Assessing the Efficiency of Environmental DNA and Metabarcoding for Conservation of the Barred Grass Snakes (*Natrix helvetica*)

Nasser S. S. Ben Qasem

Supervisor

Professor Rob Freckleton

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Biosciences
Department of Ecology and Evolutionary Biology

Submission Date

25/10/2024

Acknowledgement

I express my deepest gratitude to my supervisors, Professor Rob Freckleton and Professor Terry Burke, for their unwavering support, guidance, and expertise throughout my research journey. Their mentorship has been instrumental in shaping my work, and their encouragement has inspired me to push forward at every step. I am profoundly grateful to the government of Saudi Arabia and King Saud University for their generous support and funding, which made this research possible. I am also grateful to the University of Sheffield for providing an excellent academic environment and resources that enabled me to conduct my research at the highest level.

To my family, I extend my heartfelt thanks. To my mother, brothers, and sisters, thank you for your patience, encouragement, and steadfast belief in me throughout this journey. I am especially grateful to my father, whom I sadly lost during the first year of my studies; his memory has been a source of strength and motivation for me to persevere.

I would also like to thank the NEOF team for their camaraderie, collaboration, and shared passion for research. Working alongside such a dedicated group of individuals has been a privilege, and their insights and support have enriched my experience immensely.

This thesis is a testament to the collective guidance, encouragement, and assistance I have received, and I am sincerely thankful to everyone who has shared this journey with me.

Abstract

The rapidly developing technology of DNA analysis has gained much significance in biology, offering non-invasive methods to study the conservation ecology of elusive, cryptic, and endangered species. In the last decade, environmental DNA survey (eDNA) and metabarcoding methods have been increasingly used to study aquatic, semi-aquatic, marine, and terrestrial organisms. Moreover, most investigations into environmental DNA and metabarcoding have focused on amphibians and fish while limited to reptiles. This thesis explores the efficiency of environmental DNA (eDNA) for detecting the presence of barred grass snakes (*Natrix helvetica*) and metabarcoding for analysing diet habits, highlighting these methods as practical tools for conserving reptile species.

Environmental DNA (eDNA) is genetic material, such as urine, faeces, saliva, slough skin, and gametes of target species, that is released into the environment. This genetic material is then extracted from environmental samples such as water, soil, sediment, and air. We developed and validated a novel species-specific qPCR assay to detect *N. helvetica* from two environmental samples (water from the ponds and roller swabs of entire artificial cover objects (ACO)). The results proved the assay's efficiency, specifically and sensitivity in detecting *N. helvetica* at low concentrations of eDNA. At the same time, water eDNA samples yielded reliable detections compared to roller samples. However, we suggest that using an ACO that is flat in shape with a non-porous surface could enhance eDNA recovery and improve detection rates. To further understand the dietary habits of *N. helvetica*, the metabarcoding approach, in which different universal primers were used to detect multiple species of prey, was performed using faecal samples collected from two distinct populations. Our results found amphibians as the primary dietary component, supplemented by small mammals, birds, and insects, with dietary variations observed across age, sex, and populations. We recommend that future

monitoring and conservation strategies incorporate eDNA and metabarcoding, as these tools offer precise, cost-effective alternatives to traditional surveys for reptiles, contributing to biodiversity protection and habitat management.

CHAPTER ONE

THESIS STRUCTURE

Human activity has greatly endangered species' diversity, and today, many efforts must be made to conserve them. For this purpose, it is important to have effective means of analysing species' distribution.

Traditional ecological monitoring techniques still present some obstacles for investigators due to problems related to classification accuracy, the difficulty of detecting many cryptic species, and the need for detailed surveys, which have consequences for timeliness. As a result, there is an urgent need for alternative and functional techniques for large-scale biodiversity monitoring (Thomsen and Willerslev, 2014).

The rapidly developing technology of DNA analysis has gained much significance in conservation biology, such as eDNA, which is part of a suite of ‘next generation’ approaches to ecological monitoring and acoustic monitoring. However, the concern is that not always scientists can find specimens of the required species for a more thorough investigation. Specifically, environmental DNA (eDNA) is helpful when specimens are unavailable, while environmental analysis can deliver results. This chapter discusses the efficacy of utilising eDNA for studying conservation biology and detecting semi-aquatic snakes.

Environmental DNA is the extraction of genetic material not directly from a specimen but from environmental samples such as water, soil, air, and grass. This approach is useful when a targeted species is not available while the environment contains vivid traces of its inhabitance (Adams et al. 2019).

Pawlowski et al. (2020) argue that although the term eDNA is typically restricted to macro-organisms whose physical bodies have not fully remained in the environment, the method can also be applied to smaller species such as microbes and meiofauna. Lear et al. (2017) emphasise that the methodology of eDNA is most applied to extracting the DNA of species that live in aquatic ecosystems (freshwater habitats and marine habitats), while it applies to terrestrial species. Environmental DNA can be used for biological purposes

such as analysing biodiversity and species' environmental distribution, monitoring and detecting rare and endangered species, population genetic analysis, and ecological relationships and diet interaction (Bang & Corlett 2020; Sigsgaard 2019). There are two different methods to analyse environmental DNA: the barcoding method, in which specific species primer will be utilised to amplify the fragmented DNA of the targeted species using polymerase chain reaction and the metabarcoding method, in which multi-species will be detected by universal primer (Lear et al. 2017 & Herder et al. 2014).

Environmental DNA involves extracting DNA samples from water, soil, and air environment elements instead of extracting these samples directly from the target organisms. 70% of research into environmental DNA concentrates on amphibians and fish taxa of organisms. Additionally, eDNA testing and research on semi-aquatic reptiles account for 6% of all research into environmental DNA studies (Ficetola & Taberlet 2019; Adams et al. 2019). The investigations using eDNA to study snakes are limited due to factors: **(1)** Snakes have impermeable skin surfaces that impede the deposition of DNA material into the environment **(2)** Physical and biological environmental factors for the sample collection sites, such as the time of the day affects the temperatures of the environment and thus the presence or absence of the target snakes in the selected environment **(3)** The methodological limitations such as the efficiency of the qPCR assay, eDNA extraction protocol, the quality and quantity of eDNA, and sample size (Halstead et al. 2017; Adams et al. 2019; Ficetola & Taberlet 2019; Katz et. 2020; Ratsch et al. 2020). On the other hand, amphibians have epithelial cells that are easily left on environmental surfaces and habitats. Additionally, amphibians are excreted through urine and epithelial cells; hence, eDNA is easier to deal with than snakes (Adams et al. 2019).

It is more complex to study eDNA in the terrestrial environment than in aquatic environments. Most research into eDNA concentrates on aquatic and semi-aquatic

environments compared to terrestrial environments. Despite the term, eDNA is also performed using soil samples, and the result obtained can be used to understand surface and underground organisms. A significant fraction of amphibians and reptiles are surface and underground environments, and thus, research also concentrates on these examples of environments. Using soil samples is potentially more difficult due to the potential for DNA degradation in a terrestrial environment than in an aquatic environment. In essence, some factors that influence the validity and efficiency of eDNA research include the target environment. Aquatic and semi-aquatic environments are better equipped to conserve DNA than terrestrial environments (Ficetola & Taberlet 2019).

Few research studies have been conducted to evaluate the validity of environmental DNA testing on semi-aquatic snakes compared to traditional ecological methods. These investigations target rare and endangered species of snakes, including semi-aquatic snakes, such as South Florida's Burmese Python (Piaggio et al. 2014; Hunter et al. 2015; Kucherenko et al. 2018), California's Giant Greater snakes (Halstead et al. 2017), Easter Massasagua that inhabits semi-fossorial crayfish burrows (Baker et al. 2018), California's water snakes (Rose et. al), Kirtland's snake (Ratsch et al 2020), and Louisiana Pinesnakes (Katz et al 2020).

Among the most significant exploitations of eDNA involved South Florida's research determining the presence or absence of Burmese Pythons, *Python bivittatus*, which commonly inhabits coastal habitats, marshy plains, and abandoned or unused semi-aquatic environments. Piaggio et al. (2014) developed a specific species primer (Mitochondrial Cytochrome *b*) to detect the Burmese Pythons using the traditional PCR method. All primers have been labelled with fluorescent dye 6-FAM to visualise the fragments on the genetic analyser. During an eDNA experiment, 15 ml of water samples were taken from different regions of South Florida. The environmental DNA samples were extracted using a sodium acetate solution, filtration and Qiagen's QIAamp DNA kit to precipitate the eDNA in the

selected samples. The results indicated evidence of Burmese Pythons in the environment and that the specific species primer could detect them. Moreover, Hunter et al. (2015) attempted to expand the study of Burmese Pythons using eDNA. In this research, the Quantitative PCR method was utilised as the first time instead of conventional PCR to increase the detection and decrease the negative results in eDNA samples. To create a specific species primer (three primes: forward, reverse primers and fluorescent-labelled primer), the NADH dehydrogenase subunit 4 (ND4) gene was selected using TaqMan qPCR. For the eDNA procedure, 950 ml of water samples were extracted using sodium acetate solution, filtration, and PowerWater DNA Isolation kit. The specific species assay successfully detected the Burmese Pythons in the environment. Unlike the previous investigations, Kucherenko et al. (2018) present the first time using terrestrial environmental samples such as soil to examine the efficacy of using eDNA in the field sites of Florida Gulf Coast for Burmese Pythons as well as to determine how much accumulation and degradation time is essential for terrestrial snakes, Red Cornsnakes, to leave enough eDNA behind. For environmental DNA primer, three mitochondrial regions were chosen (Cytochrome oxidase I (COI) and 12S rRNA for Red Cornsnakes,*Pantherophis guttatus*, and cytochrome b (cyt-*b*) for Burmese Pythons). All primers were amplified using traditional PCR. In the field experiment, it was necessary to acquire soil samples from such field sites as a telemetry-monitored refuge and a telemetry-absent refuge. As a result, python eDNA was 66.7% present in the telemetry-monitored sites. At the same time, no eDNA from Burmese Pythons were detected within the telemetry-absent sites. In a laboratory experiment, three species of Red Cornsnakes were placed in three containers under control conditions to determine the accumulation time. They collected 0.5 g of soil from a sterile plastic tube from each container after 5, 10, 24, and 168 hours of snake occupancy. Snakes were removed to determine the degradation time, and all three containers remained in the laboratory for a week under the

control conditions. Then, they collected soil samples from the containers every 24 hours. The results demonstrate that it is possible to detect the eDNA of Red Corn snakes (*Pantherophis guttatus*) 3.5 hours after the snakes come into contact with soil and approximately 6 days after they leave. Also, the estimates depict that the accumulation rate of Red Cornsnake consisted of 12.6 μg. Therefore, soil usage in terms of eDNA technology has a great potential to become an extremely advantageous detection tool, particularly for detecting terrestrial snakes.

Another research involved testing aquatic samples from marshy and canal habitats to detect *Thamnophis gigas*, also known as the Giant Greater Snake (Halstead et al. 2017). The research involved extracting environmental DNA samples from water tanks in the laboratory under control conditions containing snake faecal matter, snakeskin, and live snakes to evaluate the efficiency of the eDNA survey in comparison with traditional ecological surveys such as the capture and recapture method. In the field study, One litre of water sample from each environment in California was centrifuged and filtered using two different techniques: a filter technique, the most popular method to extract eDNA for snakes and precipitation techniques. The specific species assay augmented the ND4, NADH, ND2, and Cytochrome b (Cyt-*b*) genes region of the mitochondrial DNA by utilising TagMan assay. The results demonstrated that the greater snake environmental DNA was not efficiently detectable in locations with evidence of live *Thamnophis gigas*. Moreover, the approach to eDNA analysis proved ineffective due to the poor efficiency in providing positive results for the presence of the target organisms, and they suggested that the capture and recapture method remains the best survey for snakes instead of the eDNA survey.

Additional research on environmental DNA involved research on an endangered snake species, the Massasauga Rattlesnake, also called *Sistruru scatenatus* (Baker et al .2018). The eDNA process used in the Eastern Massasauga eDNA test involved centrifugation and

filtering as some steps within the integrated process. 50 mL aquatic samples were collected from known crayfish burrows in Illinois and extracted from the target environment, and the resulting sample was centrifuged to separate solid matter from the water sample. The approach uses a TaqMan assay specific to the *Sistrurus catenatus* to target the cytochrome oxidase section of the organism's mitochondrial DNA. The process is potentially ineffective since the extracted sediments may remove DNA materials during the centrifugation process. Thus, the process itself may be a contributing factor to the inaccurate results obtained. The quality and content analysis of environmental samples were also inaccurate due to the variable nature of the samples, and thus the method used to test eDNA in the Easter Massauga environment had flaws and potential areas of weakness. The approach proved ineffective since it required samples to be collected within a 1-meter radius of a given point where the target organism left its DNA. In contrast to the initial methodology, this approach yielded two out of a hundred positive results of *Sistrurus cantenatus* presence with a one-meter radius of influence. The former and traditional approach yielded twelve out of the hundred tested samples for Massasauga rattlesnake. Furthermore, the effectiveness of eDNA testing depends on factors such as the population of the target organisms within the selected ecosystem. For example, environmental samples collected from Eastern Massagua indicated a lack of environmental DNA even when the target organisms sat near the target environment.

Rose et al. (2019) analysed the distribution of introduced semi-aquatic snake species *Nerodia fasciata and Nerodia sipedon* in California. Comparing results with trapping techniques (Plastic Aquatic Funnel Traps). They reveal that both methods illustrate a similar difference between the species' presence. In this study, a qPCR assay has been utilised to target mitochondrial ND2 and COI regions. Then, environmental DNA sampling was compared to aquatic through testing water samples from 30 sites where *N. fasciata* inhabited and 61 sites where *N. sipedon* could be found. On the grounds of environmental DNA

sampling and trapping survey, the results of the study show that there was not any distribution of *Nerodia fasciata and Nerodia sipedon* beyond the localised populations. *N. sipedon* was found at one site among 61 to have shown positive outcomes in the environmental DNA assay. Five water samples were positive for *Nerodia sipedon* eDNA, whereas three eDNA samples were negative for *Nerodia fasciata* eDNA. The results show a lower eDNA efficiency than aquatic trapping while detecting *N. fasciata* and *N. sipedon* snake populations in California. Some challenges have been encountered during calculating detection probabilities, particularly about *N. fasciata*. Consequently, there was some difference in estimating parameters for *N. sipedon*.

Ratsch et al. (2020) developed the specific species environmental DNA assay for Kirtland's Snakes by targeting the mitochondrial COI fragments and determining the degradation time. This study used an artificial coverboard as a traditional ecological survey to compare the results with an environmental DNA survey. For environmental collections, five samples were collected, including water, sediment, and two types of soils from crayfish burrows and under cover-board. The results demonstrated that the artificial coverboard, an example of a traditional ecological survey, provided higher detection than the environmental DNA survey.

The study by Katz et al. (2020) presents how eDNA was used to extract the genetic material of a snake Louisiana Pine snake (*Pituophis ruthveni*), one of the rarest snakes in the United States. As the species has reduced in its distribution, it is on the verge of extinction, and finding suitable habitats for its further preservation can be empowered by eDNA methods. Therefore, two quantitative polymer chain reaction assays have been designed to target COI and APT-6 mitochondrial regions through in silico, in vitro, and in situ validation, which helped reveal how eDNA extraction technique and genetic marker may affect assays. As the research results show, two distinct haplotypes (A and B) of each gene of the Louisiana

Pinesnake have been singled out. Thus, each haplotype varies within a single polymorphic nucleotide position. Under *Silico* assay validation, it has been found that nine and 27 non-target species appeared to be ATP6 and COI matching. Under in *Vitro* assay validation, it has been detected that assays enhance Louisiana Pinesnake DNA. Finally, *Situ* validation has shown that primers amplify Louisiana Pinesnake eDNA based on bedding and soil samples. Moreover, both qPCR (COI and ATP-6) assay were extremely accurate for detecting Louisiana Pinesnake under control conditions and field study.

The essence of eDNA methodology is extracting genetic material from the environment where specimens were abundant before. There are several specific approaches to extracting eDNA from the environment, such as Filtration and Ethanol Precipitation. Each method has its features and applications.

The filtration approach involves the passing of water through a filter of specified pore size to capture eDNA, and it is the most common method for eDNA extraction of macroorganisms or microorganisms, being used in 78% of collection cases (Tsuji et al., 2019). It can use different filters, such as cellulose or glass microfiber. The eDNA samples are then extracted from the filter using different extraction kits. The ethanol extraction approach precipitates nucleic acids from a water sample. It is noted that the selection of the best approach depends on the amount of water in question. The filtration approach helps recover more eDNA samples from the water. When the volume of the water sample is high, the best technique to employ is the filtration method. On the other hand, when the amount of water sample is limited, it is useful to employ the ethanol precipitation approach (Hinlo et al., 2017).

The ethanol precipitation approach captures a wide range of eDNA sizes. On the contrary, the filtration approach relies on the pore size and is hence limited to the size of eDNA it can capture. The ethanol precipitation and filtration approaches are largely affected

by filter materials and the extraction method. Hence, the success and efficacy is recorded differently in different studies. The efficiency of the filtration approach also depends on the turbidity of the water sample to be filtered. The ethanol precipitation approach is highly suitable for small volumes of water, about 15mL, making it very useful in samples with a high abundance of eDNA. As such, the extraction approach may affect the overall eDNA capture, though inhibitors and pollutants may also affect it (Spens et al., 2017).

Then, the eDNA material has to be amplified. Polymerase Chain Reaction (PCR) is the most common method of amplification. PCR is used with primers to reveal the presence of eDNA. With PCR, DNA material can be detected via gel electrophoresis or digital PCR (ddPCR). Alternatively, a quantitative PCR (qPCR) method can be applied, as it rarely provides false positive outcomes (U. S. Geological Survey, 2013). Sequences from databases like GenBank can help detect the necessary genetic material of targeted snakes. PCR or qPCR can help detect the presence of the target material in samples. To create a primer assay, the mitochondrial DNA material provides better results than nuclear DNA due to its larger availability.

Apart from extracting and analysing eDNA material, it is also important to interpret the data properly and accurately. Some recommendations may assist researchers in extracting eDNA material. According to the U. S. Geological Survey (2013), it is important to conduct a test with a physically available specimen so that the environmental assessment results are accurate. Probes for the eDNA analysis are taken from the environment and can be treated in different ways to be preserved, for example, via freezing. Katz et al. (2020) explain the data analysis procedures in eDNA analysis using qPCR. They point out that the first probes can be analysed statistically by SYBR™ Green and TaqMan®. Apart from statistical analysis, there are approaches to barcoding, as primers are used to mark specific areas of DNA for further

analysis. The qPCR analysis helps to detect the availability of the DNA material of a target species (Pawlowski et al. 2020).

Amphibians are also studied with the help of eDNA, as these species can leave traces in soil and water. Takahara et al. (2020) stress the need to obtain environmental information to conserve endangered species. Researchers studied three frog species, endemic to Japan, such as *Babina subaspera, Odorrana splendida, and O. amamiensis,* comparing results with acoustic detection. They underscore the suitability of eDNA in the case of these species, as an acoustic method can detect only male calling species, while eDNA offers greater flexibility. Ficetola et al. (2008) also prove the efficiency of eDNA in detecting traces of amphibians. The researchers targeted the cases of *Rana catesbeiana* and *Lithobates catesbeianus* and confirm that eDNA is a potential method for detecting traces of these species in different environments. eDNA is a prospective method for studying amphibians, as it can trace species both in soil and in water.

Grass snakes in Europe were classified as *Natrix natrix*, with different subspecies based on morphology (Kabisch, 1999). However, molecular studies have reclassified western populations, including those in the UK, as *Natrix helvetica* (Kindler et al., 2017). This species is genetically distinct from *Natrix natrix* and *Natrix astreptophora*, with five recognized subspecies: *N. h. helvetica* (Western Europe and the UK), *N. h. cetti* (Sardinia), *N. h. corsa* (Corsica), *N. h. lanzai* (Italy), and *N. h. sicula* (Sicily and parts of Calabria) (Kindler & Fritz, 2018). Understanding this taxonomic distinction is crucial for conservation efforts, as *Natrix helvetica* represents a unique evolutionary lineage with potential ecological and dietary adaptations. Recognizing these distinctions allows for more accurate assessments of its habitat use, prey selection, and population dynamics, which are essential for informing conservation strategies and ensuring effective long-term monitoring of the species.

The structure of the thesis:

1. **Chapter two**

Title: Capture Mark Recapture Survey of Barred Grass Snakes

This chapter assessed the population size of *Natrix helvetica* and evaluated the effectiveness of the Capture-Mark-Recapture (CMR) method in estimating the population size. The CMR study was conducted at Holbrook Marsh Ponds, a restored wetland designed primarily to support breeding populations of the Great Crested Newts (*Triturus cristatus*), a species of conservation concern in the UK. Restoration efforts at Holbrook included enhancing water quality, establishing suitable vegetation, and creating a balanced ecosystem structure to support diverse wildlife species. While newts were the primary focus of this restoration, monitoring other species, such as the semi-aquatic barred grass snake (*Natrix helvetica*), is essential to assess the project's broader ecological success. The main objective of this chapter is to confirm the presence of barred grass snakes at Holbrook, particularly given the absence of direct evidence in prior surveys. This chapter seeks to provide insights into the population dynamics of *N. helvetica* and underline the value of CMR in monitoring cryptic species within restored wetland habitats.

2. **Chapter three**

Title: The Efficiency of Using Environmental DNA To Detect Barred Grass Snakes

I explore using the environmental DNA (eDNA) approach to detect the presence of *Natrix helvetica* at Holbrook Marsh Ponds. Following the confirmation of barred grass snake populations in this area, this chapter aims to evaluate the effectiveness of eDNA surveys in accurately identifying this species in wetland environments. The eDNA approach presents a non-invasive alternative to traditional survey methods, offering potential advantages in ease, efficiency, and minimal disturbance to the habitat. The objectives of this chapter are: (1) to develop a novel quantitative PCR (qPCR) assay that is specifically for the detection of barred grass snakes, a validation process using a comprehensive protocol; and (2) to compare the efficacy of two different types of eDNA samples in detecting *N. helvetica*. By analysing which

eDNA samples are more efficient for detection, this study aims to establish protocols for future eDNA surveys, potentially providing a robust tool for conservationists to monitor this species more accurately. This chapter's findings will inform eDNA methodology and contribute to broader efforts to monitor and conserve the barred grass snake and other reptile species within its natural habitat.

3. **Chapter four**

Title: Dietary Assessment of Barred Grass Snakes Using Metabarcoding Approach

This chapter evaluates the diet composition of Natrix helvetica through a metabarcoding approach. Understanding the dietary habits of barred grass snakes is crucial for conservation, particularly regarding prey availability, habitat use, and ecological interactions within their environment. Metabarcoding, a DNA-based method, identifies prey items from the faecal samples, offering an in-depth view of the species' dietary diversity. The primary objectives of this chapter are to apply metabarcoding techniques to identify prey taxa consumed by *N. helvetica* and to assess any dietary variation across different age classes, sexes, or populations. This approach enables the detection of a wide range of prey items, including amphibians, fish, invertebrates, and small mammals, offering insights into the snakes' ecological roles and feeding behaviours. By mapping these dietary patterns, this study aims to provide a comprehensive picture of barred grass snake feeding ecology and contribute valuable information for habitat management and conservation strategies.

4. **Chapter Five**

General Discussion.

References

Baker, S., Niemiller, M., Stites, A., Ash, K., Davis, M., Dreslik, M. & Phillips, C., (2018). Evaluation of environmental DNA to detect Sistrurus catenatus and Ophidiomyces ophiodiicola in crayfish burrows. *Conservation Genetics Resources*, *12*(1):13-15. https://doi.org/10.1007/s12686-018-1053-9

Beng, K. & Corlett, R., (2020). Applications of environmental DNA (eDNA) in ecology and conservation: Opportunities, challenges and prospects. *Biodiversity and Conservation*, *29*(7): 2089-2121. https://doi.10.1007/s10531-020-01980-0

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

Indahsari, L., Fatchiyah, F. & Kurniawan, N., (2019). DNA barcoding as a method for species identification: Case study in Ahaetulla snake. *Journal of Physics: Conference Series*, *1374* (2019) 012011 https://doi:10.1088/1742-6596/1374/1/012011

Kabisch, K. (1999). *Natrix natrix* (Linnaeus, 1758) – Ringel natter, pp. 513 – 580 in: Böhme, W. (ed.) Handbuch der Reptilien and Amphibien Europas. Band 3/IIA: Schlangen II. Wiebelsheim, Aula-Verlag.

Katz, A., Harper, L., Sternhagen, E., Pearce, S., Melder, C., Sperry, J. & Davis, M., (2020). Environmental DNA is effective in detecting the federally threatened Louisiana Pinesnake (Pituophis ruthveni). *Environmental DNA*, *00*:1–17. https://doi.org/10.1002/edn3.126

Kelly, R., Shelton, A. & Gallego, R., (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Scientific Reports*, *9*, (12133) https://doi.org/10.1038/s41598-019-48546-x

Kindler, C., Chèvre, M., Ursenbacher, S., & Fritz, U. (2017). Hybridization patterns in two contact zones of grass snakes reveal a new Central European snake species. Scientific Reports, 7, 7378. https://doi.org/10.1038/s41598-017-07847-9

Kindler, C., Graciá, E., & Fritz, U. (2018). Extra-Mediterranean glacial refuges in barred and common grass snakes (Natrix helvetica, N. natrix). Scientific Reports, 8, 1821. https://doi.org/10.1038/s41598-018-20218-2

Kindler, C., & Fritz, U. (2018). Phylogeography and taxonomy of the barred grass snake (*Natrix helvetica*), with a discussion. Vertebrate Zoology, 68(3), 269–281. https://doi.org/10.3897/vz.68.e31615

Kucherenko, A., Herman, J., III, E. & Urakawa, H., (2018). Terrestrial snake environmental DNA accumulation and degradation dynamics and its environmental application. *Herpetologica*, *74*(1):38-49. https://doi.org/10.1655/Herpetologica-D-16-00088

Pawlowski, J., Apothéloz‐Perret‐Gentil, L. & Altermatt, F. (2020). Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. *Molecular Ecology*, *29*(22): 4258-4264. https://doi.org/10.1111/mec.15643

Rose, J., Wademan, C., Weir, S., Wood, J. & Todd, B., (2019). Traditional trapping methods outperform eDNA sampling for introduced semi-aquatic snakes. *PLOS ONE*, *14*(7), p.e0219244. https://doi.org/10.1371/journal.pone.0219244

Takahara, T., Iwai, N., Yasumiba, K. & Igawa, T., (2020). Comparison of the detection of 3 endangered frog species by eDNA and acoustic surveys across 3 seasons. Fres*hwater Science, 39*(1): 18-27. https://doi.org/10.1086/707365

Udyawer, V., Barnes, P., Bonnet, X., Brischoux, F., Crowe-Riddell, J., D'Anastasi, B., Fry, B., Gillett, A., Goiran, C., Guinea, M., Heatwole, H., Heupel, M., Hourston, M., Kangas, M., Kendrick, A., Koefoed, I., Lillywhite, H., Lobo, A., Lukoschek, V., McAuley, R., Nitschke, C., Rasmussen, A., Sanders, K., Sheehy, C., Shine, R., Somaweera, R., Sweet, S. & Voris, H. (2018). Future directions in the research and management of marine snakes. *Frontiers in Marine Science*, *5*. 399.

https://doi.org/10.3389/fmars.2018.00399

U. S. Geological Survey. (2013). *Application of environmental DNA for inventory and monitoring of aquatic species*. Fact Sheet 2012–3146, January 2013. https://www.arlis.org/docs/vol1/F/835572905.pdf

Zhang, S., Zhao, J. & Yao, M., (2020). A comprehensive and comparative evaluation of primers for metabarcoding eDNA from fish. *Methods in Ecology and Evolution*, *11*(12): 1609-1625. https://doi.org/10.1111/2041-210X.13485

CHAPTER TWO

CAPTURE MARK RECAPTURE SURVEY OF BARRED GRASS SNAKES

Table of Contents

1. Introduction

Conservation of threatened and endangered species relies on understanding their population structure (Sutherland, 2006). This understanding is essential for planning conservation efforts, including creating and managing protected areas, which are important refuges for species facing habitat destruction and climate change (Pukazhenthi et al., 2005; Runge, 2011; Mi et al., 2023). Moreover, reptile and amphibian species in the vertebrate lineage have declined and are currently threatened because of the destruction of habitats and the effects of climate change (Ficetola et al., 2018). According to WWT (Wildfowl & Wetlands Trust) and Freshwater Habitats Trust (2019), around 50% of ponds disappeared in the UK over the 20th century, with a significant impact on biodiversity and ecosystem health, and among the remaining ponds, around 80% were in poor condition (Pond Conservation: The Water Habitats Trust, 2010). Therefore, conducting surveys that preserve species' habitats and support conservation processes is paramount.

Natrix helvetica is a non-venomous snake widespread in Europe and a native species to the United Kingdom (Kindler et al., 2017). The barred grass snake (*Natrix helvetica*), is classified as a distinct species, differentiating it from the populations in central and eastern Europe, which are classified as *Natrix natrix.* It occurs in a range of habitats, from woodlands to grasslands and from marshes to the margins of running water related to ponds, lakes, and rivers. The lack of systematic monitoring, combined with the high mobility of the species, makes it difficult to evaluate its conservation status accurately. Although they are considered widespread species across England, substantial declines are suspected, especially where egg-laying and foraging sites have been lost (Edgar et al., 2010). Under the Wildlife and Countryside Act 1981, the barred grass snake has special protection in the UK, whereby it is an offence to kill, injure, or sell them. In the UK, most studies about the population of the barred grass snake have been conducted using conventional methods such as visual surveys or using artificial cover objects in the last decades, focusing on the southeast (Sewell et al., 2012)

and east (Sewell et al., 2015) of England and Jersey (Ward et al., 2017). The barred grass snake have been conducted using conventional methods such as visual surveys or using artificial cover objects in the last decades, focusing on the southeast (Sewell et al., 2012) and east (Sewell et al., 2015) of England and Jersey (Ward et al., 2017).

For monitoring and sampling reptiles, the capture-mark-recapture method (CMR) using artificial cover objects (ACO) is a common conventional method to estimate the population size and dynamics of reptiles (Wiedemer et al., 2007; Ward et al., 2017). This technique is associated with ACOs to capture a small group of the target organism, mark them harmlessly, and release them back into the population at the site. Later, another small group is captured, and the number of marked animals is counted. A statistical model is then applied to determine different ecological parameters, such as population size, detection probability, and survival rate (Wilson et al., 2007; Sewell et al., 2012).

This chapter will apply the capture-mark-recapture method to measure the size of the population of the barred grass snake *Natrix helvetica*. The objectives of the study were to (1) perform a capture mark-recapture survey using artificial cover objects and (2) provide evidence for the occurrence of *Natrix helvetica* and the population structure at the site. (3) estimate the population size using an appropriate closed population model. By confirming the occurrence of *Natrix helvetica*, we could apply an environmental survey to test the efficiency of this approach.

2. Material and Methods

2.1. Field Site

The study was undertaken at Holbrook Heath & Marsh Ponds in Sheffield, South Yorkshire, United Kingdom (53° 19 '46 "N, 1° 19' 53 "W). The size of the field site is 0.81 hectares. The site has a diverse topography, combining various wetland and terrestrial elements,

creating a rich habitat that is important for biodiversity and offers niches for aquatic and terrestrial species. The site includes a network of ponds, marshes, and heathland interspersed with patches of reeds and grasses. The natural and artificially created ponds form a central feature of the landscape, providing essential aquatic environments (Figure 2.1)**.**

2.2. Artificial Cover Objects (ACO)

Artificial refuges effectively attract reptiles because they absorb heat and offer additional protection from predators. Consequently, they can be successfully used in reptile surveys (Liberman et al., 2024; Steve Langham-SARG., 2011). The ACOs used in this project were corrugated roofing materials composed of bitumen-soaked organic fibres. They have a thickness of 2.6 mm and dimensions of 500 mm x 500 mm, featuring a corrugation depth of 40 mm. The ACOs were ordered from NHBS (Natural History Book Service) via https://www.nhbs.com.

During October 20th - 25th, 2021, 66 ACOs were placed in different parts of the field site. The ACOs were located in suitable microhabitats of the barred grass snakes, such as (1) close to the ponds, (2) areas with good vegetation cover, and (3) at natural obstacles such as logs and rocks. (4) sunny locations (balance was required to ensure the cover boards do not overheat). Each ACO was labelled with a unique identification number, the number ID, the location, and contact information. The distance between the ACOs of each selected microhabitat was less than ~3 M. In addition, 20 ACOs were also placed during the 12th - 15th of March 2023 in different parts of the site, focusing on the habitat edges to increase the chance of capturing barred grass snakes. The location of ACOs was the same during both surveys in 2022 and 2023.

2.3. The Time of the Survey

Reptiles typically warm up by basking in sunshine, near dense vegetation, or under heat-conducting objects, often for several hours daily, making this the ideal time for surveys. The activity season to survey the barred grass snakes could be classified into two phases: (1) Mid-April, May, and September are the peak months for conducting surveys. Mid-April and May are the snake mating season, in which snakes are more active, making them easier to spot and less cautious around observers. Furthermore, September presents an excellent opportunity to search for hatching snakes near their breeding grounds. (Christopher McLnerny, 2015; Froglife, 1999) (2) June, July, and August are active months for the barred grass snake. Although they remain active, the air temperatures may exceed ~ 22-25 °C, so they require less time to bask and are harder to observe (Griffiths & Inns., 2003).

Three to four visits were conducted every week, depending on the weather conditions (temperatures between 11 and 20 °C, intermittent or hazy sunshine, and little or no wind). During the first phase of the activity season, the survey started in the morning from 8:00 a.m. to 10:00 a.m. and was late from 4:00 p.m. to 7:00 p.m., consistently throughout the survey. The average weather conditions for the morning and evening surveys were a temperature of 17.28 °C, a humidity of 60.03%, and a wind speed of 14.55 km/h. During the second phase, the morning survey was conducted from 6:00 a.m. to 8:00 a.m. and from 7:00 p.m. to 9 p.m. in the evening. The average weather conditions were a temperature of 19.9 °C, a humidity of 59.1%, and a wind speed of 18.6 km/h.

2.4. Capture Mark-Recapture Survey Protocol

On arrival at the site, the survey start time was recorded along with weather conditions data, including air temperature (^{O}C), humidity (%), and wind speed (km/h), using an iPhone Apple device (weather application). Before the ACO was checked, the surface temperature and time of each ACO were recorded using a digital laser thermometer (Cunsieun). The ACO was

lifted, and any snake found underneath was captured and placed into a holding bag (20 x 30 cm, weighing 23 g) made from cotton non-mesh material. The holding bag was then loosely tied to prevent the snakes from escaping while allowing air circulation to prevent them from suffocating. In the case of capturing several snakes found underneath an ACO, the calmest snake was processed first, while the others remained in the holding bag. After processing the snake, it was immediately placed underneath the same ACO.

2.4.1. Identification Using AVS Scales

The unique patterns and variations in the anterior ventral scales (AVS) of barred grass snakes stay the same from hatchling to adulthood. This means they can be used to identify each individual using the CMR method. By photographing the first sixteen scales of the AVS on each snake, unique patterns can be documented and used for identification. These dark markings on the scales remain unchanged throughout the life cycle of the snake, ensuring that once an individual has been identified, it can be reliably recognised in future encounters (Baker & Allain, 2020).

2.4.2. Sexual Identification

Sexual identification of barred grass snakes (*Natrix helvetica*) can be performed using a combination of physical characteristics to identify the sex accurately. (1) Subcaudal scales were used to identify males and females. If the subcaudal scales were between 52 and 56, it was considered female; between 68 and 72, it would be male (Griffiths & Inns., 2003). (2) The tail length and shape of the head were used as a second physical key. Males have longer tails relative to their body size and elongated heads. In contrast, females have shorter tails and more triangular heads (Robert Vaughan, 2007 Griffiths & Inns., 2003). (3) The presence of hemipenes in the male by applying pressure near the base of the tail. The bugle of the hemipenes indicated that the snake was male. The probing technique was avoided due to several potential problems and risks, especially if not performed correctly by an experienced herpetologist (Beebee & Griffiths, 2000).

The attempts to determine the sex of juveniles and sub-adults were performed, but it was sometimes difficult due to sexual maturity.

After the captured snake was photographed and the sex was identified, the measurements were applied, such as the snout to vent length (SVL), tail length (TL), and total length (TTL) using TEFCOL Retractable Tape Measure 5 m. Different investigations attempted to categorise the age class. For instance, males were considered by Madson (1983) as adults whose SVLs exceeded 43 cm and females as adults whose SVLs exceeded 55 cm. On the other hand, Merten (1995) proposed that adult females have an SVL greater than 60 cm, while adult males have an SVL greater than 40 cm. The age class of the snake in this study was determined by the snout-vent length (SVL). If the SVL was less than 35 cm, the snake was considered a juvenile, while if the SVL was greater than 45 cm, regardless of gender, the snake was considered an adult. The snake was considered a subadult if the SVL was between 35 and 45 cm.

2.5. Data Analysis

The *MARK* program was used to perform a mark-recapture analysis to calculate population size over two surveys. I utilised the full likelihood models of Otis et al., which offer multiple models to estimate population size and capture probabilities. In capture-mark-recapture studies, several models are used to estimate population parameters. The simplest is the null model (M0), which assumes no variation in capture probability among individuals or over time, meaning that every individual in the population has the same chance of being captured in each sampling occasion. In contrast, the time-specific model (Mt) allows for variability in capture probabilities across different time periods, accounting for factors like changes in weather, seasons, or other temporal conditions that could influence the likelihood of capture. The behavioural response model (Mb) focuses on individual behaviour, accounting for changes in capture probability due to being captured previously. The heterogeneity model (Mh) considers inherent differences in capture probability among individuals, in which some animals may be easier or harder to catch due to traits like size and age. The time variation and behavioural responses model (Mtb)

combines both time variation and behavioural responses, allowing for the possibility that capture probability can vary across time and change depending on whether an animal has been previously captured. This is the most complex model, capturing both temporal and behavioural factors that may affect capture probability. These models provide a structured framework for understanding and accounting for the complexities of animal capture data.

The population of barred grass snakes at the survey site was assumed to be closed based on the following assumptions: (1) No births, fatalities, immigration, or emigration occurred during the sampling survey. Consequently, any snakes captured with an SVL of less than 20 cm would be excluded. (2) Each individual had an equal and independent capture probability on each trapping occasion. Even though that assumption is not appropriate for reptiles in general, models such as M_b, M_h, M_t, and M_{tb} can help to adjust that problem. (3) The natural marker of the barred grass snakes would not be lost during the survey period (AVS scales).

To evaluate the assumption of population closure during a study period, the Close-Test program (Stantey & Burham, 1999) was conducted. It was designed to assess the assumption of population closure in Capture-Mark-Recapture (CMR) studies. It analyses capture-recapture data, which records whether individuals are captured or not during several sampling occasions. It performs two statistical tests. The first test examines if there is a significant variation in capture probabilities between sampling occasions, which would suggest population changes like immigration or emigration, indicating the population is not closed. The second test checks for heterogeneity in individual capture probabilities, which can signal an open population if certain individuals have a higher or lower chance of being captured.

The capture history file was generated for each survey session to document the capture status of each individual across all capture occasions. Each row represented a unique individual snake (as a BGS standing for Barred Grass Snake), and columns indicated if it was captured (1) or not (0) at each trapping occasion.

I structured the file with the recommendation of Program MARK version 10.0 (White & Burnham, 1999) The best model of Otis et al. was selected by evaluating goodness-of-fit tests and the AIC to ensure it best summarised the capture-recapture data and provided the best population size estimate. To further validate and compare parameter estimates, model outputs from *ShinyRecap* (https://fellstat.github.io/shinyrecap/) (McIntyre et al., 2021) were cross-checked with results obtained from MARK software. This ensured consistency in parameter estimation and allowed for a more comprehensive evaluation of population size and dynamics.

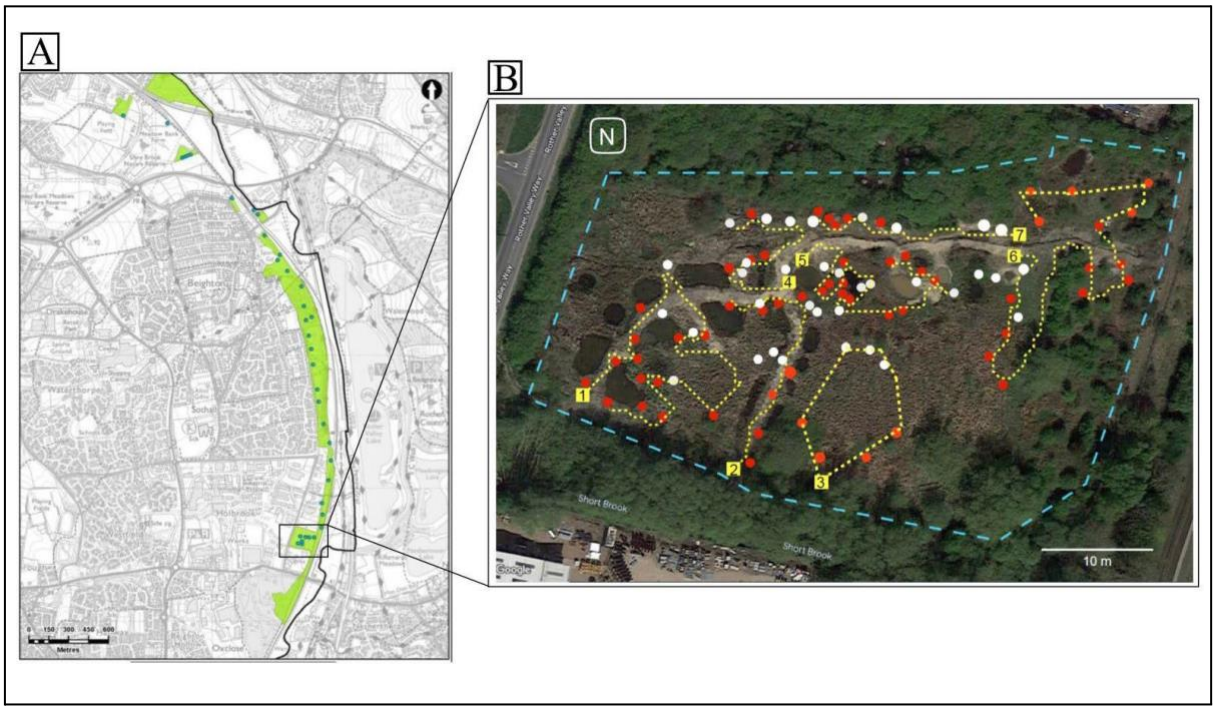

Figure 2.1. Panel A provides a map showing a broader view of Holbrook Heath & Marsh (highlighted with a black square). (B) The zoom-in of the target site: The survey route is indicated by a dashed yellow line, representing the path taken to ensure no ACOs were missed during the survey. The start point of the survey route is marked by a square yellow marker, while red circles indicate the locations of ACOs, and white circles represent the locations where grass snakes were detected under the ACOs. The area is bordered by a blue-dash line outlining the boundaries of the survey site. (Sheffield City Council).

3. Results

3.1. Capture Mark-Recapture Survey 2022 Season

Over 60 surveys, 26 barred grass snakes were captured on 12 occasions, with 15 unique individuals in which were recaptured multiple times (Figures 1.2 and 1.3-A). The age class of the 15 individuals was as follows: 4 were juveniles with an average SVL of 27.75 cm, six sub-adults with an average SVL of 38.83 cm, and five adults with an average of 50.75 cm (Figure 1.4). The class-age ratio of mature to immature was 1:2. The sex of only seven individuals were determined (four females and three males) with a male-to-female sex ratio of 0.75, and eight individuals have not reached sexual maturity. I did not observe any neonates during the survey.

The *Close-Test* program classified the population of barred grass snakes as closed, supporting the assumption that no births, fatalities, immigration, or emigration occurred during the sampling survey (X^2 = 10.51, df = 10, p-value = 0.396). Using the *MARK* program, the M0, which assumed that the capture probability is constant across all individuals and capture occasions, was selected as the best model by Otis et al. to estimate the population size (N) = 19.13 ± 3.45 with a 95% confidence interval of 15.99 to 32.19 (Figure 1.5), and the estimated capture probability (P) = 0.113 ± 0.029 with a 95% confidence interval that ranges from 0.067 to 0.184. Moreover, the *ShinyRecap* R package supported that M0 was the best model to estimate the population size (N = 19.13 ± 4 with a 95% confidence interval of 15 to 30) based on the AIC = 96.30 and BIC = 97.71 among all Otis et al. models (Table 2.1).

3.2. Capture Mark-Recapture Survey 2023 Season

During the CMR survey in 2023 (more than 45 surveys), 26 barred grass snakes were captured on nine occasions, with 19 unique individuals, of which four were also captured during the 2022 survey (Figures 2.2 and 2.3-B). The average SVL of the six adults was 54.83 cm. For six sub-adults, the average was 38.82 cm and 30.28 cm for seven juveniles. The class-age ratio of mature to immature was 1:2.17 (Figure 2.4). The sex ratio of males to females among the captured snakes was 0.22 (two males and seven females), however, ten individuals remained undetermined. Three neonates were observed at the field site, so I did not include them in the CMR data analysis.

The result of the *close test* enhanced the assumption that during the sampling period, there were no births, deaths, immigration, or emigration (X^2 = 8.94, df = 7, p-value = 0.256). The M0 model was an appropriate estimator for the population size (N = 34 ± 9.393, with a 95% confidence interval of 23 to 68) (Figure 2.5), and the estimated capture probability (p) was 0.085 ± 0.028 with a 95% confidence interval of 0.042 to 0.160. In *ShinyRecap*, the M0 was selected as the appropriate model based on AIC (72.67) and BIS (74.56). The population size matched the M0 in the Mark program (Table 2.1).

The unique patterns and variations in the first 16 anterior ventral scales (AVS) of all recaptured individuals were consistent during the survey, even though one recaptured female was in the sloughing cycle, which may cause misidentification of the AVS (Figure 2.6).

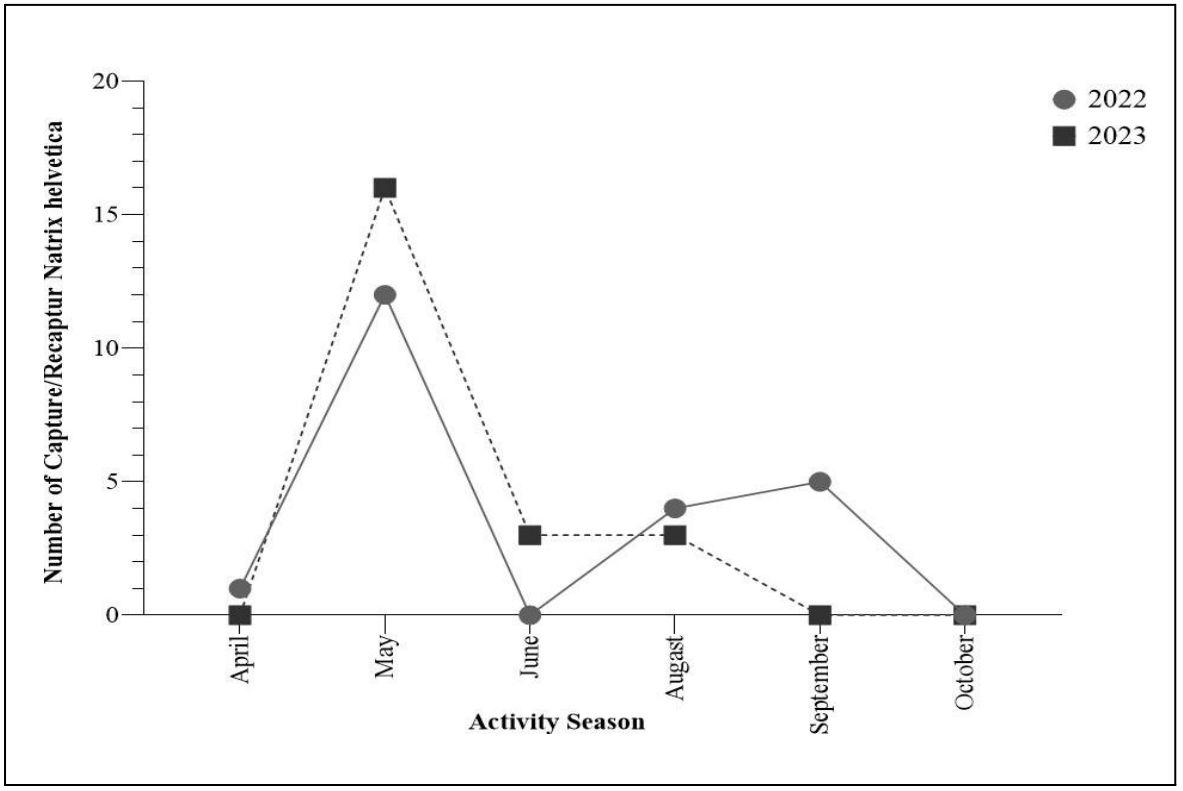

Figure 2.2. The number of captures and recaptures of *Natrix helvetica* during the activity season for 2022 and 2023. The x-axis represents the months of the activity season, and the y-axis indicates the number of captures and recaptures. Circles and solid lines represent the data points for 2022, while squares and dashed lines represent the data points for 2023.

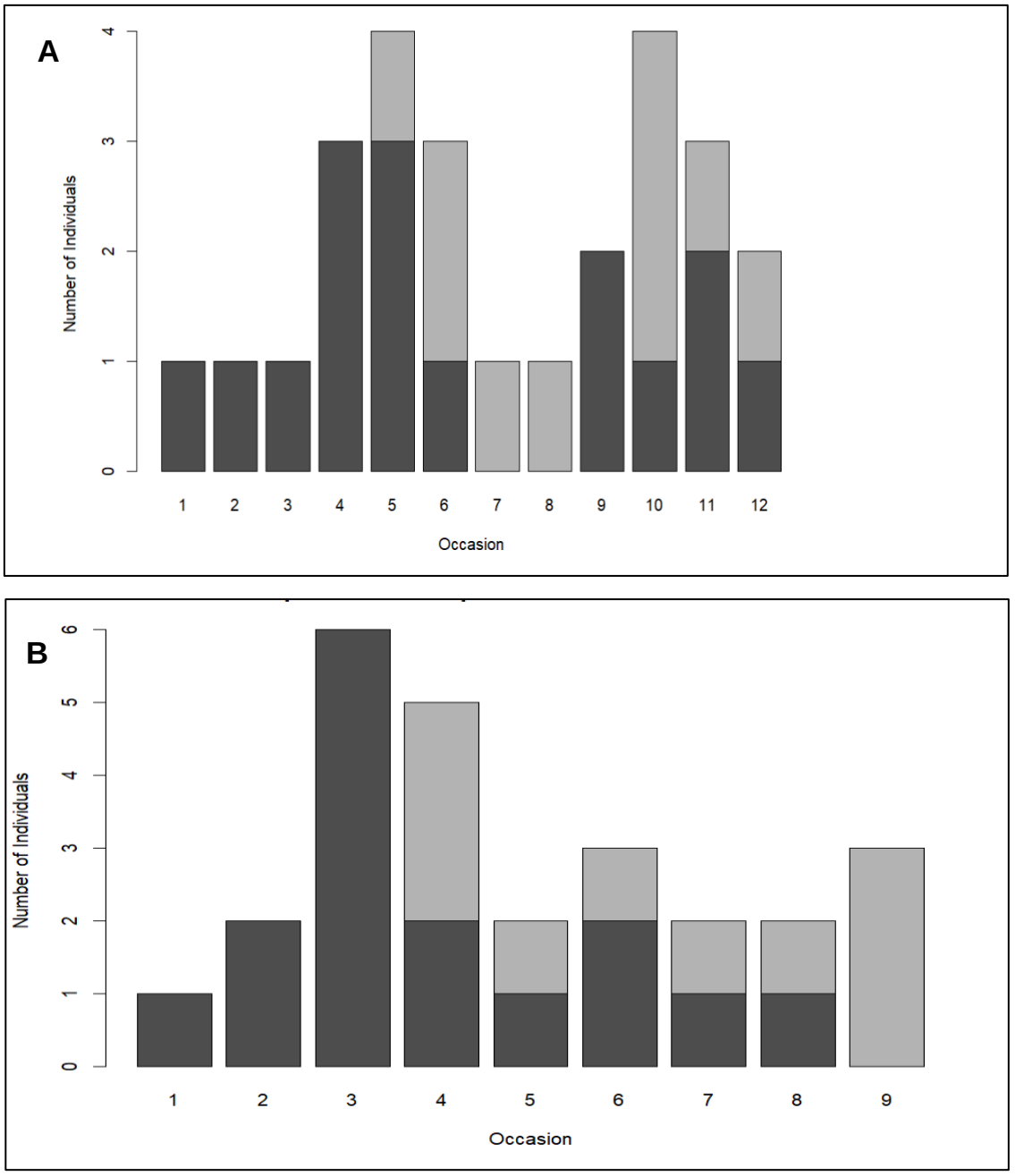

Figure 2.3. The number of individuals captured and recaptured across occasions during each survey in 2022 (A) and 2023 (B). Each bar is split into two segments: the darker grey represents captures, and the lighter grey indicates recaptures. The x-axis labels each occasion, while the y-axis is the number of individuals.

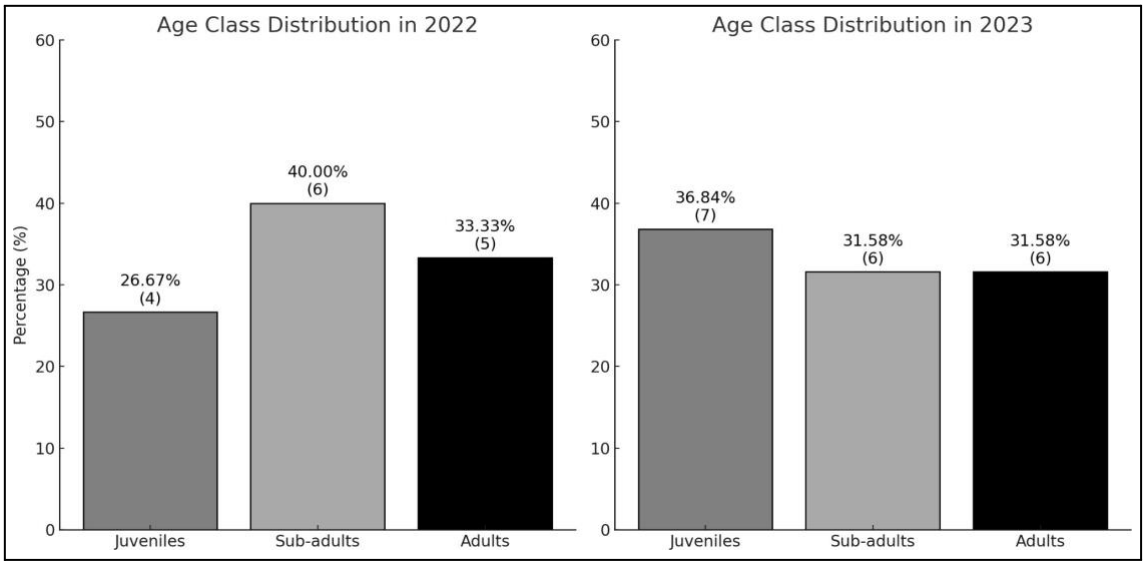

Figure 2.4. The figure consists of two bar charts comparing the age-class distribution of *Natrix helvetica* between 2022 and 2023, displayed in grayscale and black. Each bar is labelled with the percentage and the number of unique individuals captured.

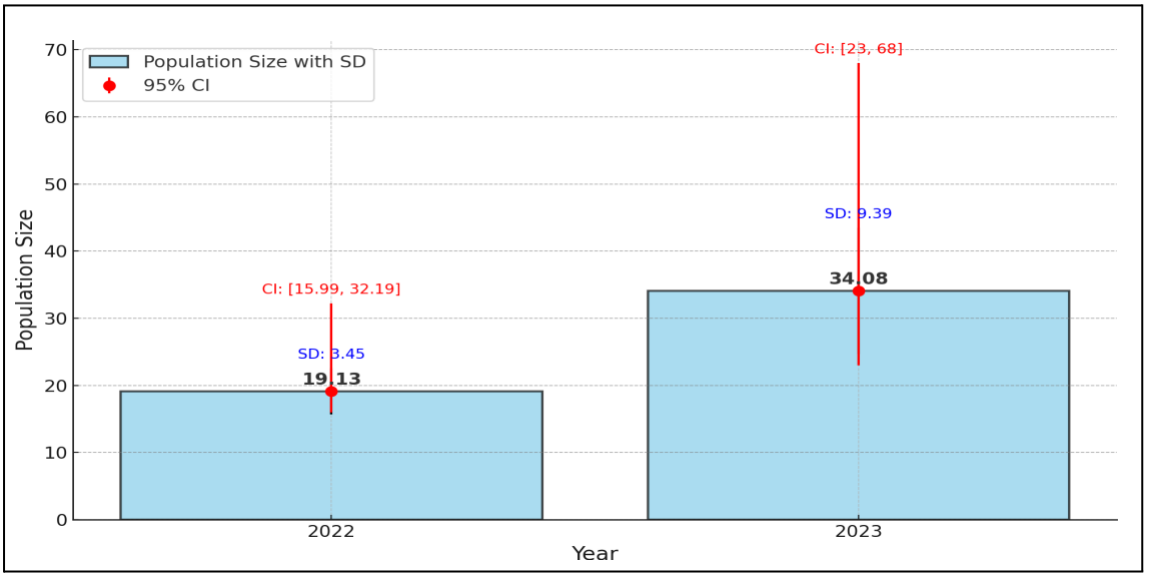

Figure 2.5. The estimated population size of the Barred Grass Snake for the 2022 and 2023 seasons, along with the associated standard deviation (SD) and 95% confidence intervals (CI) based on the M0 model (Null model).

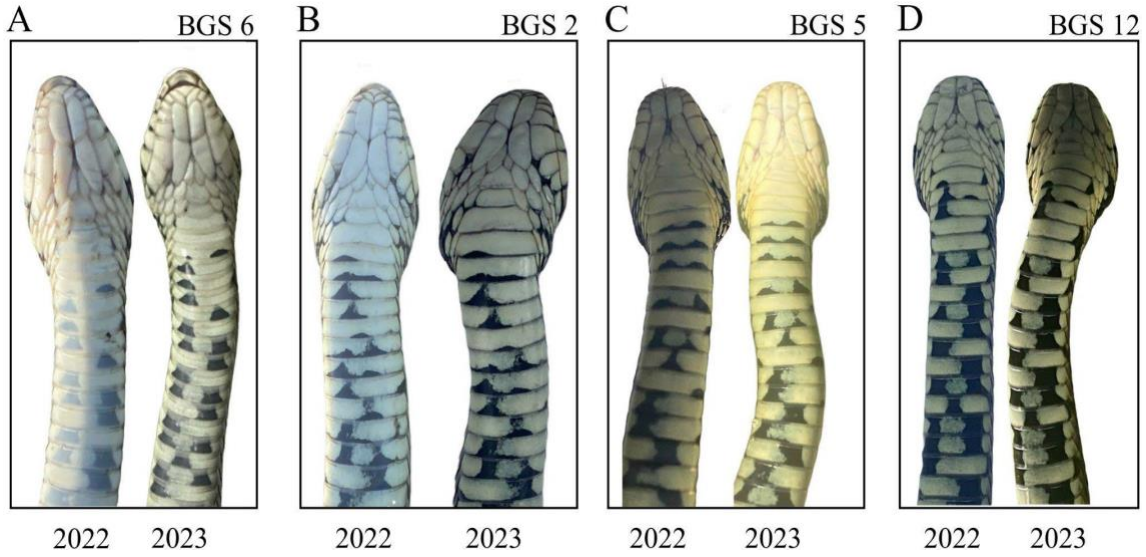

Figure 2.6. The displaying of the anterior ventral scales (AVS) of four barred grass snakes (labelled BGS 6, BGS 2, BGS 5, and BGS 12) that were re-captured in both the 2022 and 2023 surveys. Each panel (A, B, C, and D) provides a comparison of the AVS patterns from the 2022 survey (left) with the corresponding patterns observed in the 2023 survey (right). BGS 6, shown in panel A, was in a sloughing cycle during one of the surveys, which may cause slight alterations in the appearance of its scales. Despite this, the unique patterns and variations in the AVS of all the snakes, including BGS 6, demonstrate consistency between the two years.

Table 2.1. The tables present the model selection results (MARK) for the 2022 and 2023 season surveys, comparing different capture-recapture models based on various statistical criteria. For each season, several models, including the basic M0 model, the heterogeneity model (Mh), the behavioral response model (Mb), the time variation model (Mt), and a combined model (Mtb), are evaluated. The Akaike Information Criterion corrected (AICc) values are used to determine the best-fitting model, with lower AICc values indicating a better fit. The Delta AICc, AIC weight, and model likelihood are also provided to assess the relative strength of each model compared to the others.

2022 Season Survey

Model	AIC_c	Delta AIC_c	AIC Weight	Model Likelihood	No. of Parameters	Deviance	-2log(L)	Model criteria Value
M_0	**93.5631**	**0.0000**	**0.51521**	**1.0000**	**2**	**66.8249**	**89.4953**	**1.00**
M_h	93.5631	0.0000	0.51521	1.0000	2	66.8249	89.4953	0.81
M_b	93.6859	0.1228	0.48453	0.9405	3	64.8792	87.5495	0.36
M_t	109.4522	15.8891	0.00018	0.0003	13	58.5891	81.2595	0.00
M_{tb}	111.1275	17.5644	0.00008	0.0002	14	57.9117	80.5820	0.43

2023 Season Survey

Model	AICc	Delta AICc	AIC Weight	Model Likelihood	No. of Parameters	Deviance	-2log(L)	Model criteria Value
M_0	**60.639**	**0.00**	**0.36544**	**1.0000**	**2**	**44.860**	**56.5677**	**1.00**
M_h	60.639	0.00	0.36544	1.0000	2	44.860	56.5677	0.79
M_b	62.689	2.05	0.13115	0.3589	3	44.837	56.5449	0.26
M_t	70.547	9.91	0.00258	0.0071	10	37.464	49.1715	0.00
M_{tb}	72.809	12.17	0.0023	0.0023	11	37.440	49.1482	0.28

4. Discussion

The capture-mark-recapture (CMR) survey conducted at Holbrook Marsh and Heath to estimate the population size of barred grass snakes *(Natrix helvetica)* provided valuable insights into the presence of this species at a restoration site primarily aimed at benefiting Great crested newt growth. Artificial cover objects (ACOs) combined with the CMR method proved effective in detecting and monitoring *Natrix helvetica*. The ACOs provided essential microhabitats that attracted snakes, facilitating their capture and recapture. The study successfully recorded 15 unique individuals in 2022 and 19 in 2023, demonstrating the utility of this approach in a relatively small and diverse habitat.

The population estimates derived from the CMR data using the M0 (null model) indicated a population size of 19.13 in 2022 and 34.08 in 2023. The population increase observed between 2022 and 2023 may be due to (1) favorable habitat conditions from restoration site efforts, such as topography around ponds, including fallen logs and dense vegetation, which created microhabitats that offer both sunny basking spots, shaded refuges, and hibernation sites. (2) The average SVL of the barred grass snakes in 2022 and 2023 were 39.11 cm and 41.31 cm, respectively, indicating that the population is increasing and showing signs of healthy growth and development. Also, the presence of neonates in 2023 suggests successful reproduction. (3) The number of ACOs increased in the 2023 survey to cover the open patches within vegetation and at the edge of the site, unlike in the 2022 survey, in which I placed most of the ACOs around the ponds. Although the surveys were conducted over a relatively short time, and the objective was to determine the population size using a CMR survey before applying the eDNA survey (see Chapter 3), the findings provided a first impression regarding the population structure and the stable habitats.

The daily capture probabilities in 2022 and 2023 were 0.113 and 0.085, respectively. The low probability value was expected for the reptile, in general, and barred grass snakes due to the abundance of captured individuals at the target site. Moreover, 66% of the captured snakes in 2022 and 73% in 2023 were juvenile and sub-adults, with an average SVL (combination of both age classes) of 31.5 and 35 cm, respectively, indicating that the ecology and behaviour of the adult, sub-adult, and juveniles may affect the probability of detection. Compared to another investigation by Sewell et al. (2015), the study found a detection probability of 0.17. It used the captured mark-recapture data to estimate population sizes ranging from 53 to 576 snakes over multiple years, which was higher than the detection probability of this study. However, they surveyed the barred grass snakes as a long-term survey for nine years. Moreover, Kéry (2002) studied the grass snakes in France and Switzerland for five years, and the detection probability ranged from 0.11 in small populations to 0.25 in medium and large populations; however, rather than using individual detection, he relied on the presence or absence of species.

. The high number of juveniles and sub-adults captured in both surveys highlights the effectiveness of the survey effort, the number of visits, and the use of ACOs during the study because the size of the snakes can influence detection probabilities, with smaller or juvenile snakes being less likely to be detected (Lock & Griffiths 2022). 86 ACOs were placed at the 0.81-hectare target site, averaging 106 ACOs per hectare. This contrasts with Sewell et al. (2015), who used 22 ACOs on an 11-hectare site, averaging 2 ACOs per hectare.

The ratio of mature to immature in this project (1:2 and 1:2.17 during the 2022 and 2023 surveys) was higher than the study of Mertens et al. (1995), which was 1.6:1 and 1.1:1 during the 1990 and 1991 surveys. That suggests a relatively young population with more immature individuals, whereas Mertens et al. (1995) found a higher proportion of mature grass snakes, *Natrix natrix*.

The proportion of males and females sex in 2023 (0.22) was lower than in 2022 (0.75), with the most recaptured adults in both surveys were female during May, which suggests that the mature females may have been gravid. This likely affected their behaviour and mobility, making it easier for them to recapture them (Madsen 1987). The abundance of artificial nesting sites at Holbrook Heath and Marsh may attract females for egg-laying and provide the ideal thermal and humidity conditions crucial for egg incubation.

The anterior ventral scales (AVS) of barred grass snakes were consistent during the survey in 2022 and 2023 suggesting that the AVS is a key identification among individuals. For example, four individuals (two sub-adults and two adult females) were captured in 2022 and recaptured again in 2023, resulting in unique AVS between seasons (Figure 1.5). Although the sloughing cycle was observed, the dark black mark of the AVS did not affect the identity of the individuals, especially juveniles and sub-adults. Baker & Allain (2020) documented the stability of AVS patterns in barred grass snakes. They confirmed that AVS patterns remain unchanged from hatchling to adulthood, allowing for reliable individual identification across years. This finding supports the reliability of using AVS patterns for long-term monitoring and aligns with the observation that AVS patterns were consistent across the two survey years.

It is essential to compare the relative effectiveness, advantages, disadvantages, and potential biases of various sampling techniques to accurately assess herpetofauna populations globally. ACO surveys (Artificial Cover Objects) are widely used in herpetofauna studies, yet their effectiveness can vary significantly depending on the species and environmental conditions (Libeman et al.,2024 ;Kolanek & Bury, 2021).

Although the CMR survey using ACO provided evidence of the occurrence of the barred grass snake, several limitations should be acknowledged : (1) The surveys were conducted over a relatively short period of time. The barred grass snake populations can

exhibit significant temporal variations due to seasonal changes, annual weather patterns, and habitat alterations. A long-term study would provide more comprehensive data, capturing year-to-year variations and offering a better understanding of population dynamics and habitat use over time. (2) The study captured more juveniles and sub-adults than adults, which might introduce a bias in the population size and structure analysis. Juveniles and sub-adults may have different habitat use and behaviour than adults, potentially affecting detection probabilities and population estimates. (3) An abundance of vegetation can further limit the effectiveness of ACOs, as it may obscure the covers or make them less attractive to snakes seeking open, sunny basking spots. Additionally, dense vegetation can make it physically difficult for researchers to place and check ACOs, reducing the overall efficiency of the survey method. By addressing the identifying limitations and expanding research efforts, we can further enhance the conservation of *Natrix helvetica* and other species dependent on restored wetland habitats, such as Holbrook Heath and Marsh Ponds in Sheffield, South Yorkshire.

5. References

Amphibian and Reptile Conservation Trust. (2021). *Establishing a survey: National Reptile Survey Protocol (Version 2021.2)*. Amphibian and Reptile Conservation Trust.

Baker, J. M. R., & Allain, S. J. R. (2020). Photographic identification of individual Barred Grass Snakes, *Natrix helvetica*. *Transactions of the Norfolk and Norwich Naturalists' Society*, *53*(1), 67–70.

Beebee, T. J. C., & Griffiths, R. A. (2000). *Amphibians and reptiles: A natural history of the British herpetofauna*. HarperCollins.

Edgar, P., Foster, J., & Baker, J. (2010). *Reptile Habitat Management Handbook*. Amphibian and reptile conservation.

Griffiths, R. A., & Inns, H. (2003). Chapter 1: Surveying. In T. Gent & S. Gibson (Eds.), *Herpetofauna Workers' Manual* (pp. 1-11). Joint Nature Conservation Committee.

Hunter, A. (2016). *Living Waters Project (part 2). Partnership HLF grant with Froglife.* Sheffield City Council, Culture & Environment Directorate.

Kindler, C., & Fritz, U. (2018). Phylogeography and taxonomy of the barred grass snake (Natrix helvetica), with a discussion of the subspecies category in zoology. Vertebrate Zoology, 68(3), 269-281 https://doi.org/10.3897/vz.68.e31615

Kolanek, A., & Bury, S. (2021). The detection of elusive reptiles under artificial cover objects is species- and year-specific. *Polish Journal of Ecology, 68*(4), 342-347. https://doi.org/10.3161/15052249PJE2020.68.4.007

Kéry, M. (2002). Inferring the absence of a species: A case study of snakes. *The Journal of Wildlife Management, 66*(2), 330-338. https://www.jstor.org/stable/3803165

Liberman, J., Klemens, M. W., & Pinder, A. (2024). Artificial cover objects as a tool for the survey and management of herpetofauna. *Journal of Herpetology*, *38*(2), 215-223.

Lock, M. M. G., & Griffiths, R. A. (2022). Detectability of reptiles in standardised surveys: A test using grass snake *Natrix helvetica* models. *The Herpetological Journal, 32*(4), 183-189. https://doi.org/10.33256/32.4.183189

Madsen, T. (1987). Cost of reproduction and female life-history tactics in a population of grass snakes, *Natrix natrix*, in southern Sweden. *Oikos*, *49*(2), 129-132.

Marco, A., Blaustein, A. R., & Pauli, B. D. (2008). The effects of pesticides on amphibian and reptile populations. *Ecotoxicology, 17*(4), 657-673. https://doi.org/10.1007/s10646-008-0219-6

McIntyre, A. F., Fellows, I. E., Gutreuter, S., & Hladik, W. (2021). Population size estimation from capture-recapture studies using shinyrecap: Design and implementation of a web-based graphical user interface. JMIR Public Health and Surveillance, 8(4), e32645. https://doi.org/10.2196/32645

Merten, V. (1995). Population structure and abundance of grass snakes. *Herpetological Journal, 5*(3), 105-113.

Mi, C., Ma, L., Yang, M., Li, X., Meiri, S., Roll, U., Feldman, A., Novosolov, M., & Du, W. (2023). Global protected areas as refuges for amphibians and reptiles under climate change. *Nature Communications, 14*, 1389. https://doi.org/10.1038/s41467-023-36987-y

Pond Conservation: The Water Habitats Trust. (2010). *Pond Conservation Annual Report 2009/2010.*

Pukazhenthi, B., Comizzoli, P., Travis, A. J., & Wildt, D. E. (2005). Applications of emerging technologies to the study and conservation of threatened and endangered species. *Reproduction, Fertility and Development*, *18*(2), 77-90.

Ribeiro-Júnior, M. A., Gardner, T. A., & Ávila-Pires, T. C. S. (2008). Evaluating the effectiveness of herpetofaunal sampling techniques across a gradient of habitat change in a tropical forest landscape. *Journal of Herpetology, 42*(4), 733-749. https://doi.org/10.1670/07-097.1

Runge, M. C. (2011). An introduction to adaptive management for threatened and endangered species. Journal of Fish and Wildlife Management, 2(2), 220-233.

Sacchi, R., et al. (2010). Photographic identification in reptiles: A matter of scales. *Amphibia-Reptilia, 31*(4), 489-502.

Sewell, D., Baker, J. M. R., & Griffiths, R. A. (2015). Population dynamics of grass snakes (*Natrix natrix*) at a site restored for amphibian reintroduction. *Herpetological Journal, 25* (July), 155–161.

Sewell, D., Guillera-Arroita, G., Griffiths, R. A., & Beebee, T. J. (2012). When is a species declining? Optimizing survey effort to detect population changes in reptiles. *PLoS ONE, 7*(8), e43615. https://doi.org/10.1371/journal.pone.0043615

Sutherland, W. J. (Ed.). (2006). Ecological Census Techniques: A Handbook (2nd ed.). Cambridge University Press

Vaughan, R. (2007). *Grass Snakes*. ISBN 978-1-4452-7804-9.

Ward, R. J., Griffiths, R. A., Wilkinson, J. W., & Cornish, N. (2017). Optimising monitoring efforts for secretive snakes: a comparison of occupancy and N-mixture models for assessment of population status. *Scientific Reports, 7*, 1-12.

Wiedemer, R. L., Wilson, D. J., Mulvey, R. L., & Clark, R. D. (2007). Sampling skinks and geckos in artificial cover objects in a dry mixed grassland-shrubland with mammalian predator control. *New Zealand Journal of Ecology, 31*(2), 169-185

Wilson, D. J., Mulvey, R. L., & Clark, R. D. (2007). Sampling skinks and geckos in artificial cover objects in a dry mixed grassland-shrubland with mammalian predator control. *New Zealand Journal of Ecology, 31*(2), 169-185. Retrieved from New Zealand Ecological Society (NZES).

CHAPTER THREE

THE EFFICIENCY OF USING ENVIRONMENTAL DNA TO DETECT BARRED GRASS SNAKES

A Table of Contents

1. Introduction

The decline in biodiversity has become a global ecological issue due to the ongoing impact of climate change and human activities. In recent decades, habitat destruction has greatly increased, decreasing the variety and abundance of species in freshwater, marine, and terrestrial ecosystems and limiting the natural environments available to organisms (Living Planet Report, 2020). As a result, the populations of vertebrates have declined across the world, with 40.7% of amphibians, 25.4% of mammals, and 13.6% of birds considered threatened (Bowyer et al., 2019; Harfoor et al., 2021; Cox et al., 2022; Gumbs et al., 2024).

Among tetrapods, reptiles have a higher proportion of species classified as threatened than birds and mammals; 21.1% of reptile species (1,829 out of 10,196 species) worldwide are assessed as threatened with extinction by the IUCN Red List of Threatened Species (Bohm et al., 2013; IUCN, 2020; Cordier et al., 2021; Cox et al., 2022). Climate change and urbanisation are among the biggest threats to reptile populations (Farooq et al., 2024). This approach leads to habitat destruction and fragmentation, severely harming species that rely on specialised ecosystems for survival. In general, monitoring and surveys of reptile species are necessary to offer important information on threatened species' population dynamics and habitat stability. Moreover, traditional surveys are essential for demonstrating how biodiversity changes and evaluating the success of conservation strategies and management actions. However, it is challenging to monitor many taxa with conventional methods owing to the ecology of the organisms.

While valuable, conventional ecological monitoring surveys still need to overcome significant obstacles that often lead to incomplete data. For example: (1) identification accuracy requires taxonomic experience to distinguish the target organisms from closely

related ones, and (2) many reptile species are cryptic and secretive, so detectability will be challenging because they are hidden during regular survey periods and require detailed surveys, which can be time-consuming (Adams et al., 2019; Ratsch et al., 2020; Pawlowski et al., 2020). The conventional surveys for ecological conservation monitoring, such as the Capture-Mark-Recapture (CMR) method, are sensitive to the detection and identification procedures, leading to inconsistent data for herpetological management. This underscores the urgent need for alternative and functional techniques for large-scale biodiversity monitoring and determining their occurrence (Thomsen & Willerslev, 2014; Kyle et al., 2022).

Rapidly advancing DNA analysis technology in conservation biology has generated considerable interest and application in identifying and monitoring biodiversity, such as metabarcoding and barcoding (Fediajevaite et al., 2021; Allen et al., 2023). However, the major shortcoming of standard DNA analysis is the need to acquire samples directly from the target organism, which can be challenging for endangered species. Environmental DNA survey methods present a solution to overcome this issue. They introduce a more efficient and effective way to detect cryptic, rare species, potentially revolutionising biodiversity conservation efforts (Adams et al., 2019).

Environmental DNA (eDNA) is genetic material, such as urine, faeces, saliva, sloughed skin, and gametes of target species, that is released into the environment. This genetic material is extracted from environmental samples such as water, soil, sediment, and even air. Methods for analysing eDNA include (1) real-time quantitative polymerase chain reaction (qPCR) involving a species-specific primer/probe assay of the target species and (2) metabarcoding, in which multiple species are detected using a universal primer set (Adams et al. 2019; Ficetola & Taberlet, 2019; Nordstrom et al., 2023).

In the last decade, environmental DNA has been increasingly used to detect aquatic, semi-aquatic, marine, and terrestrial organisms. Seventy per cent of research into environmental DNA has been focused on amphibians and fish (Adams et al., 2019). Environmental DNA testing and research on reptiles accounts for 6% of all environmental DNA studies (Ficetola & Taberlet, 2019; Adams et al., 2019). Investigations using eDNA to study reptiles have been limited due to factors such as: (1) Reptiles have impermeable skin surfaces that may impede the deposition of DNA material into the environment. (2) Methodological limitations, such as the efficiency and specificity of the qPCR assay to detect the target species, eDNA extraction protocol, and collecting eDNA samples. (3) Physical and biological environmental factors affecting sample collection, such as temperature, UV, and the biology of the target species. By contrast, amphibians and fish epithelial cells are easily deposited on environmental surfaces and in aquatic habitats. If eDNA is to become a general tool for monitoring biodiversity, these limitations need to be overcome.

Few studies have been conducted to evaluate the possibility of environmental DNA detection of semi-aquatic, terrestrial snakes, and fossorial snakes. These investigations have targeted rare and endangered species of snake using a qPCR species-specific assay approach with various results, including semi-aquatic snakes, such as South Florida's Burmese Python (Piaggio et al., 2014; Hunter et al., 2015; Kucherenko et al., 2018), California’s Giant Greater snakes (Halstead et al., 2017), the Eastern Massasauga that inhabits semi-fossorial crayfish burrows (Baker et al., 2018), California’s water snakes (Rose et al 2020.), Kirtland’s snake (Ratsch et al., 2020), Louisiana Pine-Snakes (Katz et al., 2020), and the sharp-tailed snake (Matthias et al., 2021).

The barred grass snake, *Natrix helvetica,* is found in the United Kingdom and Western Europe. It is the largest snake in the United Kingdom. The United Kingdom's Wildlife and Countryside Act has protected the barred grass snake since 1981. During spring and summer,

the snakes reside around water bodies to feed on readily available young fish and newts (The Woodland Trust, 2021). This association with water may offer opportunities for the deposition and detection of eDNA.

Because the barred grass snake was only recently classified in 2017 as a separate species from the grass snake, *Natrix natrix,* which is not present in Britain, limited data exist to support its ecological monitoring for conservation purposes (Foster et al., 2021). To my knowledge, no investigation has used an environmental DNA survey to detect *Natrix helvetica* in the United Kingdom or Europe.

This study will assess the efficiency of using an eDNA survey to detect the barred grass snake using a species-specific qPCR assay as a model for reptiles in the UK. Moreover, I wish to test its potential to fill the gap in eDNA knowledge for surveying reptiles, particularly *Natrix helvetica* and the closely related species that are decreasing in population, such as the Dice Snake, *Natrix tessellata* and Viperine Snake, *Natrix maura,* and other snakes with similar ecologies such as the smooth snake, *Coronella austriaca*, an endangered species in the UK. as well as Sand Lizard, *Lacerta agilis*, which is categorised as a threatened species in the UK.

My study objectives were to (1) develop a novel real-time quantitative polymerase chain reaction (qPCR) assay specific to the barred grass snake, (2) collect eDNA samples such as water samples from the ponds (aquatic samples) using the filtration method and swab the artificial cover objective using a roller (terrestrial samples), (3) validate the developed assays *in silico* to evaluate primer and probe specificity *in vitro* using sloughed skin of barred grass snakes and available tissue of non-target species, and *in situ* on environmental samples.

2. Materials and methods

2.1 Field-site study

The study was conducted in Holbrook Heath & Marsh in Sheffield, South Yorkshire, United Kingdom (53° 19 '46 "N, 1° 19' 53 "W). The site's total size is 2.48 hectares, with 1.5 hectares chosen as the project's focus. In response to the complete drying up of all ponds in the summer of 2015, Natural England artificially created around 30-45 ponds in this refuge area to create a habitat suitable for the Great Crested Newt, *Triturus cristatus*. The site had reeds and grassland surrounded by woodland, mainly birch trees and willows. The topography of the site provides a habitat for Barred Grass Snakes (*Natrix helvetica*), which require aquatic environments for food (including amphibians), hibernation (leaf heaps and log piles), and thermoregulation. (Figure 3.1). (See Chapter 2 for more details on the CMR survey)

2.2 Period of eDNA Sample Collection

Water samples were gathered during barred grass snake activity periods in three phases. The initial phase involved collecting eDNA samples from 25 ponds between 29th September and 15th October 2021, during which barred grass snakes were not visually observed in the field. During the second phase, which occurred between June 3rd-10th and July 11th-12th, 2022, a total of five eDNA water samples were collected, one from each of the five ponds, after conducting a Capture Mark Recapture survey using Artificial Cover Objects (ACOs) between April and June 2022 to confirm the presence of *Natrix helvetica* at the main site. The five ponds selected in the second phase were close to the ACOs under which barred grass snakes were found. In the final phase, eDNA samples were collected from five ponds where barred grass snakes were observed swimming less than 72 hours before sampling in June and July 2023.

Figure 3.1. Habitat Mapping and Sample Collection at Holbrook Marsh Ponds. The field site map illustrates multiple sampling areas, including the locations of the ponds where the water sample was collected (indicated in blue) and roller swab mats (indicated by purple circles).

To collect terrestrial samples, we swabbed the entirety of the under surface of each of 13 Artificial Cover Objectives (ACO), made from corrugated roofing material 50 cm X 50 cm, where grass snakes had been observed less than two weeks before sampling on 10 August 2022 (5 ACOs) and 17 June 2023 (8 ACOs). Additionally, these ACOs were selected based on previous sightings of more than one individual barred grass snake, and at different times.

2.3 Environmental DNA Collection (In Situ) and Extraction

2.3.1 Capturing eDNA and Filtration Procedure of Water Samples

On arrival at the site, all equipment was washed and dried using 10% bleach to prevent cross-contamination. The collection kit for each pond includes the following items: a pair of sterile gloves, a sterile 30 mL ladle, a sterile Whirl-Pak plastic bag with a capacity of 1 L, a Sterivex HV 0.45-µm filter, and a BD Plastipak 50 mL syringe. The water samples for eDNA

analysis were collected using a filtration approach following Troth et al. (2020), with some modifications. A 30 mL blue ladle was utilised to collect 20 sub-samples of water from different points around each pond (a total of 600 mL) to conduct the sampling. During this process, a ladle was used to mix the pond water column gently, ensuring thorough mixing from the surface to close to the pond bottom without disturbing the sediment, which may reduce the efficiency of eDNA analyses. I collected eDNA water at the pond's edge to reduce cross-contamination between ponds and avoided entering the water.

Moreover, 20 sub-samples were located around the pond margin to target areas where barred grass snakes display and find prey using vegetation such as newts. Subsequently, each of the sub-samples was emptied into a Whirl-Pak bag. Before filtration, the Whirl-Pak was gently shaken for 10-15 seconds to mix the eDNA throughout the water sample. Then, using a syringe (BD Plastipak), I carefully pulled 60 ml of water from the Whirl-Pak bag. Then, I attached the Sterivex HV 0.45 μM filter unit to the syringe and started filtering the water. This step was repeated 10 times until all the water samples (600ml) were successfully passed through the Sterivex filter unit. The filter samples from each pond were labelled with the date they were collected and assigned a unique identification number. These samples were taken to the lab at the end of the day's sampling and then stored at a temperature of -20°C in preparation for the next step, which involves extracting environmental DNA (eDNA). I also collected negative control (600ml of sterile shop-bought water) filtered at each site, and at the same time, the pond water samples were taken to monitor possible cross-contamination.

2.3.2 Isolation and Quantification of eDNA Water Samples

eDNA extractions were carried out for each collection phase, including the field negative and extraction negative controls. I used a modified Qiagen DNeasy Blood and Tissue Kit protocol and a QIA shredder with some adjustments to extract the eDNA samples (Troth et al., 2020). The eDNA extraction was performed in a separate laboratory (clean room) dedicated to this

procedure. Before starting the eDNA extraction process, the filter hood was thoroughly cleaned with 10% of bleach to remove potential contaminants. Once cleaned, the hood and work area were sterilised with UV light wavelength For 45 minutes to ensure that the work area was free from any contaminating DNA. All necessary equipment for opening the filter units and extracting the environmental DNA (eDNA) was prepared and ready for use, such as a tweezer, scalpel, pipe cutter, and plastic plate. All equipment was placed in the hood during UV sterilisation.

In the first stage, I cracked open the filter unit with a pipe cutter and carefully removed the filter paper from the plastic unit. After that, I sliced the filter paper into small pieces on a sterile petri dish using a sterile razor blade and placed the small filter pieces into 2 mL Eppendorf tubes (one Eppendorf per filter paper). To prevent DNA contamination, I cleaned the tweezers, scalpel, and pipe cutter with 50% bleach between samples and used a new sterile petri dish for each sample. Next, I set the heat block to 56°C and added 450 µl of ATL (provided buffer) and 50 µL of proteinase K to each 2 mL Eppendorf tube. For some samples, I used tweezers to ensure that all filter paper pieces were fully immersed in the buffer. I thoroughly vortexed the 2 mL Eppendorf tubes to ensure the filter paper remained submerged in the ATL and proteinase solutions. Finally, I incubated the samples at 56°C overnight for 14 -16 hours.

In the second stage, each 2 mL Eppendorf tube containing the filter sample and lysis solution was vortexed for 15 seconds, and I pipetted as much buffer as possible into a new 2 mL Eppendorf tube. Then, the remaining filter paper was placed into a Qiashredder spin column, and I gave them a quick spin down (5K for 2 seconds) and pipetted the buffer mixture into a 2 mL Eppendorf tube with the rest of the lysis solution. After that, the Qiashredder spin column was placed in a centrifuge and spun down for 2 minutes at 11,000 rpm. Again, the liquid passing through the column into the collection tube was combined

with the rest of the solution within the 2 mL Eppendorf tube. Then, I disposed of the Qiashredder and used the empty microfuge tubes for lysis. After that, a 500 µL AL buffer (which helps bind the DNA to the silica membrane during purification) was added to the 2 ml microfuge tubes containing the lysis solution, instantly vortexed, and incubated at 56°C for 10 minutes. During that time, sufficient buffer AE was placed in a Falcon 50ml centrifuge tube and in the oven at 56°C to warm it to use in the final step. Next, 500 µL of absolute ethanol was added to samples and vortexed immediately. For each sample, I added 600 µL of the solution to a DNeasy spin column and centrifuged it at 8000 rpm for one minute. The spin column was placed in a new collection tube, repeating this step until all the solution passed through the spin column. I placed each spin column in a new collection tube, added 500 µL of AW1 (wash buffer), and centrifuged it at 8000 rpm for one minute. We repeated this step with AW2 (wash buffer), but the samples were centrifuged at 11,000 rpm for three minutes. The spin column was placed in 1.5 ml tubes labelled with the sample name on the lead and side. I eluted the eDNA with 50 µL buffer AE (elution buffer), incubated it at room temperature for 5 minutes, and then spun it at 8000 rpm for one minute. Finally, all eDNA extraction samples, negative fields, and negative eDNA extraction controls were stored in the freezer (-20C) for further analysis. The concentration of all eDNA extraction samples was measured using Qubit™ 4 Fluorometer dsDNA HS standard. The DNA was visualised on a 0.8% agarose gel to assess the quality and compare the amount of the eDNA extracted from the filter samples.

2.3.3 Aggregation of eDNA by the Roller Method for ACO Samples

The roller approach is chlorine-sterilised commercial paint rollers to collect environmental DNA (eDNA) from the surface and underneath ACOs. This approach has successfully detected the DNA of the terrestrial lizard *Scincella lateralis* (Kyle et al., 2022). Furthermore, it was utilised for the first time to monitor invasive forest insects, *Lycorma delicatula*, by swabbing tree branches with a cotton roller (Valentin et al., 2021). The roller

method has not been previously applied to detect any snakes. In this study, the roller technique was applied to detect barred grass snakes as outlined in (Kyle et al., 2022) with some modifications. The kit roller method for each ACO sample consists of a chlorine-sterilised commercial paint roller (10.2 cm length - 3.55 cm diameter), a pair of sterile gloves, and a sterile bag of 500 ml. On arrival at the site, the weather conditions were recorded. I selected the artificial cover objective based on the above criteria (see section 2.2, period of eDNA sample collection). After identifying the ACO, the roller was attached to a pole after being cleaned with 10% bleach to prevent cross-contamination and moistened with double-distilled water to swab the entire topside and bottom of an ACO in contact with the soil. To avoid cross-contamination, each cover object was sampled with a new roller. After swabbing, the roller was put in a sterile bag and kept in a cooler (about 5°C) to protect DNA throughout transportation to the laboratory for eDNA extraction. I included a field negative control during each sampling day, which involves following the standard technique but not collecting samples. This was done to look for any potential in-field contamination of the samples.

2.3.4 Filtration and Extraction of eDNA Roller Samples

After collecting the roller samples within 3-5 hours, each roller was rinsed by filling approximately 250 ml of deionised water into sterile bags in a separate laboratory. This was done to bring the collected environmental DNA (eDNA) into a watery solution. The liquid was then filtered through a Sterivex HV 0.45 μM filter attached to a 60 ml syringe to recover the eDNA. The Sterivex filter units were stored and frozen at -20°C until DNA extraction (Valentin et al., 2021; Kyle et al., 2022). During the eDNA filter extraction, I included a negative control to verify potential in-lab sample contamination; the DNeasy PowerSoil Pro extraction - QIAGEN was considered to be essential due to (1) including multiple PCR inhibitor removal and DNA purification steps (2) The kit can deal with the soil transferred

from cover objects to rollers (Katz et al., 2020; Kyle et al., 2022). In the laboratory assigned for DNA extraction purposes, the process began with releasing the Sterivex filter from the plastic unit by following the protocol described above (eDNA Water Samples). The small fragments of the filter were then added to a PowerBead tube with 800 ul of CD1 (lysis buffer) and vortex briefly to mix. The PowerBead tube was horizontally placed on a vortex adapter for 10 min. After that, the PowerBead was centrifuged at 15.000 X g for 1 min. The supernatant was transferred to a clean 2 ml microcentrifuge tube. 200 µL of CD2 (precipitation buffer that helps remove contaminants) was added to the 2 ml tube, vortexed for 5 seconds, and then centrifuged at 15,000 X g for 1 min. I transferred the supernatant (~ 600 µL) to a new 2 ml microcentrifuge tube, avoiding the pellet, and then I added 600 µL of CD3 (binding buffer that adjusts the salt concentration) and vortexed for 5 seconds. I loaded 650 µL of the lysate onto the Spin Miniprep Column (QIAGEN) and centrifuged at 15,000 X g for 1 min; this step was repeated until all the lysate passed via the MB spin column. After that, the Spin Miniprep Column (QIAGEN) was placed into a new 2 ml collection tube, and I added 500 µL of EA solution (wash buffer) centrifuged at 15,000 X g for 1 min. Then, I discarded the flow-through and put the Spin Miniprep Column (QIAGEN) again into the 2 ml collection tube. C5 solution (500 µL) (wash buffer) was added to the Spin Miniprep Column (QIAGEN) and centrifuged at 15,000 X g for 1 min. The flow-through was discarded, and the Spin Miniprep Column (QIAGEN) was placed into a new 1.5 ml elution tube. Finally, 50 ul of the C6 solution (elution buffer) was added to the white filter membrane and centrifuged at 15,000 X g for 1 min. The eDNA extraction samples were quantified by Qubit™ 4 Fluorometer dsDNA HS (High Sensitivity) (Thermo Fisher Scientific). DNA was visualised on a 0.8% agarose gel with ultraviolet (UV) light to assess the successful extraction of the eDNA samples. All eDNA extraction samples and negative controls, including field and eDNA extraction, were stored in the freezer (-20°C) for further analysis (qPCR).

2.4 Development and Validation of a barred grass snake Species-specific qPCR Assay

Because there was no published species-specific qPCR assay available for the barred grass snake at the start of this study (2021), a qPCR assay was developed. We focused on the mitochondrial DNA Cytochrome b gene (Cytb). A limited number of mtDNA sequences for the Cytb locus were available in the NCBI database. I, therefore, collaborated with the team at the Centre of Genomic Research at the University of Liverpool and Dr Mark Whitehead, who sequenced and assembled the complete mitochondrial genome of a UK female barred grass snake, *Natrix helvetica*, from an individual found dead in Kent, UK (snake supplied by Dr George Day). This genome sequence, i.e. from a UK-born individual, was required since the barred grass snake mitochondrial sequences available in public sequence databases originate from a different non-UK species that were previously mistakenly classified as barred grass snakes but which are now known not to correspond to the species found in the UK (Kindler & Fritz., 2018).

2.4.1 Assay Design

I compiled a list of non-target species, including vertebrates coexisting in the same geographic area as the location from where barred grass snake samples were taken, including all native UK reptile species and the closely related but geographically distinct species of barred grass snakes at the genus *Natrix* level. Subsequently, I obtained the Cytb sequences from each species on the specified list alongside several different individuals of barred grass snakes (representing different geographic variations within the UK and abroad) from the National Centre for Biotechnology Information (NCBI) sequence database, GenBank, (https://www.ncbi.nlm.nih.gov/genbank/) (Langlois et al., 2021) I aligned these Cytb mitochondrial sequences using Geneious sequence alignment software (version 2023.1) (https://www.geneious.com) to identify suitable fragments that showed high nucleotide

divergence among species and yet low intraspecific nucleotide variation within the barred grass snake species (Klymus et al., 2020). Then, I visually determined the qPCR primer and probe region for the gene based on the recommendations of The MIQE Guidelines (Bustin et al., 2009; Wilcox et al., 2013) using Geneious (version 2023.1). To prevent cross-amplification from non-target species and enhance the specificity of the assays, I ensured that the forward and reverse primers had more mismatches with the many non-target species than the barred grass snake and that these mismatches were located on the last three base pairs at the 3' end of the primers rather than at the 5' end (since this is expected to cause failure to amplify). I also ensured the location of the probe bind site was close to the primer bind site, avoiding overlapping (Katz et al., 2020; Klymus et al., 2020). Furthermore, essential parameters were considered during the designing of the assay, such as the length of the primers and probe being within 18-30 bp, the melting temperature (Tm) of the forward and reverse primers being between 58-65°C (~ 1–5°C difference in the Tm between forward and reverse primer). The Tm of the probe was kept higher than the Tm of the primers; the GC content was between 35% - 65%, having GC clamp in the last five bases of the 3' end of the primers for a strong bond. For the target amplicon (the product of qPCR), I determined that the amplicon length should be between 100-200 bp. This range was chosen because environmental DNA (eDNA) is typically found in smaller fragments resulting from natural degradation processes (Harrison et al., 2019; Ruppert et al., 2019; Langlois et al., 2021; Sahu et al., 2023).

2.4.2 In Silico Validation

For in silico validation, different bioinformatics tools were used to get an initial impression of the quality of the developed assays and amplicons. The suitability evaluation of the assay was performed with the sequence manipulation tool https://www.bioinformatics.org/sms2/pcr_primer_stats.html (Stothard.,2000). The specificity

without probes was evaluated using the NCBI Primer Blast tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) (Ye. J et al., 2012) to detect possible cross-amplification with non-target species. I selected "nr" as the database of interest for Primer-Blast parameters. I also used NCBI Nucleotide Blast (Blastin) to examine the primers in combination with a probe for specificity by following Klymus et al. (2020) using the formula (forward primer, 12 Ns, the probe, 12 Ns, the complement of the reverse primer) then searching against the "nr" database. Moreover, the IDT OligoAnalyzer Tool (https://eu.idtdna.com/pages/tools/oligoanalyzer) was applied to scrutinise the primers and probe for the potential formation of primer dimers and hairpins. The melt curve software (*uMelt*) was used to predict the melt temperature and denaturation profiles of the amplicon (https://www.dna-utah.org/umelt/quartz/um.php), which was used to compare with the positive control of the barred grass snake. during in vitro validation using the SYBR green method (Dwight et al., 2011). Finally, to accurately predict cross-amplification in the qPCR assay (the forward, reverse, and probe primers in combination), the eDNAssy mechanism tool developed by Kronenberger et al., 2022 (https://nationalgenomicscenter.shinyapps.io/eDNAssay/) was chosen to predict the false negative detection error when a template that is present will not amplify, by analysing the number and locations of mismatches within primers and probes. The optimal threshold of the eDNAssy tool was assigned to 0.55 (default setting), which means assignment probabilities less than 0.55 were predicted not to amplify, and those greater than 0.55 were predicted to amplify (Tournayre et al., 2023). After the in silico validation, the primers and probe were ordered from Integrated DNA Technologies (IDT). The probe was labelled with 5' 6-FAM fluorophore as a reporter dye and IDT 3' ZEN-Iowa Black FQ as a quencher. The double-quenched (ZEN/3'IBFQ) exhibits lower background fluorescence and accurate detection, especially in the case of longer probes. In contrast, a minor groove binding probe

(MGB), such as a single-quencher probe, was popular in environmental DNA studies. The MGB is relatively short and may potentially bind with non-target species even with a low mismatch of 2-3 base pairs in comparison with a double-quenched probe (ZEN/3'IBFQ) (Wilcox et al., 2013; Klymus et al., 2020).

2.4.3 In Vitro Validation

The developed assay was screened in vitro through experimental analysis involving (1) optimising the annealing temperature (Ta) and the final concentration of assay in qPCR reaction, (2) calculating the efficiency of qPCR primers and probe, (3) the specificity by analysing the melt curve of target amplicon and cross-amplification of non-target-species (4) the sensitivity by determining the limit of detection (LOD). In vitro validation is essential to verify the primary results obtained from the in-silico step before final assay validation. To avoid fluctuations in concentrations of mitochondrial and nuclear DNA (mixture genomic DNA) from the different tissue samples, the synthetic DNA fragment (gBlock™) that was identical to the theoretical sequence template amplicon (including the region bind of primers and probe) ordered from IDT (Table 3.1) (Klymus et al., 2017; Langlois et al., 2020; Mattias et al., 2021).

Table 3.1. The synthetic DNA fragments (gBlock™) sequences (5' – 3') corresponding to the target mitochondrial Cytb gene and the length in base pairs used in vitro validation. The forward and reverse of gBlock™ are in blue; the probe is in red. Synthetic sequences were used as standards to calculate the assay's efficiency and determine LOD and LOQ.

Target mitochondrial Gene	Synthetic DNA fragments (gBlocks™) Sequence 5' – 3'	Length (bp)
Cytochrome b (CytB)	CCTACGGATGAATAATACAAAACACCCATGCAATTGGCGCA TCAATATTTTTTATCTGTATCTATACCCACATTGCACGTGGAC TTTACTATGGCTCCTACCTAAACAAAGAAGTGTGACTATCG	125

2.4.3.1 DNA Isolation from the Tissue Samples Used In The Validation

DNeasy Blood & Tissue Kit (Qiagen) was selected to extract the total genomic DNA samples following the manufacturer's protocol, which was used in vitro validation from different populations and individuals of barred grass snakes. The collection samples of barred grass snakes were obtained from different sources: five slough skins found underneath the artificial cover board in the main site (identified using the unique scale pattern that was distinctive to the barred grass snake and confirmed by Sanger sequencing the cytochrome B (Cytb) gene, primer set: L15161-H15714, Guicking et al., 2009), five slough skins were provided by Dr Steven Allain from Norfolk, UK population, three slough skin (preserved in 75% alcohol) from Weston Park Museum, Sheffield, UK provided by Dr Alistiar McLean, and one fresh tissue from whole female grass snake found dead in the road at Doncaster, UK (53.535639,-0.970778), provided by Dr Paul J Parsons (see Appendix 3 for details). Furthermore, I collected available tissue samples of the non-target species, (1) Native reptiles of the UK provided by Dr Robert Ward (Amphibian and Reptiles Conservation), including Adder (*Vipera berus*), Smooth snake *(Coronella austriaca*), Common lizard *(Zootoca vivipara*), and Slow -worm (*Anguis fragilis*); (2) Amphibians were provided (preserved in absolute alcohol) by Dr Deborah Dawson (NEOF Visitor Facility, University of Sheffield) including Common frog (*Rana temporaria*), Common toad *(Bufo bufo*), and Great crested newt (*Triturus cristatus*); (3) Mammals (preserved in absolute alcohol) were provided by Dr Deborah Dawson, Common shrew (*Sorex araneus*), Pygmy shrew (*Sorex minutus*), Water Shrew (*Neomys fodiens*), Bank vole (*Myodes glareolus*), Water vole (*Arvicola amphibious*), Harvest mouse (*Micromys minutus*). After DNA extraction, all the genomic DNA samples were tested for quality using the Agilent 2200 TapeStation with Genomic DNA Reagents. Subsequently, the DNA samples were quantified using the Qubit™ 4 Fluorometer Invitrogen

dsDNA HS (High Sensitivity), and the purity (260/280 ratio) was assessed with the Thermo Scientific NanoDrop 8000 Spectrophotometer.

2.4.3.2 Optimization of the barred grass snake qPCR Assay

The range of optimum annealing temperature (Ta) of the qPCR primers was determined by testing 12 different PCR annealing temperatures (54.9 ℃, 55.1°C, 55.7 ℃, 56.7 ℃, 58.1°C, 59.4 ℃, 60.6 ℃, 61.9 ℃, 63.2 ℃, 64.9 ℃, 65.1°C, 66.2 ℃). PCR amplification was performed using an Eppendorf® Mastercycler X50 PCR machine. The total 20µL PCR reaction consisted of 10µL of 1X QIAGEN Multiplex PCR Master Mix (including Hot-start Taq DNA polymerase), 1µL of both forward primer (10µM) and reverse primer (10µM), so the final primer concentration was 500 µM, 6µL of ddH2O, and 2µL of 1.00E2 copy/µL of gBlock as a DNA template of barred grass snake. The PCR program was: initial activation step at 95°C for 15 min (to activate the polymerase), followed by 35 cycles of denaturation at 94°C for 15 sec, annealing at 54.9–66.2°C for 1 min (a different annealing temperature for each replicated sample (n=12)), extension 72°C for 45s (Extension), then final extension at 72°C for 10 min. After the PCR amplification, the amplified DNA fragments were visualised using 2.5% agarose gel electrophoresis and a 50 bp DNA ladder from New England Biolabs. The following factors were considered when determining the optimal Ta: the expected amplicon size, the strength of amplification, the specificity of the amplicon (whether tight bands or smeary products), the absence of primer dimers, and the presence of any additional fragments. In addition, the optimum final concentration of the assay was selected by comparing amplification after running PCRs with different concentrations of the qPCR primers: 100 nM, 200 nM, 300 nM, 400 nM, 500 nM, 600 nM, and 700 nM. The probe concentration used was 250 nM (10µM) because it was used in most studies. The 20µL reaction volume for the qPCR consisted of 10µL TaqMan™ Environmental Master Mix, 1µL each of forward and reverse primers at concentrations

ranging from 100 nM to 700 nM, 0.5µL of probe at a final concentration of 250 nM, 5.5µL ddH2O, and 2µL of 1.00E2 copy/µL of gBlock as the DNA template for the barred grass snake. The qPCR reaction was carried out with an initial hold stage at 50°C for 2 minutes and an initial denaturation at 95°C for 10 minutes, then denaturation at 95°C for 15s and annealing at 60°C for 1 minute, followed by 50 cycles using the QuantStudio 12K Flex / Fast 96-Well Block (0.1mL) from Applied Biosystems. A concentration of 500 nM was selected as the concentration of the primers to be used in the qPCR based on the earlier Ct values and the behaviour of the curve, particularly the plateau phase curve.

2.4.3.3 Determine the Efficiency of qPCR Assays

The efficiency of qPCR assay amplification was calculated by linear regression slope of the standard curve (calibration curves) using the equation *E = -1+10 (-1/slope).* In addition, the coefficient of determination R2 value was considered to represent the fit of the standard replicates from the assay on the qPCR curve (Wilcox et al., 2013; Klymus et al., 2020). The standard curve was established by creating serial dilutions (10-fold) of synthetic DNA gBlock™. As the stock, I made a 1.00E8 copy/µL of gBlock to generate six-point dilutions, 1.00E7, 1.00E6, 1.00E5, 1.00E4, 1.00E3, 1.00E2 (10-fold). The 20µL reaction volume for the qPCR consisted of 10µL TaqMan™ Environmental Master Mix, 1µL each of (10 µM) for both primers (final concentration 500 nM) and 0.5µL of the probe (10 µM) at a final concentration of 250 nM, 5.5µL ddH2O, and 2µL of the standard of the synthetic DNA gBlock in three replicates. In the PCR, I included a negative double distilled water and a positive control of barred grass snake (DNA 0.1 ng/µL). The qPCR conditions were used as described in the section on optimation assay (Section - 2.4.3.2). The analysis of the qPCR input (efficiency E % and R2) was automatically performed and calculated in the qPCR machine. For optimal results, it is ideal for the qPCR reaction efficiency to range from at least 90% to below 110% and R2 0.98 (Taylor et al., 2019; Klymus et al., 2020).

2.4.3.4 Specificity Verification of qPCR Assays

To ensure the specificity of the qPCR assay, I conducted three evaluation stages. In the first stage, melt curve analysis was used to determine the melt temperature of the target amplicon and compare it with the expected *uMelt* result in silico and the positive control (DNA of barred grass snakes). I ran six standard dilutions of the gBlock in three replicates, including NTC (sterile water) and positive control. The 20µL reaction volume for the qPCR consisted of 10µL PowerUp™ SYBR™ Green Master Mix, 1µL each of (10 µM) for both primers (final concentration 500 nM), 6µL ddH2O, and 2µL of each standard of the synthetic DNA gBlock, negative, and positive controls. The qPCR conditions: stage 1: initial hold stage at 50°C for 2 minutes and an initial denaturation at 95°C for 10 minutes, stage 2: step 1 denaturation at 95°C for 15 s step 2: 60°C annealing temperature for 1 min, followed by 50 cycles. Then, in the melt curve stage: 95°C for 15 s, 60° for 1 minute, and 95°C for 15 s using the QuantStudio 12K Flex / Fast 96-Well Block (0.1mL) from Applied Biosystems. In the second stage, I tested the specificity of the assays within different individuals of barred grass snakes for intraspecific genetic variation using the reactions and conditions (SYBR™ Green) described in stage one; the concentration of gDNA templates was at 0.1 ng/µL at this stage. During the third stage (Cross-amplification validation), I confirmed the specificity of the qPCR assay that only amplified the target organism through two methods: the use of the SYBR™ Green method to assess the specificity of the qPCR primers exclusively and the TaqMan method to evaluate the specificity of both qPCR primers and probes (using full assays). The positive control (gDNA at 0.1 ng/µL) and NTC (sterile water) were included during the validation, and all samples were run in three replicates. The amplification of the SYBR™ Green method was as the reactions and conditions (SYBR™ Green) described in-stage one. The gDNA at a concentration of 0.1 ng/µL (non-target species) was used as a template. For the TaqMan method: The qPCR reaction and conditions were similar in section 2.4.3.3.

To confirm the amplification, I used 2.5% agarose gel electrophoresis with a 50 bp DNA ladder from New England Biolabs and an Agilent 2200 TapeStation with Genomic DNA Reagents to verify the anticipated size of the amplicon and check for any possible unintended amplification (such as multiple bands) and primer-dimer formations.

2.4.3.5 Sensitivity Validation of qPCR Assays

I determined the limit of blank (LOB), the limit of detection (LOD), and the limit of quantification (LOQ) of the qPCR assay using the TaqMan approach. The LOB, LOD, and LOQ were created by serially diluting barred grass snake synthetic DNA fragments at known concentrations (copies/µL). To create low synthetic DNA concentrations, I used 1,000 copies/µL of synthetic DNA gBlock with 10 ng/µL Yeast tRNA as a carrier to overcome the possibility of the gBlock being degraded or adhering to the plastic reaction well. This initial stock was then diluted in a four-fold series with 10 ng/µL tRNA in Invitrogen UltraPure DNase/RNase-free purified water to generate seven synthetic DNA concentrations ranging from 250 copies/µL to 0.06 copies/µL (Lesperance et al., 2021; Matthias et al., 2021). After that, 2 µL of each dilution was utilised in the qPCR reactions with six technical replicates. Moreover, I also ran three replicates of each synthetic DNA concentration: 1.00E7, 1.00E6, 1.00E5, 1.00E4, 1.00E3, 1.00E2 (10-fold) copies/µL to assess the efficiency of the amplification. The Limit of Blank (LOB), Limit of Detection (LOD), and Limit of Quantification (LOQ) for the qPCR assay were determined using the default settings outlined by Lesperance et al. (2021): (1) the LOB was calculated with a false positive rate of 0.05 (5% chance of detecting false positives in blank samples) (2) the LOD was established with a false negative rate of 0.05 (95% of replicates detected the target DNA at this concentration) (3) the LOQ was defined as the lowest concentration where the coefficient of variation (CV) was less than or equal to 0.2 (maximum relative variation of 20% between replicates). The LOB, LOD, and LOQ were calculated in RStudio (Version: 2024.04.2+764) using the eLowQuant R script (https://github.com/mlespera/eLowQuant) based on the Binomial–Poisson statistical model developed by Lesperance et al. (2021).

2.4.4 *In-Situ* Validation

The in-situ validation of the qPCR assay was conducted on eDNA samples in three technical replicates. The qPCR process was conducted in a separate laboratory isolated from the DNA extraction area. Also, I used a sterile flow hood and ensured that I wore nitrile gloves and lab coats during the sample analysis. Additionally, all the qPCR reagents and DNA extraction procedures were carried out using sterile equipment, and I thoroughly sterilised all surfaces with a 10% bleach solution. The eDNA samples were amplified in 20 µL total qPCR reaction volume consisting of 10 µL TaqMan Environmental Master Mix, 1 µL each of 10 µM primer (final concentration 500 nM), 0.5 µL of 10 µM probe (final concentration 250 nM), 5.5 µL ddH_2O, and 2 µL eDNA samples, in three replicates as well as positive control (0.1 ng/µL gDNA of barred grass snake) and three replicates of each synthetic DNA concentration 1.00E7, 1.00E6, 1.00E5, 1.00E4, 1.00E3, 1.00E2 (10-fold) copies/µL as a standard curve (linear dynamic range) to assess the efficiency and performance of the assay during the amplification process. Negative controls (field, DNA extraction procedure, and qPCR), which consisted of all the components except the eDNA sample, were included to detect any contamination during qPCR amplification. The qPCR conditions included a hold stage at 50°C for 2 minutes, initial denaturation at 95°C for 10 minutes, then a denaturation at 95°C for 15 s, annealing at 60°C for 1 min, followed by 50 cycles using the QuantStudio 12K Flex / Fast 96-Well Block (0.1 mL) (Applied Biosystems). After the qPCR amplification, the amplified eDNA samples were analysed on a 2.5% agarose gel with a 50-bp DNA ladder to check for the amplicon size. After the qPCR amplification using TaqMan chemistry, the amplified eDNA samples showing a positive detection would be again tested using SYBR green chemistry and melt curve analysis in 20-µL qPCR reactions, consisting of 10 µL PowerUp SYBR Green Master Mix, 1 µL of each primer (10 µM, for final concentration 500 nM), 6 µL purified H_2O, and 2 µL of the eDNA sample,

The negative controls (here I just used qPCR negative control), and positive control (0.1 ng/µL gDNA of barred grass snake) were also involved. The qPCR conditions: stage 1: initial hold stage at 50°C for 2 minutes and an initial denaturation at 95°C for 10 minutes, stage 2: step 1 denaturation at 95°C for 15 s step 2: 60°C annealing temperature for 1 min, followed by 50 cycles. Then, in the melt curve stage: 95°C for 15 s, 60° for 1 minute, and 95°C for 15 s using the QuantStudio 12K Flex / Fast 96-Well Block (0.1mL) from Applied Biosystems.

2.4.4.1 Positive Detection of eDNA Samples

The eDNA sample is considered a positive detection if at least one replicate (1 of 3) is amplified below the detection limit. The non-template control (NTC) should not show amplification of field, DNA extraction, and qPCR NTCs. The eDNA sample must be of the accepted size determined by 2.5% agarose gel electrophoresis and a 50-bp ladder (New England Biolabs). The efficiency of the qPCR assay during the in-situ validation should be between 90-100%. The range of melt curve temperature for the eDNA positive detection should be within the range used for the positive control (depending on the quality of eDNA). The structure of the amplification curve should include the long linear phase (log exponential phase), ideally the plateau phase at the Ct ≤ 40 cycle (cut-off point). If possible, Sanger sequences should be obtained to confirm the identity of PCR products (the low DNA concentration and the quality of eDNA may affect the step). This strategy has been commonly used in eDNA studies (Hunter et al., 2015; Halstead et al., 2017; Atkinson et al., 2017; Rose et al., 2019; Ratsch et al., 2020; Katz et al., 2020; Moyer et al., 2022; Prabhakaran et al., 2023; Nordstrom et al., 2023).

3. Results

3.1 In Silico Validation

I successfully designed a -species-specific qPCR assay (BGS_CytB) for the barred grass snake (*Natrix helvetica*) to amplify 124 bp of the Cytb gene (located at 198-321 bp) as a final product (amplicon) from the total mitochondrial Cytb gene region (1116 bp) based on the length of sequence from Liverpool facility (The complete mitochondrial genome can be found in NCBI https://www.ncbi.nlm.nih.gov/nuccore/PQ553205.1/) and *Natrix helvetica* from NCBI database (The accession number: LR694441.1, LT900467.1, and LN994835.1). The forward primer (198-223 bp positions of Cytb) is 25 bp long with a GC content of 36% and a melting temperature (Tm) of 60.3°C. The reverse primer (297-322 bp positions) is also 25 base pairs long, with a GC content of 40% and a melting temperature (Tm) of 61.6°C (Table 3.2). The difference in melting temperatures between the forward and reverse primer was calculated to be 1.3°C, which falls within the acceptable range for efficient PCR amplification. Both primers had a GC clamp (two bp) at the 3' end to enhance binding specificity. During the PCR suitability test, no primer dimer or secondary structure was predicted for the designed primers. They passed all the evaluation criteria in the sequence manipulation tool (Table 3.3). The fluorescently labelled probe with a double-quencher (25 bp) was designed to have a melting temperature of 67.8°C. The GC content was 52% with a GC clamp in the 3' end. The primer dimer of the probe was very weak, and the secondary structure was absent (Table 3.3). The melting curve analysis of the 124-bp amplicon of the assay, using *uMelt* software, showed a distinct peak and predicted the temperature at 78.5°C. The single and sharp peak corresponded to a specific, well-defined melting temperature for the target amplicon (at 124 bp). There was no indication of non-specific peaks or primer dimers, confirming the assay's specificity and reliability (Figure 3.2). No non-target species were detected during the NCBI and Primer Blast analysis (for the primers without the probe), and *Natrix helvetica* had a 100% quarry cover.

The total score was 224, with an E-value of 3e-54. The forward and reverse primers were expected to be specific to 124 bp of the Cytb gene of *Natrix helvetica*. The entire BGS_CytB assay (forward and reverse primers with the inclusion of the probe) was evaluated using the eDNAssy tool to observe potential amplification from non-target species. The amplification prediction for *Natrix helvetica* was 0.972. Among the species tested, only *Natrix astreptophora* (0.779) and *Natrix natrix* (0.687) met the required threshold of 0.55 for optimal amplification. Neither of these species is native to the UK or the study site. Other taxa exhibited lower amplification predictions, with many falling below the threshold below which amplification should not occur (Figure 3.3). Moreover, the number of mismatches between the barred grass snakes and the non-target species was higher for the forward and reverse primers than for the probe, except for the *Natrix* genus (Figure 3.4-A). The proportion of mismatches at the 3' end of the forward and reverse primers and the central part of the probe are shown in Figure 3.4-B. According to the prediction of the eDNAssy tool, the BGS_CytB assay should be specific to *Natrix helvetica*.

Table 3.2. The developed specific-species qPCR assay was utilised in the study to detect the barred grass snake, *Natrix helvetica*, from eDNA samples by amplifying the mitochondrial Cytb gene. The table includes the forward and reverse primers and probe sequences (labelled with 5' 6-FAM fluorophore as reporter day and IDT 3' ZEN-Iowa Black FQ as double-quenchers, Bold), detailing the nucleotide sequences and the qPCR amplicon's final product size in base pairs (bp).

Oligo's Name	Target mtDNA gene	Sequences	Amplicon (bp)
BGS_CytB_Forward	Cytochrome b gene (CytB)	5'-CTACGGATGAATAATACAAAACACC-3'	124
BGS_CytB_Reverse		5'-CGATAGTCACACTTCTTTGTTTAGG-3'	
BGS_CytB_Probe		5'/**56-FAM**/TGCACGTGG/**ZEN**/ACTTTACTATGGCTCC/**3IABkFQ**/-3'	

Table 3.3. In silico validation of the essential parameters of Cytb assay using sequence manipulation tool and oligo-analyzer IDT. Properties of each primer include its length in base pairs (bp), melting temperature (Tm), GC content percentage, presence of a GC clamp, secondary structure, and primer dimer formation.

Oligo's Name	Length (bp)	Melting Temperature (Tm)	GC Content %	GC Clamp	Secondary Structure	Primer Dimer
BGS_CytB_Forward	25	60.3 ℃	36	2	Pass	Pass
BGS_CytB_Reverse	25	61.6 ℃	40	2	Pass	Pass
BGS_CytB_Probe	25	67.8 ℃	52	2	Pass	Pass

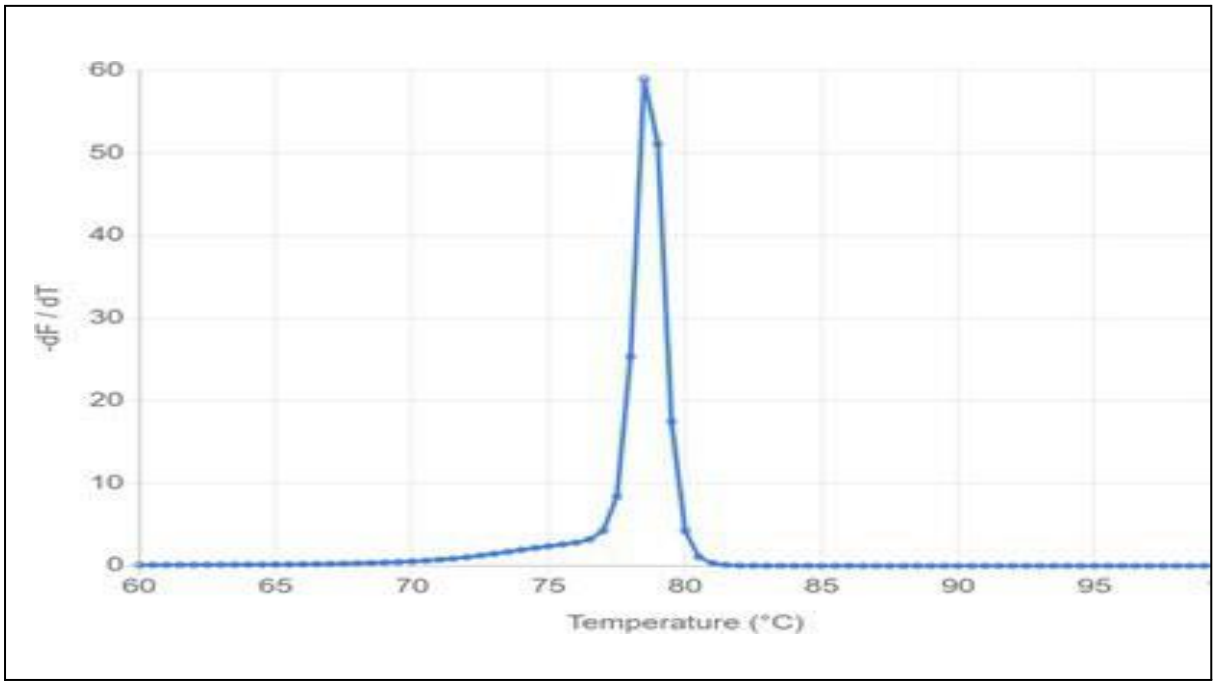

Figure 3.2. The melting curve analysis of the Cytb assay amplicon was conducted in silico validation using the uMelt software to predict the melting temperature of the amplicon. The plot displays the derivative of fluorescence (-dF/dT) against temperature, highlighting the thermal denaturation profile of the amplicon.

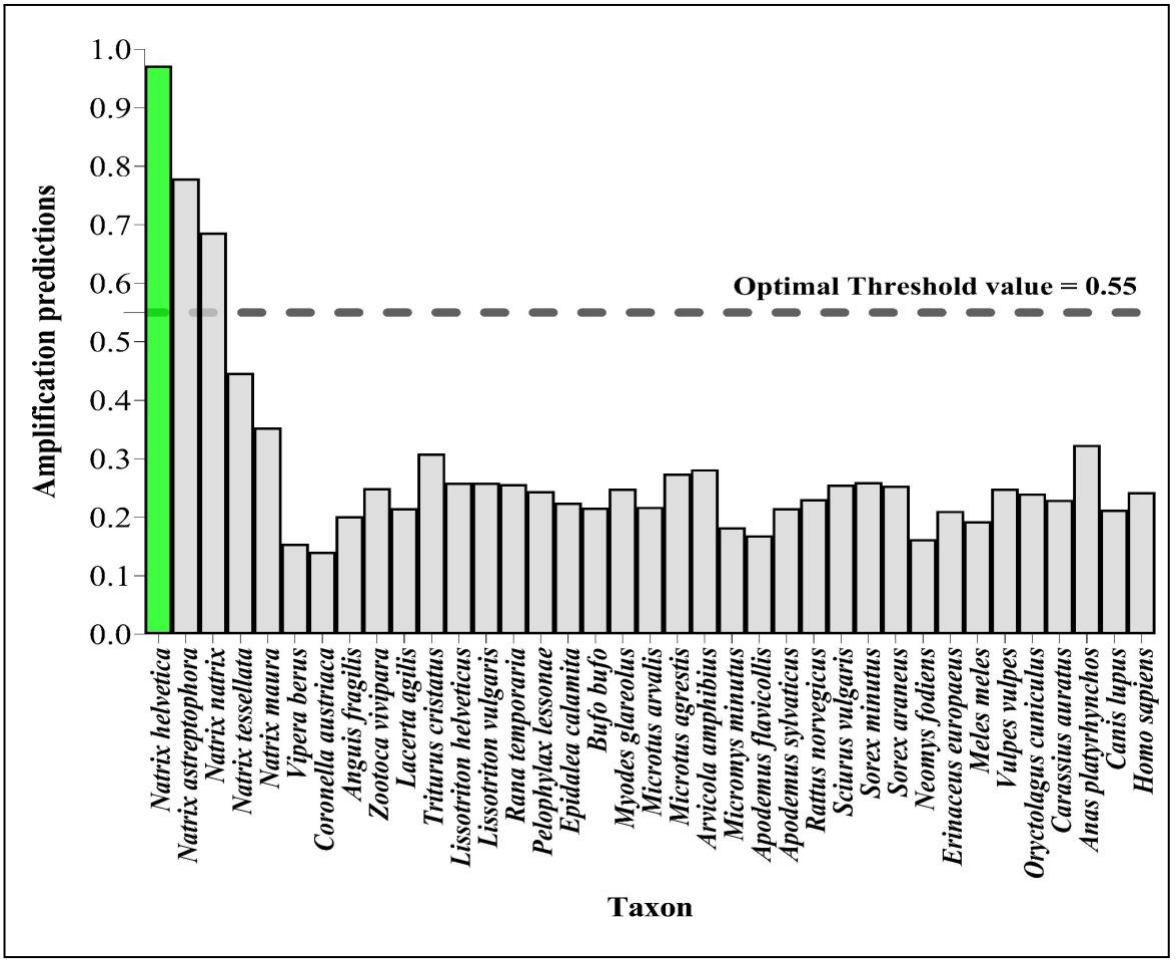

Figure 3.3. The amplification predictions for Cytb qPCR assays across the non-target taxa listed on the x-axis; each taxon is represented by one grey bar, and the target species with a green bar. The y-axis represents the amplification prediction values ranging from 0 to 1. The dashed line indicates an optimal threshold value of 0.55, a reference for evaluating the potential of cross-amplification according to the eDNAssy mechanism tool (Kronenberger et al., 2022)

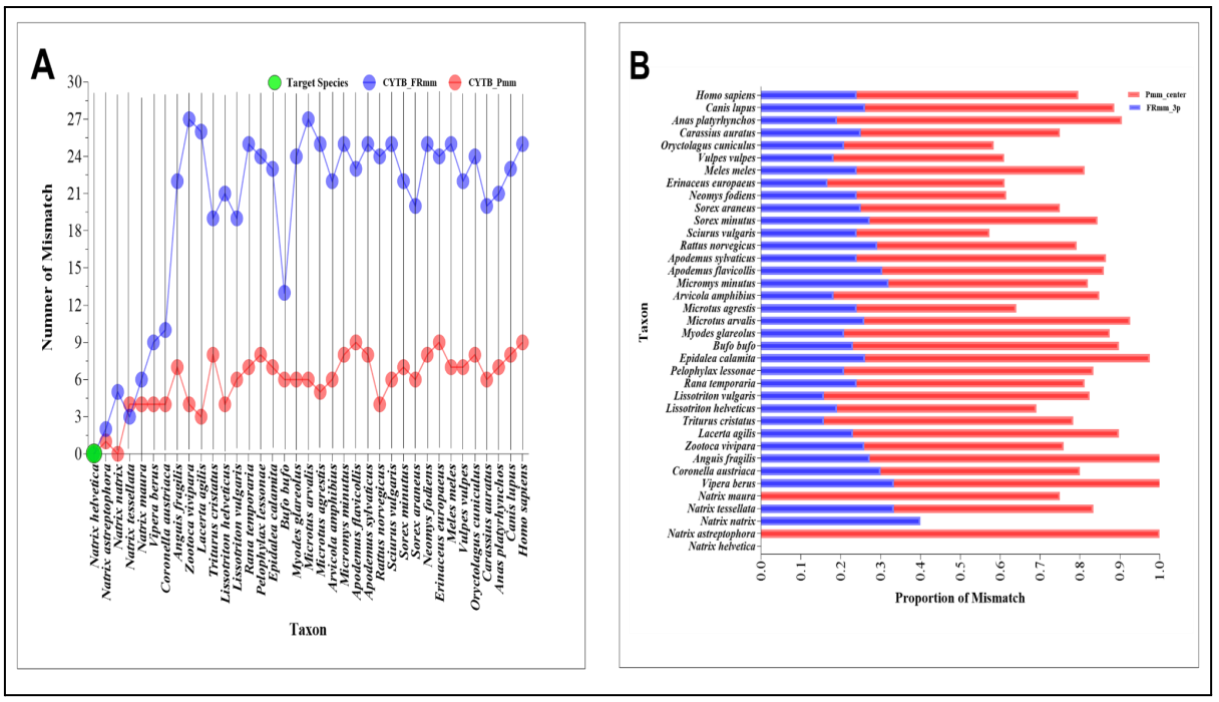

Figure 3.4. (A) A visual representation of the number of mismatches in both the forward and reverse qPCR primers (FRmm) and in the probe (Pmm) for the specific-species Cytb assay of Natrix helvetica across various species listed on the x-axis, ranging from the target species (green circle) (Natrix helvetica) to other non-target species, including reptiles, amphibians, birds, and mammals. The y-axis reflects the range of mismatches from 0 to 30. Blue-filled circles represent total mismatches for both primers (CYTB FRmm) across the different taxa, and red-filled circles represent mismatches for the probe (CYTB Pmm). **(B)** The proportion of mismatches observed at the 3' end of the last five bp in forward and reverse primers (FRmm_3p) and the proportion of mismatches in the central region of the probe (Pmm_center) across various taxa. Each bar represents a different species, highlighting the specificity of the primers for the target species, *Natrix helvetica*, which shows zero mismatches. The blue bars represent the mismatch proportions for the FRmm_3p, while the red bars represent the mismatch proportions for the Pmm_center.

3.2 In Vitro Validation

The optimisation qPCR reaction of the BGS assay was determined based on the best Ct value (earliest value) of the positive control (0.1 ng/µL), which was 24.21 (SD of Ct value 0.102) and reached the plateau phase. The calibration curve and linear regression result from the standard dilutions confirmed the high efficiency of the BGS_CytB assay and indicated that the amount of the target amplicon (final qPCR product) was doubled with each cycle during amplification (Figure 3.5). The specificity of the assay using SYBR Green to analyse the melting curve confirmed the melting temperature of the amplicon with single peaks at

77.5^{O}C, similar to the predicted temperature of 78.5 ^{O}C. Moreover, the assay's accommodation of intra-specific genetic variation of *Natrix helvetica* from different populations was assessed using melting curve analysis. The melting curve temperature ranged between 77.74-77.50^{O}C, showing no difference between the melting temperatures of the amplicon and those of the various populations, indicating that the assay is tolerant of any intra-specific genetic variation present here (Figure 3.6).

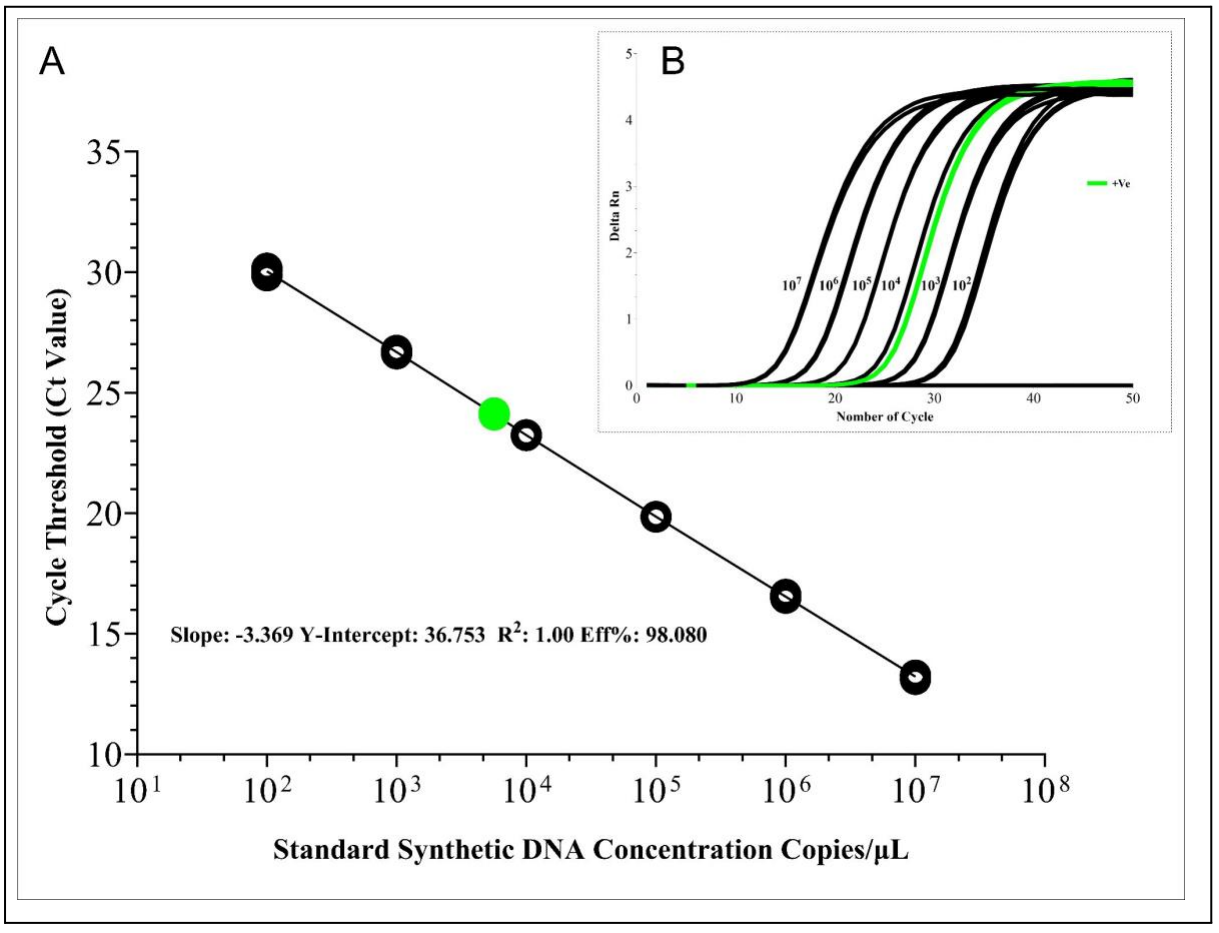

Figure 3.5. (A) The standard curves for Cytb qPCR assay by plotting the cycle threshold (Ct value) against the standard synthetic DNA concentration copies/µL. The efficiency of the assay and other metrics is provided in the grey box. **(B)** The qPCR amplification plot of the standard synthetic DNA, with each curve representing a different concentration from the standard series. The green curve and circle represent the positive control of the Barred grass snake (slough skin DNA at 0.1 ng/µL)

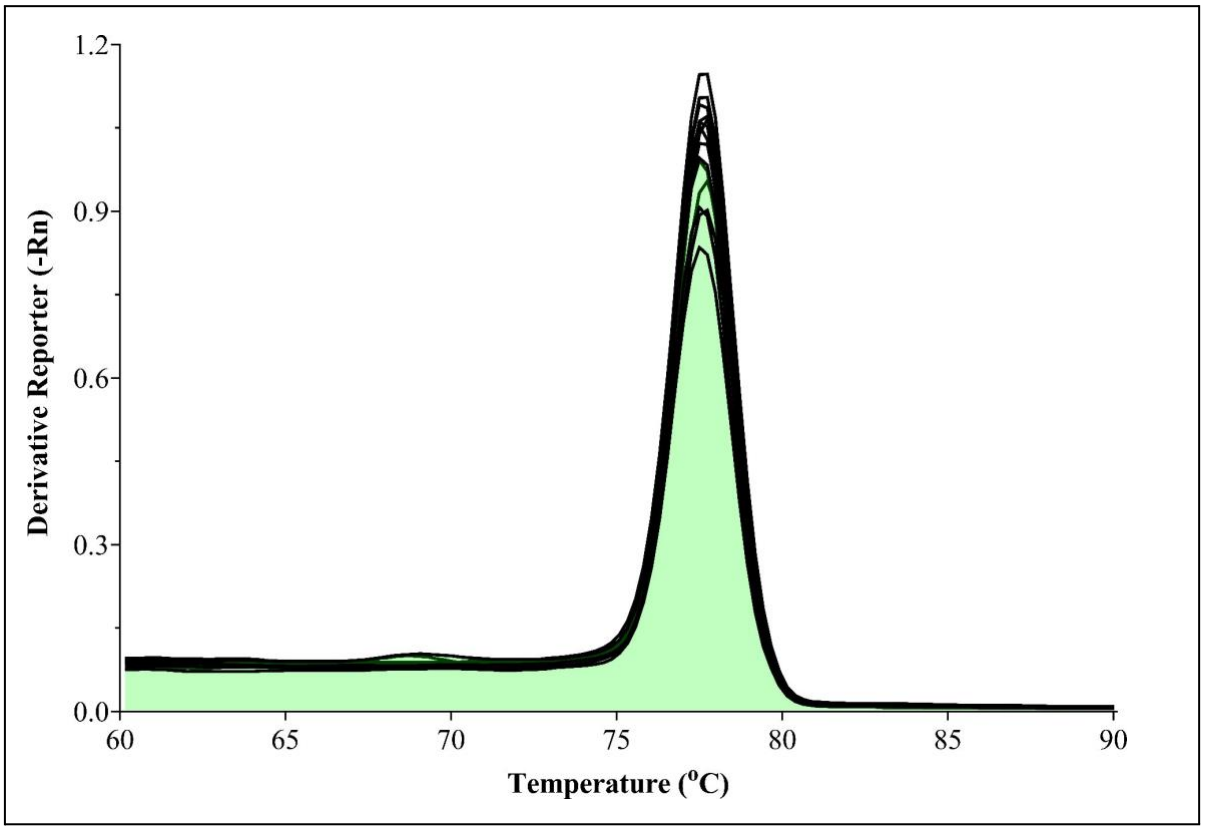

Figure 3.6. The melt curve analysis of barred grass snakes (Natrix helvetica) from various populations (Intra-specific genetic variations). The x-axis represents the temperature (°C) of the melting amplicon (final product), ranging from 60°C to 90°C, while the y-axis represents the derivative reporter (Rn), indicating the rate of change of fluorescence in response to temperature.

For cross-amplification of the non-target species, using SYBR Green the assay was shown to be specific to *Natrix helvetica* (Ct mean value = 24.951 ± 0.084) and produced a clear peak at 77.46^{O}C **(Table. 4)**. Although some species from the nan-target list amplified beyond Ct = 36.97, their melt curve analysis produced a primer-dimer between 60 and 69^{O}C as confirmed by gel electrophoresis. By Using TaqMan approaches to test the specificity of the probe against the non-target species, again only *Natrix helvetica* amplified with a clear morphology of the curve (Ct mean value = 25.70 ± 0.16) and no cross-amplification was observed at a gDNA concentration of 0.1 ng/uL. This confirmed that the probe was specific to the target species. During the in vitro specificity validation of the assay, over 60 NTCs

were used in two approaches, and only four were amplified when SYBR Green qPCR was conducted, producing only primer-dimer at ~64^{O}C. In the TaqMan approach, none of the NTCs were amplified at all. This was essential for confirming the absence of potential false positive detection (type 1 error) during specificity validation, especially for the TaqMan approach. The sensitivity of the BGS_CytB assay to detect the *Natrix helvetica* at deficient concentrations was confirmed by serial dilution of the gBlock synthetic DNA at a low standard concentration (4-fold). Using Binomial Poisson, both the intercept and no intercept models fit the data (r^2 = 0.55). The best estimate of the Poisson of the two models was the no intercept, with p-value = 0.229 (Lesperance et al., 2021). Only the first standard gBlock synthetic dilutions with no detections were analysed (less than 100% detection rate: 0.97. 0.24, and 0.06 copy/µL) and NTC (0 copy/µL) to determine the LOB, LOD, and LOQ. The calculated Limit of Detection (LOD) was 0.3 copy/uL with a 95% confidence interval of 0.2 - 0.8, corresponding to a Ct value of 38.51 based on the calibration curve and linear regression of the high standard dilutions, which was used as the cutoff Ct value during the amplification stage. The Limit of Quantification was 0.6 copy/uL with a 95% confidence interval of 0.40-1.80. The calculated Limit of Blank was 0.

Table 3.4. A list of non-target species was used in vitro for validation to test the specificity of the Cytb assay for barred grass snakes and Natrix helvetica. The DNA concentration utilised in the qPCR reaction was 0.1 ng/ µL.

Scientific Name	Common name	Observed in the site	Amplified in Vitro Validation		
Natrix helvetica	Barred Grass Snake	Yes	Yes	Yes	Yes
Vipera berus	Adder	No	No	No	No
Coronella austriaca	Smooth Snake	No	No	No	No
Anguis fragilis	Slow-Worm	No	No	No	No
Zootoca vivipara	Common Lizard	No	No	No	No
Triturus cristatus	Great Crested Newt	Yes	No	No	No
Rana temporaria	Common Frog	No	No	No	No
Bufo bufo	Common Toad	Yes	No	No	No
Sorex araneus	Common shrew	Yes	No	No	No
Neomys fodiens	Water shrew	No	No	No	No
Sorex minutus	Pygmy shrew	No	No	No	No
Myodes glareolus	Bank vole	Yes	No	No	No
Arvicola amphibius	Water vole	No	No	No	No

3.3 In-Situ Validation

The eDNA samples included 35 water samples collected in three different time phases from various ponds and 13 roller samples by swabbing the entire artificial cover boards. eDNA was successfully isolated and quantified. The eDNA concentration in water samples ranged between 16-100 ng/uL, and roller samples ranged between 42-88 ng/uL. The NTCs showed very low readings based on the Qubit (below the range of the HS standard), indicating minimal risk of cross-contamination during the eDNA extraction process and sample collection. For in situ validation, 48 eDNA samples were analysed using the TaqMan qPCR amplification method (Table 5). The efficiency of the BGS_CytB assay ranged from 96.41% to 97.23%, with slopes between -3.390 and -3.411 and R^2 values from 0.999 to 1.000. During validation, no amplification was observed in the non-template controls (NTCs) for DNA extraction, qPCR preparation, or field samples.

In the first phase, no Barred Grass Snakes were observed in the field, and none of the 25 water samples collected during this phase resulted in amplification, indicating non-detection. During the second phase, five eDNA water samples were collected from five ponds close to the artificial cover objects in which barred grass snakes were observed during a Capture Mark Recapture (CMR) survey. Four samples were successfully amplified, resulting in a positive detection rate of 60% (3 out of 5) and one inconclusive sample (20% above the detection limit). In the third phase, when barred grass snakes were observed swimming in the ponds, the detection rate increased to 80% (4 out of 5 positive detections), and one water sample had no non-detection result (Figure. 3.7, Tables 3.5 and 3.6). Moreover, during the CMR survey, 13 roller samples were collected from artificial cover objects (ACOs) where Barred Grass Snakes were found. Out of these, five roller samples showed amplification: one sample was below the limit of detection (LOD), resulting in an 8% positive detection rate (1 out of 13), while four samples were above the LOD but did not meet the criteria for confirmed positive detection, resulting in 31% inconclusive results (Figure. 3.7, Table 3.5 and 3.6). All samples that indicated detection or inconclusive detection,

the positive control, and NTCs were confirmed by gel electrophoresis (Figure 3. 8). Due to the limited availability of SYBR Green reagents during the validation, four roller samples with inconclusive detection were selected for confirmation by analysing the melting curve temperature. All three roller samples peaked at 77.37°C, consistent with positive control at 77.25°C, and one sample did not amplify (Figure 3.9). The map of ponds and ACOs that were positive for eDNA detection is found in Figure 3.10).

	Water Samples			Roller Swab Samples
	Phase 1	Phase 2	Phase 3	
Total eDNA collecting samples	25	5	5	13
No. of eDNA samples not amplified	25	1	1	8
No. of eDNA samples above the limit of detection	0	1	0	4
No. of eDNA samples below the limit of detection	0	3	4	1
eDNA Detection Rate (%)	0%	60%	80%	7.69%

Table 3.5. A summary of environmental DNA (eDNA) detection results from water (across three different phases of collecting) and roller swab samples using Cytb assay specified for barred grass snakes.

Sample ID	No. of Replicates amplified	No. of Replicates below LOD	Ct Value Mean	(SD) of Ct Value	Quantity Mean copy/reaction	Confirmation of amplicon size on Gel	Melt Curve Analysis
POND-B7-**3**	3/3	3/3	37.348	0.998	0.614	Yes	No applied
POND-C2-**3**	3/3	2/3	37.730	1.478	0.614	Yes	No applied
POND-B5-**3**	3/3	3/3	34.034	0.405	5.177	Yes	No applied
POND-C1-**3**	3/3	1/3	38.890	0.713	0.557	Yes	No applied
POND-A1-**2**	1/3	1/3	36.657	-	0.859	Yes	No applied
POND-B7-**2**	2/3	2/3	37.724	0.240	0.421	Yes	No applied
POND-D10-**2**	2/3	1/3	38.235	1.649	0.393	Yes	No applied
POND-C2-**2**	2/3	0/3	41.552	0.189	0.071	Yes	No applied
POND-A3-**1**	1/3	0/3	41.771	-		Yes	No applied
POND-B3-**1**	1/3	0/3	41.042	-		Yes	No applied
POND-B4-**1**	1/3	0/3	39.832	-		Yes	No applied
POND-B7-**1**	1/3	0/3	39.736	-		Yes	No applied
ROLLER-C2	1/3	1/3	38.036	-	0.339	Yes	No applied
ROLLER-A3	1/3	0/3	38.864	-	0.194	Yes	77.371
ROLLER-B2	1/3	0/3	39.585	-	0.296	Yes	77.371
ROLLER-C1	1/3	0/3	39.286	-	0.363	Yes	undetermined
ROLLER-D1	1/3	0/3	38.851	-	0.487	Yes	77.371

Table 3.6. A summary of in situ validation results of Cytb assay obtained from eDNA samples. The table includes details such as the number of replicates tested for each sample, the number of replicates below the limit of detection (LOD), the mean cycle threshold (Ct) value, the standard deviation (SD) of the Ct Value, and the mean quantity value and the confirmation the detection by gel electrophoresis and melt curve analysis. The Bold number indicates the phase of collecting eDNA water samples.

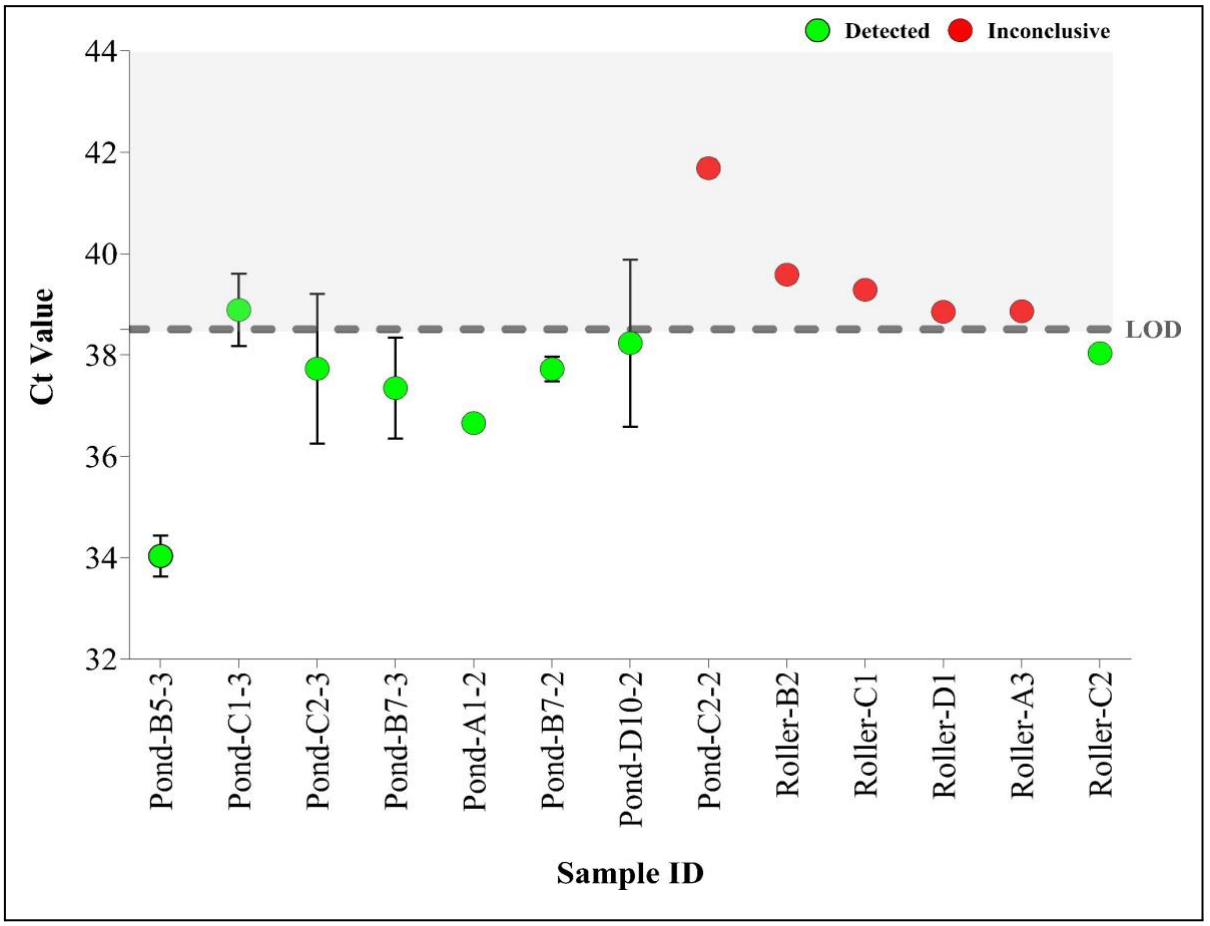

Figure 3.7. The results of in situ validation of environmental DNA (eDNA) to detect the presence of Barred grass snakes from water and roller samples. The y-axis represents the number of cycles (Ct value) at which barred grass snake DNA was detectable, with lower values meaning there is more DNA. The sample ID (mean ± SD) (three technical replicates) is shown on the x-axis. Green dots indicate detected eDNA, and a single red dot represents an inconclusive detection. The dashed gray line indicates the limit of detection.

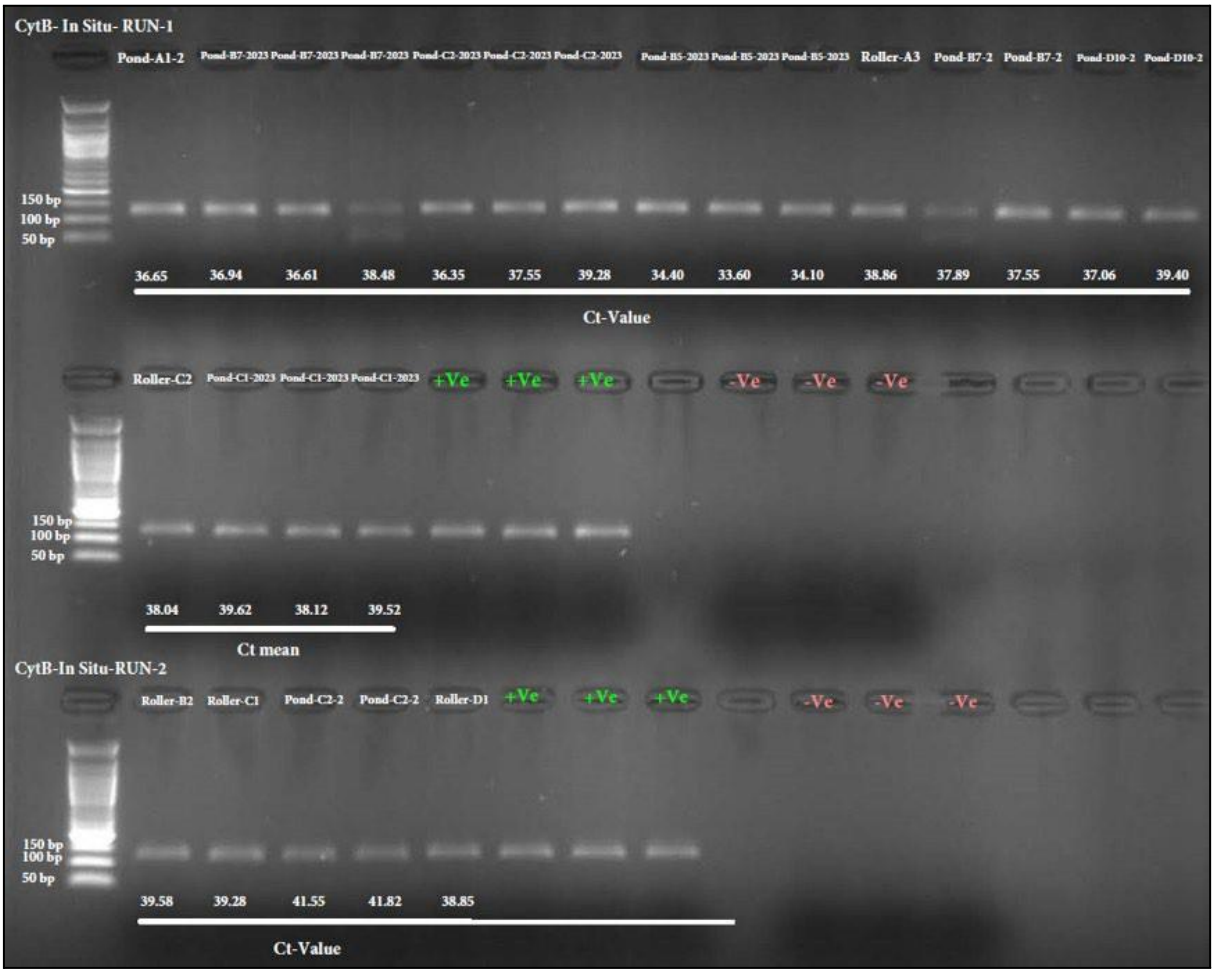

Figure 3.8. The results of qPCR amplification of eDNA samples for detecting Barred grass snakes were visualised using gel electrophoresis across two runs (RUN-1 and RUN-2). The gel includes bands for various samples alongside positive (+Ve) and negative (-Ve) controls. Each lane corresponds to a different sample ID, including ponds and roller samples. Threshold cycle (Ct) values are shown for each sample lane, indicating the PCR cycle at which fluorescence exceeds the threshold.

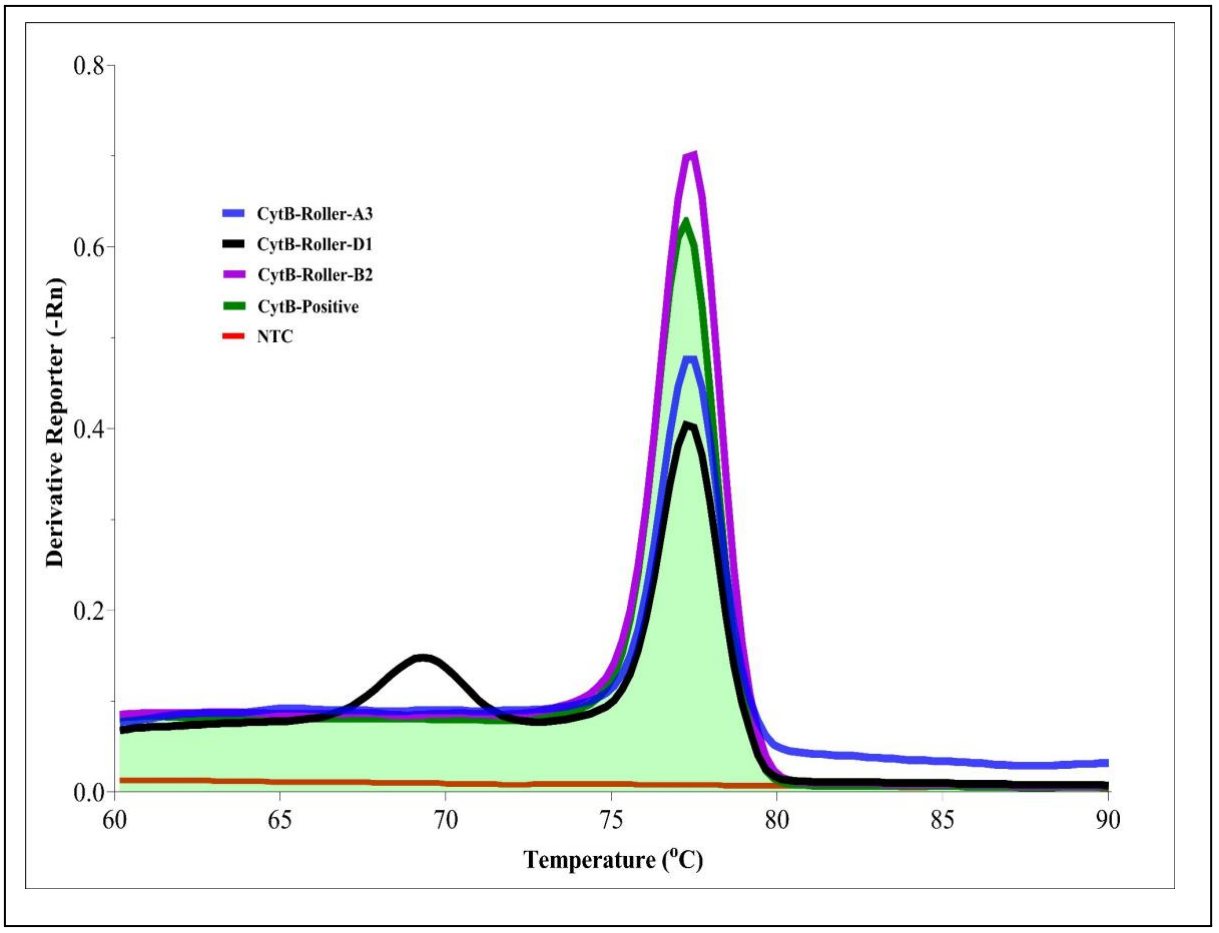

Figure 3.9. The melt curve analysis graph illustrates the derivative reporter (Rn) against the melting temperature (°C) for selected eDNA samples used to confirm the presence of barred grass snake DNA. The green shaded area is positive control. Only four eDNA roller samples that had inconclusive detection were analysed. Three roller samples, CytB-Roller-B2, CytB-Roller-B2, and CytB-Roller-D1, were analysed.

Figure 3.10. The map illustrates the outcomes from the eDNA samples collected during the in-situ validation process. The marked yellow ponds and ACOs indicate the presence of barred grass snakes, *Natrix helvetica*.

4. Discussion

Our study revealed the usefulness of a species-specific qPCR assay for detecting *Natrix helvetica* (barred grass snake) from eDNA samples. During the validation procedure, the *BGS_CytB* assay was highly specific for *Natrix helvetica*, with no cross-amplification observed from non-target species, indicating the assay's specificity to target only *Natrix helvetica* from the eDNA samples. Moreover, the assay was highly sensitive to *Natrix helvetica* at low eDNA concentrations, confirming its usefulness in monitoring this species under field conditions, even when eDNA availability was limited. However, some modifications to the eDNA sample collection methods should be considered to improve

detection rates, particularly for terrestrial environments. Optimising the roller swab technique in eDNA surveys could make it a more effective tool for monitoring *Natrix helvetica* and other semi-aquatic and terrestrial reptiles, providing a valuable tool for conservation and management efforts.

During the in-situ validation procedure, the detection rates for roller swab samples were lower than for water samples, with only one out of 13 samples yielding a positive detection (8%) and four out of 13 samples showing inconclusive results (31%), even though we confirmed the presence of *Natrix helvetica* underneath the ACOs by the CMR survey in the field site (see Chapter Two). Furthermore, the inconclusive detections indicate that the roller swab samples were below the limit of detection (0.3 copies/µL – Ct value of 38.51), resulting in a reduced confidence level for detecting the target sequence when it is present at such low concentrations (Klymus et al., 2020). Thus, the inconclusive detections from the roller swab samples were still meaningful and required further confirmation. Therefore, we conducted a melt curve analysis for the inconclusive detections, and 3 out of the 4 roller samples produced a melt curve at 77.37°C, which was consistent with the positive control at 77.25°C (Katz et al., 2020) (Figure 3.9).

To our knowledge, two studies have applied the concept of swabbing the entire underside of the ACO to collect the eDNA; both reported success in detecting reptile species (Matthias et al., 2021; Kyle et al., 2022). In contrast, we attempted to swab the entire ACO and the soil surface that was in contact with the ACO. However, several factors related to the method may explain the lower detection rates observed in our study using this method. For example, the material and structure of the ACO may influence the amount of eDNA recovered. Kyle et al. (2022) used ACOs made of corrugated metal and wood to aggregate the eDNA of *Scincella lateralis* (brown skink) via a roller method. Their study showed a

higher positive detection rate with metal cover boards (76% – 138 of 181) compared to wood (pressure-treated plywood) (51% – 92 of 181), possibly because metal may offer a smoother and less porous surface for collecting the eDNA (Alketbi & Goodwin, 2019; Kuffel et al.,2024). By contrast, Matthias et al. (2021) used asphalt roofing-shingle ACOs to detect *Contia tenuis* (sharp-tailed snake), swabbed with a finger cot instead of a roller, with a positive detection rate of 57% (13 of 23). Despite the differences in ACO material, both studies showed successful eDNA recovery, likely due to their use of smooth surfaces that allow for better aggregation of surface-deposited DNA. The ACOs used by Matthias et al. (2021) were flat in shape, which may have facilitated easier direct contact between the sharp-tailed snake and the surface, potentially enhancing eDNA deposition and recovery. In contrast, we used ACOs made of corrugated roofing materials composed of bitumen-soaked organic fibers (waterproof). The corrugated shape of the ACO may have created less contact between the snakes and the surface of the ACO.

The fibers might have also trapped eDNA within the material, particularly underneath the ACO, making it difficult to collect during swabbing and potentially reducing the overall efficiency of eDNA recovery (Harrison et al., 2019; Arsenault et al., 2024). Despite the lower success rate in our study, the application of roller swabbing on ACOs holds promise for detecting terrestrial species in other contexts. For example, Valentin et al. (2020) successfully used this technique to detect invasive insects, showing that it can be adapted to a variety of taxes. However, the variability in detection rates both in previous studies and our own suggests that further optimisation is needed. We suggest that using an ACO that is flat in shape with a non-porous surface could enhance eDNA recovery and improve detection rates. However, further investigations are needed to validate this concept (Matthias et al., 2021; Kyle et al., 2022).

The soil in contact with the ACO, along with microbial activity, may speed up the degradation of eDNA. Even though the ACO provided some protection from UV light, the moist soil conditions at the time of roller swab sample collection could have contributed to faster eDNA breakdown (Mauvisseau et al., 2022). Moisture can encourage microbial activity, which may further degrade any eDNA left on the surface, making it more difficult to detect (Harrison et al., 2019; Katz et al., 2020; Kyle et al., 2022). Moreover, moist soil could have contributed to inhibitors, such as organic matter and humic acids, which may have interfered with eDNA detection and prevented the amplification of any snake DNA, potentially leading to false negatives. As a result, using an exogenous Internal Positive Control (IPC) may minimise the risk of false negatives, even in the presence of potential inhibitors (Klymus et al., 2020; Katz et al., 2020; Matthias et al., 2021)

During the Capture-Mark-Recapture (CMR) survey, more than 65% of captured *Natrix helvetica* were juveniles or subadults with an average Snout-to-Vent Length (SVL) of 40.21 cm. Juvenile and subadult snakes, being smaller, may have less surface area in contact with ACOs, resulting in lower eDNA deposition compared to larger adults. This size factor could explain the lower detection rates observed in swab samples, as smaller individuals may deposit less detectable eDNA over a smaller area (Ratsch et al., 2020).

The biological behaviour of the target organism can influence detection rates when collecting eDNA samples. For example, *Contia tenuis* (sharp-tailed snake) is a sedentary, semi-fossorial species that uses underground habitats and is typically found concealed under artificial cover objects (ACOs) (Wilkinson et al., 2007; COSEWIC, 2009; Matthias et al. 2021). Similarly, *Scincella lateralis* (little brown skink) lives within the leaf litter layer on the forest floor (Kyle et al., 2022). These species spend more time under ACOs, which may increase the chance of eDNA deposition, improving detection rates when using the roller

swab method. In contrast, barred grass snakes (*Natrix helvetica*) are semi-aquatic and more mobile, which suggests that the higher detection rate in water samples compared to swab samples indicates that aquatic eDNA samples (e.g., water samples) may be more efficient at detecting *Natrix helvetica* than are terrestrial eDNA samples. (Matthias et al; Kyle et al., 2022)

During the first phase of water sample collection, none of the 25 samples tested positive for *Natrix helvetica* DNA. However, these samples were collected in October, the inactive season of barred grass snakes. This result suggested that eDNA sampling should be conducted during the species' active season to increase the chance of detection. Furthermore, the absence of detections, combined with no amplification of contamination controls and validation of the assay *in silico* and *in vitro*, indicates a minimal or non-existent risk of false positives in the study. By contrast, during the second phase, when water samples were collected near artificial cover objects (ACOs) where grass snakes had been observed, the detection success rate increased to 60% (3 out of 5 samples). In the third phase, when barred grass snakes were seen swimming in ponds, the detection success rate rose to 80% (4 out of 5 samples). These findings highlight the importance of collecting samples during the summer of *Natrix helvetica* activity to improve detection success.

The efficiency, specificity, and sensitivity of the species-specific qPCR assay are key factors determining the success of an eDNA survey (Wilcox et al., 2013; Klymus et al., 2020). Based on the qPCR assay validation scale (Thalinger et al., 2021), we validated the *Cytb-BGS* assay at both level 3 (essential) and level 4 (substantial). This included: (1) in silico and in vitro validation using multiple individuals of *Natrix helvetica* (accounting for geographic variation), all native reptile species in the UK, closely related species within the *Natrix* genus, and co-occurring species; (2) determination of the limit of blank (LOB), limit

of detection (LOD), and limit of quantification (LOQ); and (3) in situ validation using different eDNA samples (aquatic and terrestrial samples). The validation results suggested that the *Cytb-BGS* assay was efficient and specific to detecting *Natrix helvetica* using eDNA samples (aquatic environments) at low concentrations.

While our study successfully demonstrated the potential of eDNA for detecting *Natrix helvetica*, several areas for future research should be considered: (1) the optimisation of terrestrial sampling methods, such as roller swabbing, could enhance the detection of semi-aquatic reptiles in terrestrial environments: (2) further experimental studies are needed to evaluate the rates of eDNA deposition and degradation for *Natrix helvetica* on the soil surface (Kucherenko et al. 2018) and under ACOs (Matthias et al; Kyle et al., 2022). Moreover, expanding the use of eDNA to other reptile species, particularly those that are more terrestrial or cryptic, could provide valuable insights into their population dynamics and habitat use. Applications to species such as *Coronella austriaca* (smooth snake) or *Natrix tessellata* (dice snake) could further enhance the utility of eDNA for reptile conservation.

5. References

Adams, C. I. M., Hoekstra, L. A., Muell, M. R., & Janzen, F. J. (2019). A brief review of non-avian reptile environmental DNA (eDNA), with a case study of painted turtle (Chrysemys picta) eDNA under field conditions. Diversity, 11(4), 50. https://doi.org/10.3390/d11040050

Allen, M. C., Kwait, R., Vastano, A., Kisurin, A., Zoccolo, I., Jaffe, B. D., Angle, J. C., Maslo, B., & Lockwood, J. L. (2023). Sampling environmental DNA from trees and soil to detect cryptic arboreal mammals. Scientific Reports, 13(180), 1-11. https://doi.org/10.1038/s41598-023-27512-8

Alketbi, S. K., & Goodwin, W. (2019). The effect of surface type, collection and extraction methods on touch DNA. Forensic Science International: Genetics Supplement Series, 7, 704–706. https://doi.org/10.1016/j.fsigss.2019.10.145

Arsenault, H., Kuffel, A., Nic Daeid, N., & Gray, A. (2024). Trace DNA and its persistence on various surfaces: A long-term study investigating the influence of surface type and environmental conditions – Part one, metals. Forensic Science International: Genetics, 70, 103011. https://doi.org/10.1016/j.fsigen.2024.103011

Atkinson, S., Carlsson, J. E. L., Ball, B., Egan, D., Kelly-Quinn, M., Whelan, K., & Carlsson, J. (2018). A quantitative PCR-based environmental DNA assay for detecting Atlantic salmon (Salmo salar L.). Aquatic Conservation: Marine and Freshwater Ecosystems, 28(5), 1238–1243.https://doi.org/10.1002/aqc.2931

Böhm, M., Collen, B., Baillie, J. E. M., Bowles, P., Chanson, J., Cox, N., Hammerson, G., Hoffmann, M., Livingstone, S. R., Ram, M., Rhodin, A. G. J., Stuart, S. N., van Dijk, P. P., Young, B. E., Afuang, L. E., Aghasyan, A., García, A., Aguilar, C., Ajtic, R., ... Zug, G. (2013). The conservation status of the world's reptiles. Biological Conservation, 157, 372-385. https://doi.org/10.1016/j.biocon.2012.07.015

Bowyer, R. T., Boyce, M. S., Goheen, J. R., & Rachlow, J. L. (2019). Conservation of the world's mammals: Status, protected areas, community efforts, and hunting. Journal of Mammalogy, 100(3), 923-941. https://doi.org/10.1093/jmammal/gyy180

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

Cordier, J. M., Aguilar, R., Lescano, J. N., Leynaud, G. C., Bonino, A., Miloch, D., Loyola, R., & Nori, J. (2021). A global assessment of amphibian and reptile responses to land-use changes. Biological Conservation, 253, 108863. https://doi.org/10.1016/j.biocon.2020.108863

Cox, N., Young, B. E., Bowles, P., Fernandez, M., Marin, J., Rapacciuolo, G., Böhm, M., Brooks, T. M., Hedges, S. B., Hilton-Taylor, C., Hoffmann, M., Jenkins, R. K. B., Tognelli, M. F., Alexander, G. J., Allison, A., Ananjeva, N. B., Auliya, M., Avila, L. J.,

Chapple, D. G., Cisneros-Heredia, D. F., Cogger, H. G., Colli, G. R., de Silva, A., Eisemberg, C. C., Els, J., Fong, G. A., Grant, T. D., Hitchmough, R. A., Iskandar, D. T., Kidera, N., Martins, M., Meiri, S., Mitchell, N. J., Molur, S., Nogueira, C. C., Ortiz, J. C., Penner, J., Rhodin, A. G. J., Rivas, G. A., Rödel, M. O., Roll, U., Sanders, K. L., Santos-Barrera, G., Shea, G. M., Spawls, S., Stuart, B. L., Tolley, K. A., Trape, J. F., Vidal, M. A., Wagner, P., Wallace, B. P., & Xie, Y. (2022). A global reptile assessment highlights shared conservation needs of tetrapods. Nature, 605(7909), 285–290. https://doi.org/10.1038/s41586-022-04664-7

Dwight, Z., Palais, R., & Wittwer, C. T. (2011). uMELT: Prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application. Bioinformatics, 27(7), 1019–1020. https://doi.org/10.1093/bioinformatics/btr065

Farooq, H., Harfoot, M., Rahbek, C., & Geldmann, J. (2024). Threats to reptiles at global and regional scales. Current Biology, 34(10), 2231-2237.e2. https://doi.org/10.1016/j.cub.2024.04.007

Fediajevaite, J., Priestley, V., Arnold, R., & Savolainen, V. (2021). Meta‐analysis shows that environmental DNA outperforms traditional surveys, but warrants better reporting standards. Ecology and Evolution, 11(8), 4803-4815. https://doi.org/10.1002/ece3.7382

Ficetola, G. F., Manenti, R., & Taberlet, P. (2019). Environmental DNA and metabarcoding for the study of amphibians and reptiles: Species distribution, the microbiome, and much more. Amphibia-Reptilia, 40(2), 129–148. https://doi.org/10.1163/15685381-20191194

Foster, J., Driver, D., Ward, R., & Wilkinson, J. (2021). IUCN Red List assessment of amphibians and reptiles at Great Britain and country scale. Report to Natural England. ARC report. ARC, Bournemouth.

Geneious. (2023). Geneious sequence alignment software (version 2023.1). Biomatters Ltd. https://www.geneious.com

Guicking, D., Joger, U., & Wink, M. (2009). Cryptic diversity in a Eurasian water snake (Natrix tessellata, Serpentes: Colubridae): Evidence from mitochondrial sequence data and nuclear ISSR-PCR fingerprinting. Organisms, Diversity & Evolution, 9(3), 201–214

Gumbs, R., Scott, O., Bates, R., Böhm, M., Forest, F., Gray, C. L., ... Rosindell, J. (2024). Global conservation status of the jawed vertebrate Tree of Life. Nature Communications, 15(1), 1101. https://doi.org/10.1038/s41467-024-45119-z

Harfoot, M. B. J., Johnston, A., Balmford, A., Burgess, N. D., Butchart, S. H. M., Dias, M. P., ... Geldmann, J. (2021). Using the IUCN Red List to map threats to terrestrial vertebrates at global scale. Nature Ecology & Evolution, 5(11), 1510-1519. https://doi.org/10.1038/s41559-021-01542-9

Halstead, B. J., Wood, D. A., Bowen, L., Waters, S. C., Vandergast, A. G., Ersan, J. S. M., Skalos, S. M., & Casazza, M. L. (2017). An evaluation of the efficacy of using environmental DNA (eDNA) to detect giant gartersnakes (Thamnophis gigas). U.S.

Geological Survey Open-File Report, 2017–1123, 41 p. https://doi.org/10.3133/ofr20171123

Harrison, J. B., Sunday, J. M., & Rogers, S. M. (2019). Predicting the fate of eDNA in the environment and implications for studying biodiversity. Proceedings of the Royal Society B: Biological Sciences, 286(1914), 20191409. https://doi.org/10.1098/rspb.2019.1409

Hunter, M. E., Oyler-McCance, S. J., Dorazio, R. M., Fike, J. A., Smith, B. J., Hunter, C. T., Reed, R. N., & Hart, K. M. (2015). Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese pythons. PLOS ONE, 10(4), e0121655. https://doi.org/10.1371/journal.pone.0121655

Katz, A. D., Harper, L. R., Sternhagen, E. C., Pearce, S. E., Melder, C. A., Sperry, J. H., & Davis, M. A. (2020). Environmental DNA is effective in detecting the federally threatened Louisiana Pinesnake (Pituophis ruthveni). Environmental DNA. https://doi.org/10.1002/edn3.126

Kindler, C., & Fritz, U. (2018). Phylogeography and taxonomy of the barred grass snake (Natrix helvetica), with a discussion of the subspecies category in zoology. Vertebrate Zoology, 68(3), 269-281 https://doi.org/10.3897/vz.68.e31615

Kuffel, A., Gray, A., & Nic Daeid, N. (2024). Impact of metal ions on PCR inhibition and RT-PCR efficiency. International Journal of Legal Medicine, 135(1), 63–72. https://doi.org/10.1007/s00414-020-02363-4

Kyle, K. E., Allen, M. C., Dragon, J., Bunnell, J. F., Reinert, H. K., Zappalorti, R., Jaffe, B. D., Angle, J. C., & Lockwood, J. L. (2022). Combining surface and soil environmental DNA with artificial cover objects to improve terrestrial reptile survey detection. Conservation Biology, 36(6), e13939. https://doi.org/10.1111/cobi.13939

Klymus, K. E., Merkes, C. M., Allison, M. J., Goldberg, C. S., Helbing, C. C., Hunter, M. E., Lance, R. F., Mangan, A. M., Monroe, E. M., Piaggio, A. J., Stokdyk, J. P., & Wilson, C. C. (2019). Reporting the limits of detection and quantification for environmental DNA assays. Environmental DNA, 2(3), 271–282. https://doi.org/10.1002/edn3.29

Klymus, K. E., Ruiz Ramos, D. V., Thompson, N. L., & Richter, C. A. (2020). Development and testing of species-specific quantitative PCR assays for environmental DNA applications. Journal of Visualized Experiments (JoVE), 165, e61825. https://doi.org/10.3791/61825

Kucherenko, A., Herman, J. E., Everham, E. M., & Urakawa, H. (2018). Terrestrial snake environmental DNA accumulation and degradation dynamics and its environmental application. Herpetologica, 74(1), 38–49. http://www.jstor.org/stable/26428814

Kronenberger, J. A., Wilcox, T. M., Mason, D. H., Franklin, T. W., McKelvey, K. S., Young, M. K., & Schwartz, M. K. (2022). eDNAssay: A machine learning tool that accurately predicts qPCR cross-amplification. Molecular Ecology Resources, 22(8), 2994–3005. https://doi.org/10.1111/1755-0998.13681

Laura Matthias, M. J. Allison, C. Y. Maslovat, J. Hobbs, & C. C. Helbing. (2021). Improving ecological surveys for the detection of cryptic, fossorial snakes using eDNA on and under artificial cover objects. Ecological Indicators, 131, 108187. https://doi.org/10.1016/j.ecolind.2021.108187

Lesperance, M. L., Allison, M. J., Bergman, L. C., Hocking, M. D., & Helbing, C. C. (2021). A statistical model for calibration and computation of detection and quantification limits for low copy number environmental DNA samples. Environmental DNA, 3(5), 970–981. https://doi.org/10.1002/edn3.220

Langlois, V. S., Allison, M. J., Bergman, L. C., To, T. A., & Helbing, C. C. (2021). The need for robust qPCR-based eDNA detection assays in environmental monitoring and species inventories. Environmental DNA, 3(3), 519–527. https://doi.org/10.1002/edn3.164

Mauvisseau, Q., Harper, L. R., Sander, M., Hanner, R. H., Kleyer, H., & Deiner, K. (2022). The multiple states of environmental DNA and what is known about their persistence in aquatic environments. Environmental Science & Technology, 56(9), 5322–5333. https://doi.org/10.1021/acs.est.1c07638

Moyer, G. R., Bartron, M. L., Galbraith, H. S., Grassi, J., & Rees, C. B. (2023). Development and validation of two environmental DNA assays for American Eel (*Anguilla rostrata*). Environmental DNA, 5(1), 175–190. https://doi.org/10.1002/edn3.369

Nordstrom, B., Budd, A., Mitchell, N., Cornish, C., Byrne, M., Kuchling, G., & Jarman, S. (2024). Environmental DNA reflects spatial distribution of a rare turtle in a lentic wetland assisted colonization site. Environmental DNA, 6, e507. https://doi.org/10.1002/edn3.507

Pawlowski, J., Apothéloz‐Perret‐Gentil, L., & Altermatt, F. (2020). Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. Molecular Ecology, 29(22), 4258-4264. https://doi.org/10.1111/mec.15643

Piaggio, A. J., Engeman, R. M., Hopken, M. W., Keacher, K. L., & Goldberg, C. S. (2014). Detecting an elusive invasive species: A diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. Molecular Ecology Resources, 14(2), 374-380. https://doi.org/10.1111/1755-0998.12180

Prabhakaran, S., Arjuna, A., Sankar, G., Prasanth, K., Kaliyaperumal, R., & Suresh, V. R. (2023). Development of a species-specific qPCR assay for the detection of invasive African Clarias gariepinus. Biological Invasions, 25(5), 975–982.

Ratsch, R., Kingsbury, B. A., & Jordan, M. A. (2020). Exploration of environmental DNA (eDNA) to detect Kirtland's snake (Clonophis kirtlandii). Animals, 10(6), 1057. https://doi.org/10.3390/ani10061057

Rose, J. P., Wademan, C., Weir, S., Wood, J. S., & Todd, B. D. (2018). Traditional trapping methods outperform eDNA sampling for introduced semi-aquatic snakes. PLOS ONE, 14(7), e0219244. https://doi.org/10.1371/journal.pone.0219244

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation, 17, e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Sahu, A., Kumar, N., Singh, C. P., & Singh, M. (2023). Environmental DNA: Powerful technique for biodiversity conservation. Journal for Nature Conservation, 71, 126325. https://doi.org/10.1016/j.jnc.2022.126325

Stothard, P. (2000). The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques, 28(6), 1102-1104.

Thalinger, B., Deiner, K., Harper, L. R., Rees, H. C., Blackman, R. C., Sint, D., Traugott, M., & Goldberg, C. S. (2021). A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. Environmental DNA, 3(4), 823–836. https://doi.org/10.1002/edn3.189

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation, 183, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Taylor, S. C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M., & Fenrich, J. (2019). The ultimate qPCR experiment: Producing publication quality, reproducible data the first time. Trends in Biotechnology, 37(7), 761-764. https://doi.org/10.1016/j.tibtech.2018.12.002

Valentin, R. E., Fonseca, D. M., Gable, S., Kyle, K. E., Hamilton, G. C., Nielsen, A. L., & Lockwood, J. L. (2020). Moving eDNA surveys onto land: Strategies for active eDNA aggregation to detect invasive forest insects. Molecular Ecology Resources, 20(3), 746–755. https://doi.org/10.1111/1755-0998.13151

Wilcox, T. M., McKelvey, K. S., Young, M. K., Jane, S. F., Lowe, W. H., Whiteley, A. R., & Schwartz, M. K. (2013). Robust detection of rare species using environmental DNA: The importance of primer specificity. PLOS ONE, 8(3), e59520. https://doi.org/10.1371/journal.pone.0059520

Woodland Trust. (2021). State of the UK's woods and trees 2021. Woodland Trust. https://www.woodlandtrust.org.uk/media/51705/state-of-the-uks-woods-and-trees-2021-thewoodlandtrust.pdf

Wilkinson, K. E. (2009). COSEWIC assessment and status report on the Sharp-tailed Snake (Contia tenuis) in Canada. Committee on the Status of Endangered Wildlife in Canada, Ottawa. vii + 38 pp.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics, 13, 13

CHAPTER FOUR

DIETARY ASSESSMENT OF BARRED GRASS SNAKES USING METABARCODING APPROACH

Table of Contents

1. Introduction

Continued biodiversity loss requires a deeper understanding of how species react to disturbances and environmental changes, particularly food supply. Characterising the diets of endangered species has become a focus of ecological research to enhance conservation methods (Diaz-Abad et al., 2021; Ruppert et al., 2019). Examining stomach contents and faeces using microscopy for the remains of prey, such as bones, hair, and scales, has been the usual method for assessing diets. Although these strategies can provide an initial step in diet characterisation, they are time-consuming and require specialist taxonomic knowledge (Kurita & Toda, 2022). In addition, conventional methods are susceptible to significant biases, such as the failure to detect soft-bodied or easily digested prey due to the absence of identifiable remnants (Heo et al., 2022). Prey items with identifiable remains in faeces are typically limited to broad taxonomic identification, such as the genus or family level, whereas classification at the species level is uncommon (Homma et al., 2022). Diet studies are typically constrained by the relatively invasive technique of collecting stomach and gut contents, which requires stomach pumping or euthanasia to access and dissect the GI tract. These invasive techniques pose ethical difficulties if applied to wild predators and are, therefore, not a viable choice for endangered predators.

High-throughput sequencing enables the automated identification of multiple species from a single sample, such as an environmental sample containing the eDNA of multiple organisms. This technique is known as DNA metabarcoding, in which universal primers amplify all the species within a single taxonomic group (Alberdi et al., 2019). Using next-generation sequencing, the amplified DNA is sequenced, and the recovered sequences are then matched to a reference database to generate a list of taxonomic units. By extracting DNA from faeces, DNA metabarcoding has recently been widely employed for diet studies. Research has already been conducted on several animal groups, including mammals (Harper et al., 2020), birds (Mitchell et al., 2022), and fish (Tsuji et al., 2020). They allowed for the description of a diet with unprecedented resolution, revealing unexpected habits and

dietary shifts. Certain groups, including reptiles, are under-represented in DNA meta-barcoding studies of diet, which have only been conducted on a few lizards, turtle and snake species. Nonetheless, most of these studies either employ group-specific primers (Brown et al., 2014) or describe the diet of insectivorous predators (Kartzinel & Pringle, 2015). Group-specific primers exclusively target specific taxonomic groupings of prey, which can prevent the discovery of complete diets and the ecological significance of prey.

This study seeks to determine the nutritional content of the barred grass snake, *Natrix helvetica*, inhabiting Western Europe. Recent DNA analyses have indicated that *N. helvetica* is taxonomically distinct from those in Eastern and Southern Europe (Kindler & Fritz, 2018). *N. helvetica* has been protected in the United Kingdom since 1981 Wildlife and Countryside Act was passed. The barred grass snake was identified as a priority species by the UK Post-2010 Biodiversity Framework 2010 (The Wildlife Trusts, n.d.).

In the past, the remnants of common frogs (*Rana temporaria*), common toads (*Bufo bufo*), and great crested newts *(Triturus cristatus*) have been discovered in the stomachs of grass snakes, providing evidence that amphibians comprise a significant component of their diet (Gregory & Isaac, 2004). However, research has also indicated that reptiles are a significant part of the grass snake's diet (Brown et al., 2014). The diversity is explained by localised variation in diet based on prey availability. The diet of the closely related *Natrix natrix* is also amphibian-dominated. Nevertheless, it varies depending on the availability of different prey species (Gleed-Owen, 1994; Gleed-Owen, 1996). For example, amphibians are not always available throughout a portion of the snake's active phase, mainly after adults have scattered from breeding ponds (Kindler *et al*., 2017). As a result, snakes would be expected to switch to alternative prey sources (Filippi *et al*., 1996; Filippi & Luiselli, 2002). There is direct proof that grass snakes, in general, feed on small mammals and amphibians, as these creatures have been observed regurgitating meals (Consul et al., 2009). There is no direct evidence that snakes prey on avifauna at this time. Moreover, as predators, grass snakes are gape-restricted, meaning they can only consume anything they can ingest whole. Therefore, the

food of hatchlings and juveniles is likely different from that of adults (Gregory & Isaac, 2004). Using a minimally invasive method, the project aims to characterise the barred grass snake's preferred prey source at two sites. Dietary data are limited, with diet analyses having been carried out in York, England (Gleed-Owen, 1994), the Doncaster area of South Yorkshire, England (Gleed-Owen., 1996), the River Stour near Fordwich (Gregory & Isaac, 2004) and South Dorset (Reading & Davies, 1996). With the exception of the Brown et al. (2014) investigation, this study will be the first to use molecular analysis, particularly a meta-barcoding method, to characterise the barred grass snake diet and one of few to apply this method to a snake species. DNA meta-barcoding approaches have yet to be commonly implemented in reptile diet studies. Reptiles are currently facing global declines (Farooq et al., 2024); therefore, we intend that the methods developed in this study should be applicable to other snake species when traditional approaches are not possible.

In this study, we hypothesised that (1) the bulk of the barred grass snake diet will consist of amphibian species, with the occasional occurrence of arthropod, bird, and small mammal prey, (2) amphibians would dominate the diet of both populations, with possible variations driven by local prey availability, (3) subadult grass snakes would exhibit an ontogenetic shift, consuming smaller prey while adults would target larger prey, and (4) females would have a broader and more varied diet compared to males, driven by the reproduction needs.

The objectives of the project are to (1) utilise a metabarcoding approach to identify the prey items of barred grass snakes that are mainly consumed in two different populations: the Watermill Broad Nature Reserve, Cranwich, Norfolk, and Holbrook Marshal ponds, Sheffield; (2) identify if there are differences in diet between age classes, sexes, and populations among individual.

2. Material and Methods

2.1 Faecal Sample Collection and General Measurement

Sixty-eight faecal samples of barred grass snakes were collected from both sites as follows: (1) Fifty-four samples were from Watermill Broad Nature Reserve, Cranwich, Norfolk (2) fourteen samples from Holbrook Marshal Ponds, Sheffield, England, during the CMR surveys for both sites. The Watermill Broad (52°31′51″N, 000°37′02″E) is a large nature reserve in Norfolk, covering 10 hectares within a 130-hectare forestry area. The site was used for agriculture and it has been restored into a diverse of wetlands, grasslands, and ponds, creating an ideal habitat for amphibians, particularly frogs and toads. In contrast, Holbrook Marsh Ponds (53° 19 '46 "N, 1° 19' 53 "W) is a smaller wetland area in Sheffield. primarily designed to support the growth of great crested newts. While this site can provide some suitable habitat for barred grass snakes, it is less connected and more fragmented compared to Watermill Broad (For more details, see Chapter 2).

All samples from the Norfolk population were provided by Dr Steven Allain. To obtain reliable dietary data, the faeces were obtained directly from the snake by swabbing the surface of the cloacal opening into a 2-ml microcentrifuge tube. As an anti-predatory behaviour, barred grass snakes usually void their faecal materials, facilitating collection. Faecal samples were either (1) preserved in 70% ethanol and frozen at -20°C (Norfolk population) or immediately frozen using dry ice then transferred to a -80 to avoid DNA degradation (Holbrook population). The SVL (snout-vent length) for age class category and mass for each snake were measured using the protocol described in chapter two, and the sex was recorded. However, due to the low capture rate of juveniles and subadults, I categorised the age classes into adult (SVL greater than 40 cm) and subadult (SVL less than 40 cm) to reduce the variation within the age groups.

2.2 DNA Extraction and Quantification of Faecal Samples

QIAamp Fast DNA Stool Mini Kit (QIAGEN, Hilden, Germany) was used to extract the DNA from faecal samples. This kit has often been used in diet studies using the faecal metabarcoding approach (Liu et al., 2021). The DNAs of sixty-eight faecal samples were extracted following the standard protocol with some modifications to increase the DNA yield from the faecal samples because it was expected that the DNA quality would be deficient compared to tissue samples. The modifications, following Mitchell et al. (2022), involved increasing the sample size to 300 mg (from the standard 220 mg), using 1.4 mL of inhibitex buffer instead of the recommended 1 mL, and extending the incubation time by 10 minutes after adding the lysis buffer to improve cell lysis and boost DNA yield. If the faecal sample exceeded 300 mg, it was divided into subsets, then analysed together.

The DNAs from faecal samples were extracted in a separate laboratory room containing a laminar flow hood with UV sterilisation that was reserved for eDNA extraction to avoid any potential contamination. Two DNA extraction negative controls (NTC) containing only the reagents and purified water were included in each extraction to test for any contamination. The concentration of all eDNA extraction samples was measured using the Qubit 4 Fluorometer dsDNA HS standard. An agarose gel electrophoresis of 0.8% was also performed to assess the quality of the eDNA extraction samples. The range of DNA concentrations was between 0.11–29.2 ng/μL (mean 4.06 ng/μL). No contamination was detected in the DNA extraction procedure (negatives out of range–low).

2.3 The Primer Set Selection Criteria and In Vitro Validation

Selecting appropriate primer sets for metabarcoding to identify the biodiversity of the diet using faecal samples is crucial for obtaining accurate and comprehensive dietary data. The choice of the primer set depends on several factors, including the target prey taxa, the taxonomic resolution and the quality and quantity of the DNA. Due to the degradation of prey DNA in faecal samples, taxon-specific primers targeting a highly conserved short fragment of DNA within a taxon were considered because they are essential for successful amplification in the case of using highly degraded DNA surviving the digestive system, as well as to provide higher resolution classification (e.g., genus or species level) (Ando et al., 2020; Mondion et al. 2022). In addition, using a multi-marker strategy in which different primers amplify the same taxon in different fragments of DNA will provide more comprehensive coverage, improving both the detection of diverse species and the resolution of taxonomic identification across various levels (Sarki et al., 2022; Wang et al., 2023). A list of prey items was created from previous investigations to assist in selecting appropriate primers for metabarcoding analysis, ensuring that all the potential prey were effectively detected (DiNicolae & Zabbia, 2021; Brown et al., 2014; Gregory & Isaac, 2004).

The collection of candidate primer sets assessed was as follows: MiAmphS (150–200 bp,16S rRNA gene, Sakata et al. 2021), Amphi16S-Mod (250–300 bp 16S rRNA gene, Sakata et al. 2021), Vert16S-eDNA (200–250 bp,16S rRNA gene, Vences et al. 2016), MiMammal (171–210 bp, 12S rRNA gene, Ushio et al. 2017), Amphi_B (111 bp, Cytb gene Lacoursiere-Roussel et al. 2016), Riaz-12S-V5 (106–145 bp, 12S rRNA gene, Riaz et al, 2011) MiFish (12S rRNA gene, 163–218 bp, Miya et al. 2015), and L2513/H2714 (202–247 bp, 16S rRNA gene, Kitano et al., 2007).

2.3.1 In Vitro Validation

The candidate primer sets that described above were evaluated by using available genomic DNA from the potential prey taxa, including British amphibians (Bufo bufo, Rana temporaria, and Triturus cristatus), small mammals (Myodes glareolus, Neomys fodiens, Sorex minutus, Sorex araneus, Arvicola amphibious and Mus musculus). The gDNA of the potential prey was extracted using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's protocol. Moreover, six faecal samples that had the lowest DNA concentration (0.11–0.20 ng/μL) were included in the evaluation. Also, barred grass snake DNA (0.1 ng/ul) was included in the validation to see if the host DNA would be amplified, which might affect the amplification of prey (Pinol et al., 2023).

The total 20-μL reaction consisted of 10 μL of 1X Qiagen Multiplex PCR Master Mix (including Hot-start Taq DNA polymerase), 1.5 μL of both forward primer (5 μM) and reverse primer (5 μM), so the final primer concentration was 350 nM, 2 μL of ddH2O, and 5 μL of DNA of the potential prey (0.1 ng/μL), the barred grass snake DNA (0.1 ng/μL), and the selected faecal samples. The PCR program was: initial activation step at 95°C for 15 min (to activate the polymerise), followed by 40 cycles of denaturation at 94°C for 30 s, annealing at Tm°C* (depending on the tested primer) for 45 s, extension 72°C for 45s (extension), then final extension at 72°C for 10 min. The amplification was run using an Eppendorf Mastercycler X50 PCR machine. The PCR product sizes were analysed using 2.5% agarose gel electrophoresis and a 50-bp DNA ladder (New England Biolabs). The following three criteria were used to select the best primers: (1) 50% of faecal samples amplified with the correct amplicon size, (2) all the potential prey amplified showing the strong band with no or weak primer-dimer (specificity), (3) absent or weak amplification of the barred grass snake. As a result, Amphi16S-Mod and Riaz-12S-V5 were selected as the best primer sets based on the criteria for metabarcoding faecal samples of barred grass snakes.

The Amphi16S-Mod primer set was developed as a modified version of the Amphi16S primer set (Sakata et al., 2021), specifically designed to amplify the amphibian 16S ribosomal RNA gene. The modification, particularly in the forward primer, while leaving the reverse unchanged, includes degenerate bases to align more closely with the Vert16S-eDNA primer, which is used to detect vertebrates in environmental DNA (eDNA) samples (Vences et al., 2016). Previously, the Amphi16S primer set had not been used for faecal samples, making this the first time it was applied in a dietary assessment using metabarcoding. It was only tested on water samples during validation (Sakata et al., 2021). All potential prey were successfully amplified during the in vitro validation of the Amph16S-Mod primer set, producing clear and strong bands of the expected size during gel electrophoresis. Although the barred grass snake DNA was amplified, the band size differed from the prey species. Moreover, four out of six faecal samples (which had low DNA concentrations) were amplified at the correct size (~300 bp). The Riaz-12S-V5 primer set was the second most effective during vitro validation. This primer set was developed to target vertebrates' 12S ribosomal RNA gene and is considered one of the most well-published primers used in metabarcoding studies. Moreover, the Riaz-12S-V5 primer set has been validated in silico, in vitro, and in Moreover, the Riaz-12S-V5 primer set has been validated in silico, in vitro, and in situ, specifically suitable for detecting British vertebrates (Harper et al., 2018; Valsecchi et al., 2020). Furthermore, the Riaz-12S-V5 primer set has been successfully applied in the diet analysis of various species using faecal samples (Eusemann et al., 2024; Tosa et al., 2023; Ingala et al., 2021). The primer set (IN16STK-Mod) targeting the mitochondrial 16S rRNA gene was used to amplify invertebrate (Arthropods) DNA and modified not to amplify DNA from Squamata, particularly lizards. Moreover, the primer set was used to identify the faecal DNA of lizards and could detect most insects at the family taxonomic level (Pinho et al., 2023; Pinho et al., 2018; Kartzinel & Pringle, 2015).

2.4. Amplicon PCR and Library Preparation for Illumina Sequencing (Illumina, 2013)

2.4.1. Target-Specific Amplification (PCR-1)

The PCR1 amplification involved amplifying the target-specific region of the prey from sixty-

eight faecal samples of barred grass snakes. The three selected primer sets (Amph16S-Mod, Riaz-12S-V5, and IN16STK-Mod) were tailed with partial Illumina adapter sequences at the 5' ends of both the forward and reverse primers, which provided the complementary binding sites needed for the dual-indexed Illumina primers (unique identifier sequences) during PCR2. All three primer sets were sourced from Sigma-Aldrich. The 20-µL PCR reaction consisted of 10 µL of 1X Qiagen Multiplex PCR Master Mix (which includes Hot-start Taq DNA polymerase), 1.5 µL of forward primer (5 µM) and 1.5 µL of reverse primer (5 µM), resulting in a final primer concentration of 350 nM, along with 2 µL of nuclease-free water and 5 µL of faecal sample. As quality control, known positive controls (+ve) were included in three replicates for each primer set: spotted hyena (*Crocuta crocuta*) DNA for Amph16S-Mod and Riaz-12S-V5 primers and field cricket (*Gryllus bimaculatus*) DNA for IN16STK-Mod primers. These control species were chosen because they are not part of the diet components of barred grass snakes. Also, three to five negative controls (-ve), containing only nuclease-free water and the PCR reagents (without DNA template), were included on each PCR plate for every primer set to monitor contamination. The positive and negative controls were randomly distributed across the plate, with unique ID numbers to aid sample tracking and help identify any loading errors, potential contamination or computational errors during the data trimming or metabarcoding analysis. The PCR program was: initial activation step at 95°C for 15 min (to activate the hot start Taq DNA polymerase), followed by 40 cycles of denaturation Tm at 94°C for 30 s, annealing Tm for 45 s at 58°C for Amph16S-Mod, 55°C for Riaz-12S-V5, and 52°C for IN16STK-Mod, extension Tm at 72°C for 45s, then final extension at 72°C for 10 min. The PCR amplification was performed using an Eppendorf Mastercycler X50 PCR machine. The size of amplicons for each primer set (bp) and the strength of the products were analysed using 2.5% agarose gel electrophoresis and a 50 bp DNA ladder (New England Biolabs). Each gel resulting from the first PCR amplification was photographed for reference. After confirming target-specific amplification (PCR1) through gel electrophoresis, bead purification was performed using Promega ProNex beads to remove residual reagents and extraneous DNA

fragments, thereby isolating only the desired target amplicons for downstream processing according to the manufacturer's protocol (ProNex Size-Selective Purification, Promega). The size cutoff for the amplicons of each primer set was determined using the ProNex chemistry ratio: Amph16S-Mod at 1.3X, Riaz 12S-V5 at 1.5X, and IN16STK-Mod at 1.4X.

2.4.2. Dual-Indexing and Adapter Amplification (PCR-2)

In this PCR2 amplification, i7-R/i5-F indexes (unique identifier sequences for each sample) and P7-R/P5-F adapters were added to the target amplicon (PCR-1 of clean product). The 20-µL PCR reaction consisted of 10 µL of Meridian Bioscience MyTaq HS Mix, 0.5 µL of i5-P5/forward (at 10 µM), 0.5 µL i7-P7/reverse primer (at 10 µM), 1 µL of sterile ddH2O, and 8 µL of PCR-1 clean products. The PCR program was: initial activation step at 95°C for 15 min (to activate the polymerise), followed by 12 cycles of denaturation T at 94°C for 10 s, annealing Tm for 45 s at 65°C, extension Tm at 72°C for 30 s, then final extension at 72°C for 5 min. The amplification was run using an Eppendorf Mastercycler X50. To confirm the successful amplification of dual indexes (PCR-2), the Agilent 2200 TapeStation D1000 ScreenTap (Agilent Technologies) was used to check that the PCR-2 (target-specific + Dual-Indexes and Adapters) product increased in size by randomly selecting and comparing the pre and post-PCR-2. Twelve samples were chosen randomly, with 1 µL of the PCR-2 product and 3 µL of Buffer D1000 in each TapeStation tube.

2.4.3. Quantification and Pooling

The PCR-2 products were quantified using a BioTek Synergy LX Fluorometer. DNA concentrations were measured by combining 2 µL of each PCR-2 product with 200 µL of QuantiFluor mix (prepared using 1 µL of 20X TBE, 50 µL QuantiFluor dsDNA dye, and 19 mL purified H2O). The means and ranges of PCR-2 product concentrations were as follows: Amphi16S-Mod primer – 35.48 ng/µL (range: 16.44–66.17 ng/µL), Riaz-12S-V5 primer 30.17 ng/µL (range: 12.17–83.48 ng/µL), and IN16STK-Mod primer – 21.20 ng/µL (range: 9.03–40.49 ng/µL). In the first pooling process (referred to as Pre-library), sixty-eight PCR-2 products for each primer were divided into four pre-libraries, with each pre-library consisting of

17–24 samples, which were normalised to ensure equal DNA concentrations. The Opentrons pipetting robot was used to pool all the PCR-2 products for each primer set, resulting in four pre-libraries for each primer. All the 12 pre-libraries were purified using Promega ProNex beads using the same ratio described above. Then, the TapeStation was used to check for the presence of adapter dimers from the PCR-2. In the case of the primer dimer observed, I repeated the bead clean and used the TapeStation again. The 12 pre-libraries were quantified using qPCR so that the four pre-libraries for each primer were pooled in one final library at 30 nM. To achieve that, serial dilution for each pre-library was made at 100, 1000, and 10000-fold using a sample dilution buffer. So, each independent dilution and six KAPA Library Quantification standards were run on the QuantStudio 12K Flex Real-Time PCR System (Applied Biosystem). The total 10 µL qPCR reaction consisted of 6 µL of SYBR FAST mix + primers, 2 µL of purified H2O, and 2 µL of the kit standard, diluted pre-library, and NTC (only dilution buffer) in three replicates following the qPCR profile: 5 min at 95°C, 35 cycles of the following: 30 s at 95°C and 45 s at 60°C. The amplification efficiency was 93% with R2 = 0.99, and the slope was -3.49. After qPCR amplification, the four pre-libraries for each primer were pooled in equimolar amounts into a single tube (referred to as a library) at a final concentration of 30 nM.

The quality of the pooled libraries was assessed using the TapeStation High SensitivityD1000 (Agilent Technologies) to check for the presence of primer dimers. Primer dimers were detected in the libraries prepared with the Riaz-12SV5 and IN16STK-Mod primers. To remove these dimers, the Blue Pippin DNA Size Selection System (SageScience) was used following the manufacturer's protocol, with size cut-offs set to 200–450 bp for Riaz-12SV5 and 250–450 bp for IN16STK-Mod. The pooled libraries were reanalysed on the TapeStation High Sensitivity D1000 to confirm the successful removal of primer dimers and the selection of the appropriate size range. The TapeStation results showed that Blue Pippin successfully removed the primer dimers, and the libraries were within the desired size ranges. The exact final quantification of the three libraries was determined using qPCR amplification using the same protocol described above. The final concentration of each library was aimed to be between 4–5

nM. Based on the amplicon sizes of the different primer sets used, Amphi16S-Mod (420 bp) was processed separately at 4.5 nM using a 2x 300-bp MiSeq (Illumina) v3 run, and Riaz-12S-V5 (213 bp) with IN16STK-Mod (245 bp) was pooled together at 4 nM using a 2x 250-bp MiSeq (Illumina) v3 run. The libraries were sequenced at the UK NERC Environmental Omics Facility (NEOF), University of Liverpool.

2.5. Bioinformatics Processing and Taxonomic Identification of Sequencing Data

2.5.1. DADA2 Pipeline Procedure

The Illumina raw sequences for the diet composition of barred grass snakes were downloaded from the NEOF sequencing facility at the Centre of Genomic Research (CGR), University of Liverpool (CGR Project ID: SSP201311) and processed using the Bessemer High- Performance Computing (HPC) platform at the University of Sheffield. A DADA2 pipeline (https://github.com/khmaher/HPC_dada2), adapted from that of Callahan et al. (2016) https://benjjneb.github.io/dada2/tutorial_1_8.html) allowed processing of metabarcoding data on the University of Sheffield's Bessemer HPC.

Firstly, reads containing any uncalled or ambiguous bases ("N" characters) were removed using filterAndTrim (Callahan et al., 2016) to ensure high-quality sequences for downstream DADA2 processing, as DADA2 requires clean sequences without N characters for accurate denoising. After the sequences with Ns were removed from the data, Cutadapt (Martin, 2011) was applied to remove the forward and reverse primers (based on primers sequence), with a minimum read length (-M 10) to avoid generating zero-length reads and with two occurrences (-N 2) of the primer to be trimmed. The discard-untrimmed function was included as a default, which removed any reads from which primers were not trimmed. This function was proper when the library contained two primer sets (e.g. Riaz-12SV5 and IN16STK-Mod primer sets). To determine the appropriate truncation parameters for filterAndTrim, quality plots were generated by randomly selecting 10–12 samples per primer set. Based on these plots, the script filterAndTrim was run with the following parameters for each primer: Amphi16S-Mod (-T 220, -S 220, -G 2, -H 2, -Q 2, -L 50), Riaz-12S-V5(-T 100, -S 100, -G 2,-H 2, -Q 2, -L 50), and

IN16STK-Mod (-T 115, -S 110, -G 2, -H 2, -Q 2, -L 50). Error modelling was used to minimise false positives by accurately correcting sequencing errors. For all primers, most of the observed points closely matched estimated error rate lines, indicating a good fit.

Denoising and sequence processing steps were applied to identify Amplicon Sequence Variants (ASVs: Callahan et al., 2017), including (1) Dereplication in which identical reads of each amplicon were collapsed, (2) Denoising to identify true ASVs, using the error model generated, (3) Merging of the paired-end reads to create single, high-quality reads, and, finally, (4) removal of chimeric reads representing PCR artefacts To check for the quality of the DADA2 denoising and other steps, a summary table was generated to check the number of reads that survived each processing stage. The table provided a quick overview of where data loss may have occurred, helping identify any steps requiring optimisation.

2.5.2. Taxonomic Assignment of Amplicon Sequence Variants (ASVs)

Assigning taxonomy to the identified ASVs used a pipeline developed by Ewan Harney (https://github.com/ewan-harney/hpc_blast2megan). The first step was to run BLASTn to match the ASV query sequences against the NCBI database (https://www.ncbi.nlm.nih.gov/genbank). In the second step, MEGAN (Huson et al., 2016) uses the BLAST2LCA function tool to calculate the LCA (lowest common ancestor) based on BLAST output (from the first step) to identify the lowest taxonomic assignment that is common among the best BLAST hits (Top per cent - bit score) for each ASV. After that, a taxonomic classification for each ASV was assigned to the taxonomic categories (e.g., Kingdom, Phylum, Class, Order, Family, Species, sub-species). To run the BLAST2LCA function, the minimum percentage identity (-B) for the blast results to be considered by BLAST2LCA was 98%, and the Top per cent parameter (-T) for LCA calculation was 10, using megan-nucl-Feb2022.db as the MEGAN nucleotide database. Finally, a custom R-script in the pipeline was used to create combined files (four files) as input for downstream analysis with phyloseq (McMurdie & Holmes, 2013) in the R package (R version 4.2.1 (2022-06-23 ucrt)).

2.6. Diet Composition Analysis of Barred Grass Snakes

2.6.1 Decontamination, Filtration and Quality Control of ASVs

To identify the contamination of ASV, the decontam package in R (Davis et al., 2018) was used with two contamination detection methods: (1) prevalence-based, which detects the contamination based on how often (prevalence) an ASV appeared in negative control samples compared to real biological samples, (2) frequency-based, which detects the contamination by comparing the abundance (number of reads) of each ASV to the total read count for each sample, which serves as a proxy for DNA concentration. The phyloseq package in R (McMurdie & Holmes, 2013) was first used to organise and combine the ASV_table, taxonomy_data, and sample_metadata (the three files created from the last step of the DADA2 pipeline) into a single object, making it easier to perform contamination analysis. After that, two methods in decontam were applied: prevalence-based and frequency-based. In the prevalence method, the NTC control samples were assigned as a negative control in the metadata then decontam was used to detect the ASVs that appeared more frequently in negative control samples compared to true biological samples. An ASV was flagged as a contaminant if it was more prevalent in negative controls, with a threshold set at 0.1 (10% prevalence). For the frequency-based method, the total read counts for each sample (used as a proxy for DNA concentration) were calculated using phyloseq. The relative read abundance of each ASV was then compared to the total read counts using a linear regression model. The resulting p-value (threshold < 0.05) indicates whether the relationship between ASV abundance and total DNA concentration is consistent with a true biological signal. If the p-value is low, it suggests that the abundance pattern does not follow the expected relationship and that the ASV is likely to be a contaminant. Finally, If an ASV was flagged as a contaminant by both the prevalence-based and frequency-based methods, it was removed from the dataset. However, if an ASV was identified as a contaminant by only one of the methods, further ecological considerations were made; for example, the ASV was evaluated to determine if it might represent potential environmental contamination (e.g., human DNA or lab contaminants), or could have biological relevance to the diet of the barred grass snakes.

2.6.2. Normalization of Samples for Comprehensive Diet Analysis

Normalising the filtered ASVs of metabarcoding diet data is an important step for ensuring that the dietary data reflects the composition of the diet rather than biases that might arise by varying sequencing depths across samples. In general, there are two main common approaches to normalising the dietary data using metabarcoding: (1) Relative Read Abundance (RRA), in which the number of reads for each ASV is divided by the total number of reads in the sample, resulting in a proportional value between 0 and 1. Using the RRA approach ensures that each sample can be compared regardless of variations in sequencing depth. (2) Frequency of Occurrence (FOO) transforms ASV read counts into presence/absence data. The FOO approach is less sensitive to biases introduced during amplification and sequencing, making it a conservative approach for identifying diet components (Deagle et al., 2018).

2.6.3. Statistical Analysis (Beta Diversity) of The Diet Composition

Due to the low sample size of the Holbrook population (14 samples), the Norfolk population of barred grass snakes (54 samples) was selected for the statistical analysis to ensure more robust and reliable results. The Bray–Curtis dissimilarity matrix, calculated using RRA (Relative Read Abundance) as the data input (vegdist function - vegan package R) (Oksanen et al., 2022), along with Permutational Multivariate Analysis of Variance (PERMANOVA) with 9,999 permutations using the (adonis2 function - vegan package R) (Anderson, 2006; Oksanen et al., 2022) was used to assess the compositional dietary differences between (1) Age Classes (Adult and Subadult) and (2) Sex. Then, Principal Coordinates Analysis (PCoA) was performed using the ordinate function to visualise the dietary patterns for each group. Moreover, the Beta Dispersion Test (betadisper function vegan package R) (Anderson et al.,2006; Oksanen et al.,2022) was used to assess variability among individuals within each group and to ensure that differences in diet composition reflect true differences in group means (assumptions of PERMANOVA), rather than differences in within-group variability. The Similarity Percentage (SIMPER) analysis was used to identify the taxa driving the differences in diet composition between Age Classes and Sex, highlighting which prey contributed the most to the dissimilarity

between groups (Clarke, 1993) using a simper function in the vegan package in R with 999 permutations (Oksanen et al., 2022). We used only the Amph16S-Mod and Riaz-12SV5 primers for comparison.

3. Results

3.1. Sequencing Analysis and Quality Assessment

3.1.1. Analysis of The Total Raw Reads of Faecal Samples

A total of 18,790,194 raw sequence reads were generated from the three primer sets across 68 faecal samples during DADA2 pipeline processing. The Amphi16S-Mod primer yielded 11,338,191 raw reads, of which 48.7% (5,528,043 reads) passed the quality-filtering phase. Following denoising, 48.4% of the input reads were retained in both the forward (5,493,650 reads) and reverse (5,488,920 reads) steps, reflecting minimal data loss. After the merging step, 43.7% of the raw reads (4,951,903 reads) were remained. In the final non-chimeric step (Nochim), 4,822,360 reads were retained, representing 42.5% of the total raw and 97.4% of the merged reads. For the Riaz-12S-V5 primers, 3,750,676 raw reads were reduced to 1,831,305 reads after the filtering step (48.8% of the raw data). The forward and reverse denoising steps retained 48.6% (1,822,731 reads) and 47.3% (1,774,728 reads) of the raw reads, respectively. The merging stage resulted in 42.4% (1,589,664 reads) of the raw reads, and the final non-chimeric step retained 1,106,014 reads (29.5% of the total raw reads and 69.6% of the merged reads). The IN16STK-Mod primer started with 3,701,327 of the raw reads, and during the filtering step, 47.4% (1,754,063 reads) were retained. The forward and reverse denoising stages resulted in 46.9% (1,735,382 and 1,735,399 reads, respectively) of the raw reads. Then, 42.2% (1,561,804 reads) of the raw reads were retained during the merging step. The final non-chimeric step yielded 1,085,487 reads, corresponding to 29.3% of the raw reads and 69.5% of the merged reads. The Amphi16S-Mod primer had the highest proportion of reads at the final stage (non-chimeric step) (42.5% of raw reads: 4,822,360 reads), indicating more effective read processing compared to the other primers, which showed lower retention (29.5% (1,106,014 reads) and 29.3% (1,085,487 reads)) for Riaz-12S-V5 and IN16STK-Mod, respectively.

3.1.2. Analysis, Decontamination and Filtration of ASVs

During the taxonomic assignment procedure of the raw sequences using both BLAST2LCA and MEGAN (the last step of the DADA2 pipeline), 1,134 ASVs were identified and calculated based on LCA (lowest common ancestor) before the decontamination and filtration steps. The taxonomic assignment of the ASVs using the Amphi16S-Mod primer identified 200 unique ASVs, resulting in 2,212,845 reads (19.5 % of raw 11,338,191 reads and 45% of non-chimeric 4,822,360 reads). The taxonomic resolution of the 200 ASVs based on the calculated LCA was as follows: 156 (78%) at species resolution, 23 (11.5%) at genus resolution, 9 (4.5%) at order level, and 12 (6%) with undetermined taxonomic resolution. During the decontamination analysis using both prevalence-based and frequency-based methods, 9 ASVs were determined as contaminations. These ASVs were removed from the dataset to ensure the accuracy of the downstream analyses. Moreover, the filtration step involved the elimination of ASVs associated with the non-target taxa dietary of barred grass snakes or known sources of possible environmental and laboratory contamination. For example, *Homo sapiens* (55 ASV, 27.5% – 780,763 reads, 35%), *Natrix natrix* (11 ASVs, 5.5% – 739 reads, 0.03%), *Canis lupus* (2 ASVs, 1% –55 reads, <0.01%), *Apus apus* (2 ASVs, 1% – 2500 reads, 0.11%), *Bos taurus* (1 ASV, 0.5% – 10 reads, <0.01%). Also, as the Amphi16S-Mod targeted vertebrates (amphibians), the insects, Cloeon dipterum (2 ASVs, 1% – 815 reads, 0.03%) and Sitona lineatus (1 ASV, 0.5% – 1,700 reads, 0.07%) were removed from the dataset. After the decontamination and filtration processes, a total of 99 ASVs corresponding to 1,204,758 reads from 44 faecal samples were used for downstream analyses (Figure 4.1).

For the Riaz-12S-V5 primers, 282 ASVs (749,718 reads representing 20.0% of the total raw 3,701,327 reads and 67.8% of the non-chimeric 1,106,014 reads) were identified during the BLAST2LCA and MEGAN steps. According to the LCA calculation, 136 ASVs (48.22%) were assigned to species resolution, seven ASVs (2.48%) at genus resolution, 53 ASVs (18.79%) at family resolution, 18 ASVs (6.38%) at order resolution, five ASVs (1.77%) at class resolution,

and 63 ASVs (22.34%) with undetermined taxonomic resolution (NA). The decontamination analysis identified 16 ASVs as contaminants, which were subsequently removed from the dataset. Potential environmental and laboratory contamination was filtered out (filtration step), such as Homo sapiens (56 ASVs, 20% – 290,690 reads, 39%), Aerodramus swifts (2 ASV, 0.70% – 1405 reads,0.18%), Canis lupus (5 ASVs, 1.77% – 474 reads, 0.06%), Hebius vibakari (2 ASVs, 0.70% – 1228 reads, 0.16%), Ovis aries (1 ASV 0.35% – 504 reads, 0.06%), Bos taurus (1 ASV, 0.35% – 65 reads <0.01%), insect (83 ASVs, 29% – 25,441 reads, 3.39%), Diplopoda (13 ASVs, 4.60% – 4863 reads, 0.64%), Asarum shuttleworthii (1 ASV, 0.35% – 78 reads, <0.01%), Trichoniscus pusillus (2 ASVs, 0.70% – 1176 reads, 0.15%), Isotoma viridis (1 ASV, 0.35% – 32 reads, <0.001%), Allacma fusca (1 ASV, 0.35% – 47 reads, <0.001%), and 63 ASVs (22.34%) with undetermined taxonomic resolution (NA) 27,229 reads, 3.63%). As a result, 35 ASVs were filtered out with total reads of 309,560 from 49 faecal samples (Figure 4.1).

For IN16STK-Mod primer, 655 ASVs were created with a total of 627,577 reads (17% of raw 3,701,327 reads and 58% of non-chimeric 1,085,487 reads). The taxonomic resolution of the 653 ASVs as follows: 251 ASVs at species resolution (38%), 71 ASVs at genus resolution (11%), 149 ASVs at family resolution (23%), 19 ASVs at order resolution (3%), and 165 ASVs (25%) were classified as (NA). Among the 655 ASVs, only one ASV was assigned to contamination based on the decontamination analysis. In the filtration step, 165 ASVs were removed from the dataset because of undetermined taxonomic classification (NS). The final filtered IN16STK-Mod dataset contained 490 ASVs with 425,614 reads from 35 faecal samples (Figure 4.1).

3.2. Taxonomic Identification (ASV) of Prey Items

The filtered ASVs of Amphi16S–Mod (99 ASVs – 1,204,758 reads) in which 68 ASVs were identified at species resolution as follows based on LCA: **(1) Amphibian species**: *Bufo bufo* (15 ASV – 135,457 reads), *Lissotriton vulgaris* (10 ASV – 58,163 reads), *Triturus cristatus* (15 ASV – 108,295), **(2) Mammal species**: *Microtus agrestis* (8 ASV – 55,884 reads) and *Sorex*

minutus (3 ASV – 14,715 reads), **(3) Avian species**: *Gallus gallus* (9 ASV – 147,397 reads), *Turdus merula* (3 ASV – 74,191 reads), *Troglodytes troglodytes* (4 ASV – 3480 reads), *Columba palumbus* (1 ASV – 141 reads). Moreover, 22 ASVs were assigned at genus resolution as follows: Bufo (18 ASVs – 491,779 reads), Triturus (1 ASV – 624 reads), and Oncorhynchus (3 ASVs – 2361 reads). Finally, nine ASVs were assigned to Anura with 112,271 reads (Figures 4.2 and 4.3).

For the filtered ASVs of the Riaz–12S–V5 primer set, 28 ASVs were identified at species resolution: **(1) Amphibian species**: *Bufo bufo* (4 ASV – 457 reads), *Lissotriton vulgaris* (1 ASV – 171 reads), *Rana temporaria* (1 ASV – 171 reads), **(2) Fish species**: *Cyprinus carpio* (1 ASV – 100 reads), **(3) Avian species**: *Gallus gallus* (8 ASVs – 57146 reads), *Troglodytes troglodytes* (3 ASV – 7,510 reads), *Turdus philomelos* (1 ASV – 147 reads), **(4) Mammal species**: *Microtus agrestis* (6 ASV – 174,572 reads) and *Sorex minutus* (1 ASV –49,461 reads), *Myodes glareolus* (2 ASVs – 140 reads), and *Neomys fodiens* (1 ASV – 7 reads). In addition, five ASVs were at family resolution: Leuciscidae (2 ASVs – 766 reads), Bufonidae (1 ASV – 7 reads), Salamandridae (1 ASV – 16 reads), and Columbidae (1 ASV – 549 reads) (Figures 4.2 and 4.4.). Finally, only one ASV was assigned to Order resolution: Passeriformes, with a total read of 18,500.

For the filtered ASVs using the IN16STK–Mod primer, 164 ASVs were identified at species resolution, representing 40 different species with a total of 133,028 reads. Moreover, 68 ASVs (68,665 reads) were at genus resolution, 112 ASVs (173,137 reads) at family resolution, and 16 ASVs at order resolution (44,089 reads).

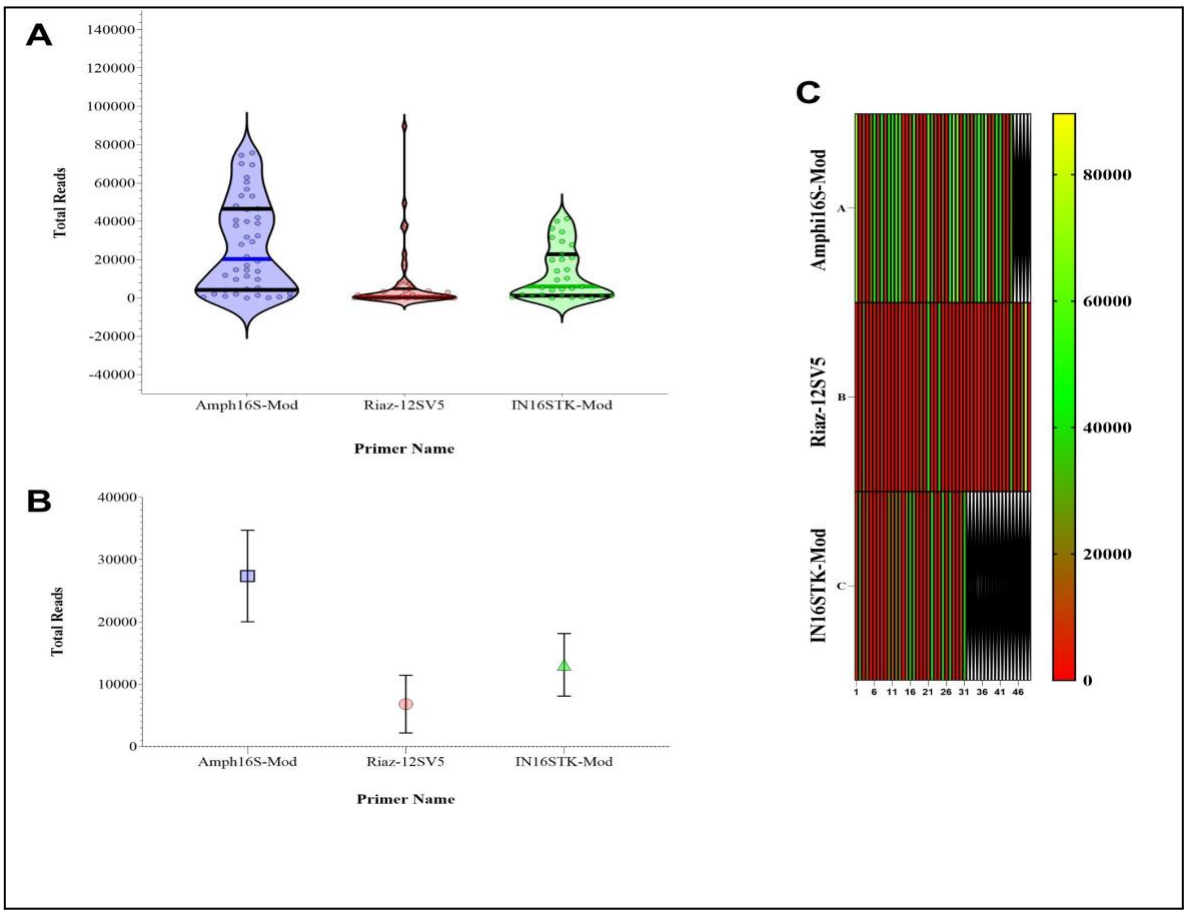

Figure 4.1. The distribution and means of total reads (after the decontamination and filtration process) for three different primer sets, Amph16S-Mod, Riaz-12SV5, and IN16STK-Mod, were used in the study. **Panel A**: The violin plot shows various faecal samples, highlighting the variability across different primers. The horizontal lines within each violin plot represent the median values and interquartile ranges (IQR), indicating the central tendency and spread of read counts. Amph16S-Mod has the highest median read count with a median of 20,344 reads (IQR: 4,162–46,508), while Riaz-12SV5 shows a much lower median read count of 518 (IQR: 93–4,842). IN16STK-Mod falls in between, with a median read count of 5,975 (IQR: 1,193–22,772). **Panel B** represents the mean read counts for each primer set with 95% confidence intervals (CI). The mean read count for Amph16S-Mod is 27,381 (95% CI: 20,040–34,721), which is higher compared to Riaz-12SV5 and IN16STK-Mod, both of which have lower mean read counts of 6,836 (95% CI: 2,213–11,459) and 13,142 (95% CI: 8,127–18,156), respectively. **Panel C** illustrates the read distribution for each sample using a colour gradient, where the colour scale on the right side ranges from 0 to 80,000 reads. Each bar corresponds to a sample, organised by primer name, and provides a visual overview of the variability in total reads across all three primers.

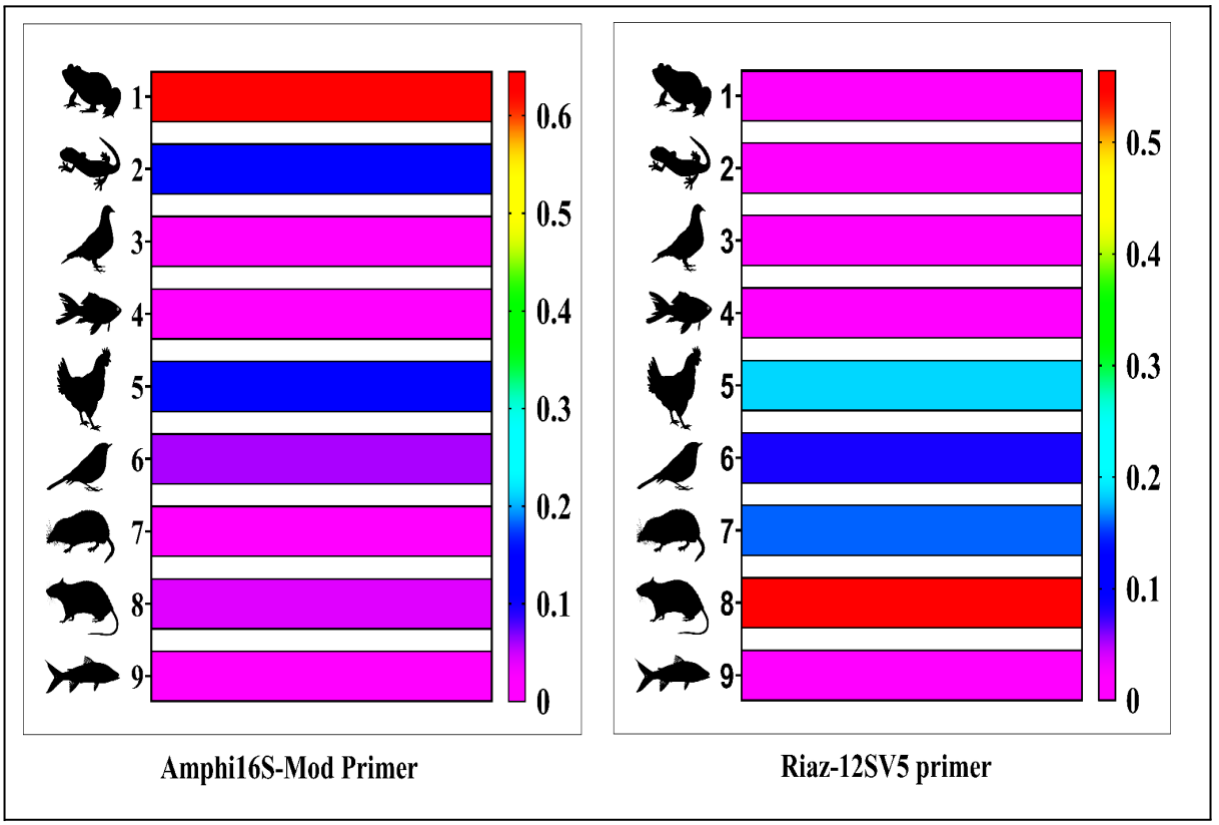

Figure 4.2. The efficiency and specificity of detection and taxonomic resolution of vertebrate taxa between Amphi16S-Mod and Riaz-12SV5 primer sets using RRA. The RRA at order level (aggregated the LCA to the higher taxonomic unit for border view) shows more in detecting amphibians: Anura (0.64) and - Caudata (0.12), and Galliformes (0.11), while the Riaz-12SV5 provides better detection for mammals: Rodentia (0.56) and- Eulipotyphla (0.16), and birds: Galliformes (0.18) - Passeriformes (0.08).

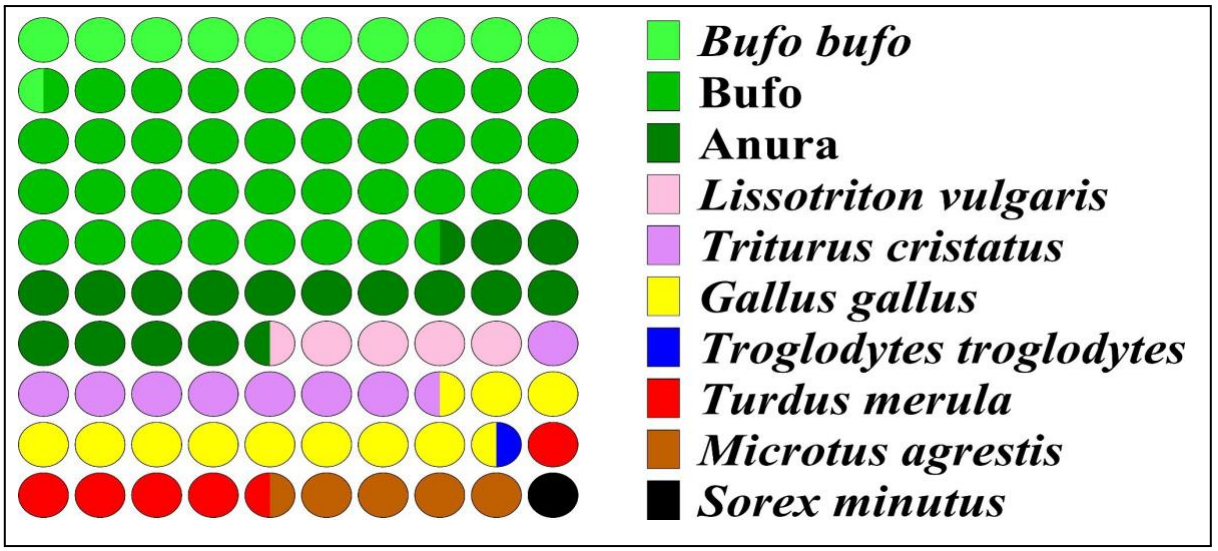

Figure 4.3. The taxonomic resolution of vertebrate taxa of Amphi16S-Mod primer set using RRA based on LCA unite. The Amphi16S-Mod primer se appears highly efficient in detecting amphibian taxa across different taxonomic units. The genus level (Bufo) has the highest RRA following the order Anura, and at the species level unit, three amphibian species: *Bufo bufo*, *Lissotriton vulgaris*, and *Triturus cristatus*. Also, two mammalian species and three avian species were detected by Amphi16A-Mod, indicating the success of modification of the primer set to target vertebrates (not only amphibians)

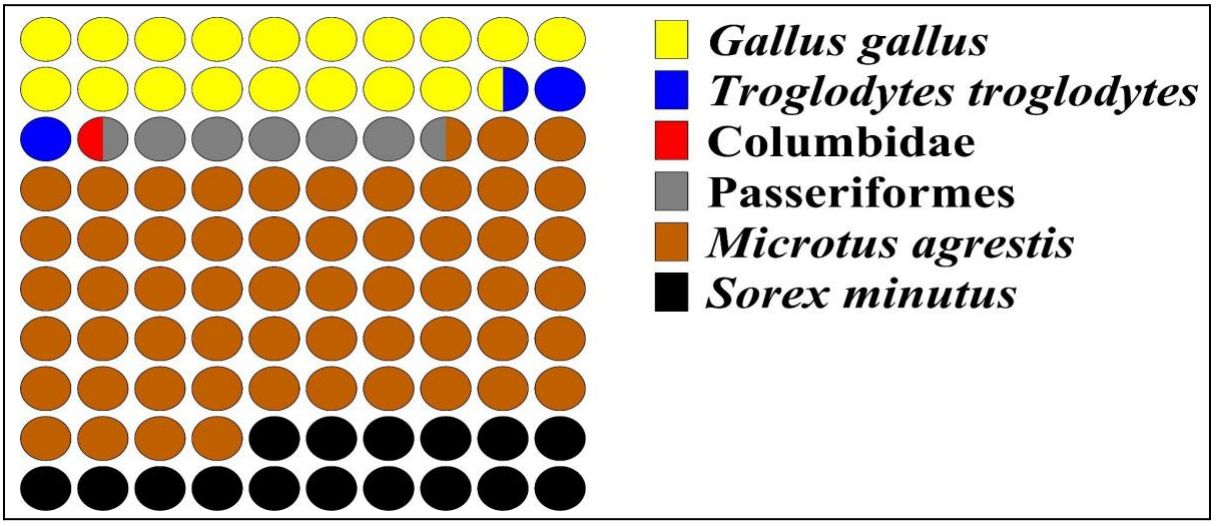

Figure 4.4. The taxonomic resolution of vertebrate taxa of Riaz-12SV5 primer set using RRA based on LCA unite. The Riaz-12SV5 primer set shows a different detection pattern, with a broader representation of bird species, including *Gallus gallus* and *Troglodytes troglodytes*, *Columbidae*, and at the order level for Passeriformes. Mammalian taxa such as *Microtus agrestis* and *Sorex minutus* are also well represented, similar to the Amph16S-Mod primer results, but there are fewer detections of amphibians, showing a bias toward avian and mammalian taxa.

3.2. Dietary Composition Analysis (Vertebrate Prey Items)

3.2.1. Comparison of Dietary Composition Between Two Populations of *Natrix helvetica*

The dietary composition of the two barred grass snake (*Natrix helvetica*) populations (Norfolk and Holbrook) at the order level differs with respect to prey abundance (RRA) and frequency (FOO). The Norfolk population (37 individuals) primarily fed on amphibians, with Anura (frogs and toads) making up more than half of the total diet composition (RRA = 0.57) and Caudata (newts) contributing 0.08. Moreover, The RRA for Rodentia (0.12), Galliformes (0.11), and other prey taxa suggest that small mammals and birds are occasional or secondary prey items when amphibians are unavailable. However, the Galliformes were the most frequent prey (FOO = 0.64), followed by Anura (FOO = 0.27) and Rodentia (FOO = 0.12) (Figure 4.5)

By contrast, the Holbrook population (9 individuals) shows a more diverse diet, with a higher relative abundance of Rodentia (RRA = 0.27), Caudata (RRA= 0.25), and Galliformes (RRA = 0.24). The Frequency of Occurrence (FOO) indicated that the Galliformes was found to be the more frequent prey item (FOO = 0.70), followed by Rodentia (FOO = 0.1), Eulipotyphla (FOO = 0.1), and Salmoniformes (FOO = 0.1) (Figure 4.5).

The PERMANOVA analysis indicated no significant difference in prey preferences between the two populations at the order level (P = 0.18) and species level (P = 0.10). However, the Betadisper test at the species level suggested significant differences in dietary variation within the population (P = 0.05), whereas no significant differences were observed at the order level (P = 0.40). The SIMPER analysis at the species level identified that *Sorex minutus* (P = 0.0332) and *Gallus gallus* (P = 0.0375) significantly contributed to the dietary differences between the Norfolk and Holbrook populations, with both species being consumed more in Holbrook (Figure 4.6).

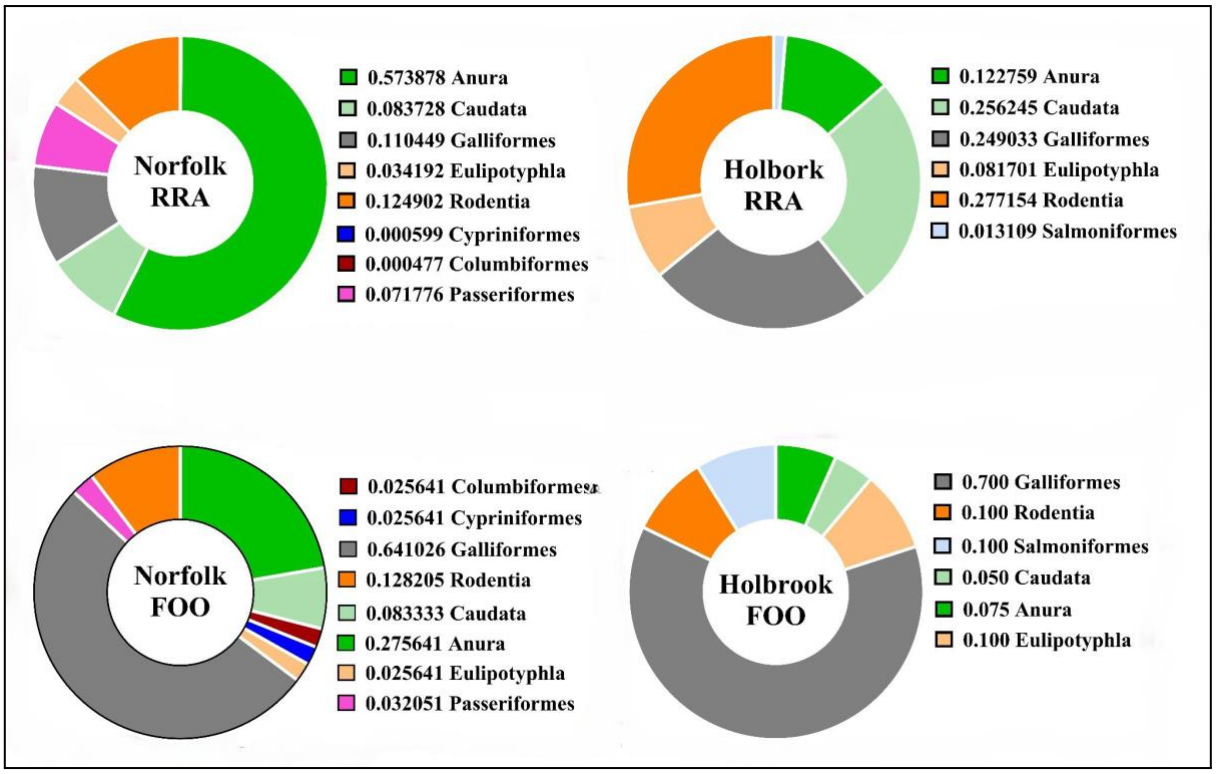

Figure 4.5. The Relative Read Abundance (RRA) and frequency of occurrence (FOO) prey at the order unit between two populations, Norfolk and Holbrook.

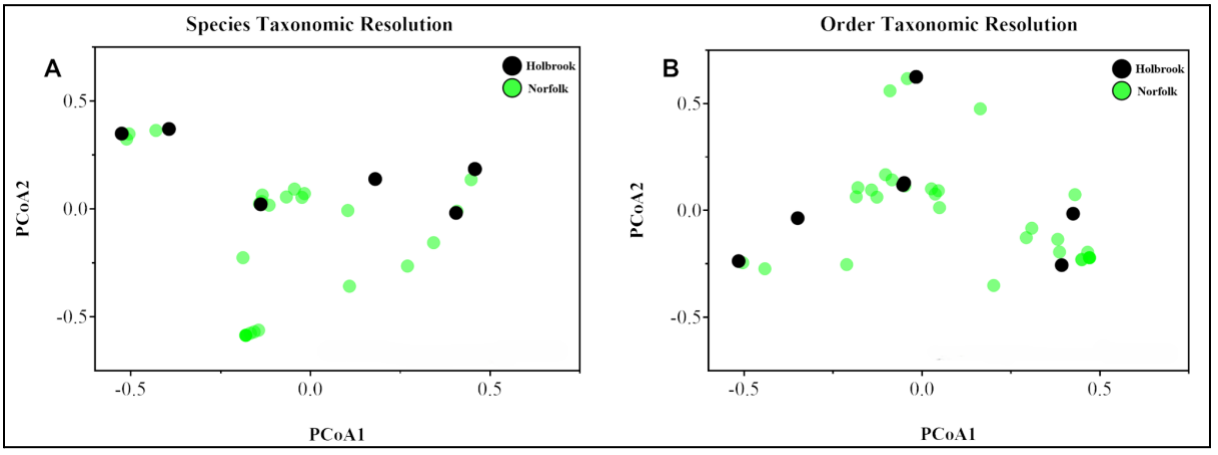

Figure 4.6. The figure presents a Principal Coordinates Analysis (PCoA) of Bray-Curtis dissimilarity, illustrating population-based dietary differences in *Natrix helvetica* between two sites, Holbrook (black) and Norfolk (green), at both the species **(Panel A)** and order **(Panel B)** taxonomic resolutions. **In Panel A** (species level), PERMANOVA p = 0.10 indicates a non-significant difference in dietary composition between the two populations. The Betadisper p = 0.05 suggests that the variability within each population is close to being significantly different, indicating that one population may have slightly more diverse diets than the other. **In Panel B** (order level), PERMANOVA p = 0.18 shows no significant difference in diet composition between the populations, and Betadisper p = 0.40 indicates no significant difference in dietary variability between the sites at this broader taxonomic level. Overall, the analysis suggests that there may be some dietary differences between populations at the species level, but these differences are less at the order level.

3.2.2. Dietary Composition Comparison Between Sex and Age Classes in the Norfolk Population

The dietary composition of the Norfolk population of *Natrix helvetica* was compared between sexes and age classes to identify differences in prey preference. The Bray–Curtis dissimilarity matrix showed no significant difference in dietary composition between males and females at the species resolution (PERMANOVA P = 0.382), suggesting an overlap in species-level prey consumption. Moreover, the Betadisper test (P = 0.13) indicated no significant difference in dietary variation (dispersion) within the group. However, at the order resolution, there was a slight difference between males and females in dietary composition (PERMANOVA P = 0.056) (Figure 4.7). The Betadisper test (P = 0.04) suggested that the slight difference observed was due to variation in dietary composition of females compared to males. In the SIMPER analysis, the Anura (*Bufo bufo*) appeared to be the most consumed prey for both males (RRA = 0.612) and females (RRA = 0.335). The ratio between them was 1.146, meaning males consumed about 14.6% more Anura than females, although the difference was insignificant (0.47). Moreover, Galliformes also showed relatively high consumption in males (RRA = 0.211) compared with females (RRA = 0.163) (P = 0.001). The order Caudata showed higher consumption by females (RRA = 0.259) compared to males (RRA = 0.00295), in which females consumed *Triturus cristatus* more often than males (P = 0.01) (Figures 4.8 and 4.9). Finally, Galliformes were the most frequently consumed order in males and females, as well as Anura and Rodentia.

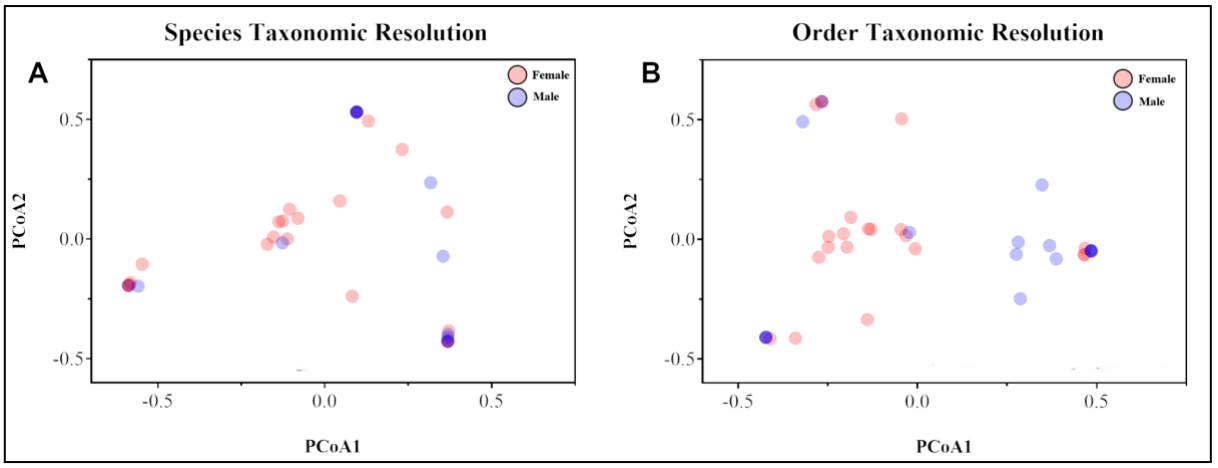

Figure 4.7. The figure compares the diets of males and females of *Natrix helvetica* (barred grass snakes). **Panel A** is the diet differences at the species level, while **Panel B** is at the order level (a broader classification). **In plot A**, there is no significant difference between the diets of males and females at the species level (PERMANOVA p = P = 0.38), and there is also no significant difference in how varied their diets are (Betadisper P = 0.13). **In plot B**, at the order level, there is a slightly significant difference between males and females (PERMANOVA P = 0.05). However, there is a significant difference in how varied the diets are (Betadisper (P = 0.04). This means that, at the order level, females tend to have a more diverse range of prey items compared to males.

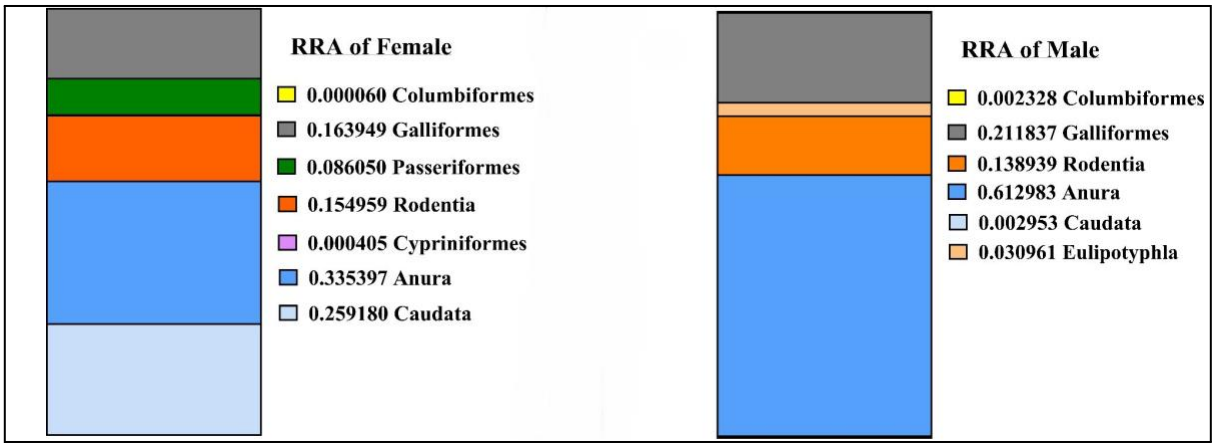

Figure 4.8. Comparison of Relative Read Abundance (RRA) of Dietary at Orders Between Females and Males *Natrix helvetica* in the Norfolk Population. The RRA of dietary orders for females showed a higher representation of Anura (0.335) and Caudata (0.259) in the diet compared to other orders. Rodentia (0.155) and Passeriformes (0.086) were also present, indicating a varied diet. The RRA of dietary orders for males highlighted Anura (0.613) as the dominant prey item, followed by Galliformes (0.212) and Rodentia (0.139) and (0.212). Eulipotyphla (0.031) suggests a broader range of small mammalian prey in the male diet.

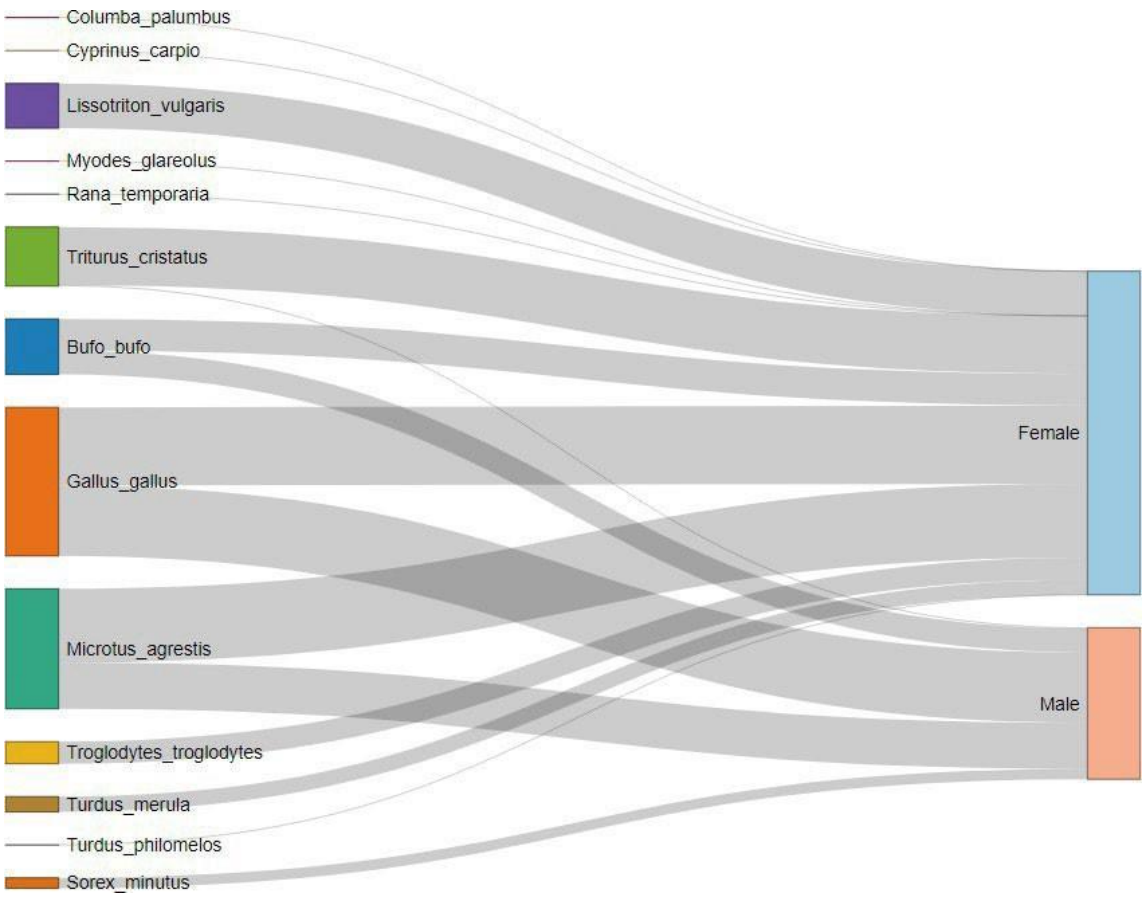

Figure 4.9. Sankey diagram illustrating the diet composition of males and females of *Natrix helvetica* based on species-level prey identification. Each flow represents a connection between a prey species (left side) and the sex of the snakes that consumed it (right side). The width of the flows corresponds to the RRA of each prey species found in the diet of males or females.

The age classes (adult and subadult) of barred grass snakes did not show significantly different dietary composition at the order level (PERMANOVA P = 0.19), and there was no significant difference in the variation of dietary composition within age class (Betadisper P = 0.89). However, when the Bray–Curtis dissimilarity matrix at species resolution was calculated for the RRA, there was a significant difference in dietary composition in age class (PERMANOVA P = 0.01), and the variation of dietary composition within age class was not significantly different (Betadisper P = 0.30) (Figures 4.10 and 4.11). In the SIMPER analysis of species resolution, *Bufo bufo* was the most consumed prey in adults (RRA = 0.22) and subadults (RRA = 0.384). *Gallus gallus* was consumed more in adults (0.43) than in

subadults (RRA = 0.02), but this was not statistically significant (P = 0.13). On the other hand, *Triturus cristatus* and *Sorex minutus* were only found in the diets of subadults (RRA = 0.223) and (RRA = 0.07), respectively. These species showed statistically significant differences (P = 0.01), suggesting they are an important prey item for subadults (Figure 4.10).

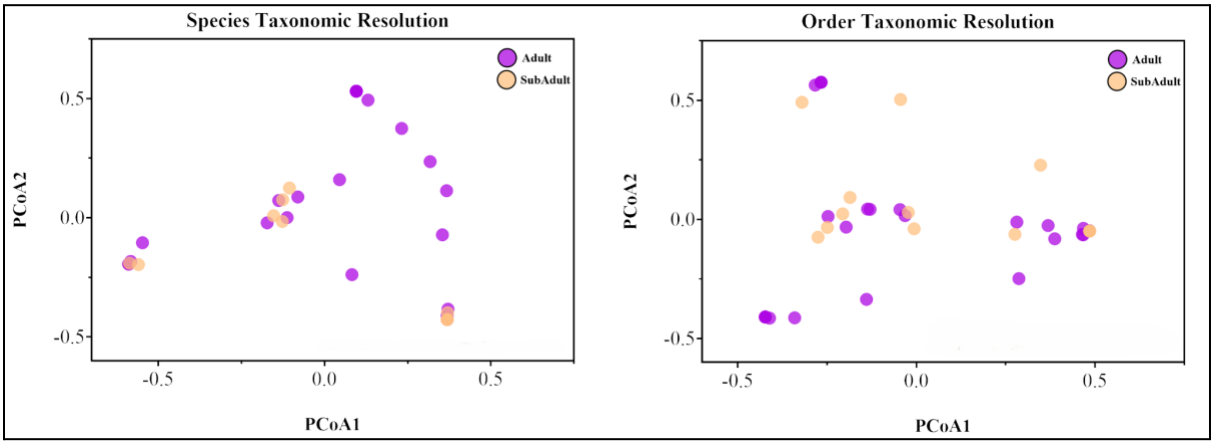

Figure 4.10. The figure presents a Principal Coordinates Analysis (PCoA) of Bray-Curtis dissimilarity to examine age-class-based dietary differences in Natrix helvetica at the species and order) taxonomic resolutions. The colours represent two age classes: adults (purple) and subadults (orange). At species taxonomic resolutions, PERMANOVA indicated a significant difference in diet composition between age classes (p = 0.01), suggesting that adults and subadults differ significantly in their diets at the species level. The Betadisper (P = 0.30) suggests that the variability (spread) within the groups is not significantly different. At order taxonomic resolutions, PERMANOVA P = 0.19 indicates no significant difference in diet composition between adults and subadults at the order level, and Betadisper P = 0.89 shows no significant difference in variability between the groups. This suggests that while age classes show significant differences in diet at the species level, there are no differences at the broader order level.

.

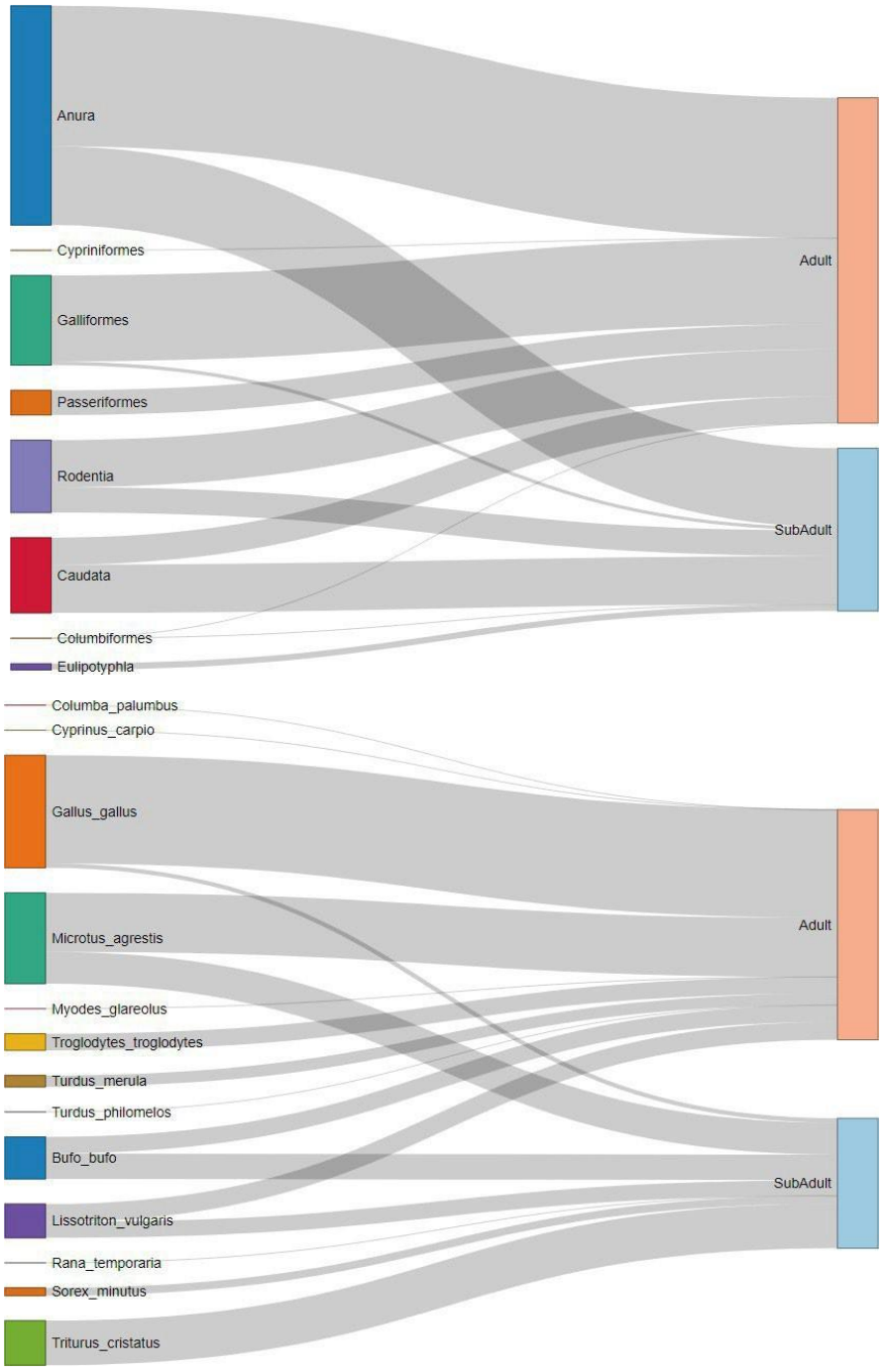

Figure 4.11. The Sankey diagram illustrates the diet composition of adults and subadults of *Natrix helvetica* based on species and order level prey identification. Each flow represents a connection between a prey species (left side) and the age class of the snakes that consumed it (right side). The width of the flow corresponds to the RRA of each prey species found in the diet of adults or subadults.

4. Discussion

The project utilised a metabarcoding approach to investigate the dietary composition of *Natrix helvetica* in two different populations, Norfolk and Holbrook, and compared the dietary composition between age classes and sex in the Norfolk population. We hypothesised that (1) amphibians would dominate the diet of both populations, with possible variations driven by local prey availability, (2) subadult grass snakes would exhibit an ontogenetic shift, consuming smaller prey while adults would target larger prey, and (3) females would have a broader and more varied diet compared to males, driven by the reproduction needs.

4.1. Detected dietary items of *Natrix helvetica* in Two Populations

The metabarcoding results of our study indicated that the detected dietary items of barred grass snakes between the two populations were different, despite PERMANOVA indicating that these differences were not statistically significant at both species and order levels. The lack of statistical significance could be due to the limited power due to the small sample size, particularly in the Holbrook population. The species accumulation curves supported the limitation of the small sample size, as the Holbrook curve plateaued much earlier than Norfolk, suggesting that a smaller sample size may have limited the detection of additional species. A larger sample size may be necessary to fully capture the extent of dietary variation and increase the robustness of the statistical analyses. Nonetheless, the observed differences in detected dietary items between the two populations suggested that local prey availability and habitat structure may play an important role in determining the diet of *Natrix helvetica*. For instance, in Norfolk, amphibians were the main prey group, with 60% of the relative read abundance in species identified in faecal DNA. Within amphibians, Anura (*Bufo bufo*) comprised 44% of prey items. This was associated with Norfolk's wetland habitat, which supported an abundance of amphibians, confirming our hypothesis that amphibians would dominate the diet of the Norfolk population.

By contrast, the species identified in faecal DNA of barred grass snakes in Holbrook was more diverse, including amphibians (38%), small mammals (36%), and birds (25%). This suggests that the ponds that were created at Holbrook offer an opportunity for variation in prey types, which may not have the same abundance in the more specialised wetland habitat in Norfolk. Moreover, the Holbrook Marsh Ponds site was developed primarily to support the growth of newt populations, especially *Triturus cristatus*. This is reflected in the higher RRA of Caudata in Holbrook (25%) than in Norfolk (8%). It was expected that *Triturus cristatus* would be the main prey item for the barred grass snakes in the Holbrook population, based on our initial hypothesis. However, opposite to our expectations, *Lissotriton vulgaris* was the only newt species detected as prey. This can indicate that the population of *Lissotriton vulgaris* at Holbrook is more abundant than *Triturus cristatus*, which was detected as prey in the Norfolk population.

The dietary patterns of faecal DNA observed in the two barred grass snake populations suggest that ecological factors, such as habitat type and prey availability, might play an important role in shaping their diets (Gregory & Isaac, 2004). Our findings are consistent with previous studies that have demonstrated the influence of habitat type and prey availability on the diet composition of closed-related species (*Natrix natrix*). For example, Luiselli et al. (2005) observed dietary differences between populations of European grass snakes (*Natrix natrix*) across six different habitats in Italy. In two of these populations, Caudata and rodents were the main prey items, demonstrating how local prey availability and habitat conditions can shape the diet of grass snakes.

Our metabarcoding results supported the conclusion that *Natrix helvetica* may not be a specialist predator, meaning that its diet is not restricted to a narrow range of prey. While amphibians were the preferred target prey, likely due to the semi-aquatic ecology of *Natrix helvetica*, the species demonstrated the ability to consume a variety of other prey, including

small mammals as secondary prey, as well as birds and fish on occasion. This dietary flexibility suggests that *Natrix helvetica* adapts its feeding behaviour based on local prey availability, further supporting its role as a generalist predator (Gregory & Isaac, 2005). Moreover, the lack of dietary restrictions for amphibians could be an advantage, as amphibians are among the fastest-declining vertebrate groups. Thus, *Natrix helvetica* might be less affected by these declines.

Previous dietary analysis for *Natrix helvetica* was conducted either using a conventional approach based on morphological identification of prey remains from faecal samples or stomach contents (Reading & Davies, 1996; Gregory & Isaac, 2004; Luiselli et al., 2005), or through molecular techniques such as PCR-based methods using a specific-species taxa and Sanger sequencing (Brown et al., 2014). This supported our findings that amphibians were the preferred prey of *Natrix helvetica* at the order level, with *Bufo bufo*, *Lissotriton vulgaris*, and *Triturus cristatus* identified at the species level. Also, Filippi et al (1996) found that most of the diet consisted of anurans (frogs, toads, and their tadpoles).

There was evidence of barred grass snakes consuming juvenile birds; for example, Di Nicola & Bruni (2020) reported a barred grass snake preying on a juvenile European starling (*Sturnus vulgaris*). Schiefenhövel (2017) documented predation on a juvenile yellowhammer (*Emberiza citrinella*). Rogers (1901) and Gregory & Isaac (2004) also noted that birds remain in the stomachs of barred grass snakes. Luiselli & Rugiero (1991) recorded *Turdus merula*, and Farane et al (2010) observed *Gallus domesticus* as prey of the barred grass snakes. To our knowledge, and based on our metabarcoding results, this study reported for the first time that *Troglodytes troglodytes*, *Columba palumbus,* and *Turdus philomelos* were consumed by barred grass snakes, as well as *Gallus gallus* and *Turdus merula* as recorded in previous studies. Because metabarcoding cannot identify the sex or size of the prey, we assume that

the bird species detected in our study would have been predated as juveniles or eggs, based on previous reports also, gape capacity is size limiting.

Faraone et al. (2010) observed *Gallus domesticus* as prey for *Natrix helvetica*, noting that the grass snake consumed it as carrion. This finding suggested that *Natrix helvetica* may occasionally scavenge on dead animals. Contamination is always a consideration in molecular studies. Thus, the steps taken in our study, such as the decontamination approach and filtration process, helped reduce this possibility. Moreover, *Gallus gallus* was detected via metabarcoding of water samples in Holbrook (Ali Al-Amrie – personal communication), demonstrating that chicken DNA can be present in environmental samples. At the Norfolk site, a large chicken farm is located within a few hundred meters of the study area, with a stream flowing past the farm into the site. Given these factors, it is likely that *Gallus gallus* DNA is abundant within the study site. While this does not rule out the possibility of *Natrix helvetica* consuming *Gallus gallus*, it raises the possibility that DNA detected in faecal samples may be due to environmental exposure rather than direct predation. We recommend that future studies conduct additional validation to confirm *Gallus gallus* as a diet item of barred grass snakes. However, our current findings align with existing literature on grass snake diets.

Natrix helvetica is sympatric with smooth snakes (*Coronella austriaca*) and adders (*Vipera berus*) in some habitats in the UK. However, the potential for dietary competition is likely low due to different dietary preferences (Luiselli, 2006). The use of molecular techniques (traditional Sanger sequencings) by Brown et al. (2014) found that the diet of *Natrix helvetica* (sample size = 14) was dominated by amphibians, particularly *Lissotriton vulgaris*, small mammals (*Neomys fodiens*) and reptiles, which indicated some potential competition with juveniles of smooth snakes, particularly for reptilian prey in habitats where amphibians are less abundant. However, our study did not detect any reptilian prey, suggesting that the potential for direct competition between *Natrix helvetica* and *Coronella austriaca* for reptilian prey is likely minimal.

4.2. Detected dietary items of *Natrix helvetica* By Male and Female in Norfolk Site

The detected dietary items of *Natrix helvetica* in our metabarcoding results showed no significant difference between males and females at order and species level (PERMANOVA P = 0.05 and 0.38, respectively). The females were more diverse in their diet than males at the order level, supported by the Betadisper test (P = 0.04). For example, the orders Passeriformers and Cypriniformes were consumed only by females. Gregory & Isaac (2004) observed that the birds were found to be consumed only by females, which is consistent with our findings. Also, Luiselli et al (2005) documented that Passeriformes were found only in female diets. All these results, including ours, suggest that the morphological size and/or reproduction needs may encourage or provide an advantage to females to capture a wider range of prey types. For example, the heads of the females are wider than those of males, enabling the capture of larger prey (Borczyk, 2014; Tamagnini et al., 2018). Another explanation for the diversity of diets in females may be the nutritional requirements for egg production and post-laying recovery, which create greater metabolic needs for females compared to males. In general, females usually feed less or may even stop feeding during their reproductive cycle in comparison with non-gravid females. By contrast, non-gravid females appear to be more opportunistic and diverse in their prey selection due to their higher energy needs during and after the reproductive cycle (Filippi et al (1996); Reading & Davies, (1996); Gregory & Isaac (2004).

On the other hand, the Relative Read Abundance (RRA) of Anura in our results was higher in males (0.61) than in females (0.33), with *Bufo bufo* being the primary species consumed by males. However, the RRA of Caudata (specifically *Lissotriton vulgaris* and *Triturus cristatus*) was higher in females (0.26) compared to males (0.002). This pattern of diet composition is similar to the findings of Luiselli & Shine (1996) and Luiselli et al. (2005), but different from Reading & Davies (1996), who reported that females consumed more anurans than males. While the study of Filippi et al. (1996) has reported that males tend to have a more diverse diet compared to females, the primary difference was Anura, unlike

our results which suggested that females have greater dietary diversity. However, the difference between our results and the study of Filippi et al. (1996) may be due to geographical and ecological differences between populations or the methodologies used. As a result, this again may support the conclusion that prey availability and habitat use play an important role in shaping the dietary habits of grass snakes.

We may conclude that the dietary diversity observed in female Natrix helvetica is likely due to both size differences and reproductive needs. Larger body size in females allows them to handle a wider range of prey, while the increased energy needs associated with egg production and post-laying recovery drive them to consume a more varied diet to meet their nutritional requirements. Our results suggest another support of the possible size-based differentiation in prey selection, with females preferring more diverse diets of amphibians (newts and toads), while males focus more on toads in the Norfolk population.

4.3. Detected dietary items of *Natrix helvetica* by Age class in the Norfolk population

The significant difference in diet composition between adults and subadults at the species level (PERMANOVA, P = 0.01) supports the hypothesis that Natrix helvetica may be subject to ontogenetic shifts in prey selection. Moreover, *Triturus cristatus* and *Sorex minutus* were consumed only by subadults (SIMPER, P = 0.01), suggesting that the small preys are important for subadults due to their smaller body size and hunting limitations. The average SVL of the subadults was 28 cm, indicating that most of the subadults were at the juvenile stage, thus indicating the relevance of these prey species in the earlier stages of development. Another piece of evidence is that our Capture-Mark-Recapture survey at the Holbrook site revealed that snakes with SVL < 40 cm primarily consumed Caudata and Eulipotyphla, highlighting the importance of these prey groups during the subadult life stage. These findings suggested that these two orders, including the *Triturus cristatus*, *Lissotriton vulgaris*, and *Sorex minutus*, are important prey items for subadults, indicating potential dietary specialisation or preference during the life stage. To our knowledge, our metabarcoding results are the first to describe the diet composition of subadults of *Natrix helvetica*, which fills a gap

in our knowledge of this species (Gregory & Isaac (2004). There was an overlap in prey between adults and subadults for Anura and Rodentia. However, the adults consumed more diverse diets, particularly Passeriformes and Galliformes, which required a strategy of handling prey. For example, adults have wider heads than subadults, which provide increased space for jaw muscles, allowing them to have a stronger bite and capture larger prey (Borczyk, 2014). As a result, our findings support (1) the patterns of ontogenetic shifts that larger snakes eat both larger prey and continue to include smaller prey (Arnold, 1993) and (2) there was not only an ontogenetic shift in prey size of the subadults as they grew, but also a change in the types of prey species consumed (taxonomical dietary composition) (Filippi et al. 1996).

4.4. Limitations and Recommendations for Future Research Using Metabarcoding

The results indicated that the metabarcoding approach is a sufficient method to determine and explore the dietary composition of Natrix helvetica, not only at the broad taxonomy level but also by providing specific species resolution, in comparison with conventional methods such as the analysis of prey remains using morphological identification, which requires taxonomic knowledge to accurate classify the prey at the species level. Moreover, the result of using multiple gene markers that targeted the same taxa in our study enhanced the detection of the same prey species across different taxonomic ranks, such as when one primer set failed to classify the prey at a lower taxonomic level or amplified the faecal samples when they were highly degraded (Heo et al., 2022: Swinehart., et al 2023).

The Amphi16S-Mod primer set was more efficient in detecting amphibians compared to the Riaz-12S-V5 one, which was better at detecting mammals and birds. As expected, the Riaz-12S-V5 primer produced fewer total reads than the Amphi16S-Mod primer because the primary taxa consumed by *Natrix helvetica* were expected to be amphibians based on literature. Even though the Amphi16S-Mod primer set was modified, particularly in the forward primer of the Vert16S-eDNA, to detect a broader range of vertebrate taxa, its primary strength remained in efficiently detecting amphibians. The advantage of using Amphi16S-Mod is to detect amphibians for dietary analysis and investigate genetic diversity within amphibian populations.

For example, in our metabarcoding, the *Bufo bufo* was assigned to 15 ASVs, *Triturus cristatus* to 15 ASVs, and *Lissotriton vulgaris* to 10 ASVs) indicating the presence of mtDNA variation within these species (Tsuji et al., 2020; Sakata et al., 2022). To our knowledge, the Amphi16S-Mod primer set has not been used in environmental analysis; thus, we highly recommend using this primer set if the target taxa are amphibians. Another advantage of using the Amphi16S-Mod primer set, based on our results, is its effectiveness with highly degraded faecal samples. Despite the low average DNA concentration (~3.20 ng/μL), the primer set still successfully amplified amphibian DNA with species resolution, showing robustness in detecting target species even when DNA quality is bad.

During the analysis of the taxonomic assignment of prey (ASVs), we noticed that many ASVs were contaminated using all the primer sets, either during the processing of the samples in the laboratory (e.g. human DNA) or during the collection of the faecal samples (environmental contamination, e.g. *Bos taurus*, *Apus apus*, and *Canis lupus*). Thus, we recommend including more than two negative controls to track any contamination. For example, in our case, we struggled to decide if the ASV assigned to Galliformes (*Gallus gallus*) were caused by contamination during the lab procedure or sample collection or if it was a part of the *Natrix helvetica* diet. We decided that the Galliformes (*Gallus gallus*) were a part of the *Natrix helvetica* diet because (1) none of the ASVs were observed in the five negative controls, (2) only two ASVs out of nine were flagged as contamination during decontamination procedure (due to the low read numbers), (3) the two primers could detect Galliformes (*Gallus gallus*) with an acceptable number of reads, and (4) using literature for more confirmation (Faraone et al. 2010). However, further investigation should be conducted regarding the Galliformes (*Gallus gallus*) as a part of the diet because it was the most frequent bird taxa. Moreover, Ando et al. (2020) mentioned that only 35% of the metabarcoding studies used negative controls, which may have affected the findings and conclusions in some; thus, using negative controls is highly recommended.

The use of the metabarcoding approach for diet assessment of snakes has been limited compared to studies of mammals and fish. To our knowledge, four studies used the metabarcoding approach based on NGS to analyse dietary composition, showing efficiency in identifying a broader range of prey at different taxonomic levels compared to traditional methods, allowing for more accurate information about feeding habits and ecological aspects (Mondino et al. 2022; Swinehart et al., 2023). However, general challenges associated with the metabarcoding approach remain difficult to address. For example, the metabarcoding could not provide ecological information regarding the prey, such as age classes and sexes (Bonin et al., 2020). As a result, combining traditional methods, such as visual analysis of faeces or stomachs, with metabarcoding may solve this limitation.

Our results may be biased because (1) the Amphi16S-Mod and Riaz-12S-V5 primer sets were analysed together even though they targeted different taxa, and (2) since we did not sample across different seasons at the same sites, we were unable to conduct seasonal analyses without accounting for the season of collection, which may have prevented us from detecting differences in diet composition between the sexes. Future research on the diet of *Natrix helvetica* using DNA metabarcoding should incorporate additional molecular markers that target short, variable target sequencing suitable for identifying specific mammalian and bird families or genera of interest, as well as markers that target invertebrate taxa. Moreover, developing a local comprehensive reference database may improve the taxonomic resolution. In addition, future sampling should aim to assess diet composition across different seasons at the same sites to investigate seasonal environmental changes and how nutritional needs during growth and reproductive stages affect dietary patterns.

5. References

Alberdi, A., Aizpurua, O., Bohmann, K., Gopalakrishnan, S., Lynggaard, C., Nielsen, M., & Gilbert, M. T. P. (2019). Promises and pitfalls of using high-throughput sequencing for diet analysis. *Molecular Ecology Resources*, *19*(2), 327–348. https://doi.org/10.1111/1755-0998.12960

Anderson, M. J. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, *62*(1), 245–253. https://doi.org/10.1111/j.1541-0420.2005.00440.x

Anderson, M. J., Ellingsen, K. E., & McArdle, B. H. (2006). Multivariate dispersion as a measure of beta diversity. *Ecology Letters*, *9*(6), 683–693.

Ando, H., Mukai, H., Komura, T., Dewi, T., Ando, M., & Isagi, Y. (2020). Methodological trends and perspectives of animal dietary studies by noninvasive fecal DNA metabarcoding. *Environmental DNA*, *2*(3), 391–406.

Arnold, S. J. (1993). Foraging theory and prey-size–predator-size relations in snakes. In R. A. Seigel & J. T. Collins (Eds.), *Snakes: Ecology and behavior* (pp. 87–115). McGraw-Hill.

Bonin, M., Dussault, C., Taillon, J., Lecomte, N., & Côté, S. D. (2020). Combining stable isotopes, morphological, and molecular analyses to reconstruct the diet of free-ranging consumers. *Ecology and Evolution*, *10*(13), 6397–6410.

Borczyk, B. (2014). Allometry of head size and shape dimorphism in the grass snake (*Natrix natrix* L.). *Turkish Journal of Zoology*, *38*(6), 734-738.

Brown, D. S., Ebenezer, K. L., & Symondson, W. O. (2014). Molecular analysis of the diets of snakes: Changes in prey exploitation during development of the rare smooth snake *Coronella austriaca*. *Molecular Ecology*, *23*(15), 3734-3743.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583.

Clarke, K. R. (1993). Non-parametric multivariate analysis of changes in community structure. *Australian Journal of Ecology*, *18*(1), 117–143.

Consul, A., Eger, S., & Kwet, A. (2009). The grass snake, *Natrix natrix natrix* (Squamata: Colubridae), as a predator of the Great Ramshorn Snail, *Planorbarius c. corneus* (Gastropoda: Planorbidae). *Salamandra*, *45*(1), 50-52.

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, *6*(1), 226.

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., Kartzinel, T. R., & Eveson, J. P. (2019). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*, *28*(2), 391–406.

Di Nicola, M. R., & Zabbia, T. (2021). *Natrix helvetica* (Barred Grass Snake): Diet. *Herpetological Review*, *52*(4), 877–878.

Díaz-Abad, L., Bacco-Mannina, N., Madeira, F. M., & others. (2022). eDNA metabarcoding for diet analyses of green sea turtles (*Chelonia mydas*). *Marine Biology*, *169*(18). https://doi.org/10.1007/s00227-021-04002-x

Eusemann, P., Rees, J., Kuhlenkamp, V., Lippitsch, P., & Schumann, H. (2024). Dietary analysis of wolf (*Canis lupus*)—A comparison of markers and methods. *Conservation Genetics Resources, 16*(2), 217-220.

Faraone, F., et al. (2010). In Atti VIII Congresso Nazionale della Societas Herpetologica Italica (Proceedings of the 8th National Congress of the Italian Herpetological Society) (pp. 247–252). Ianieri Edizioni

Farooq, H., Harfoot, M., Rahbek, C., & Geldmann, J. (2024). Threats to reptiles at global and regional scales. *Current Biology, 34*(10), 2231-2237.e2. https://doi.org/10.1016/j.cub.2024.04.007.

Filippi, E., Capula, M., Luiselli, L., & Agrimi, U. (1996). The prey spectrum of *Natrix natrix* (Linnaeus, 1758) and *Natrix tessellata* (Laurenti, 1768) in sympatric populations. *Herpetozoa, 8*(3/4), 155-164.

Filippi, E., & Luiselli, L. (2002). Crested newts, *Triturus carnifex* (Laurenti, 1768) form the bulk of the diet in high-altitude Grass snakes *Natrix natrix* (Linnaeus, 1758) of the central Apennines. *Herpetozoa, 15*, 83-85.

Gleed-Owen, C. (1994). The stomach contents of a grass snake, *Natrix natrix*, identified from skeletal remains. *British Herpetological Society Bulletin, 50*, 34-36.

Gleed-Owen, C. (1996). Amphibian remains identified from the gut of a Grass Snake and a polecat. *British Herpetological Society Bulletin, 57*, 21-23.

Gregory, P. T., & Isaac, L. A. (2004). Food habits of the grass snake in southeastern England: Is *Natrix natrix* a generalist predator? *Journal of Herpetology, 38*(1), 88-95.

Harper, L. R., Handley, L. L., Hahn, C., Boonham, N., Rees, H. C., Gough, K. C., Lewis, E., Adams, I. P., Brotherton, P., & Hänfling, B. (2018). Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*). *Ecology and Evolution, 8*(12), 6330-6341.

Harper, L. R., Watson, H. V., Donnelly, R., Hampshire, R., Sayer, C. D., Breithaupt, T., & Hänfling, B. (2020). Using DNA metabarcoding to investigate diet and niche partitioning in the native European otter (*Lutra lutra*) and invasive American mink (*Neovison vison*). *bioRxiv*.

Heo, Y.-J., Jo, H., Kim, J. Y., Kim, G.-Y., Joo, G.-J., & Kim, H.-W. (2023). Application of DNA metabarcoding for identifying the diet of Asian clam (*Corbicula fluminea*, Müller, 1774). *Sustainability, 15*(1), 441.

Homma, C., Inokuchi, D., Nakamura, Y., Uy, W. H., Ohnishi, K., Yamaguchi, H., et al. (2022). Effectiveness of blocking primers and a peptide nucleic acid (PNA) clamp for 18S metabarcoding dietary analysis of herbivorous fish. *PLoS ONE, 17*(4), e0266268. https://doi.org/10.1371/journal.pone.0266268

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN community edition - Interactive exploration and analysis of large-scale microbiome sequencing data. *PLOS Computational Biology, 12*(6), e1004957.

Ingala, M. R., Simmons, N. B., Wultsch, C., Kramvis, K., Provost, K. L., & Perkins, S. L. (2021). Molecular diet analysis of neotropical bats based on fecal DNA metabarcoding. *Ecology and Evolution, 11*(8), 4378-4391.

Kartzinel, T. R., & Pringle, R. M. (2015). Molecular detection of invertebrate prey in vertebrate diets: Trophic ecology of Caribbean island lizards. *Molecular Ecology Resources, 15*(4), 903-914.

Kindler, C., Graciá, E., & Fritz, U. (2018). Extra-Mediterranean glacial refuges in barred and common grass snakes (*Natrix helvetica*, *N. natrix*). *Scientific Reports, 8*, 1821.

Kindler, C., & Fritz, U. (2018). Phylogeography and taxonomy of the barred grass snake (*Natrix helvetica*), with a discussion of the subspecies category in zoology. *Vertebrate Zoology, 68*(3), 269-281. https://doi.org/10.3897/vz.68.e31615

Kitano, T., Umetsu, K., Tian, W., & Osawa, M. (2007). Two universal primer sets for species identification among vertebrates. *International Journal of Legal Medicine, 121*(5), 423-427.

Kurita, T., & Toda, M. (2022). Comparison of morphological identification and DNA metabarcoding for dietary analysis of faeces from a subtropical lizard. *Wildlife Research*. https://doi.org/10.1071/WR22054

Lacoursière-Roussel, A., Dubois, Y., Normandeau, E., & Bernatchez, L. (2016). Improving herpetological surveys in eastern North America using the environmental DNA method. *Canadian Journal of Fisheries and Aquatic Sciences, 73*(6), 991-999.

Luiselli, L., & Agrimi, U. (1991). Composition and variation of the diet of *Vipera aspis francisciredi* in relation to age and reproductive stage. *Amphibia-Reptilia, 12*(2), 137-144.

Luiselli, L. (2006). Resource partitioning and interspecific competition in snakes: The search for general geographical and guild patterns. *Oikos, 114*(2), 193-211.

Luiselli, L., Capula, M., Capizzi, D., Filippi, E., Trujillo Jesus, V., & Anibaldi, C. (2005). Geographic variation in the diet composition of the grass snake (*Natrix natrix*) along the mainland and an island of Italy: The effects of habitat type and interference with potential competitors. *Herpetological Journal, 15*(3), 221-230.

Luiselli, L., Capula, M., & Shine, R. (1997). Food habits, growth rates, and reproductive biology of grass snakes, *Natrix natrix* (Colubridae) in the Italian Alps. *Journal of Zoology, 241*(3), 371-380.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal, 17*(1), 10-12.

Massey, A. L., Rehberg, G. H., Harper, T., Varman, T., Jensen, J. M., Allen, J. M., & Levi, T. (2021). Comparison of mechanical sorting and DNA metabarcoding for diet analysis with fresh and degraded wolf scats. *Ecosphere, 12*(7), e03682.

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE, 8*(4), e61217.

Mitchell, L. J., Horsburgh, G. J., Dawson, D. A., Maher, K. H., & Arnold, K. E. (2022). Metabarcoding reveals selective dietary responses to environmental availability in the diet of a nocturnal, aerial insectivore, the European Nightjar (*Caprimulgus europaeus*). *Ibis, 164*(1), 60-73.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Kondoh, M., & Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *Royal Society Open Science, 2*(7), 150088.

Mondino, A., Crovadore, J., Lefort, F., & Ursenbacher, S. (2022). Impact of invading species on biodiversity: Diet study of the green whip snake's (*Hierophis viridiflavus*, L. 1789) in Switzerland. *Global Ecology and Conservation, 38*, e02239.

Oksanen, J., Simpson, G. L., Blanchet, F. G., & Kindt, R. (2022). Vegan community ecology package version 2.6-2 April 2022. The Comprehensive R Archive Network.

Pinho, C. J., Darwish, M., Smíd, J., Carranza, S., & Vasconcelos, R. (2023). Green matters: Dietary assessment of a reptile community using DNA metabarcoding. *Global Ecology and Conservation, 47*, e02667.

Pinho, C. J., Santos, B., Mata, V. A., Seguro, M., Nascimento, R., & Lopes, R. J. (2018). What is the giant wall gecko having for dinner? Conservation genetics for guiding reserve management in Cabo Verde. *Genes, 9*(11), 532.

Reading, C. J., & Davies, J. L. (1996). Predation by grass snakes (*Natrix natrix*) at a site in southern England. *Journal of Zoology, 239*(1), 73-82.

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: Inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research, 39*, e145.

Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation, 17*, e00547.

Sakata, M. K., Kawata, M. U., Kurabayashi, A., Kurita, T., Nakamura, M., Shirako, T., ... & Minamoto, T. (2022). Development and evaluation of PCR primers for environmental DNA (eDNA) metabarcoding of Amphibia. *Metabarcoding and Metagenomics, 6*, e76534.

Sarkis, C. M., Hoenig, B. D., Seney, E. E., Gaspar, S. A., & Forsman, A. M. (2022). Sea snacks from DNA tracks: Using DNA metabarcoding to characterize the diet of green turtles (*Chelonia mydas*). *Integrative and Comparative Biology, 62*(2), 223-236. https://doi.org/10.1093/icb/icac080

Swinehart, A., Partridge, C., Russell, A., Thacker, A., Kovach, J., & Moore, J. (2023). Diet of a threatened rattlesnake (eastern massasauga) revealed by DNA metabarcoding. *Ecology and Evolution, 13*(4), e10029.

Tamagnini, D., Stephenson, J., Brown, R. P., & Meloro, C. (2018). Geometric morphometric analyses of sexual dimorphism and allometry in two sympatric snakes: *Natrix helvetica* (Natricidae) and *Vipera berus* (Viperidae). *Zoology, 129*, 25-34.

Tosa, M. I., Lesmeister, D. B., Allen, J. M., & Levi, T. (2023). Multi-locus DNA metabarcoding reveals seasonality of foraging ecology of western spotted skunks in the Pacific Northwest. *Ecosphere, 14*(8), e4536.

Tsuji, S., Shibata, N., Sawada, H., & Ushio, M. (2020). Quantitative evaluation of intraspecific genetic diversity in a natural fish population using environmental DNA analysis. *Molecular Ecology Resources, 20*(5), 1323-1332.

Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., Murata, K., Nikaido, M., Sado, T., Sato, Y., Takeshita, M., Iwasaki, W., Yamanaka, H., Kondoh, M., & Mira, M. (2017). Environmental DNA enables detection of terrestrial mammals from forest pond water. *Molecular Ecology Resources, 17*(6), e63-e75.

Valsecchi, E., Bylemans, J., Goodman, S. J., Lombardi, R., Carr, I., Castellano, L., & Galimberti, A. (2020). Novel universal primers for metabarcoding environmental DNA

surveys of marine mammals and other marine vertebrates. *Environmental DNA, 2*(4), 460-476.

Vences, M., Lyra, M. L., Perl, R. G. B., Bier, H. C., Bletz, M. C., Stanković, D., Lopes, C. M., Jarek, M., Bhuju, S., Gerfers, R., Haddad, C. F. B., & Steinfartz, S. (2016). Freshwater vertebrate metabarcoding on Illumina platforms using double-indexed primers of the mitochondrial 16S rRNA gene. *Conservation Genetics Resources, 8*, 323-327.

Wang, Z., Liu, X., Liang, D., Wang, Q., Zhang, L., & Zhang, P. (2023). VertU: Universal multilocus primer sets for eDNA metabarcoding of vertebrate diversity, evaluated by both artificial and natural cases. *Frontiers in Ecology and Evolution, 11*, 1164206. https://doi.org/10.3389/fevo.2023.1164206

CHAPTER FIVE

GENERAL DISCUSSION

This thesis explored the efficiency of using eDNA-based two different approaches: (1) a Specific-species qPCR assay (Barcoding) for monitoring and a Metabarcoding approach to analyse the dietary composition of *Natrix helvetica*. Focusing on two populations in Norfolk and Holbrook, we sought to understand how local environmental factors, prey availability, and snake biology influence these aspects. Our results contribute to the growing body of research on the conservation and ecological needs of *Natrix helvetica* while addressing several methodological challenges in monitoring and dietary analysis.

The CMR surveys conducted at Holbrook Marsh and Heath in 2022 and 2023 revealed a growing population of *Natrix helvetica*. The population estimates of 19.13 in 2022 and 34.08 in 2023 likely reflect the favourable habitat conditions of recent restoration efforts. These efforts, such as introducing artificial cover objects (ACOs), provided critical microhabitats, including basking spots, shaded refuges, and hibernation sites, which are essential for the survival and reproduction of *Natrix helvetica*. The presence of neonates in 2023 further supports the conclusion that this habitat is conducive to successful reproduction. Moreover, the high recapture of juveniles and sub-adults in both surveys suggests a relatively young population, which may result from favourable conditions for juvenile recruitment (Luiselli et al., 2006).

However, CMR has its limitations. The relatively short duration of our study limits our ability to draw long-term conclusions about population trends. Seasonal and annual variations in population size could affect estimates. Thus, long-term monitoring is necessary to provide more accurate estimates of population dynamics. In addition, using ACOs may introduce bias in detecting certain age groups or individuals, as smaller snakes are often less likely to be detected than adults (Lock & Griffiths, 2022).

The application of eDNA as a monitoring tool for *Natrix helvetica* provided valuable information into the potential and limitations of this non-invasive method. Our species-specific qPCR assay was highly effective in detecting *Natrix helvetica* in aquatic environments, confirming the utility of eDNA for monitoring semi-aquatic reptiles. However, terrestrial roller

swab detection was less successful, with lower detection rates and higher inconclusive results.

One key factor that we observed influencing the success of eDNA detection in terrestrial environments was the material and structure of the ACOs. Previous studies have shown that smooth, non-porous surfaces, such as metal or asphalt shingles, are more effective at capturing surface-deposited DNA (Kyle et al., 2022). In contrast, the corrugated bitumen ACOs used in our study may have reduced the contact area between the snakes and the surface, leading to lower eDNA recovery. Moreover, the porous nature of the material may have trapped eDNA, making it difficult to collect during swabbing. Future studies should consider using ACOs with smoother, flat surfaces to improve eDNA recovery in terrestrial environments.

Another factor that may have contributed to the lower detection rates in roller swab samples is the biology and behaviour of *N. helvetica*. As a semi-aquatic species, *N. helvetica* spends time in the water, where eDNA dispersal is more uniform and easier to detect. In contrast, terrestrial environments may present challenges for eDNA collection, as eDNA degradation occurs more due to microbial activity, UV exposure, and environmental inhibitors such as humic acids. Despite these challenges, the roller swab method still holds potential for detecting terrestrial reptiles, provided that modifications are made to improve eDNA recovery.

The metabarcoding analysis of diet composition revealed no significant differences in prey composition between the Norfolk and Holbrook populations. Amphibians dominated the diet of the Norfolk population, with *Bufo bufo* being the most frequently consumed species. This result aligns with the wetland habitat of Norfolk, which supports an abundance of amphibians, particularly anurans. In contrast, the Holbrook population, including amphibians, small mammals, and birds, displayed a more diverse diet. This variation in diet suggests that local prey availability plays a role in shaping the diet of *Natrix helvetica*, with snakes in different habitats adapting their feeding behaviour to the prey types available. Despite the ecological differences between the two sites, the dietary differences were not statistically significant at the species or order level, as indicated by PERMANOVA results. This lack of significance may be due to the small sample size, particularly in the Holbrook population, where the species accumulation curve plateaued early, suggesting limited species detection.

Future studies should aim to increase sample sizes and extend the study duration to capture a more complete picture of the dietary habits of these populations.

The ontogenetic shifts in diet observed in *N. helvetica* further highlight its feeding ecology. Sub-adults were found to consume smaller prey, such as newts (*Triturus cristatus*) and small mammals (*Sorex minutus*), while adults consumed larger prey, including birds and larger amphibians. This shift in prey size and type is consistent with the idea that snakes can handle larger prey as they grow due to their increased body size and stronger bite force (Borczyk, 2014). The dietary flexibility of *N. helvetica* suggests that it is a generalist predator, capable of adapting its diet to local prey availability, which may be an advantage in changing environments.

Sex-based differences in diet were also observed, with females displaying a more diverse diet than males. This is likely due to the reproductive needs of females, which require higher energy to support egg production and post-laying recovery. Females also tend to have wider heads than males, allowing them to capture a broader range of prey sizes (Luiselli et al., 2005). These findings are consistent with previous studies on sexually dimorphic feeding behaviour in reptiles, where females often exhibit greater dietary diversity due to their higher energetic demands during reproduction (Reading & Davies, 1996).

The findings of this study have important implications for the conservation of *N. helvetica* and other semi-aquatic reptiles. The observed population growth at Holbrook suggests that habitat restoration efforts, particularly the creation of ponds and other wetland features, can provide critical habitats for *N. helvetica* and support their reproductive success. Conservation strategies should continue to focus on maintaining and enhancing these habitats, particularly in areas where amphibian populations are declining.

Using eDNA to monitor *N. helvetica* offers a promising non-invasive method for tracking populations, especially in aquatic environments. However, to improve the effectiveness of eDNA in terrestrial environments, future studies should explore alternative sampling methods and materials for ACOs and refine eDNA extraction and amplification techniques. Additionally, combining eDNA with other methods, such as CMR or radio

telemetry, could provide a more comprehensive understanding of population dynamics and habitat use.

Regarding dietary analysis, metabarcoding has proven to be a powerful tool for identifying a broad range of prey species, providing insights into the feeding ecology of *N. helvetica*. Future research should aim to incorporate seasonal sampling to capture potential changes in diet over time and investigate the dietary habits of *N. helvetica* in other regions. By expanding the scope of dietary studies, we can better understand how local prey availability and habitat structure influence the feeding behaviour of this species.

Finally, the integration of CMR, eDNA, and metabarcoding in my thesis demonstrates the value of using multiple methods to study the ecology of elusive species. Each method has its strengths and limitations, but when used together, they provide a more comprehensive picture of population size, habitat use, and dietary habits. This integrative approach should be adopted in future conservation studies to maximise the accuracy and effectiveness of monitoring efforts.

References

Borczyk, B. (2014). Allometry of head size and shape dimorphism in the grass snake (*Natrix natrix* L.). Turkish Journal of Zoology, 38(6), 734-738.

Kyle, K. E., Allen, M. C., Dragon, J., Bunnell, J. F., Reinert, H. K., Zappalorti, R., Jaffe, B. D., Angle, J. C., & Lockwood, J. L. (2022). Combining surface and soil environmental DNA with artificial cover objects to improve terrestrial reptile survey detection. Conservation Biology, 36(6), e13939. https://doi.org/10.1111/cobi.13939

Lock, M. M. G., & Griffiths, R. A. (2022). Detectability of reptiles in standardised surveys: A test using grass snake Natrix helvetica models. The Herpetological Journal, 32(4), 183-189. https://doi.org/10.33256/32.4.183189

Luiselli, L. (2006). Resource partitioning and interspecific competition in snakes: The search for general geographical and guild patterns. Oikos, 114(2), 193-211.

Reading, C. J., & Davies, J. L. (1996). Predation by grass snakes (*Natrix natrix*) at a site in southern England. Journal of Zoology, 239(1), 73-82.