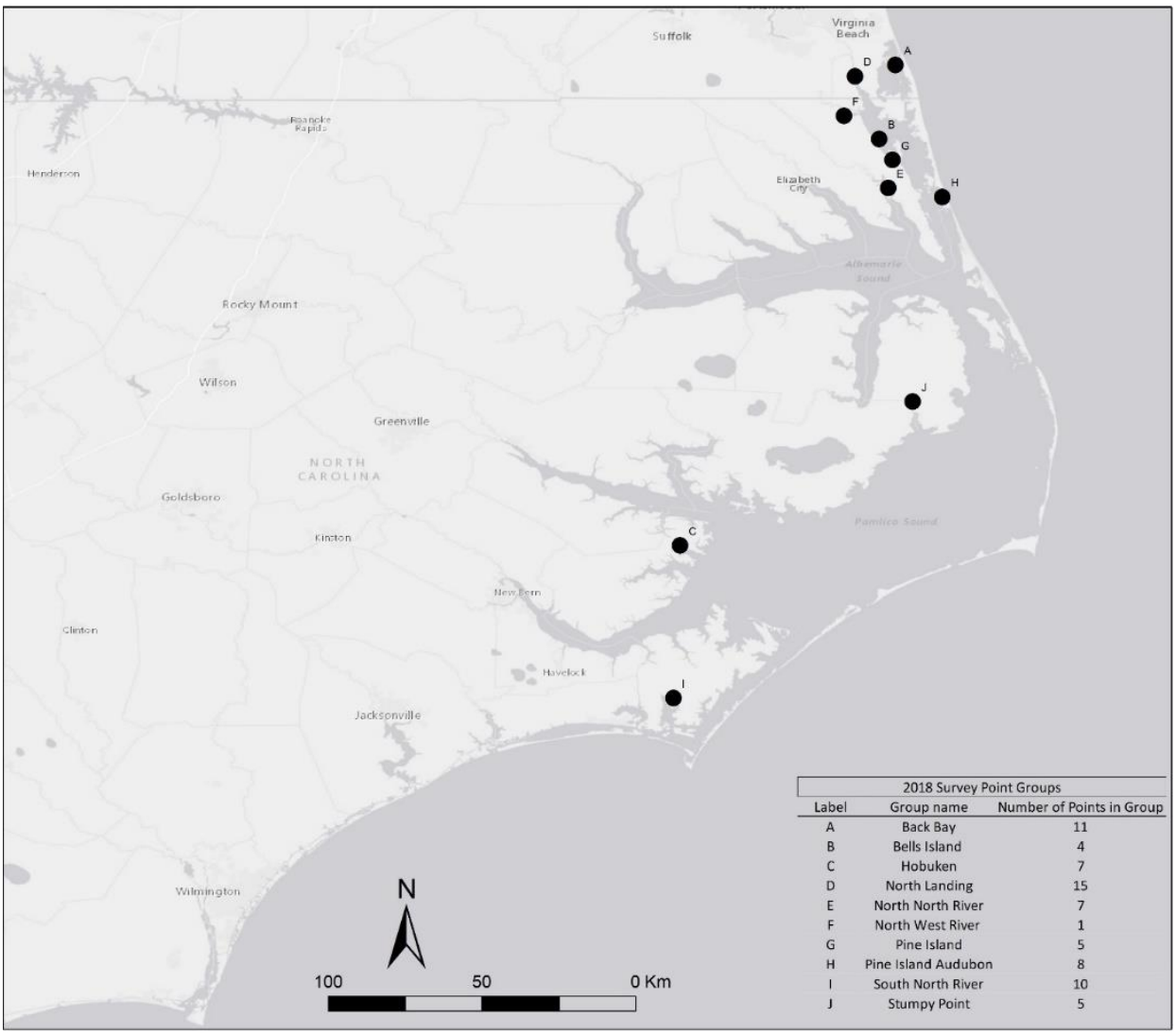


Figure 1-1. Locations of Black Rail surveys conducted during the 2018 breeding season. The letter next to each point corresponds to the group label in the table listing the number of individual survey points at each site.

Figure 1-2. Locations of Black Rail surveys conducted during the 2019 spring/summer breeding season. The letter next to each point corresponds to the group label in the legend listing the number of individual survey points at each site.

Comparison of DNA extraction methods

Four water environmental samples were extracted using both the Qiagen DNeasy tissue kit and PowerSoil kit according to the quick start protocol. One filter was used for each water sample, half of the filter was extracted using the tissue kit while the other half was filtered using the PowerSoil kit. When these eDNA samples were tested by spiking standard PCR reactions, all the DNeasy tissue kit extractions, but only two of the PowerSoil kit extractions, resulted in inhibition. Two water environmental samples were extracted using both the PowerWater kit and PowerSoil kit. In a side by side comparison using the extractions to spike standard PCR reactions, the PowerSoil kit was the most effective for removing enzyme inhibitors. Both of the PowerWater kit extractions showed inhibition and both PowerSoil extractions showed amplification of the spiked DNA. The salting out method was also tested for inhibitor removal. Out of four environmental water samples that were extracted using the salting out method, only one sample showed amplification of the spiked DNA. The PowerSoil kit was used to extract all subsequent eDNA samples.

Black Rail qPCR diagnostic test: validation using eDNA from sites of positive detections

Comparison between standard PCR and qPCR with SYBR green. Using conventional PCR the BLRA COI2 primer set did not detectably amplify the non-target species. A PCR of a dilution series of purified product of this primer set resulted in an estimated detection limit of 1.18E-6ng or ~5000 copies of target DNA (calculated using the equation, number of copies = (amount (ng) * 6.022×10^{23}) / (template length (bp) * 1×10^{9} * 650). A qPCR using this primer set and a SYBR Green reporter showed that eleven eDNA samples tested had Cq at or slightly below a standard that contained 2.36×10^{-8}ng Black Rail genomic DNA that was run and measured at the same time. This supported the move to qPCR versus standard PCR. A dilution series test for this

primer set and SYBR green qPCR resulted in a significantly lower estimated detection limit, $2.36x10^{-8}$ng or ~100 copies of target DNA. However, when tested for cross-species amplification, this protocol using SYBR green resulted in amplification with genomic DNA extracted from the blood of non-target species (Clapper Rail, Yellow Rail, Common Gallinule, Virginia Rail, Sora). All tested non-target species amplified above or within the standard curve except for Clapper Rail which did not have detectable amplification.

TaqMan dual probe assay test. When tested for cross-species amplification, the assay showed low but detectable amplification of King Rail and Common Gallinule DNA. This assay used one primer set and two probes that have similar sequences. I was concerned that the probes were so similar that they would compete for the vanishingly small amount of target expected in the eDNA samples and also about the possibility of mis-annealing with the non-target sequence (cross-amplifying). The dual probe assay was therefore abandoned.

TaqMan single probe assay specificity test. Cross-species amplification was significantly reduced with the single probe assay: full strength genomic DNA from tissue of only one of the sympatric rails amplified, the Common Gallinule (*Gallinula galeata*). This cross-species amplification was detectable, but to a lesser extent compared to the Black Rail samples, and was consistent when replicated 3 times with DNA from 2 different Common Gallinule individuals. Later in the study, I visualized the products of this qPCR on an agarose gel and discovered that when Black Rail genomic DNA from tissue samples was used as template, it produced a second product that was slightly larger than the target. The TaqMan chemistry prevented this additional product from being detected in the qPCR assay, but it would have negatively impacted the efficacy of the reaction.

Affinity Plus assay specificity test. When the Affinity Plus assay was tested for cross-species amplification with a full panel of non-target species, full strength Yellow Rail (*Coturnicops noveboracensis*) genomic DNA from tissue showed comparatively low but detectable amplification. Cross-species amplification was consistent with different individuals of this species tested multiple times. I tested this primer pair with traditional PCR and Black Rail genomic DNA from tissue to see if it produced accessory products. Visualization on an agarose gel revealed that the primers produced a very large band and possibly a second smaller band near the target band.

Modified Affinity Plus assay specificity test. When the modified Affinity Plus assay was tested for cross-species amplification with the full panel of non-target species, the assay showed comparatively low but detectable amplification from full-strength genomic DNA extracted from tissue of all sympatric non-target species except Clapper Rail.

Combination assay specificity and sensitivity test. The BLRA COI2 purified product dilution series conducted with traditional PCR showed a loss of detectability around ~10,900 copies (Figure 1-3), while the qPCR showed detectable fluorescence for dilutions as low as ~109 copies (Figure 1-4). The results of the five replicate dilution series for the combination assay produced a limit of quantification (LOQ) of 901 template copies and a limit of detection (LOD) of 109 copies. Diluting the master mix improved reaction outcomes: almost no NTCs or non-target samples produced any signal, but the reaction success of the standards was relatively unchanged. Only one out of six eDNA plates run produced a signal from a single NTC replicate and the Cq for this signal was after the LOD threshold. I found either no amplification with non-target species template, or amplification below the LOD, therefore considered negative. The two environmental negative controls assayed on two different plates did not produce a signal.

Combination assay eDNA results. Out of 74 eDNA samples that were tested using the qPCR diagnostic test, 8 (10.8%) had at least three of the six replicates with a Cq before or at the LOQ, and these were the same six that amplified consistently above the LOD. In addition, 22 (30%) samples had at least two replicates with a Cq at or before the LOD, and 35 (47%) of them had at least one with a Cq at or before the LOD (Table 1-7, raw data in appendix Table A2). Out of the 74 eDNA samples 18 of them were collected during audio surveys with positive audio Black Rail detections associated with them, 9 of these had at least one positive replicate in relation to the LOD.

Table 1-7. Summary of eDNA positive replicate qPCR results with the combination assay. All eDNA samples had 6 replicates. The thresholds were the limit of quantification (LOQ) which was 1,090 copies of target DNA for this assay, and the limit of detection (LOD) at 109 copies of target DNA. The cumulative percent is the percent of samples with that number or more positive replicates.

	Limit of Detection		Limit of Quantification	
# Replicates with Signal	# Samples	Cumulative % with Signal	# Samples	Cumulative % with Signal
6	8	11%	3	4%
5	1	12%	1	5%
4	4	18%	1	7%
3	5	24%	3	11%
2	4	30%	0	11%
1	13	47%	0	11%
No signal	39	53%	66	89%

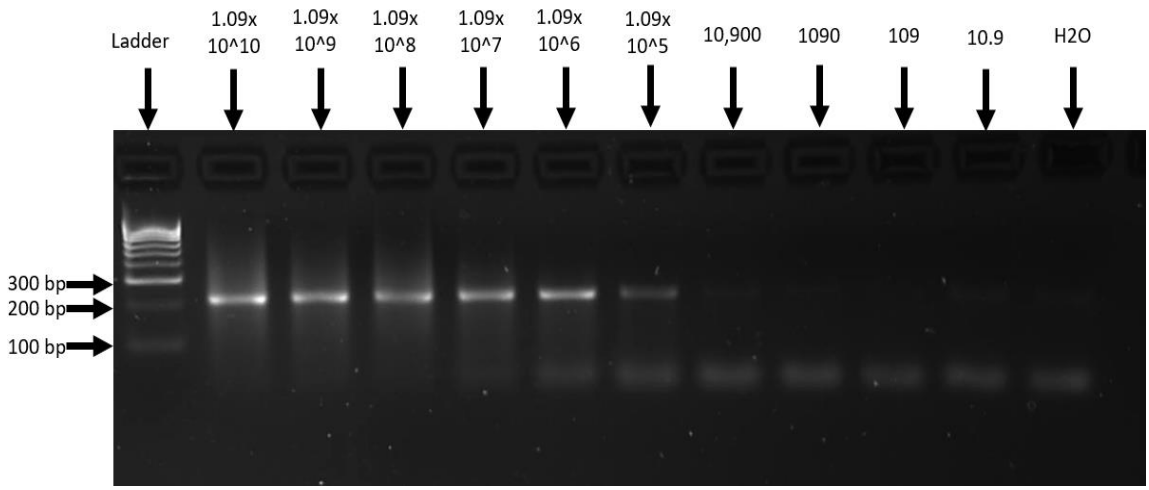


Figure 1-3. Black Rail combination assay dilution series. PCR products were amplified from purified product template ranging in estimated copy number from 1.0914x10^10 [2.52ng] to 10 copies [2.52x10^-9ng]. Electrophoresis on a 3% agarose gel was conducted for 90 minutes at 70V.

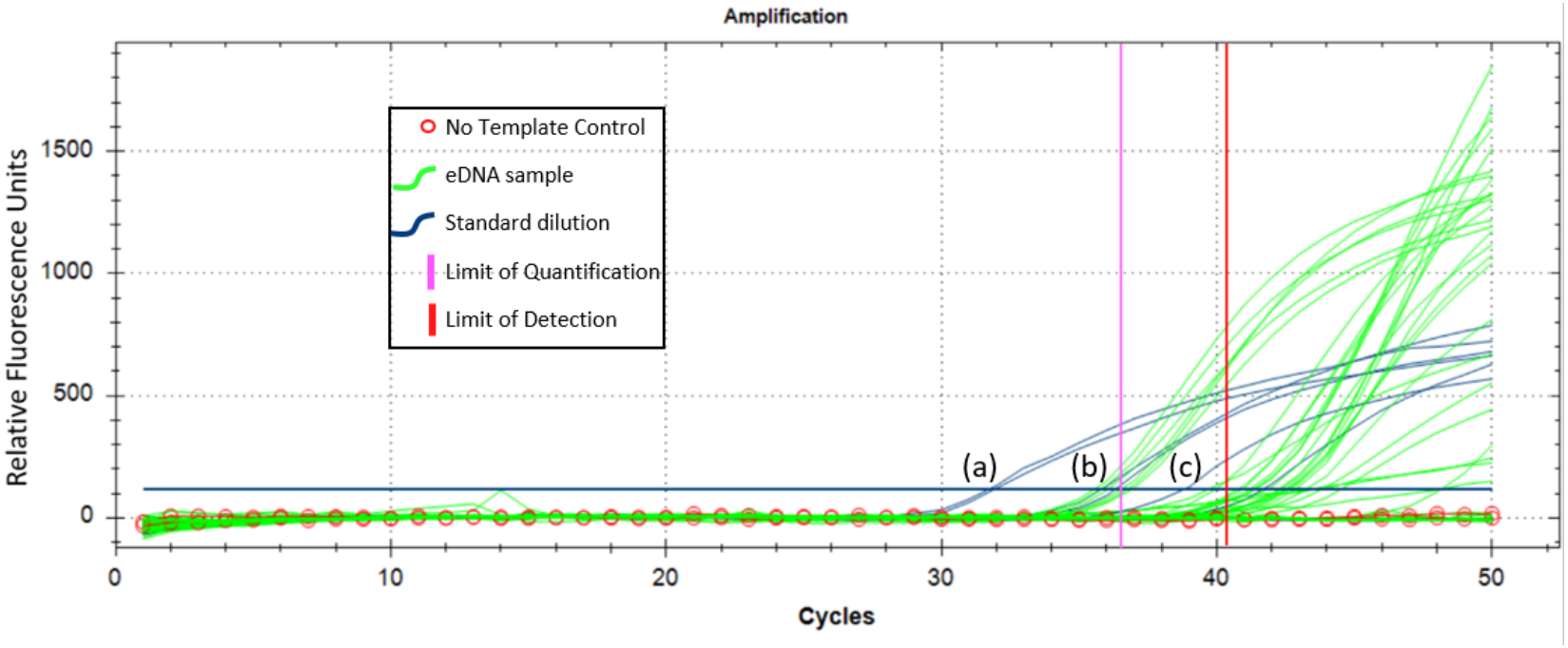


Figure 1-4. BLRA combination assay qPCR plot showing eDNA amplification. Amplification curves (eDNA – green, Black Rail DNA standards – blue) are in relative fluorescence units. The horizontal blue line is the baseline fluorescence threshold. The point at which the amplification curve crosses the line is the Cq or the cycle at which signal was first detected. Vertical lines indicate the inferred LOQ (red) and LOD (pink), and reference points for them: paired blue curves are duplicate standard dilutions of (a) ~10,900 copies (b) ~1,090 copies and (c) ~109 copies of target DNA. A standard dilution of ~10 copies did not show detectable fluorescence.

Statistical model results: determining the effect of environmental and sample handling on Black Rail eDNA detectability

With bootstrap confidence intervals, if 0 (the average of the distribution of means after resampling the full data set) is included in the 95% distribution of means from the resampling of data with variable of interest, then the variable is considered non-significant because the average is no different from the null. The original model for the detection type had to be slightly modified. Samples with more than one detection type such as the two samples that had visual detection that also had audio detection and the three camera and audio detections had to be simplified to audio detections because the number of samples in these dual detention groups were insufficient for statistical analysis. Out of the variables tested, the sample's state of origin (NC, SC, or FL) and the sample type (soil or water) had *p*-values < 0.05 but significance was not corroborated with the bootstrap confidence intervals (Table 1-8). The only variable with both $p < 0.05$ and confidence intervals not including 0 was the estimated water depth, defined as the depth of any water in the vegetation surrounding the sample collection point within a 50 m radius, regardless if the water was sampled or not. The positive estimate for the model indicates that water depth has a positive correlation with detection rate. To get a better understanding of this result I created three categories of water depth, 0-5 cm, 5-10 cm, and >10 cm. I then calculated the average positive detection rate (using the LOD as the positive threshold) for each category and found that the 5-10 cm and >10 cm categories had higher average detection rates (Mean $LOD_{5\text{-}10cm}$ = 37%, Mean $LOD_{>10cm}$ = 37%) than the 0-5cm category (Mean $LOD_{0\text{-}5cm}$ = 20%) (Table 1-9). They also had a higher percentage of samples with at least one positive replicate in relation to the LOD. All of the other variables (number of days between collection and DNA extraction, the storage temperature of water samples before filtering, field detection method

(audio, visual, camera trap, none), number of days since last camera detection before sample collection, water sample salinity, and air temperature) were non-significant with *p*-values < 0.05 and confidence intervals that included 0 (Table 1-8). Paired water and soil samples were found to have a non-zero difference between means (mean LOD_{water} = 29%, mean LOD_{soil} = 18%, *N* = 17) (Table 1-9), but this was not significant (*p* = 0.259).

Table 1-8. Generalized linear mixed models for Black Rail eDNA detectability. Models were considered significant if $p < 0.05$ and the bootstrap confidence intervals did not include 0.

				95% Bootstrap Confidence Interval	
Fixed Effect	Random Effect(s)	Category	Pr(>|z|)	0.0250	0.9750
Days between collection and DNA extraction †	Sample location/sample group	N/A	0.0511	-0.0598	0.0032
Water sample storage temperature §	sample group*	4°C	0.8700	-3.2030	2.4171
Sample state of origin	Sample group*	NC	0.0355	-0.2815	15.7573
		SC	0.0379	-0.5022	15.3692
Detection type	Sample group*	Camera	0.5980	-1.3293	2.2017
		No Detection	0.6910	-3.1494	1.2914
Days between collection and last camera detection ‡	Camera sample group	N/A	0.5662	-0.4927	0.1289
Water sample salinity §	sample group*	N/A	0.7880	-0.2956	0.1984
Air temperature	sample group*	N/A	0.0638	-0.1447	0.0008
Estimated water depth at sample location	sample group*	N/A	0.0254	0.0004	0.3707

* Due to small sample size some models returned a singular fit and model complexity needed to be reduced by dropping the sample location as a random effect.

†Model originally returned a scaling error which was fixed by dividing the number of days between collection and DNA extraction by 10.

‡ A subset of the data containing only samples with camera detections was used for this model.

§ A subset of the data containing only water samples were used for this model.

To determine if there were any trends, and for reference when a variable was significant, I calculated the average detection rate for each variable. If the variable was continuous, I created range categories.

Table 1-9. Average positive detection rate of Black Rail eDNA for different environmental variables.

Variable	Category	Average Positive Rate	# Samples in Category	# Positive Samples
Days between collection and extraction	<180 days	7%	12	4
	180-365 days	27%	46	22
	>365 days	27%	16	9
Water sample storage temperature	4°C	21%	8	6
	-20°C	54%	19	8
State of origin	North Carolina	21%	35	13
	South Carolina	25%	34	18
	Florida	30%	5	4
Detection type	Audio	25%	38	18
	Camera trap	25%	25	13
	No Detection	15%	11	4
Days between collection and camera detection	3 or less days	19%	18	8
	more than 3 days	33%	10	7
Water sample salinity	10 or less ppt	27%	14	7
	more than 10 ppt	35%	13	7
Air Temperature	≤ 26°C	29%	33	16
	27°F -29°F	23%	21	11
	>29°F	16%	20	8
Estimated water depth	0 to 5cm	20%	59	24
	5 to 10cm	37%	10	6
	10 to 15cm	37%	5	3
Paired water and soil samples	Soil	18%	17	7
	Water	29%	17	8

Discussion

I developed a diagnostic assay that can specifically detect Black Rail DNA in small amounts from environmental samples and does not amplify DNA of sympatric North American rail species. The assay was validated using eDNA samples collected from sites of positive detection of Black rails in the wild and did not amplify negative control samples.

Several different qPCR assays were vetted and despite attempts to maximize mismatches in both the primers and the probe many of the assays showed cross-species amplification with non-target species. After some refinement and ingenuity, a working assay was developed. The combination qPCR assay (the BLRA COI2 primers and Affinity Plus probe) was able to reliably detect amounts of Black Rail DNA down to 109 copies of target DNA as shown by multiple dilution series. This detection limit was within the same magnitude of sensitivity as described for another species-specific bird eDNA study (300 copy number limit of detection in Day et al., 2019). My qPCR assay with fluorescent quencher reporter outperformed traditional PCR using the same primers which had a detection limit of ~10,900 copies. Its specificity was demonstrated by the lack of a signal in qPCR reactions with concentrated genomic DNA from confamilial species as template.

Environmental samples that were considered positive for the presence of Black Rail DNA were obtained by conducting audio surveys. The surveys had vanishingly small returns with only 2% detection rate for both years and 4% occupancy rate for 2019 (12 out of 115 points). This is in spite of targeting mostly areas where Black Rails have been detected in the past. Unfortunately, similar numbers have reported on the east coast by other recent Black Rail surveyors. A 2015 breeding season survey in South Carolina reported a 5% occupancy rate (17 out of 344 survey sites occupied) (Roach & Barrett, 2015) and a 2015 survey effort in Virginia

had a 1.7% occupancy rate (2 out of 114 points) (Wilson et al., 2015). A 2015 survey effort in North Carolina reported a 4.5% occupancy rate (5 out of 109 locations) (Wilson et al., 2016).

Most of the samples collected for this project were considered environmental positive controls because they were collected at sites of recent detections and were expected to produce signal when assayed. Several samples did so reliably (~10%), but a lack of amplification could be due to several factors. The eDNA collected could have contained Black Rail DNA but may have been degraded and so did not have intact copies of the section of COI gene that is targeted in this study. It is also possible that the extracted eDNA contained the target Black Rail sequence but so few copies of it that it did not meet the threshold and thus was considered negative for presence. Finally, the eDNA sample may simply not have contained any Black Rail DNA which could be attributed to the species small size correlating with small amounts of shed DNA, and individual movements contributing to sampling stochasticity.

A few previous studies have found effects of sample storage on eDNA detectability. For example, in previous studies higher DNA yields were obtained from sample filters that were frozen before extraction compared to filters that were preserved in alcohol (Hinlo et al., 2017). The same study also looked at DNA yield from water samples stored at different temperatures over time. The results suggested that for short term storage (less than 14 days), refrigeration (4ºC) provided a greater yield but for storage longer than 14 days freezing (-20°C) gave a greater yield. Although there was no formal testing of the two storage methods, the Black Rail eDNA samples in this study showed no statistically significant difference in detectability with water samples stored at 4°C and filtered with in 24 hours and samples that were frozen and then later thawed directly before filtering.

Many of the environmental conditions and sample handling variables were shown to have no statistically significant effect on eDNA detectability of the samples in this study. A similar result was described in the great crested newt study where eDNA was sampled in ponds in Great Britain: most environmental factors tested did not have an effect (Biggs et al., 2015). The only factor they found influenced great crested newt eDNA detectability was the habitat suitability score for the species. This is promising and suggest that there are few confounding environmental variables to eDNA detectability.

In my study, 34 samples (17 pairs) were collected as water/soil pairs from the same point at the same time providing the opportunity to test the effect of sample type on detectability. One study tested this by comparing the amounts of Bighead Asian Carp (*Hypophthalmichthys* spp.) eDNA detected in surface water samples to substrate samples (Turner et al., 2015). The study found DNA was 8-1800 times more concentrated in sediment compared to water and lasted up to 132 days after the species was removed (Turner et al., 2015). A paired design statistical test for Black Rail eDNA samples suggested that this pattern did not apply to the samples in this study and soil samples were just as likely to amplify as water samples.

The only environmental variable that significantly affected detectability of samples in this study was estimated water depth near the sampling site. Samples with an estimated water level above 5 cm had a higher average percent positive detection rate. This suggest that deeper water plays a positive role in eDNA detectability for these samples. Deeper water is often more mobile and could spread the DNA a farther distance from its source. This could dilute it over a wider area and provide a greater likelihood of sampling the DNA. It has been demonstrated that eDNA can move rapidly and be detectable when water movement is sufficient. A study looking at the transport distance of invertebrate eDNA showed detection as far as 10 km from the source

(Deiner & Altermatt, 2014). Deeper water may also protect DNA from degradation by buffering temperature or sunlight, both of which have been shown to influence eDNA decay. A study looking at the effect of temperature, UV-B, and acidity on eDNA found that, under controlled lab conditions, temperature was most strongly correlated with eDNA degradation and the addition of UV-B light compounded this effect (Strickler et al., 2015).

Future studies should analyze more eDNA samples taken from areas where the habitat is suitable for Black Rail but has been without any recent detections. The utility of an eDNA assay is to add detections of Black Rails where they have not recently or previously been detected by other means. Also, analyzing samples from known or suspected non-breeding areas could be useful in understanding the Black Rails' non-breeding ecology. Formal testing of the effects of environmental factors on detectability would greatly improve the understanding of the method and lead to improved quality of the assay. In particular, for this tool to be transferable, a better understanding of water bird eDNA persistence as it applies to marsh environments will be needed. The assay did not show perfect detection and could be optimized further, including optimization of eDNA concentration and extraction processes.

In developing the Eastern Black Rail Conservation Plan, it is more important than ever to be able to survey the distribution and occupancy patterns of Black Rails. Development of a reliable and cost-effective diagnostic method to detect these imperiled marsh birds, represents an important addition to the conservation toolbox, and complements other methods employed to monitor occupancy and persistence of rails in an area. Environmental DNA offers an alternative to behavioral detection using field sampling techniques that are less invasive and unconstrained by time of day or season.

Literature Cited

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., Foster, J., Wilkinson, J. W., Arnell, A., Brotherton, P., Williams, P., & Dunn, F. (2015). Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (*Triturus cristatus*). *Biological Conservation*, *183*, 19–28. https://doi.org/10.1016/j.biocon.2014.11.029

BirdLife International. (2019). *IUCN red list of threatened species: Black rail*. IUCN Red List of Threatened Species. https://www.iucnredlist.org/en

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE Guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, *55*(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797

Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2016). GenBank. *Nucleic Acids Research*, *44* (Database issue), D67–D72. https://doi.org/10.1093/nar/gkv1276

Conway, C. J. (2011). Standardized North American marsh bird monitoring protocol. *Waterbirds*, *34*(3), 319–346.

Conway, C. J., & Sulzman, C. (2007). Status and habitat use of the California Black Rail in the southwestern USA. *Wetlands*, *27*(4), 987–998. https://doi.org/10.1672/0277-5212

Davidson, L. M. (1992). Black Rail. In: Schneider, K. J. ,& Pence. D. M. (Eds.), *Migratory non game birds of management concern in the Northeast* (pp. 119–134). U.S. Fish and Wildlife Service.

Day, K., Campbell, H., Fisher, A., Gibb, K., Hill, B., Rose, A., & Jarman, S. (2019). Development and validation of an environmental DNA test for the endangered Gouldian finch. *Endangered Species Research*, *40*, 171–182. https://doi.org/10.3354/esr00987

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS ONE*, *9*(2), e88786. https://doi.org/10.1371/journal.pone.0088786

Djurhuus, A., Port, J., Closek, C. J., Yamahara, K. M., Romero-Maraccini, O., Walz, K. R., Goldsmith, D. B., Michisaki, R., Breitbart, M., Boehm, A. B., & Chavez, F. P. (2017). Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. *Frontiers in Marine Science, 4*, 314. https://doi.org/10.3389/fmars.2017.00314

Dunn, N., Priestley, V., Herraiz, A., Arnold, R., & Savolainen, V. (2017). Behavior and season affect crayfish detection and density inference using environmental DNA. *Ecology and Evolution, 7*(19), 7777–7785. https://doi.org/10.1002/ece3.3316

Feng, W., Bulté, G., & Lougheed, S. C. (2020). Environmental DNA surveys help to identify winter hibernacula of a temperate freshwater turtle. *Environmental DNA, 2*(2), 200–209. https://doi.org/10.1002/edn3.58

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters, 4*(4), 423–425. https://doi.org/10.1098/rsbl.2008.0118

Flores, R. E., & Eddleman, W. R. (1995). California Black Rail use of habitat in southwestern Arizona. *The Journal of Wildlife Management*, *59*(2), 357–363. https://doi.org/10.2307/3808949

Hand, C. E., Znidersic, E., & Tegeler, A. K. (2019). First documentation of eastern Black Rails (*Laterallus jamaicensis jamaicensis*) breeding in South Carolina, USA in more than a century. *Waterbirds*, *42*(2), 237–241. https://doi.org/10.1675/063.042.0212

Harper, K., Goodwin, K., Harper, L., LaCasella, E., Frey, A., & Dutton, P. (2020). Finding crush: Environmental DNA analysis as a tool for tracking the green sea turtle *Chelonia mydas* in a marine estuary. *Frontier in Marine Science, 6*, 810. https://doi.org/10.3389fmars.2019.00810

Harper, L. R., Handley, L. L., Hahn, C., Boonham, N., Rees, H. C., Lewis, E., Adams, I. P., Brotherton, P., Phillips, S., & Hänfling, B. (2020). Generating and testing ecological hypotheses at the pondscape with environmental DNA metabarcoding: A case study on a threatened amphibian. *Environmental DNA, 2*(2), 184–199. https://doi.org/10.1002/edn3.57

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S., & Francis, C. M. (2004). Identification of birds through DNA barcodes. *PLoS Biology*, *2*(10), e312. https://doi.org/10.1371/journal.pbio.0020312

Hinlo, R., Gleeson, D., Lintermans, M., & Furlan, E. (2017). Methods to maximize recovery of environmental DNA from water samples. *PLoS ONE*, *12*(6), e0179251. https://doi.org/10.1371/journal.pone.0179251

Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). "Sight-unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*, *4*(2), 150–157. https://doi.org/10.1111/j.1755-263X.2010.00158.x

Kerlinger P., & Wiedner D. S. (1990). Vocal behavior and habitat use of Black Rails in south Jersey. *Records of New Jersey Birds*, *16*, 58–62.

Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, *23*(10), 1289–1291. https://doi.org/10.1093/bioinformatics/btm091

Kutyavin, I. V., Afonina, I. A., Mills, A., Gorn, V. V., Lukhtanov, E. A., Belousov, E. S., Singer, M. J., Walburger, D. K., Lokhov, S. G., Gall, A. A., Dempcy, R., Reed, M. W., Meyer, R. B., & Hedgpeth, J. (2000). 3′-Minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. *Nucleic Acids Research*, *28*(2), 655–661.

Legare, M. L., Eddleman, W. R., Buckley, P. A., & Kelly, C. (1999). The Effectiveness of Tape Playback in Estimating Black Rail Density. *The Journal of Wildlife Management*, *63*(1), 116-125. https://doi.org/10.2307/3802492

Livak, K. J., Flood, S. J. A., Marmaro, J., Giusti, W., & Deetz, K. (1995). Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *Genome Research*, *4*(6), 357–362.

Mächler, E., Deiner, K., Spahn, F., & Altermatt, F. (2016). Fishing in the water: Effect of sampled water volume on environmental dna-based detection of macroinvertebrates. *Environmental Science & Technology*, *50*(1), 305–312. https://doi.org/10.1021/acs.est.5b04188

Monge, O., Dumas, D., & Baus, I. (2020). Environmental DNA from avian residual saliva in fruits and its potential uses in population genetics. *Conservation Genetics Resources*, *12*(1), 131–139. https://doi.org/10.1007/s12686-018-1074-4

Nadeau, C. P., Conway, C. J., Piest, L., & Burger, W. P. (2013). Multi-species call-broadcast improved detection of endangered Yuma Clapper Rail compared to single-species call broadcast. *Wetlands*, *33*(4), 699–706. https://doi.org/10.1007/s13157-013-0425-x

Neiman, M., & Taylor, D. R. (2009). The causes of mutation accumulation in mitochondrial genomes. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1660), 1201 1209. https://doi.org/10.1098/rspb.2008.1758

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, *302*(1), 205–217. https://doi.org/10.1006/jmbi.2000.4042

Owczarzy, R., You, Y., Groth, C. L., & Tataurov, A. V. (2011). Stability and Mismatch Discrimination of Locked Nucleic Acid–DNA Duplexes. *Biochemistry*, *50*(43), 9352 9367. https://doi.org/10.1021/bi200904e

Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary, P. R., Litvin, S. Y., Nickols, K. J., Yamahara, K. M., & Kelly, R. P. (2016). Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, *25*(2), 527–541. https://doi.org/10.1111/mec.13481

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roach, N. S., & Barrett, K. (2015). Managed habitats increase occupancy of Black Rails (*Laterallus jamaicensis*) and may buffer impacts from sea level rise. *Wetlands*, *35*(6), 1065–1076. https://doi.org/10.1007/s13157-015-0695-6

Spear, L. B., Terrill, S. B., Lenihan, C., & Delevoryas, P. (1999). Effects of temporal and environmental factors on the probability of detecting California Black Rail. *Journal of Field Ornithology*, *70*(4), 465–480.

Stedman, S. M., & Dahl, T. E. (2013). *Status and trends of wetlands in the coastal watersheds of the conterminous United States 2004 to 2009*. U.S. Department of the Interior, Fish and Wildlife Service and National Oceanic and Atmospheric Administration, National Marine Fisheries Service.

Strickler, K. M., Fremier, A. K., & Goldberg, C. S. (2015). Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation*, *183*, 85–92. https://doi.org/10.1016/j.biocon.2014.11.038

Stuart, G. H. (1920). Nesting of the little Black Rail in Atlantic County, N. J. *The Auk, 37*(2), 292–293. https://doi.org/10.2307/4072725

Takahara, T., Minamoto, T., & Doi, H. (2013). Using environmental DNA to estimate the distribution of an invasive fish species in ponds. *PLoS ONE, 8*(2), e56584. https://doi.org/10.1371/journal.pone.0056584

Taylor, B., & Van Perlo, B. (1998). *Rails: A guide to the rails, crakes, gallinules and coots of the world*. Pica Press. Sussex, U.K.

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., & Willerslev, E. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS ONE, 7*(8), e41732. https://doi.org/10.1371/journal.pone.0041732

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., Orlando, L., & Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA: Species monitoring by environmental DNA. *Molecular Ecology, 21*(11), 2565–2573. https://doi.org/10.1111/j.1365-294X.2011.05418.x

Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Bioloigcal Conservation, 183*, 93 102. https://doi.org/10.1016/j.biocon.2014.11.017

U.S. Fish and Wildlife Service. (2018). 12-month petition finding and threatened species status for Eastern Black Rail with a section 4(d) rule. Federal Register Agency Docket No. FWS-R4-ES-2018-0057 pp.50610-50630.

Ushio, M., Murata, K., Sado, T., Nishiumi, I., Takeshita, M., Iwasaki, W., & Miya, M. (2018). Demonstration of the potential of environmental DNA as a tool for the detection of avian species. *Scientific Reports, 8*(1), 4493. https://doi.org/10.1038/s41598-018-22817-5

Valiere, N., & Taberlet, P. (2000). Urine collected in the field as a source of DNA for species and individual identification. *Molecular Ecology, 9*, 2149–2154.

Watts, B. D. (2016). Status and distribution of the Eastern Black Rail along the Atlantic and Gulf Coasts of North America. *The Center for Conservation Biology Technical Report Series*, CCBTR-16-09. College of William and Mary/Virginia Commonwealth University, Williamsburg, VA. 148 pp.

Wilson, M.D., Fletcher, S. M., & Watts, B. D. (2015). Re-Survey and Population Status Update of the Black Rail in Virginia. *Center for Conservation Biology Technical Report Series,CCBTR-15-004., College of William and Mary and Virginia Commonwealth University*.

Wilson, M.D., Watts, B. D., & Poulton, D. (2016). Black Rails in North Carolina. *Center for Biology Technical Report Series, CCBTR-16-01, College of William and Mary and Virginia Commonwealth University*.

Chapter 2: Mapping Habitat Suitability for Black Rail on the Atlantic Coastal Plain Using *MaxEnt*

Summary:

Modeling a species' distribution can be a powerful tool for predicting the location of suitable habitat for that species. This is especially useful for data-deficient species of conservation concern. The Black Rail (*Laterallus jamaicensis*) is a small marsh bird that is globally listed as threatened and is being considered for listing under the Endangered Species Act. I created a habitat suitability model for the Eastern Black Rail focusing on the Atlantic coastal plain using eBird data contributed by citizen scientists and environmental variable data from the Esri databank using a maximum entropy model framework. The map generated by the *MaxEnt* model indicated habitat suitability in areas known for Black Rail occupation in southern Florida, and along the coasts of Georgia, South Carolina, North Carolina, Maryland, and New Jersey. The environmental factors that best predicted Black Rail presence were flooded areas of shrub and herbaceous vegetation, proximity to water, and flat plains. These environmental variable associations were congruent with descriptions of high marsh, a habitat in which Black Rails have been found in this part of their range. Other habitat association studies conducted in other parts of the species' range have focused on smaller areas, and used presence data collected through species-specific callback surveys, but identified similar habitat characteristics. I found a weaker than predicted association with water depth, based on previous empirical studies emphasizing the importance of shallow water for the species. However, water level preference was in the expected direction: the model predicted shallower water favoring species presence. In terms of statistical power, this habitat suitability model was comparable with other *MaxEnt* models created for birds. Correlation with known areas of Black Rail occupation and previous

habitat associations confirms the validity of the model and importance of high marsh habitat for the species. The map generated by this model will inform land management decisions and habitat restoration efforts identified by the Eastern Black Rail Working Group.

Introduction:

The Black Rail is a small marsh bird found in the southeastern coastal plain of North America, with few inland patches (Taylor and van Perlo 1998). There are two North American subspecies, the California Black Rail (*Laterallus jamaicensis coturniculus*) found in the western range of the species, which is almost exclusively in California and Arizona, and the Eastern Black Rail (*Laterallus jamaicensis jamaicensis*) found in the eastern part of the species range more spread out along the Atlantic Coast (Spautz et al., 2005; U.S. Fish and Wildlife Service, 2018a). The California subspecies has been reported to have a disjunct and fragmented distribution (Richmond et al., 2008). Although the eastern subspecies has a wider range, its distribution has been shown to be rather fragmented, exacerbated by the shrinking of its range in recent years (Watts, 2016; U.S. Fish and Wildlife Service, 2018a).

The Black Rail, like most rails, has been described as secretive in nature and weak flying, favoring a speedy foot retreat rather than a revealing flight (Davidson, 1992; Stuart, 1920). Their concealing habitat, small size, and cryptic behavior make them difficult to detect visually, and audio call-back surveys is currently the main detection method for Black Rails (Conway, 2011). This has been a barrier to gaining information about their habitat requirements and ecology.

Both North American subspecies are considered globally Near Threatened (BirdLife International, 2019). There are multiple lines of evidence to support that the eastern population of Black Rails has significantly declined in the last few decades. In 2016, the Atlantic Coast population size was estimated at only 355-815 breeding pairs (Watts, 2016). The Eastern Black Rail Working Group comprised of members from the U.S. Fish and Wildlife Service, state agencies, academic biologists and other interested parties across the species' range, has proposed

that the Eastern Black Rail be listed as Threatened under the Endangered Species Act (U.S. Fish and Wildlife Service, 2018b). Due to its concerning population declines, identifying areas of appropriate habitat for the species has become a priority for the Atlantic Coast Joint Venture. The recently completed Black Rail Conservation Action Plan lists as explicit objectives the need to identify and expand suitable habitat for Black Rails (Atlantic Coast Joint Venture, 2020).

There is a paucity of data on the ecology and behavior of the Eastern Black Rail (but see Hand (2019), for exception). A better understanding of Black Rail habitat requirements is urgently needed in support of the conservation management plan. Watts (2016) compiled an extensive report on the status of the Eastern Black Rail and identified five main habitat categories where Black Rails have been detected: tidal saltmarshes, impoundments, grassy fields, freshwater wetlands, and coastal prairie. Despite their diversity, these habitats share common features of hydrology (shallow water with consistent and regular flooding) and the presence of dense, early successional vegetation (Watts, 2016). In addition to these characteristics, topographic diversity has recently associated with Black Rail occupancy (Atlantic Coast Joint Venture, 2020).

A few recent studies have specifically addressed habitat features associated with Black Rail occupancy during the breeding season. A radiotelemetry and nest success study in Florida concluded that Black Rails selected as nesting habitat areas of low water level formed by salt pans (Legare & Eddleman, 2001). A habitat association study in South Carolina found that Black Rails were commonly found in managed impoundments and the habitat characteristics most associated with Black Rail occupancy were proximity to forest and higher proportion of marsh surrounding the detection site (Roach & Barrett, 2015). The authors inferred that the proximity to

forest was less about the woody vegetation but that the birds favored the sloped landscape and shallow water associated with trees near marsh. There was also a weaker association with vegetation height above half a meter. This kind of information is useful for knowing what kind of habitat to look for but does not indicate where suitable habitat might be.

Ecological niche modeling, habitat suitability modeling, or species distribution modeling are all terms for approaches that make use of large-scale remote sensing data across landscapes to determine areas that have similar features associated with areas known to be occupied by a target species (Corsi et al., 2000; Pearce & Boyce, 2006; Phillips et al., 2006, 2004). Using values of a large number of biotic and abiotic variables at locations where members of a species of interest have been observed, a model is constructed that generates maps estimating the probability of finding that species at other locations based upon similar values for the same ecological variables (Corsi et al., 2000; Pearce & Boyce, 2006; Phillips et al., 2006, 2004). Thus, ecological niche modeling is a powerful way to predict additional sites where rare species may be found.

There are many ways to model species distribution. If presence and absence data are available for the study area, generalized linear models (GLM) and generalized additive models (GAM) can be used. When absence data are not available GLMs and GAMs are not accurate because they depend on reliable absence data and more complex methods must be used. One method called the Genetic Algorithm for Rule-Set Prediction (GARP) uses artificial intelligence to create rules unique to the data set for the model and is therefore able to treat background points as absences (Stockwell & Noble, 1992). Another type of modeling called Maximum Entropy (MaxEnt) uses machine learning to estimate the most geographically uniform

distribution of the target species within constraints derived from environmental characteristics at known occurrence points (Phillips et al., 2006, 2004). A set of positive species occurrences, either a separate file or a subset of the occurrence data entered, is used to train the model and set the constraints before testing on the rest of the data which produces the end output. MaxEnt (Phillips et al., 2006) has been shown to be the most effective model when using presence only data when compared to other models (Elith et al., 2011; Phillips et al., 2004). One of the advantages it has over GARP is that MaxEnt predictions are continuous, while GARP predictions are discrete, allowing for better resolution of habitat suitability. It can be used to determine not only areas of habitat suitability but also the relative importance of each factor entered into the model to the presence of the target species.

MaxEnt has been used successfully to determine areas and qualities of suitable habitat for rare and cryptic bird species. MaxEnt was used to determine the microhabitat of two potential bio-indicator sympatric bird species in Chile the Black-throated Huet-huet (*Pteroptochos tarnii*) and the Ochre-flanked Tapaculo (*Eugralla paradoxa*) (Moreno et al., 2011). Both species inhabit the same general temperate forest ecotone. The variables entered into the model were slope, aspect, elevation, distance to waterway, distance to trail and distance to ecotone. The MaxEnt model revealed that *Pteroptochos tarnii* had few, mostly topographic, requirements such as a slope greater than 30% being the most influential for the model. Whereas, *Eugralla paradoxa* had slightly different requirements such as low elevation and being 70 m from trails being important to their microhabitat (Moreno et al., 2011). MaxEnt was also used to look at nesting habitat suitability for Red-crowned Crane (*Grus japonensis*) at Zhalong National Nature Reserve in China in order to inform conservation and management efforts (Na et al., 2018). The study found that water depth between 15 cm and 30 cm and distance greater than 2500 m from roads

and ditches were the best predictors of crane nesting habitat. The model identified habitat suitability in areas of the reserve that had not been surveyed for cranes due to inaccessibility. In another study, MaxEnt was used to create species distribution models for 27 endemic bird species in Taiwan (Wu et al., 2012). The authors considered all these species to be data deficient, rare, or cryptic. The environmental variables used were distance to river, distance to sea, mean elevation, forest density, mean NDVI, annual precipitation, road density, mean slope, human population density, and ecoregion. Despite the scarcity of occupancy data for these species, MaxEnt was still able to produce quality models and add to the understanding of their ecology.

Few previous studies have attempted to use ecological niche modeling to predict current Black Rail distributions. One study evaluated habitat suitability for the California Black Rail in the Sacramento–San Joaquin Delta using a MaxEnt species distribution model with presence only data sampled for the study from audio surveys, with vegetation type and tidal status as the predictor variables. California Black Rail presence was most often correlated with tall emergent vegetation interspersed with riparian shrubs and indicated areas of suitability that correlated with historical records (Tsao et al., 2015). A more recent study investigating Eastern Black Rails along the Texas gulf coast used a large dataset of callback surveys to determine occupancy of Black Rails across a number of contiguous and non-contiguous refuges (Haverland, 2019). The model predicted rail presence mainly in high-marsh habitats with minimal tidal influence and >50% herbaceous vegetative cover. Gulf cordgrass (*Spartina spartinae*) was the dominant species of vegetation that correlated most highly with Black Rail occupancy in this system (Haverland, 2019). A major strength of this study was that it used current presence and absence data based on extensive surveying conducted by the author and her collaborators. This detailed level of ground-truthing over a series of contiguous coastal refuges allowed for greater

understanding of not only where the birds prefer to be but also adjacent areas that were not supporting Black Rails.

Since an ecological niche model has been helpful in identifying predictors over smaller ranges in California and Texas, I reasoned that it has potential to inform our understanding of habitat use in the Atlantic coastal plain that comprises a large portion of the species range. Importantly, a model could help identify additional areas of habitat suitability. The maps generated from the model will inform future surveys and identify candidate areas for habitat protection or restoration efforts.

Black Rails are difficult to find, and occupancy data are scarce along the East Coast. One solution for expanding the dataset of Black Rail observations is utilizing citizen science data. The Cornell Lab of Ornithology's eBird citizen science initiative is a great source of bird presence data and has been shown to be a reasonable proxy for professionally collected data in species distribution modeling (Coxen et al., 2017; Walker & Taylor, 2017). A model created from eBird data and a model created from satellite tracking data of Band-tailed Pigeons (*Patagioenas fasciata*) in New Mexico were compared and found that both models had good accuracy and high overlap between habitat suitability scores (Coxen et al., 2017). An evaluation of models made with either Breeding Bird Survey (where data are usually collected by experienced technicians) or eBird data (where the data are largely contributed by citizen scientists) for 22 bird species in Ontario found that the eBird data models agreed with the Breeding Bird Survey models (Walker & Taylor, 2017). Similarly, citizen science data from the Swedish Species Observation System were used to produce a habitat suitability model for

Siberian Jay *Perisoreus infaustus* that was found to be comparable to one generated using systematically-collected Swedish Bird Survey data (Bradter et al., 2018).

The goals of this study were to find areas of suitable habitat for Black Rail in the eastern part of its mainland North American range (primarily on the Atlantic coastal plain). Finding new habitat is of critical importance as the Eastern Black Rail Working group is developing a Species Action Plan focused on this region. The strategy of my modeling approach was to quantify biotic and abiotic characteristics at sites where Black Rails have been detected and to determine what environmental factors are associated with Black Rail occupancy. Environmental data such as distance to water, land cover, terrain, bio-climate, water depth, and lithology were added to the model as predictor variables. Black Rail presence data were acquired from eBird. Other sites were then identified that have similar characteristics, indicative of their possible suitability.

Based on previous studies on Black Rail (e.g. Haverland, 2019, Legare & Eddleman, 2001; Roach & Barrett, 2015; Tsao et al., 2015), I expected that the model would heavily associate characteristics indicative of high marsh with Black Rail habitat suitability such as flooded vegetation, proximity to open water, shallow water level, low relief topography, low slope, loose soils, wet and warm climate, and low human population. This study used citizen science data in conjunction with maximum entropy modeling techniques to test my hypothesis, and to create a model to better understand the habitat requirements and find areas of suitable habitat for the imperiled Black Rail.

Methods:

Data sources and acquisition

I downloaded from eBird (eBird, 2018), Black Rail detection data spanning the years 1851 to 2018 from around the world. This dataset contained not only points recorded during the activity of the citizen science initiative but also retroactively added historical points from records verified by staff of the Cornell Lab of Ornithology. I made a separate file of only decimal degree (dd) coordinates and imported the .csv file into ArcMap 10.7 (Esri, Redland, USA) with the light grey canvas basemap. I then displayed the XY coordinate data with a WGS_1984 geographic coordinate system. I used the select by rectangle tool in ArcMap on the layer with the Black Rail points and selected the East Coast region of the United States by demarcating a rectangle, making an extent of -87.822860 dd to -70.788500 dd Longitude and 44.175883 dd to 24.885264 dd Latitude. I set this as the processing extent for the environment and the output coordinate system as WGS_1984_Web_Mercator_Auxiliary_Sphere. The raster analysis cell size was set to 928, the resolution of the geographically coarsest dataset. This was approximately equivalent to a square kilometer (928 m by 928 m). I added several data layers from the Esri data bank to the map (Table 2-1) in order to process them for use in the model.

Table 2-1. Environmental predictor variables used to develop habitat suitability models for Black Rail on the East Coast using occurrence records derived from eBird.

Environmental Predictor	Description	Data Source
World Land Cover ESA 2010	The surface of the earth classified into 36 classes focused on vegetation type, including agriculture, forests, grasslands, artificial surfaces, and other categories.	Esri, ESA (https://www.arcgis.com/home/)
World Distance to Water	Distance in meters from surface water calculated using Euclidian distance	Esri, USGS, ESA (https://www.arcgis.com/home/)
World Ecological Facets Landform Classes	Topographic classification based on Hammond Landform Classification combining slope, relief, and profile	Esri (https://www.arcgis.com/home/)
World Lithology	Classification of underlying soils and rock	Global Lithological Map Database v1.0 (https://www.arcgis.com/home/)
U.S. Soils Water Table Depth	shallowest depth to water in the soil at any time of the year in centimeters in the United States and territories	USDA NRCS, Esri (https://www.arcgis.com/home/)
World Population Estimate 2016	global estimate of human population for 2016 based on human settled area footprint	Esri (https://www.arcgis.com/home/)
Terrain: Slope in Degrees	slope values calculated dynamically from the elevation data (within the current extents) using the server-side slope function applied to a Terrain layer	Airbus, USGS, NGA, NASA, CGIAR, NLS, OS, NMA, Geodatastyrelsen, GSA, GSI and the GIS User Community (https://www.arcgis.com/home/)
World Bioclimate	Climate classification combining temperature and aridity categories	Esri, USGS, Metzger and others 2012 (https://www.arcgis.com/home/)

File manipulations and processing

To enable all these data layers to conform to the same parameters set in the environment, I used the copy raster tool on all data layers. This ensured that all the data layers had the same extent and raster cell size so that the data were comparable during analysis. When the layers were copied, the raster grids were resampled to a 928 m cell size using the nearest neighbor

method. This method defines the value of the larger cell by assigning the value of the smaller cell closest to the center of the larger cell when being resampled.

The ‘World Land Cover ESA 2010’ dataset describes categories of vegetation cover types. This was chosen for the model because both previous Black Rail habitat suitability models found vegetation type and density to be important to Black Rail occupancy (Tsao et al., 2015; Haverland, 2019). The habitat association studies for the species also concluded that vegetation was important (Legare & Eddleman, 2001; Roach & Barrett, 2015). I reclassified the dataset using “classname” as the reclassification field to eliminate any value number gaps in the outputs.

The ‘World Distance to Water’, which gives each cell a value indicating its distance from surface water, was included because Black Rails are known to be associated with wetlands, and these often surround the perimeters of lakes and form living shorelines. The ‘World Distance to Water’ layer had no data in raster cells categorized as being zero meters from open water, so I used the ‘Is Null’ tool in ArcMap to create a layer where cells with data had a value of 0 and cells without data have a value of 1. I then used the Con tool to merge the copied ‘World Distance to Water’ layer to the null layer to create a layer where every cell has a value and cells that have water have a value of 0, meaning that cell is zero meters away from water.

The ‘World Ecological Facets Landform Classes’ was used to represent the topography. This data set uses the Hammond’s landform classification which created categories based on slope, local relief and, profile. Topography is a significant characteristic of the landscape and is used in some form in many ecological models (Haverland, 2019, Coxen et al., 2017; Hu & Liu,

2014; Moreno et al., 2011; Wu et al., 2012). I reclassified the data using "classname" as the reclassification field to eliminate any value number gaps in the outputs.

Underlying 'Lithology' and soil type can influence vegetation and hydrology and has been shown to be influential in habitat models for birds in less vegetated areas. In a study looking at habitat preference of Cream-colored Courser (*Cursorius cursor*), Palomino et al. (2008) reasoned that if there is not much vegetation to influence habitat selection then factors such as topography and lithology are likely to play a larger role. The results showed that there was indeed a correlation between the Courser and areas with less than 23% rock cover and the authors encouraged other studies to include fine-grained habitat characteristics like soil type and topography in habitat selection models.

'Terrain: Slope in Degrees' was added to the model because slope influences hydrology which is known to be important to Black Rail occupancy (Richmond et al., 2008, 2010). The Joint Venture identified slope as a potential important quality for finding new areas to manage for Black Rail (Atlantic Coast Joint Venture, 2020). The 'Terrain: Slope in Degrees' dataset had to be reclassified because the value number was based on the average elevation difference between adjacent cells so if the dimensions of the raster cell is changed the value changes. To correct for this, I reclassified the copied 'Terrain: slope in Degrees' layer to have a continuous 0 to 90-degree scale.

The 'World Population Estimate 2016' dataset was added to the model to determine if there was an association between human occupation and Black Rails, which is important when considering land management. The estimates in this data set are based on the surface footprint of

human residence. This layer has no data in raster cells categorized as having zero population, so I used the 'Is Null' tool in ArcMap to create a layer where cells with data had a value of 0 and cells without data have a value of 1. I then used the Con tool to merge the copied layer to the null layer to create a layer where every cell has a value and cells that have no human residents have a value of 0.

Previous models found that Black Rails favored high marsh with shallow water, so the 'U.S. Soils Water Table Depth' was used in the model to help inform habitat suitability. This layer had no data in raster cells categorized as having Zero water depth, so I used the 'Is Null' tool in ArcMap to create a layer where cells with data had a value of 0 and cells without data have a value of 1. I then used the Con tool to merge the copied layer to the null layer to create a layer where every cell had a value and cells that did not have water had water a value of 0.

Finally, the World Bioclimate data was added to the model. Climate is a large part of the habitat environment and many models use some form of climate data when modeling a species distribution over a wide area (Coxen et al., 2017; Hu & Liu, 2014; Moreno et al., 2011; Phillips et al., 2006, 2004; Wu et al., 2012). This dataset uses temperature and aridity descriptions to categorize climate conditions (e.g. warm and dry).

I converted all of the modified raster layer files to ASCII files so they could be used in the MaxEnt program. In order to get a file of Black Rail presence points with the same coordinates as the environmental files, I added two new double type columns to the attribute table of the selected Black Rail points and filled one with the calculated geometry of the x coordinate and the other with the y coordinate.

Because the presence data are from a citizen science initiative, it is likely that multiple people recorded and submitted occurrences of the same bird from the same time and place. To eliminate obvious replicates from the Black Rail presence file, entries with the same coordinates were removed and a new file with only unique coordinates was created using the program R (R Core Team, 2019). The model is therefore not weighted by Black Rail population density.

The unique presence file was then used as the sample file and the ASCII files of all the environmental data were used as the environmental layer inputs in the MaxEnt program (Philip et al., MaxEnt software version 3.4.1). The random test percentage was set to 50 percent with all other default settings. Response curves and jackknife plots were created for the model to determine the importance of the environmental variables and to look at the relationship between Black Rail presence and the environmental variables.

Model selection

To ensure the model used was the most optimal several versions were tested using different combinations of environmental variables and were assessed using the test data "Area Under the Receiver Operating Characteristic Curve" (AUC) value. The environmental variables and their association with Black Rail presence were assessed by looking at the Jackknife plot, percent contribution, and permutation importance to determine which group of variables was the most predictive of Black Rails. I first tried a model with all of the above variables and the three variables with the lowest percent contribution and permutation importance were selected out (Table 2-2). Models were made with these three variables being omitted alone or in combinations, to see if their omission or inclusion affected the quality of the model (Table 2-3).

The model with the highest test data AUC value was selected for final interpretation. The final model's response curves were used to determine what values or categories of the environmental variables were most closely associated with Black Rail presence.

Results:

The eBird data set had 1597 observations of Black Rail within the bounds of the area set for this study (-87.822860 dd to -70.788500 dd Longitude and 44.175883 dd to 24.885264 dd Latitude). Of these observations, 366 of them had observation dates before 2002 when the eBird database was created (Appendix, Table A3). As mentioned before, some of the locations had multiple observations, some likely submitted for the same bird at the time by multiple people. Under a systematic formal data collection protocol with the explicit intent of determining a species presence through surveying an area multiple times, multiple observations would help strengthen a spatial model by adding meaningful information about species density. With the eBird dataset, there is no formal data collection protocol. Therefore, repeat entries for the same coordinates was stochastic, rendering multiple entries meaningless for the habitat suitability model. For this reason, extraneous observations with the same coordinates were eliminated, leaving a list of 434 unique coordinates that were entered into the habitat suitability model. Of these unique points, 132 of them had detections exclusively before 2002, 302 points had observations exclusively after 2002, and 30 points had observations both before and after 2002 (Appendix, Table A3).

I ran a model using all the environmental variables, and the three variables with the lowest contribution and importance were selected out (Table 2-2). The Water Table Depth, Estimated Population, and Bioclimate data were the three variables with the least effect on the model. Different variations of the model were attempted excluding these variables sequentially, alone and in different combinations, to see if the model could be improved without their data input. Excluding the bioclimate data produced the model with the highest test AUC (Table 2-3). This was considered the final model (Table 2-4) and was subject to further analyses to characterize the contributions of individual variables.

Table 2-2. Analysis of variable contributions for a model including all environmental variables. The bottom three variables, with the lowest contribution and importance, were selected out for testing of the effect of their exclusion and inclusion on the quality of the model.

Variable	Percent contribution	Permutation importance
Distance to Water	36.7	27.1
Land Cover Classes	26.5	11
Ecological Landform Classes	11.8	11.7
Slope	10.3	33.9
Lithology	9.9	5.1
Bioclimate	2.8	3.4
Estimated Population	1.6	5.4
Water Table Depth	0.5	2.4

Table 2-3. Variable exclusion model test using the three variables with the least influence on the model. Different models included (+) or excluded (-) the bioclimate, water depth, or population data. The quality of the model was assessed using the test AUC and the model with the highest value (highlighted) was considered the final model.

Model	Bioclimate	Water depth	Population	Training AUC	Test AUC
All variables included	+	+	+	0.892	0.853
Without population	+	+	-	0.885	0.851
Without water depth	+	-	+	0.89	0.853
Without bioclimate	-	+	+	0.882	0.874
Without water depth and population	+	-	-	0.883	0.852
Without bioclimate and population	-	+	-	0.878	0.871
Without bioclimate and water depth	-	-	+	0.883	0.872
Without bottom three variables	-	-	-	0.878	0.869

Table 2-4. Percent contribution and permutation importance of each environmental variable included in the final model.

Variable	Percent contribution	Permutation importance	Data source layer
Land Cover Classes	42.1	20.8	World Land Cover ESA 2010
Distance to Water	26.9	23.5	World Distance to Water
Ecological Landform classes	12.3	23.5	World Ecological Facets Landform Classes
Lithology	9.4	12.8	World Lithology
Slope	8.1	12.4	Terrain: Slope in Degrees
Population	0.7	4.4	World Population Estimate 2016
Water Table Depth	0.6	2.6	U.S. Soil Water Table Depth

Models were assessed using the AUC value. The AUC is a threshold independent measure with a value range of 0 to 1. An AUC value 1 indicates a model with perfect predictive ability and a value of 0.5 indicates a model with predictive abilities no better than random (Phillips et al., 2006, 2004). The receiver operating characteristic curve shows an AUC value of 0.874 for this model, indicating the model's predictive ability is significantly better than random (Figure 2-1).

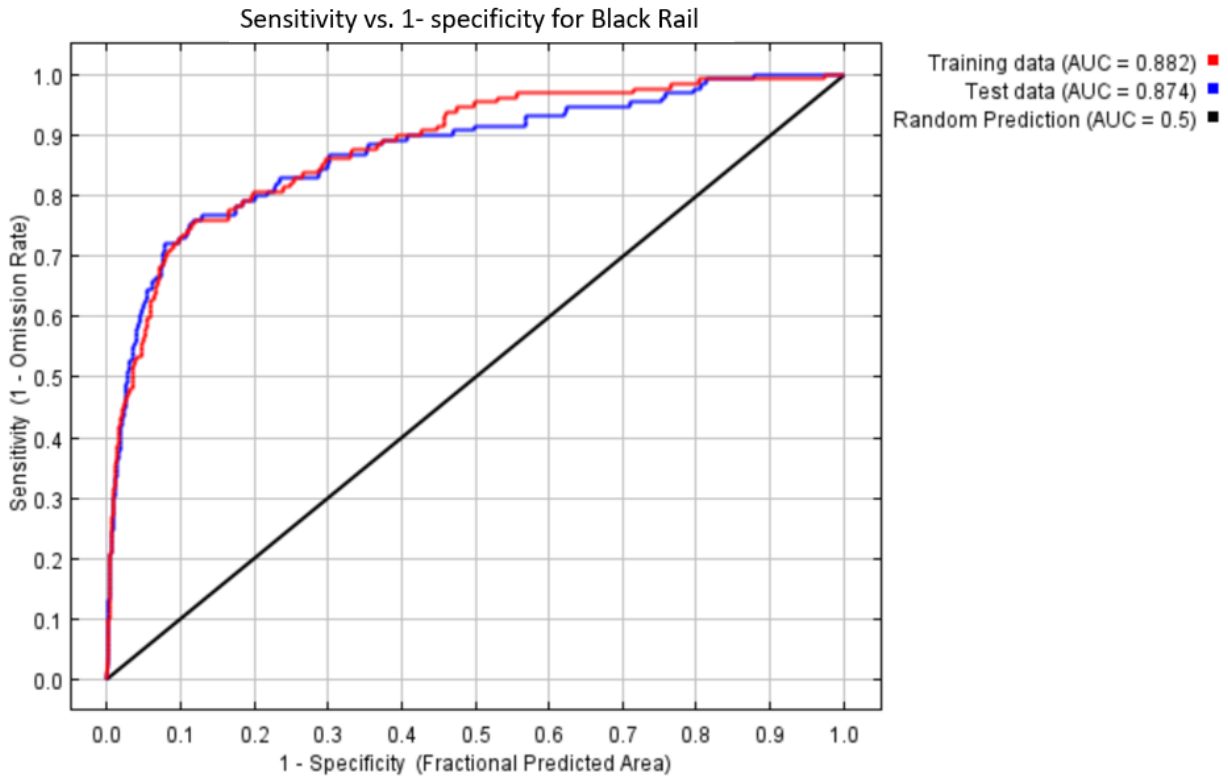


Figure 2-1. Receiver operator characteristic curve with area under the curve (AUC) value for the final Black Rail MaxEnt model. The training data (red) were used to set the constraints of the model and the test data (blue) were used to create the rest of the model output.

A jackknife plot was generated for the final model (Figure 2-2). A high gain or value for the "with variable only" model indicates that the variable is informative by itself. A variable that

reduces the gain or value the most when it is excluded, “without variable”, indicates a variable that contributes the most information not found in the other variables.

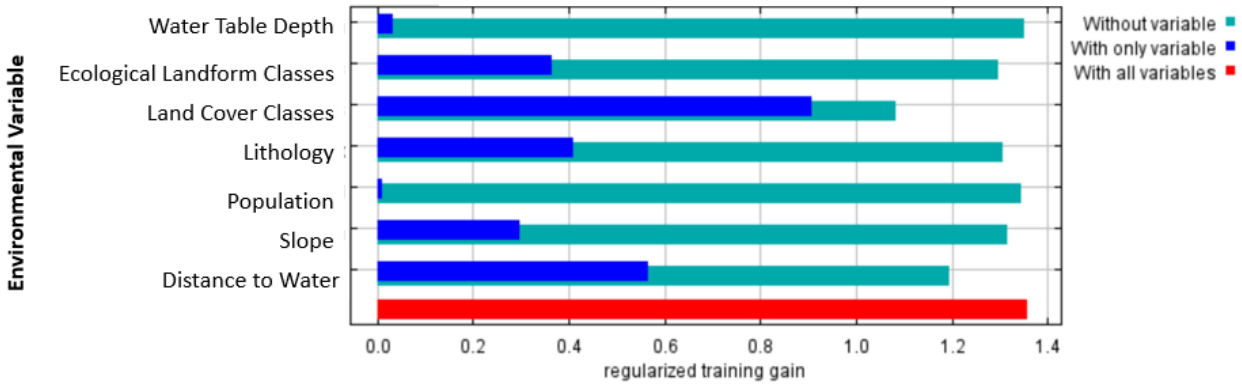


Figure 2-2. Jackknife of regularized training gain for Black Rail. The change in model gain, or effectiveness, is illustrated when different variables are excluded (green) or used exclusively (blue). The gain when including all variables is shown in red.

For each variable, two response plots were generated that show the prediction probability change, in ‘cloglog’ units, with different values or categories of the environmental variables. A high cloglog value indicates that the value or category of the environmental variable is positively associated with target species presence. One plot (e.g. Figure 2-8), called the marginal plot, is generated by looking at the change in just the one environmental variable while keeping all the other variables at their average sample value. The other plots (e.g. Figures 2-3 through 2-7 and 2-9), referred to here as independent plots, are the results of a model using just that individual variable and showing the prediction probability associated with each value or category.

Of the variables included in the final model, land cover data was the most predictive of Black Rail presence. According to the jackknife plot of regularized training gain, this variable yielded the most useful information independently, including the greatest amount of information

not explained by the other variables (Figure 2-2). The land cover data, which describes the vegetation structure, also had the highest percent contribution and the second highest permutation importance value (Table 2-4). The land cover classes that had the highest responses are “shrub or herbaceous cover, flooded”, “tree cover, flooded, saline water”, “shrubland”, and “water bodies” respectively (Figure 2-3).

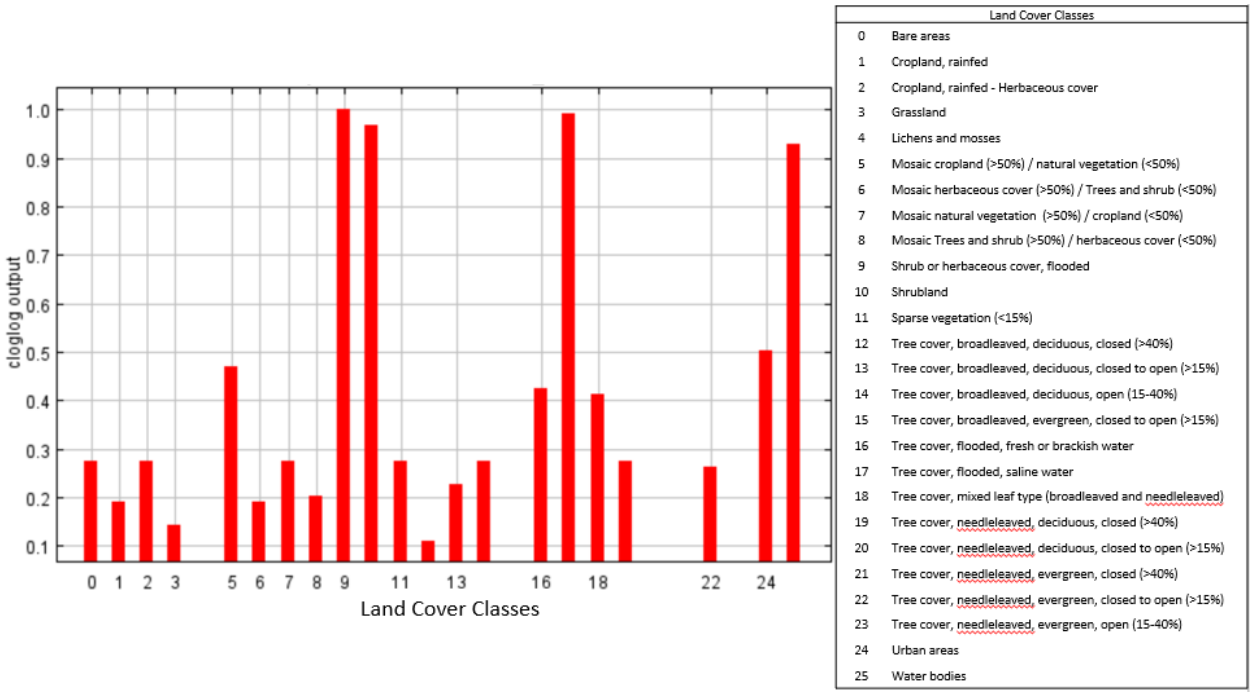


Figure 2-3. The response of Black Rail presence to land cover classes. These are the results from an independent model using only the Land Cover Classes ESA 2010 data. The ‘cloglog’ output is a measure of change in the predicted probability of presence of Black Rail.

The distance to water data had the next highest percent contribution and the highest permutation importance indicating its importance to the model. The response plot for the distance to water shows a precipitous drop in predictability the farther from water (Figure 2-4). The next most important variable according to the percent contribution and jackknife plot is the ecological landform classes (Table 2-4, Figure 2-2). Looking at the response histogram the one major category that stands out is the flat or nearly flat plains (Figure 2-5).

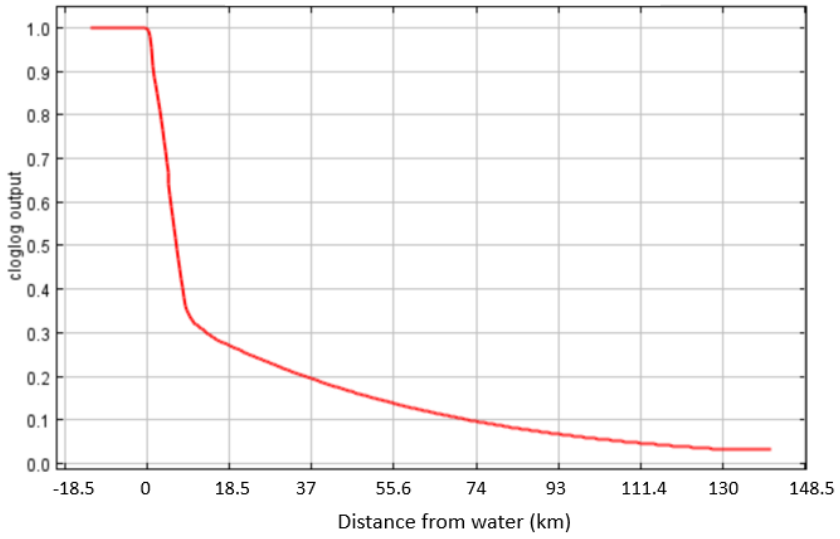


Figure 2-4. The response of Black Rail presence to distance from water. This independent plot is the result from an independent model using only the World Distance to Water data. The cloglog output is a measure of change in the predicted probability of presence of Black Rail.

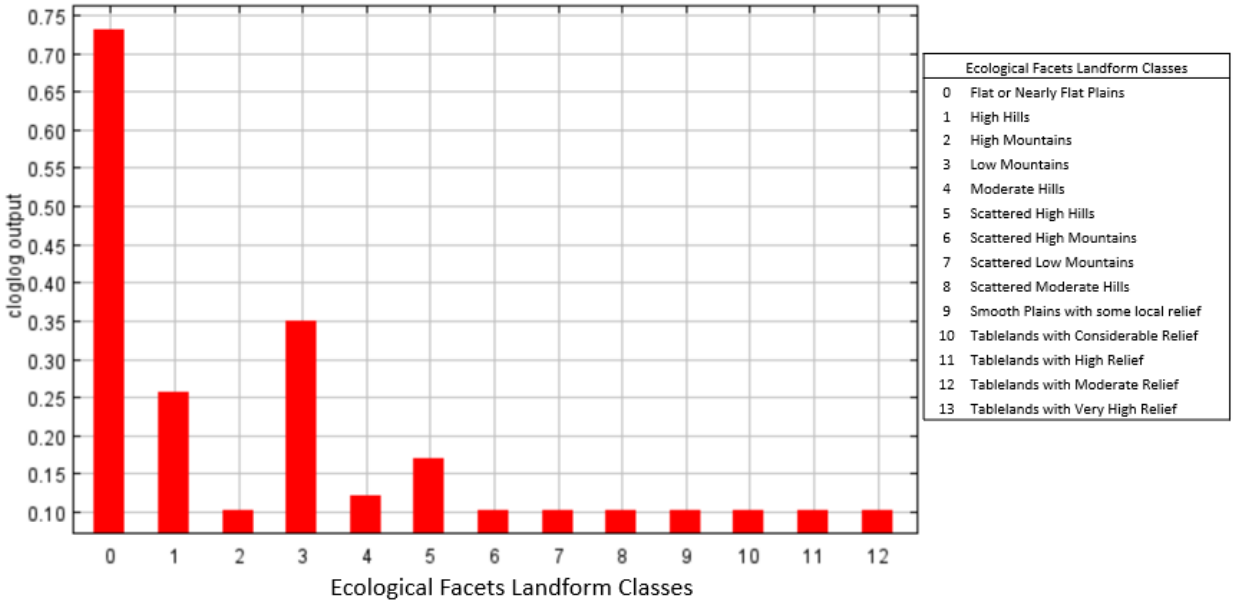


Figure 2-5. The response of Black Rail presence to ecological landform classes. This independent plot is the result from an independent model using only the World Ecological Landform Classes data. The cloglog output is a measure of change in the predicted probability of presence of Black Rail.

The “World Lithology” data which describes soil types came fourth in importance with fairly low percent contribution (Table 2-4). The soil types with the strongest response were non-defined, intermediate volcanic, and unconsolidated, respectively (Figure 2-6).

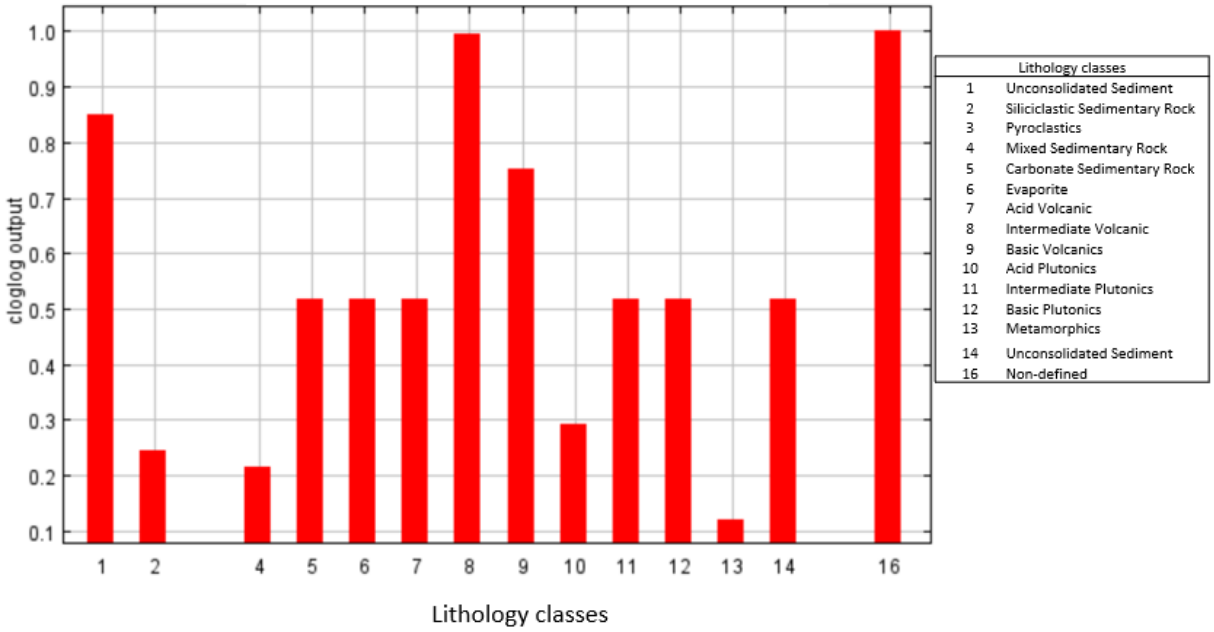


Figure 2-6. The response of Black Rail presence to lithology. This independent plot is the result from an independent model using only the World Lithology data. The cloglog output is a measure of change in the predicted probability of presence of Black Rail.

The “Terrain: Slope in degrees” had percent contribution and permutation importance score just below the Lithology data (Table 2-4). The response curve shows a steep drop at the lowest slopes (~0-3°) and then a steady decline as slope increases (Figure 2-7).

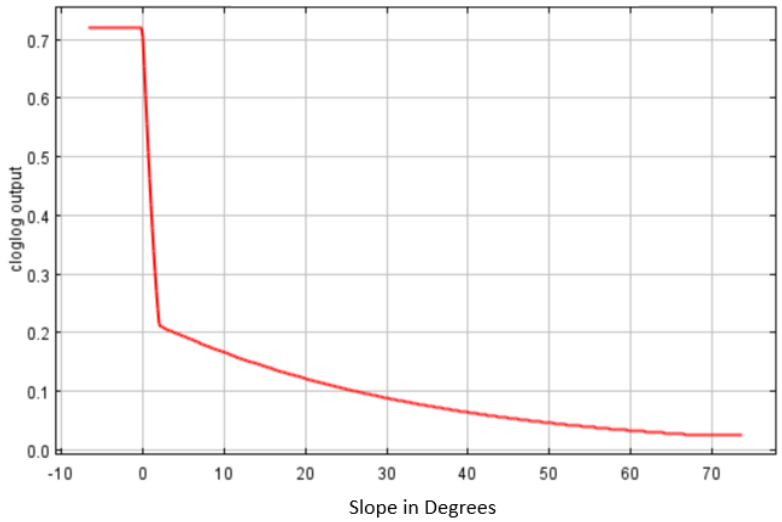


Figure 2-7. The response of Black Rail presence to slope. This independent plot is the result from an independent model using only the Terrain: Slope in Degrees data. The cloglog output is a measure of change in the predicted probability of presence of Black Rail.

With a significant drop in both the percent contribution and permutation importance the "World Estimated Population 2016" data is second to last in importance. As the relative population of humans increases the response continually decreases (Figure 2-8).

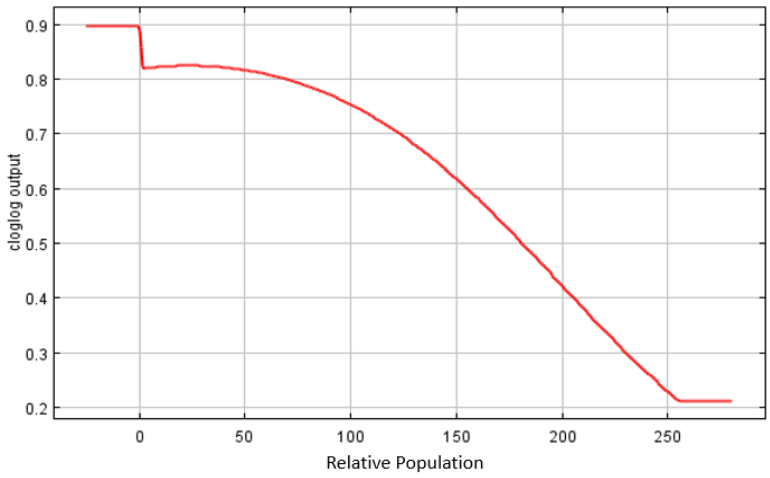


Figure 2-8. The response of Black Rail presence to the relative estimated human population. This marginalized plot is the result from a model where changes in the response to "World Estimated Population 2016" data was examined while all other variables were kept at their average sample value. The cloglog output is a measure of change in the predicted probability of presence of Black Rail. The population value is a relative estimate based on the footprint of where people live: a higher number means more people and a lower number mean fewer people.

Surprisingly, among the ecological variables included in the model, Water Table Depth had the lowest percent contribution and lowest permutation importance (Table 2-4). As water depth increased, predicted Black Rail presence decreased exponentially before marginally increasing again (Figure 2-9).

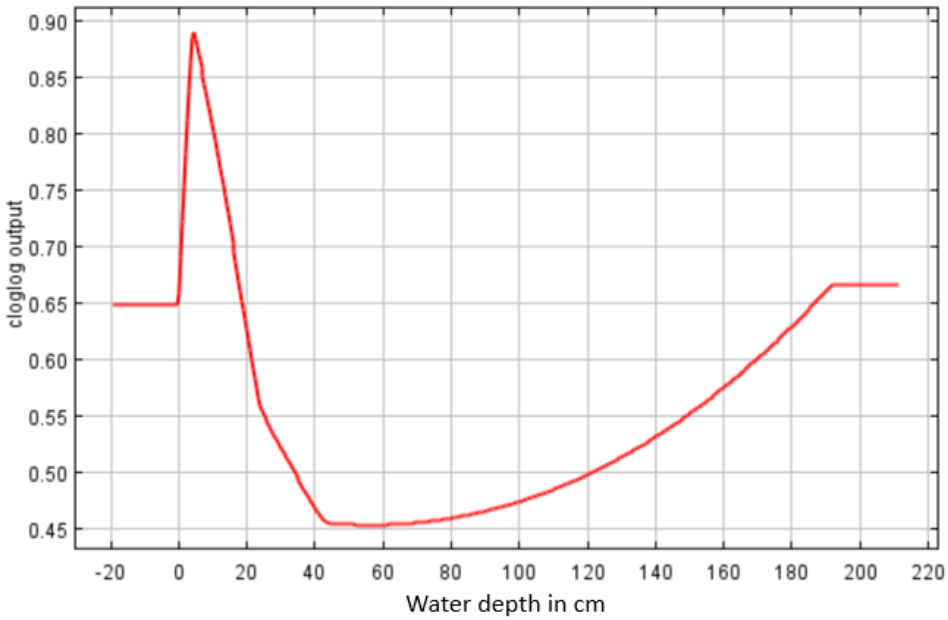


Figure 2-9. The response of Black Rail presence to water depth. This independent plot is the result from an independent model using only the U.S. Water Table Depth data. The cloglog output is a measure of change in the predicted probability of presence of Black Rail.

The map created by the MaxEnt model shows that suitable habitat is generally on the coast with a few small pockets inland (Figure 2-10). There are a few noticeable hotspots such as southern Florida, and the coasts of Georgia, South Carolina, North Carolina, Maryland, and New Jersey.

Coastal North Carolina has long been considered a stronghold for Eastern Black Rails on the Atlantic coast, and Cedar Island was considered the center of this population (Watts, 2016). I performed targeted audio callback surveys for Black Rail during the 2018 and 2019 breeding seasons (see Chapter 1). Only 12 out of 620 surveys (2%), conducted over both years, detected Black Rails, and this comprised only 6 out of 115 survey points for 2019, a 4% occupancy rate. Therefore, I focused specifically on the map of North Carolina generated by the model. As expected, the most suitable habitat was concentrated around the Pamlico Sound (Figure 2-11). There is a notable small inland pocket near Mebane, North Carolina.

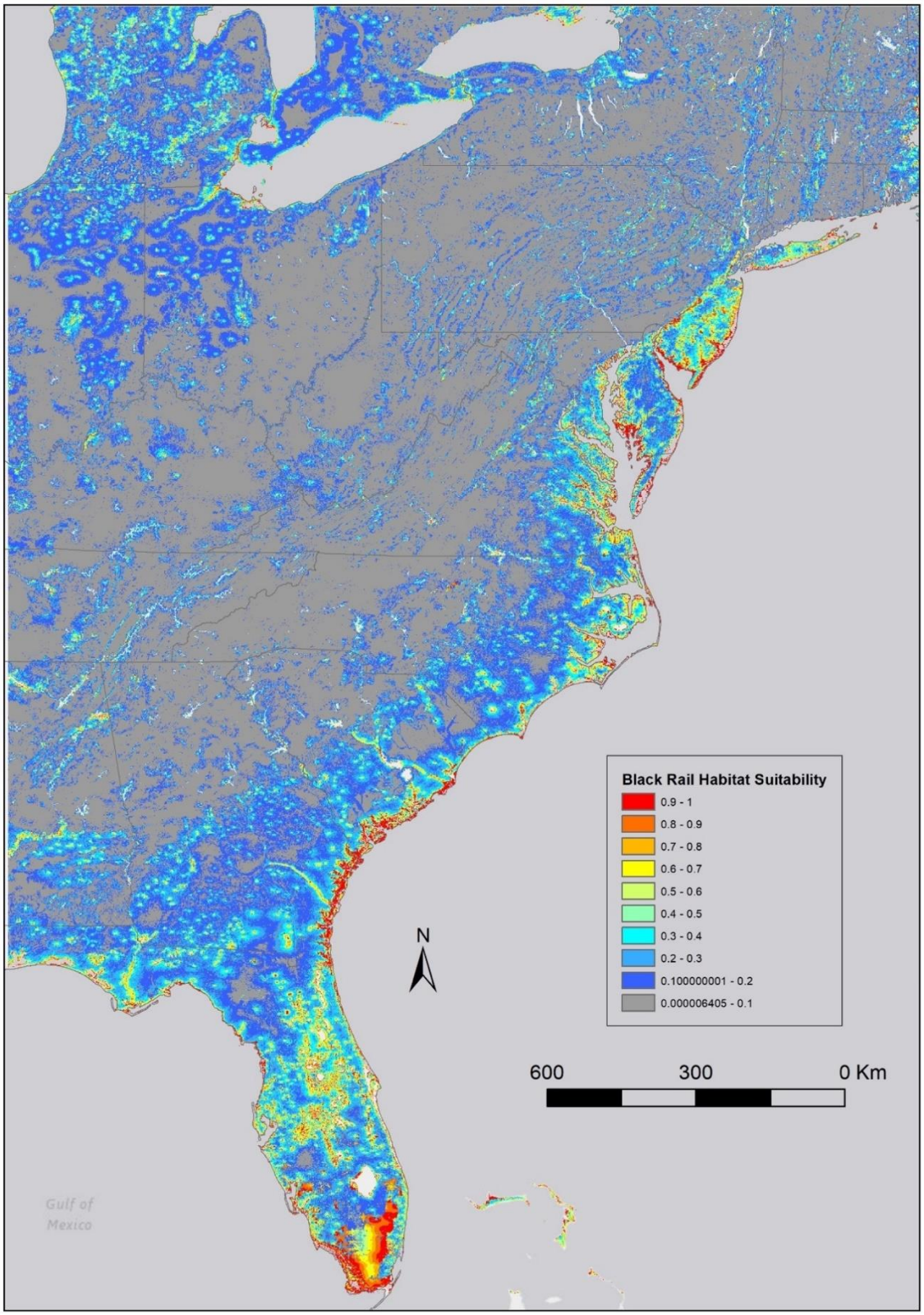


Figure 2-10. Eastern Black Rail habitat suitability across the Atlantic coast of the U.S. created using the optimized MaxEnt model. The legend scale indicates probability of suitable habitat. A value of 1 indicates the most suitable habitat. White terrestrial areas are areas with no data.

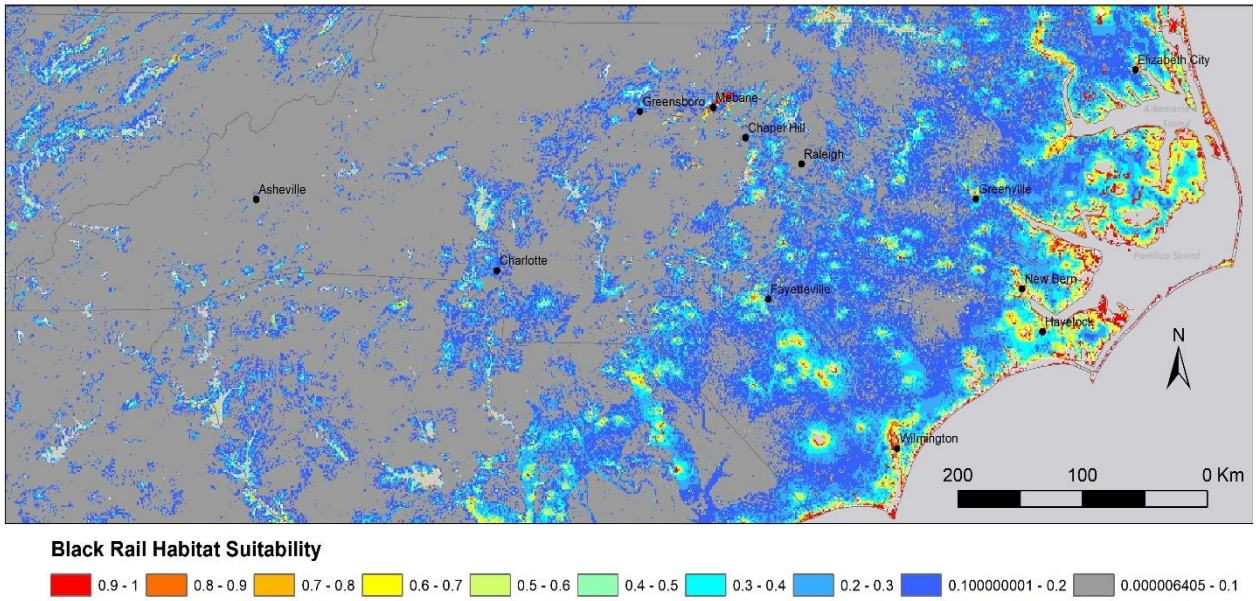


Figure 2-11. Detailed habitat suitability for Black Rail in North Carolina and surrounding regions with reference points. The scale in the legend indicates probability of suitable Black Rail habitat, with the highest probability in red. White terrestrial areas are areas with no data.

Discussion

The final model for Black Rail habitat suitability included the land cover, ecological land facets, distance to water, lithology, slope, human population, and water depth data. The AUC for the model (0.874) was well within the acceptable range (>0.7) and consistent with other MaxEnt models for birds (Haverland, 2019 (AUC=0.67); Hu & Liu, 2014 (AUC= 0.81); Moreno et al., 2011 (AUC= 0.87 and 0.99); Tsao et al., 2015 (AUC=0.92); Wu et al., 2012 (multiple AUC values ranging from 0.7-0.98)).

The factor most associated with Black Rail presence in this model was land cover, and specifically flooded shrub and herbaceous vegetation. All of the land cover categories that had the highest responses in the model, “shrub or herbaceous cover, flooded”, “tree cover, flooded, saline water”, “shrubland”, and “water bodies” support the hypothesis that flooding in association with vegetation – characteristics generally associated with marsh - would be important. These are all categories that fit the predicted high marsh habitat that Black Rail were reported to occupy in another habitat association study that looked at vegetation type on the Gulf coast in Texas (Haverland, 2019).

As predicted for a wetland bird, distance to water was also important to the model. The model indicated a higher predicted presence closer to water, and it is reasonable to assume that marsh habitat would be considered either 0 km away or very close to water. In California, using habitat data collected at each survey site, Black Rails were shown to have a negative correlation with distance from channels (Tsao et al., 2015). This shows that even on a smaller scale and considering only suspected Black Rail habitat that Black Rail are associated with proximity to

open water. Agreement between this and other Black Rail habitat association studies strengthens the confidence in the model despite its large scale and use of citizen science data.

Among the ecological landform classes, flat land was most predictive, also supporting my hypothesis considering that, in a remote-sensing framework, marshes typically have little to no above-water topography. Even in areas with surrounding topography it has previously been shown that Black Rail choose relatively level areas. The population in the Sierra Nevada foothills were found mostly in places with relatively little slope or elevation (Richmond et al., 2010).

The high response for unconsolidated lithology is understandable because flooded soils associated with marshes generally have loose sediment (Gornitz et al., 1994). Unconsolidated soils make up a large portion of the Atlantic coastal plain and are most vulnerable to erosion, adding to the sensitivity of salt marsh habitat (Gornitz et al., 1994). The two other significant lithology types identified by the model, intermediate volcanic (intermediate referring to pH) and non-defined, had a higher predictability than the unconsolidated category. This prompted scrutiny of the lithology data in ArcMap. After looking at the lithology data and the Black Rail occurrence points a curious pattern appeared. There were comparatively few raster cells with undefined (1,115 cells or 0.03%) or intermediate volcanic lithology (2,025 cells or 0.06%) in the East Coast area entered into this model, but in the few places Black Rails have been recorded these lithology types are overrepresented. This may be a coincidence, an artifact of very small sample size for the lithology types, or perhaps there is something about the lithology types that are preferred by Black Rail. This warrants further investigation. Although the results of the lithology response are interesting, the importance of lithology to the model should be considered.

The percent contribution and permutation importance of the lithology data is low, so although the results are interesting it has comparatively little influence on the model.

California Black Rails live in a severely water-limited environment in the Sierra Nevada foothills, and their occupancy is strongly related to presence of even small tracts of marsh (depressional, fluvial, fringe marsh or slope) (Richmond et al., 2010). That flooding is an important habitat requirement is evidenced by rapid colonization occurring within a year of marsh creation (Richmond et al., 2008). On the East Coast, water is not limiting to the same extent, but water depth remains important. Previous studies provide evidence that a moderate slope facilitates the essential hydrology required for Black Rail habitat (Nadeau & Conway, 2015; Atlantic Coast Joint Venture, 2020). Persistent shallow water is difficult to achieve in flat areas that are flooded, but the presence of a gentle slope in a permanent wetland means that while the water may rise and recede, there should always be an area within the favorable range of depth. The percent contribution and permutation importance for slope in my final model emphasized less importance of sloped landscape. However, having less slope overall being more predictive of Black Rails is consistent with my original hypothesis, given that this is the typical condition in marshes. This pattern closely matches the findings of the study that looked at habitat characteristics including slope for a California population; Black Rails were found in areas with an average slope of only 3.7 degrees (Richmond et al., 2008). Moreover, the moderate degree of slope required locally for suitable hydrology was probably not compatible with the scale of this model.

Based on previous studies, I predicted that water depth would be comparatively influential in the model, but the results showed that in fact water depth had the least impact

among the variables in the final model. The response plot using only water depth indicated that Black Rail presence has the highest response with shallower water (~5 cm, as expected from previous studies). Black Rails have been reported to inhabit fringe marshes (marsh along the edge of bodies of water) (Richmond et al., 2008), which could explain the correlation with deeper water. These smaller areas of low water depth near deeper water could have become obscured when the raster cell size was modified to larger than the original data and caused them to have a higher water depth value explaining the additional (albeit smaller) association with higher water depth values.

Human population density is a concern when dealing with sensitive species, especially when considering land management (i.e. where to invest in land preservation and restoration projects). The response results for population in this model supported the hypothesis that there would generally be fewer Black Rails in more densely populated areas. Although the response presented the pattern expected for a sensitive marsh dwelling species, the low percent contribution and low permutation importance shows that human population numbers are not inherently a major influence on Black Rail occupation. Nevertheless, the inverse relation between Black Rails and humans warrants some review. Black Rails have been declining in coastal areas in the southeast where human development has increased substantially over the same period (Crossett et. al., 2004; Crossett et. al., 2013). It is unclear whether there is a causal effect of human disturbance. Future studies should investigate this relationship further.

North Carolina is a historical stronghold for Black Rail and has been identified as population center targeted for management by the Atlantic Coast Joint Venture (Watts, 2016; U.S. Fish and Wildlife Service, 2020) Recent surveys in North Carolina have shown a drastic

decrease in the number of Black Rails detected (Wilson et al., 2016). These surveys detected very few individuals at even fewer locations. In many cases, places that used to be hotspots have had significantly fewer or no detections in the last few years (Watts, 2016; Wilson et al., 2016). My own surveys during the 2018 and 2019 breeding seasons reflect these trends with very low detection and occupancy rates. The model showed that there is still suitable habitat in North Carolina on both the coast and farther inland. These inland sites may become especially important as refuges for Black Rails as coastal marshes disappear due to sea level rise or human development (Nicholls, 2004). The Joint Venture has an explicit goal to increase non-tidal habitat for Black Rail because of the predicted loss of coastal wetlands (Atlantic Coast Joint Venture, 2020). The habitat suitability map generated by this model identifies areas that should be inspected for Black Rail occupancy and considered areas of potential for future management.

This model is the first to look at habitat suitability for Black Rails for more than one state on the East Coast. It adds to a growing list of studies showing that models using distribution data collected through citizen science can have significant predictive ability. Areas that the model predicted would have high habitat suitability were consistent with historically known locations (Watts, 2016; U.S. Fish and Wildlife Service, 2018a; Atlantic Coast Joint Venture, 2020). Furthermore, the model corroborated the same general habitat requirements for Black Rail found in California, Texas and other parts of the species' range (Tsao et al. 2015; Haverland 2019), namely flooding in association with herbaceous and shrubby vegetation, also apply to the wider East Coast range. This model better elucidated the environmental variables that are most predictive of Black Rail occupancy. The most influential variables were those which characterize high marsh, confirming the importance of this habitat type to Black Rails. The map product will

be useful for finding areas that are likely to support Black Rails now and in the future, and help inform land management plans in support of species conservation.

Literature Cited

Atlantic Coast Joint Venture Black Rail Working Group. (2020). Eastern Black Rail Conservation Plan. *Unpublished Report*.

BirdLife International. (2019). *IUCN red list of threatened species: Black Rail*. IUCN Red List of Threatened Species. https://www.iucnredlist.org/en

Bradter, U., Mair, L., Jönsson, M., Knape, J., Singer, A., & Snäll, T. (2018). Can opportunistically collected Citizen Science data fill a data gap for habitat suitability models of less common species? *Methods in Ecology and Evolution*, *9*(7), 1667–1678. https://doi.org/10.1111/2041-210X.13012

Conway, C. J. (2011). Standardized North American marsh bird monitoring protocol. *Waterbirds*, *34*(3), 319–346.

Corsi, F., de Leeuw, J., Skidmore, A. (2000). Modeling species distribution with GIS. In: Boitani, L., Fuller, T. (Eds.), Research Techniques in Animal Ecology. Columbia University Press, New York, 389–434.

Coxen, C. L., Frey, J. K., Carleton, S. A., & Collins, D. P. (2017). Species distribution models for a migratory bird based on citizen science and satellite tracking data. *Global Ecology and Conservation*, *11*, 298–311. https://doi.org/10.1016/j.gecco.2017.08.001

Crossett, K.M., T.J. Culliton, P.C. Wiley, and T.R. Goodspeed. (2004). Population Trends Along the Coastal United States: 1980-2008, Coastal Trends Report Series (vol 55).US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Management and Budget Office, Special Projects. Silver Spring, MD.

Crossett, K., Ache, B., Pacheco, P., and Haber, K. (2013). National Coastal Population Report Population Trends from 1970 to 2020. *NOAA State of the Coast Report Series, US Department of Commerce, Washington*.

Davidson, L. M. (1992). Black Rail. In K. J. Schneider & D. M. Pence (Eds.), *Migratory non game birds of management concern in the Northeast* (pp. 119–134). U.S. Fish and Wildlife Service.

eBird. 2018. eBird: An online database of bird distribution and abundance [web application]. eBird, Cornell Lab of Ornithology, Ithaca, New York. Available: http://www.ebird.org. (Accessed: Date [e.g., May 1, 2018]).

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, *17*(1), 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Gornitz, V. M., Daniels, R. C., White, T. W., & Birdwell, K. R. (1994). The Development of a Coastal Risk Assessment Database: Vulnerability to Sea-Level Rise in the U.S. Southeast. *Journal of Coastal Research*, special issue *12*, 327–338.

Hand, C. E., Znidersic, E., & Tegeler, A. K. (2019). First documentation of eastern Black Rails (*Laterallus jamaicensis jamaicensis*) breeding in South Carolina, USA in more than a century. *Waterbirds*, *42*(2), 237–241. https://doi.org/10.1675/063.042.0212

Haverland, A. A. (2019). Determining the status and distribution of the eastern Black Rail (*Laterallus jamaicensis*) in coastal Texas. *Texas State University*. dissertation

Hu, J., & Liu, Y. (2014). Unveiling the conservation biogeography of a data-deficient endangered bird species under climate change. *PLoS ONE*, *9*(1), e84529. https://doi.org/10.1371/journal.pone.0084529

Legare, M. L., & Eddleman, W. R. (2001). Home range size, nest-site selection and nesting success of Black Rails in Florida. *Journal of Field Ornithology*, *72*(1), 170–177.

Moreno, R., Zamora, R., Ramón Molina, J., Vasquez, A., & Ángel Herrera, M. (2011). Predictive modeling of microhabitats for endemic birds in South Chilean temperate forests using maximum entropy (Maxent). *Ecological Informatics*, *6*(6), 364–370. https://doi.org/10.1016/j.ecoinf.2011.07.003

Na, X., Zhou, H., Zang, S., Wu, C., Li, W., & Li, M. (2018). Maximum entropy modeling for habitat suitability assessment of Red-crowned Crane. *Ecological Indicators*, *91*, 439–446. https://doi.org/10.1016/j.ecolind.2018.04.013

Nadeau, C. P., & Conway, C. J. (2015). Optimizing water depth for wetland-dependent wildlife could increase wetland restoration success, water efficiency, and water security. *Restoration Ecology*, *23*(3), 292–300. https://doi.org/10.1111/rec.12180

Nicholls, R. J. (2004). Coastal flooding and wetland loss in the 21st century: Changes under the SRES climate and socio-economic scenarios. *Global Environmental Change*, *14*, 69–86.

Palomino, D., Seoane, J., Carrascal, L. M., & Alonso, C. L. (2008). Competing effects of topographic, lithological, vegetation structure and human impact in the habitat preferences of the Cream-coloured Courser. *Journal of Arid Environments*, *72*(4), 401 410.

Pearce, J. L., & Boyce, M. S. (2006). Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, *43*(3), 405–412. https://doi.org/10.1111/j.1365 2664.2005.01112.x

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S. J., Dudík, M., & Schapire, R. E. (2004). A maximum entropy approach to species distribution modeling. *Twenty-First International Conference on Machine Learning ICML '04*, 83. https://doi.org/10.1145/1015330.1015412

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richmond, O. M., Chen, S. K., Risk, B. B., Tecklin, J., & Beissinger, S. R. (2010). California Black Rails depend on irrigation-fed wetlands in the Sierra Nevada foothills. *California Agriculture*, *64*(2), 85–93. https://doi.org/10.3733/ca.v064n02p85

Richmond, O. M., Tecklin, J., & Beissinger, S. R. (2008). Distribution of California Black Rails in the Sierra Nevada foothills. *Journal of Field Ornithology*, *79*(4), 381–390. https://doi.org/10.1111/j.1557-9263.2008.00195.x

Roach, N. S., & Barrett, K. (2015). Managed habitats increase occupancy of Black Rails (*Laterallus jamaicensis*) and may buffer impacts from sea level rise. *Wetlands*, *35*(6), 1065–1076. https://doi.org/10.1007/s13157-015-0695-6

Spautz, H., Nur, N., & Stralberg, D. (2005). California Black Rail (*Laterallus jamaicensis coturniculus*) distribution and abundance in relation to habitat and landscape features in the San Francisco Bay estuary. Pages 20–24 in J. C. Ralph and T. D. Rich, editors. Bird conservation implementation and integration in the Americas: Proceedings of the Third International Partners in Flight Conference. http://www.treesearch.fs.fed.us/pubs/31463

Stockwell, D. R. B., & Noble, I. R. (1992). Induction of sets of rules from animal distribution data: A robust and informative method of data analysis. *Mathematics and Computers in Simulation*, *33*(5–6), 385–390. https://doi.org/10.1016/0378-4754(92)90126-2

Stuart, G. H. (1920). Nesting of the little Black Rail in Atlantic County, N. J. *The Auk*, *37*(2), 292–293. https://doi.org/10.2307/4072725

Taylor, B., & Van Perlo, B. (1998). *Rails: A guide to the rails, crakes, gallinules and coots of the world*. Pica Press. Sussex, U.K.

Tsao, D. C., Melcer, Jr., R. E., & Bradbury, M. (2015). Distribution and habitat associations of California Black Rail (*Laterallus jamaicensis cortuniculus*) in the Sacramento–San Joaquin Delta. *San Francisco Estuary and Watershed Science*, *13*:4.

U.S. Fish and Wildlife Service. (2018a). Species status assessment report for the Eastern Black Rail (*Laterallus jamaicensis jamaicensis)*, Version 1.2. June 2018. Atlanta, GA

U.S. Fish and Wildlife Service. (2018b). 12-month petition finding and threatened species status for Eastern Black Rail with a section 4(d) rule. Federal Register Agency Docket No. FWS-R4-ES-2018-0057 pp.50610-50630.

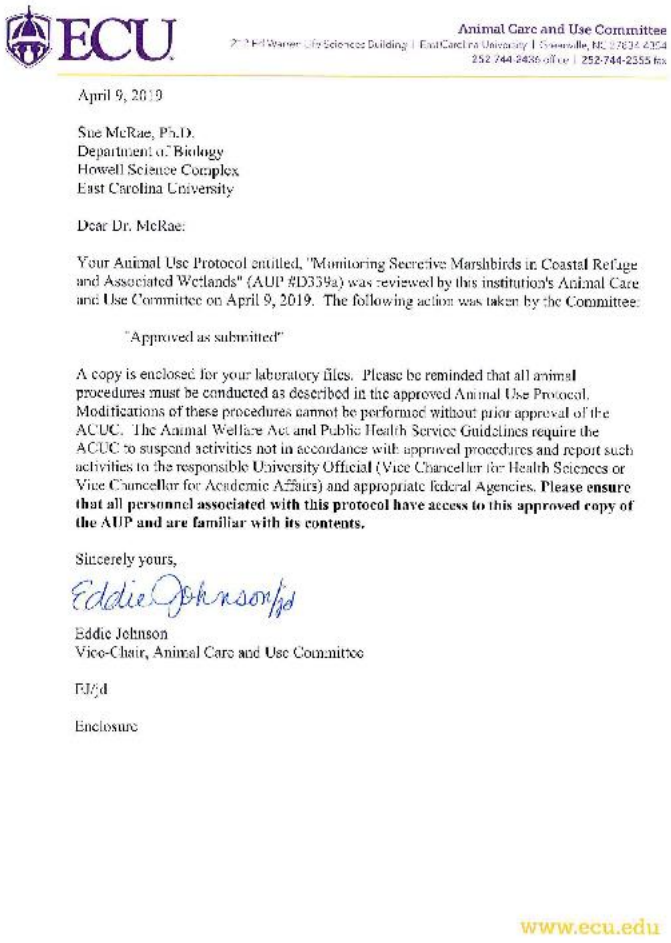
Walker, J., & Taylor, P. D. (2017). Using eBird data to model population change of migratory bird species. *Avian Conservation and Ecology*, *12*(1):4.

Watts, B. D. (2016). Status and distribution of the Eastern Black Rail along the Atlantic and Gulf Coasts of North America. *The Center for Conservation Biology Technical Report Series*, CCBTR-16-09. College of William and Mary/Virginia Commonwealth University, Williamsburg, VA. 148 pp.

Wilson, M.D., Watts, B. D., & Poulton, D. (2016). Black Rails in North Carolina. *Center for Biology Technical Report Series*, CCBTR-16-01, College of William and Mary and Virginia Commonwealth University, Williamsburg, V.A.

Wu, T.Y., Lee, P.F., Lin, R.S., Wu, J.L., & Walther, B. A. (2012). Modeling the distribution of rare or cryptic bird species of Taiwan. *Taiwania*, *57*(4), 342-358.

APPENDIX



ECU

Animal Care and Use Committee
252-744-2436 office | 252-744-2355 fax

April 9, 2019

Sue McRae, Ph.D.
Department of Biology
Howell Science Complex
East Carolina University

Dear Dr. McRae:

Your Animal Use Protocol entitled, "Monitoring Secretive Marshbirds in Coastal Refuge and Associated Wetlands" (AUP #D339a) was reviewed by this institution's Animal Care and Use Committee on April 9, 2019. The following action was taken by the Committee:

"Approved as submitted"

A copy is enclosed for your laboratory files. Please be reminded that all animal procedures must be conducted as described in the approved Animal Use Protocol. Modifications of these procedures cannot be performed without prior approval of the ACUC. The Animal Welfare Act and Public Health Service Guidelines require the ACUC to suspend activities not in accordance with approved procedures and report such activities to the responsible University Official (Vice Chancellor for Health Sciences or Vice Chancellor for Academic Affairs) and appropriate federal Agencies. **Please ensure that all personnel associated with this protocol have access to this approved copy of the AUP and are familiar with its contents.**

Sincerely yours,

Eddie Johnson/jd

Eddie Johnson
Vice-Chair, Animal Care and Use Committee

EJ/jd

Enclosure

www.ecu.edu

Figure A1. Approved Animal use protocol for this study.

GGTACCCTCCTAGGAGATGACCAAATTTATAATGTAATCGTCACTGCCCATGCCTTCGTAATGATCTTCT
TTATAGTAATACCGATCATAATTGGAGGATTCGGTAATTGACTAGTCCCCCTCATAATCGGAGCCCCAGA
TATGGCCTTTCCACGCATAAACAACATAAGCTTCTGACTCCTCCCTCCTTCCTTCCTACTCCTCCTAGCA
TCCTCCACAGTCGAAGCAGGGGCAGGAACAGGATGAACAGTATACCCCCCAATAGCTGGCAACCTAGCCC
ACGCAGGAGCTTCAGTAGACCTAGCCATTTTCTCACTCCACCTGGCAGGTGTATCATCCATCCTGGGCGC
AATCAACTTTATCACGACCGCCATTAACATGAAACCCCCAGCCCTCTCCCAATATCAAACCCCCCTATTC
GTATGATCTGTCCTCATCACCGCCGTCCTCCTATTACTATCCCTTCCCGTCCTAGCCGCA

Figure A2. King Rail (*Rallus elegans*) cytochrome oxidase c subunit 1 consensus sequence using the Rail COI primer set.

TACTTAATCTTTGGGGCCTGAGCCGGAATAATTGGCACCGCCCTAAGCCTGCTTATTCGAGCAGAACTCG
GACAGCCCGGAACCCTCCTAGGAGATGACCAAATTTATAATGTAATCGTCACTGCCCATGCCTTCGTAAT
GATCTTCTTTATAGTAATACCGATCATAATTGGAGGATTCGGTAATTGACTAGTCCCCCTCATAATCGGA
GCCCCAGATATGGCCTTTCCACGCATAAACAACATAAGCTTCTGACTCCTCCCTCCTTCCTTCCTACTCC
TCCTAGCATCCTCCACAGTCGAAGCAGGGGCAGGAACAGGATGAACAGTATACCCCCCACTAGCTGGCAA
CCTAGCCCACGCAGGAGCTTCAGTAGACCTAGCCATTTTCTCACTCCACCTGGCAGGTGTATCATCTATC
CTGGGCGCAATCAACTTCATCACGACCGCCATTAACATGAAACCCCCAGCCCTCTCCCAATATCAAACCC
CCCTATTCGTATGATCTGTCCTCATCACCGCCGTCCTCCTATTACTATCCCTTCCCGTCCTAGCC

Figure A3. Clapper Rail (*Rallus crepitans*) cytochrome oxidase c subunit 1 consensus sequence using the Rail COI primer set.

King Rail (*Rallus elegans*)	----------	0
Clapper Rail (*Rallus crepitans*)	TACTTAATCTTTGGGGCCTGAGCCGGAATAATTGGCACCGCCCTAAGCCTGCTTATTCGAGCAGAACTCGGACAG	75
King Rail	---GGTACCCTCCTAGGAGATGACCAAATTTATAATGTAATCGTCACTGCCCATGCCTTCGTAATGATCTTCTTT	72
Clapper Rail	CCCGGAACCCTCCTAGGAGATGACCAAATTTATAATGTAATCGTCACTGCCCATGCCTTCGTAATGATCTTCTTT	150
King Rail	ATAGTAATACCGATCATAATTGGAGGATTCGGTAATTGACTAGTCCCCCTCATAATCGGAGCCCCAGATATGGCC	147
Clapper Rail	ATAGTAATACCGATCATAATTGGAGGATTCGGTAATTGACTAGTCCCCCTCATAATCGGAGCCCCAGATATGGCC	225
King Rail	TTTCCACGCATAAACAACATAAGCTTCTGACTCCTCCCTCCTTCCTTCCTACTCCTCCTAGCATCCTCCACAGTC	222
Clapper Rail	TTTCCACGCATAAACAACATAAGCTTCTGACTCCTCCCTCCTTCCTTCCTACTCCTCCTAGCATCCTCCACAGTC	300
King Rail	GAAGCAGGGGCAGGAACAGGATGAACAGTATACCCCCCAATAGCTGGCAACCTAGCCCACGCAGGAGCTTCAGTA	297
Clapper Rail	GAAGCAGGGGCAGGAACAGGATGAACAGTATACCCCCCACTAGCTGGCAACCTAGCCCACGCAGGAGCTTCAGTA	375
King Rail	GACCTAGCCATTTTCTCACTCCACCTGGCAGGTGTATCATCCATCCTGGGCGCAATCAACTTTATCACGACCGCC	372
Clapper Rail	GACCTAGCCATTTTCTCACTCCACCTGGCAGGTGTATCATCTATCCTGGGCGCAATCAACTTCATCACGACCGCC	450
King Rail	ATTAACATGAAACCCCCAGCCCTCTCCCAATATCAAACCCCCCTATTCGTATGATCTGTCCTCATCACCGCCGTC	447
Clapper Rail	ATTAACATGAAACCCCCAGCCCTCTCCCAATATCAAACCCCCCTATTCGTATGATCTGTCCTCATCACCGCCGTC	525
King Rail	CTCCTATTACTATCCCTTCCCGTCCTAGCCGCA	480
Clapper Rail	CTCCTATTACTATCCCTTCCCGTCCTAGC---C	555

Table A1. Consensus sequence alignment of King Rail and Clapper Rail COI sequences. These were created using the Rail COI primer set (Forward: ACCTAATCTTTGGGGCCTGA, Reverse: GGGTCGAAGAATGTGGTGTT) on tissue samples from two individuals.

Table A2. Black Rail eDNA samples, environmental variables at time of collection and diagnostic test amplification results.

Sample name	Sample type	paired?	Sample group	Collection date	Location name	State	Detection type	Detection type simplified	Estimated water depth cm	Air temperature °F	Salinity ppt	Sample Storage °C	Days between collection and camera detection	Days between collection and extraction	Total positive replicates LOQ	Percent positive LOQ	Total positive replicates LOD	Percent positive LOD
BLRA 18.001	water	Yes	A	4/17/2018	North River	NC	audio	audio	15	43	17	4°C		226	0	0	2	0.333333333
BLRA 18.001	soil	Yes	A	4/17/2018	North River	NC	audio	audio	15	43		-20°C		285	0	0	0	0
BLRA 18.B04	soil		B	5/6/2018	BIWMA	SC	camera	camera	0	70		-20°C	0	266	6	1	6	1
BLRA 18.003	water	Yes	C	5/7/2018	North River	NC	audio	audio	4	64	17	4°C		589	4	0.666666667	6	1
BLRA 18.003	soil	Yes	C	5/7/2018	North River	NC	audio	audio	4	64		-20°C		265	0	0	3	0.5
BLRA 18.B02 w	water	Yes	D	5/29/2018	BIWMA	SC	audio	audio	1	81	6.42	-20°C		729	0	0	0	0
BLRA 18.B02 s	soil	Yes	D	5/29/2018	BIWMA	SC	audio	audio	1	81		-20°C		509	0	0	0	0
BLRA 18.B01	water		D	5/29/2018	BIWMA	SC	audio	audio	3	80	3.1	-20°C		259	3	0.5	4	0.666666667
BLRA 18.SM1	soil		E	6/1/2018	St. Marks	FL	audio	audio	4	73	0.6	-20°C		506	0	0	3	0.5
BLRA 18.B03	soil		E	6/1/2018	BIWMA	SC	audio	audio	1	80		-20°C		256	0	0	6	1
BLRA 18.SJ	soil		E	6/4/2018	St. Johns	FL	None	None	1	87		-20°C		503	0	0	1	0.166666667
BLRA 18.B05	soil		F	6/6/2018	BIWMA	SC	camera	camera	0	77		-20°C	0	501	0	0	0	0
BLRA 18.004	soil	yes	G	6/7/2018	North River	NC	audio	audio	5	74		-20°C		234	5	0.833333333	5	0.833333333
BLRA 18.004	water	yes	G	6/7/2018	North River	NC	audio	audio	5	74	17	4°C		558	3	0.5	3	0.5
BLRA 18.SVN2	soil		H	6/16/2018	St. Vincent	FL	audio	audio	2	76		-20°C		185	0	0	1	0.166666667
BLRA 18.005	water		H	6/16/2018	North River	NC	audio	audio	4	81	17	-20°C		381	0	0	6	1
BLRA 18.006	water		H	6/16/2018	North River	NC	audio	audio	6	81	17	-20°C		274	0	0	6	1
BLRA 18.SM2	soil		I	6/18/2018	St. Marks	FL	audio	audio	4	79	0.8	-20°C		743	0	0	4	0.666666667
BLRA 18.YW001	soil		I	6/18/2018	BIWMA	SC	audio	audio	0	87		-20°C		489	0	0	1	0.166666667
BLRA 18.SNV3	soil		J	6/19/2018	St. Vincent	FL	audio	audio	0	83		-20°C		742	0	0	0	0
BLRA 18.R03	water	yes	K	7/10/2018	BIWMA	SC	camera	camera	8	90	3.06	-20°C	0	217	6	1	6	1
BLRA 18.R03	soil	yes	K	7/10/2018	BIWMA	SC	camera	camera	8	90		-20°C	0	467	0	0	1	0.166666667
BLRA 18.R10	water	yes	K	7/10/2018	BIWMA	SC	camera	camera	8	86	3	-20°C	7	217	3	0.5	6	1
BLRA 18.R10	soil	yes	K	7/10/2018	BIWMA	SC	camera	camera	8	86		-20°C	7	467	0	0	1	0.166666667
BLRA 18.R820	water	yes	K	7/10/2018	BIWMA	SC	camera	camera	8	89	3	-20°C	3	192	0	0	0	0
BLRA 18.R820	soil	yes	K	7/10/2018	BIWMA	SC	camera	camera	8	89		-20°C	3	467	0	0	0	0
BLRA 18.R32	water	yes	K	7/10/2018	BIWMA	SC	camera	camera	8	92	3.24	-20°C	2	156	0	0	0	0

BLRA 18.R32	soil	yes	K	7/10/2018	BIWMA	SC	camera	camera	8	92		-20°C	2	467	0	0	0	0
BLRA 18.R86	soil	yes	L	7/11/2018	BIWMA	SC	camera	camera	15	77		-20°C	7	249	6	1	6	1
BLRA 18.R86	water	yes	L	7/11/2018	BIWMA	SC	camera	camera	15	77	3	-20°C	7	249	0	0	3	0.5
BLRA 19.RMN	soil		M	4/18/2019	Cedar Island	NC	audio	audio	0	73		-20°C		327	0	0	0	0
BLRA 19.001	soil		N	5/6/2019	Goose Creek	NC	audio	audio	0	63		-20°C		387	0	0	0	0
BLRA 19.003	water	yes	O	5/7/2019	North River	NC	audio	audio	4	82	13.8	-20°C		226	0	0	3	0.5
BLRA 19.002	soil	yes	O	5/7/2019	North River	NC	audio	audio	4	82	13.8	-20°C		226	0	0	1	0.166666667
BLRA 19.004	soil		P	5/28/2019	Swan Quarter	NC	audio	audio	0	78	8.5	-20°C		205	0	0	0	0
BLRA 19.B01 S	water		Q	6/3/2019	BIWMA	SC	camera	camera	3	81	5	-20°C	0	359	0	0	1	0.166666667
BLRA 19.B02	water		Q	6/3/2019	BIWMA	SC	camera	camera	3	81	5.54	-20°C	0	359	0	0	0	0
BLRA 19.B04	water		R	6/6/2019	BIWMA	SC	camera/audio	audio	3	75	9	-20°C	0	136	0	0	1	0.166666667
BLRA 19.B05	water		R	6/6/2019	BIWMA	SC	camera/audio	audio	3	76	9	-20°C	0	356	0	0	0	0
BLRA 19.B03	water		R	6/6/2019	BIWMA	SC	camera/audio	audio	6	74	9.4	-20°C	0	136	0	0	2	0.333333333
BLRA 19.007	water	yes	S	6/13/2019	North River	NC	audio	audio	3	77	20.1	-20°C		129	0	0	0	0
BLRA 19.005	soil	yes	S	6/13/2019	North River	NC	audio	audio	3	77		-20°C		189	0	0	0	0
BLRA 19.008	water	yes	S	6/13/2019	North River	NC	audio	audio	3	77	20	-20°C		236	0	0	0	0
BLRA 19.006	soil	yes	S	6/13/2019	North River	NC	audio	audio	3	77		-20°C		189	0	0	0	0
BLRA 19.012	soil	yes	T	6/15/2019	Cedar Island	NC	audio	audio	1	86		-20°C		269	0	0	1	0.166666667
BLRA 19.010	soil	yes	T	6/15/2019	Cedar Island	NC	audio	audio	2	86		-20°C		187	0	0	0	0
BLRA 19.011	water	yes	T	6/15/2019	Cedar Island	NC	audio	audio	4	86	11.6	-20°C		234	0	0	0	0
BLRA 19.009	water	yes	T	6/15/2019	Cedar Island	NC	audio	audio	4	86	11	4°C		127	0	0	0	0
BLRA 19.013	soil		U	6/21/2019	Goose Creek	NC	audio/visual	audio	0	75	8	-20°C		121	0	0	0	0
BLRA 19.014	soil		V	6/22/2019	Belhaven	NC	audio/visual	audio	0	76	3	-20°C		120	0	0	0	0
BLRA 19.015	water		W	7/10/2019	Belhaven	NC	None	None	12	85	6.3	4°C		102	0	0	0	0
BLRA 19.017	soil	yes	X	7/14/2019	Goose Creek	NC	None	None	4	71		-20°C		240	0	0	0	0
BLRA 19.016	water	yes	X	7/14/2019	Goose Creek	NC	None	None	4	71	8.3	4°C		98	0	0	0	0
BLRA 19.019	soil		X	7/14/2019	Goose Creek	NC	None	None	5	86		-20°C		205	0	0	0	0
BLRA 19.B07	soil		Y	7/16/2019	BIWMA	SC	camera	camera	0	82		-20°C	16	96	0	0	1	0.166666667
BLRA 19.B08	soil		Z	7/17/2019	BIWMA	SC	camera	camera	0	89		-20°C	0	237	0	0	2	0.333333333
BLRA 19.B09	soil		Z	7/17/2019	BIWMA	SC	camera	camera	0	89		-20°C	8	237	0	0	0	0
BLRA 19.B11	soil		AA	7/18/2019	BIWMA	SC	camera	camera	3	85	11.6	-20°C	1	236	0	0	1	0.166666667
BLRA 19.B10	soil		AA	7/18/2019	BIWMA	SC	camera	camera	5	79		-20°C	4	154	0	0	0	0

BLRA 19.B12	soil		AB	7/19/2019	BIWMA	SC	camera	camera	3	80		-20°C	4	235	0	0	2	0.333333333
BLRA 19.B14	soil		AC	7/22/2019	BIWMA	SC	camera	camera	0	87		-20°C	2	232	0	0	0	0
BLRA 19.B13	soil		AC	7/22/2019	BIWMA	SC	camera	camera	0	87		-20°C	0	197	0	0	0	0
BLRA 19.B15	soil		AD	7/23/2019	BIWMA	SC	audio	audio	0	89		-20°C	0	231	0	0	0	0
BLRA 19.020	soil		AE	7/29/2019	Cedar Island	NC	None	None	0	81		-20°C		225	0	0	4	0.666666667
BLRA 19.023	soil		AE	7/29/2019	Cedar Island	NC	None	None	0	85		-20°C		225	0	0	0	0
BLRA 19.021	soil		AE	7/29/2019	Cedar Island	NC	None	None	3	82		-20°C		225	0	0	4	0.666666667
BLRA 19.022	water		AE	7/29/2019	Cedar Island	NC	None	None	3	82	11	-20°C		303	0	0	0	0
BLRA 19.024	water	yes	AE	7/29/2019	Cedar Island	NC	None	None	4	85	11	-20°C		303	0	0	1	0.166666667
BLRA 19.025	soil	yes	AE	7/29/2019	Cedar Island	NC	None	None	4	85		-20°C		225	0	0	0	0
BLRA 19.026	water	yes	AE	7/29/2019	North River	NC	audio	audio	5	80	17	-20°C		303	0	0	0	0
BLRA 19.027	soil	yes	AE	7/29/2019	North River	NC	audio	audio	5	80		-20°C		225	0	0	0	0
BLRA 19.B16	soil		AF	7/30/2019	BIWMA	SC	camera	camera	0	85		-20°C	2	224	0	0	0	0
BLRA 19.B17	soil		AF	7/30/2019	BIWMA	SC	camera	camera	1	85		-20°C	6	82	0	0	1	0.166666667
BLRA 19.B18	soil		AG	7/31/2019	BIWMA	SC	camera	camera	0	82		-20°C	6	223	0	0	0	0

Table A3. Distribution of eBird Black Rail observations and unique locations among states and in relation to the creation of eBird in 2002. The "before 2002" column under "unique locations" refers to locations with observations exclusively before 2002, the "after 2002" column refers to locations with observations exclusively after 2002, the "both" column refers to locations with observations both before and after 2002.

	eBird Observational Records			Unique Locations					
State or Province	Total Observations	Before 2002	After 2002	Total Unique Locations	Before 2002	After 2002	Both	Points with a Single Observation	Points with Multiple Observations
Ontario	4	4	0	3	3	0	0	2	1
Connecticut	15	15	0	10	10	0	0	6	4
District of Columbia	3	3	0	1	1	0	0	0	1
Delaware	44	22	22	11	2	6	3	5	6
Florida	344	47	297	99	19	76	4	59	40
Georgia	6	0	6	3	0	3	0	2	1

Indiana	61	0	61	21	0	21	0	17	4
Massachusetts	74	1	73	10	1	9	0	4	6
Maryland	224	104	120	38	24	8	6	17	21
Michigan	15	2	13	5	1	4	0	1	4
North Carolina	169	61	108	63	26	29	8	39	24
New Hampshire	10	0	10	6	0	6	0	4	2
New Jersey	307	38	269	36	10	21	5	18	18
New York	26	19	7	16	9	7	0	11	5
Ohio	62	3	59	22	3	19	0	14	8
Pennsylvania	24	10	14	16	10	6	0	13	3
Rhode Island	13	0	13	6	0	6	0	4	2
South Carolina	96	20	76	19	4	11	4	11	8
Tennessee	1	1	0	1	1	0	0	1	0
Virginia	62	15	47	18	6	12	0	12	6
West Virginia	37	1	36	30	1	29	0	27	3
Total	1597	366	1231	434	131	273	30	267	167