Trophic ecology of the Eurasian otter

(Lutra lutra)

A thesis submitted to Cardiff University for the degree of Doctor of Philosophy
(PhD) in the School of Biosciences by

Lorna Elizabeth Drake

September 2020
Acknowledgements

I would like to thank my funders - KESS II (Knowledge Economy Skills Scholarship) and the Wildlife Trust of South and West Wales who partnered this project. I would also like to thank my supervisors; Liz Chadwick for her support and guidance over the four years, Bill Symondson for always having his door open to me whether I was stressed, upset or happy, and Robbie McDonald for making time for me whenever I got the chance to travel back down to Cornwall.

I would like to thank my family for their emotional (and financial) support, particularly my mum and dad who listened to my worries and believed in me when I didn’t believe in myself, as well as my cousin Christian who gave me somewhere to live during my unfunded period.

Thank you to the Wildlife Trust of South and West Wales for welcoming me into their team and giving me opportunities to engage with projects outside of academia. Thank you to everyone who has contributed to the project, including Cardiff University Otter Project members for collecting samples during post mortems, Jez and Fred for their help with data analyses and Angela for putting so much time and effort into helping me solve my metabarcoding issues.

Thank you to members of the molecular ecology lab, who provided a welcoming, friendly atmosphere during my time at Cardiff and provided words of wisdom throughout the PhD. Thank you to the members of Team Otter, especially Nia who has been an amazing source of support and knowledge since I started my PhD, for making meetings enjoyable and helping out with the Otters in the City event each year. Thank you to the members of team Bill for being great friends and suffering through the turmoils of DNA metabarcoding alongside me, I couldn’t have done this without your support and help with trouble shooting issues. I would also like to thank those who dragged me away from the office to climb or dance, giving me a chance to relax and switch off from work.

Special thanks go to Indie, Emma, Jordan, Becca, Dan and Alex. You all supported me as I went through my toughest moments, listened to me moan, distracted me when I needed it and celebrated in my successes with me. Without you guys I could not have completed this project and I will forever be grateful to you!
# Table of Contents

*Declarations* .......................................................................................................................... ii  

*Acknowledgements* .................................................................................................................. iii  

*List of Figures* ............................................................................................................................ ix  

*List of Tables* ............................................................................................................................. xii  

*Thesis Summary* .......................................................................................................................... 1  

**Chapter One – General Introduction** .................................................................................... 3  

1.1 Freshwater ecosystems ........................................................................................................... 3  

1.2 Food webs and species interactions ....................................................................................... 4  

1.3 Methods for investigating trophic interactions ...................................................................... 5  

1.3.1 Traditional methods ............................................................................................................ 5  

1.3.2 Stable isotope analysis ....................................................................................................... 6  

1.3.3 DNA metabarcoding ......................................................................................................... 10  

1.4 Eurasian otter: ecology and diet ............................................................................................ 12  

1.5 Aims and hypotheses ............................................................................................................. 15  

**Chapter Two – There’s something fishy going on: a stable isotope analysis of marine-derived nutrients in the diet of Eurasian otters (Lutra lutra)** ................................................................ 17  

2.1 Abstract ............................................................................................................................... 17  

2.2 Introduction ........................................................................................................................... 17  

2.3 Methods ............................................................................................................................... 19  

2.3.1 Sample and data collection ............................................................................................... 19  

2.3.2 Location data ................................................................................................................... 20  

2.3.3 Body Condition Index ...................................................................................................... 22  

2.3.4 Stable isotope analysis ..................................................................................................... 22  

2.3.5 Statistical analysis .......................................................................................................... 23  

2.4 Results .................................................................................................................................. 25  

2.4.1 Variation in $\delta^{15}$N ...................................................................................................... 26  

2.4.2 Variation in $\delta^{13}$C ....................................................................................................... 29  

2.4.3 Nitrogen and carbon isotopic signatures ......................................................................... 30  

2.5 Discussion ............................................................................................................................ 33  

2.5.1 Distance from the coast .................................................................................................... 33  

2.5.2 Anadromous fish abundance ......................................................................................... 35
Chapter Four – True or false: identifying artefacts in metabarcoding data and methods to reduce their prevalence.................................................................128

4.1 Abstract ..................................................................................................................128

4.2 Introduction..............................................................................................................128

4.3 Methods..................................................................................................................131

4.3.1 Otter faecal samples and data collection ...............................................................131
4.3.2 DNA metabarcoding .............................................................................................131
4.3.3 Sequence analysis .................................................................................................135
4.3.4 Minimum Sequence Copy Thresholds (MSCTs) ..................................................136

4.4 Results.....................................................................................................................138

4.4.1 No filter (‘No Filter’)..............................................................................................139
4.4.2 Remove singleton reads (‘Singletons’) .................................................................144
4.4.3 Remove read counts less than 10 (‘< 10’)............................................................144
4.4.4 Remove maximum taxon contamination (‘Max Contamination’) .........................144
4.4.5 Proportion of total read count (‘Total %’)............................................................144
4.4.6 Proportion of read count per sample (‘Sample %’) ...............................................145
4.4.7 Proportion of read count per taxon (‘Taxon %’)....................................................145
4.4.8 Combining methods ..............................................................................................146
4.4.9 Repeats ................................................................................................................148
4.4.10 NMDS analysis ..................................................................................................148

4.5 Discussion ..............................................................................................................150

4.5.1 Identifying artefacts .............................................................................................150
4.5.2 Performance of Minimum Sequence Copy Thresholds (MSCTs) .........................152
4.5.3 Combining MSCTs ..............................................................................................153
4.5.4 Choosing an appropriate threshold ......................................................................153
4.5.5 Conclusions ........................................................................................................155

4.6 Acknowledgements ...............................................................................................155

4.7 Supplementary Information ...................................................................................156

4.7.1 Literature review of the use of minimum sequence copy thresholds ................156
4.7.2 Preliminary metabarcoding experiment ................................................................159
4.7.3 Bioinformatics scripts ..........................................................................................163
4.7.4 Scripts for visualising metabarcoding data in R following application of minimum sequence copy thresholds .................................................................169
4.7.5 Additional figures for visualising effectiveness of different minimum sequence copy thresholds .........................................................................................181
Chapter Five – Otterly delicious: Investigating spatio-temporal variation in the diet of a recovering population of Eurasian otters (Lutra lutra) through DNA metabarcoding and morphological analysis of prey remains

5.1 Abstract .............................................................................................................. 188
5.2 Introduction......................................................................................................... 188
5.3 Methods ............................................................................................................. 191
  5.3.1 Sample and data collection............................................................................ 191
  5.3.2 Morphological analysis............................................................................... 192
  5.3.3 DNA metabarcoding analysis...................................................................... 192
  5.3.4 Comparison of methods.............................................................................. 194
  5.3.5 Statistical analysis....................................................................................... 194
5.4 Results ................................................................................................................ 200
  5.4.1 Morphological analysis............................................................................... 200
  5.4.2 DNA metabarcoding analysis...................................................................... 200
  5.4.3 Comparison of methods.............................................................................. 201
  5.4.4 Dietary variation.......................................................................................... 204
5.5 Discussion .......................................................................................................... 211
  5.5.1 Comparison of methods.............................................................................. 211
  5.5.2 Dietary Composition.................................................................................... 211
  5.5.3 Spatial variation............................................................................................ 212
  5.5.4 Temporal variation....................................................................................... 213
  5.5.5 Biotic variation............................................................................................. 215
  5.5.6 Limitations.................................................................................................... 216
  5.5.7 Conclusions.................................................................................................. 216
5.6 Acknowledgements ............................................................................................ 217
5.7 Supplementary Information .............................................................................. 218
  5.7.1 Correlation between landscape variables.................................................. 218
  5.7.2 Scripts for analysing data in R .................................................................... 218
  5.7.3 Taxonomic identifications and filtering of dietary data............................... 225

Chapter Six – General Discussion ......................................................................... 243
6.1 Aims and Overview ............................................................................................ 243
6.2 Summary of findings and comparison of SIA with NGS .................................. 244
  6.2.1 Associations between otter diet and spatial changes in nutrient influxes and prey availability ................................................................. 244
  6.2.2 Associations between otter diet and temporal changes in prey availability 246
6.2.3 Identifying differences in diet between otters of different demographic groups..........................................................247
6.2.4 Optimising metabarcoding protocols to decrease the impact of artefacts on dietary data .................................................................248
6.2.5 Comparison of methods for analysing diet of otters..........................249
6.2.6 Interpreting changes in otter diet over different timescales ......................250

6.3 Implications for species conservation and ecosystem monitoring......251
6.3.1 Monitoring habitats ........................................................................251
6.3.2 Impacts of variable diets for otter populations .................................251
6.3.3 Predation impacts on prey species ....................................................252
6.3.4 Predation impacts on stocked fish and conflict with aquaculture .............253

6.4 Future directions ..............................................................................253
6.4.1 Expanding analyses to other populations.............................................253
6.4.2 Ecosystem level assessments .............................................................253
6.4.3 Assessing otter health in relation to dietary variation.............................254
6.4.4 Individual dietary variation assessments ................................................255
6.4.5 Investigating eel-otter dynamics .........................................................255
6.4.6 Improving metabarcoding analyses ....................................................255

6.5 Conclusions ......................................................................................256

Bibliography ...........................................................................................257
List of Figures

Figure 1.1. Illustration of isotopic signatures in $\delta^{15}$N and $\delta^{13}$C from primary producers (plants and zooplankton), herbivores and carnivores ................................................................. 9

Figure 2.1. Otter sampling locations. Distribution of Eurasian otter (*Lutra lutra*) carcasses collected 1993 – 2007 and analysed for carbon and nitrogen stable isotopes using (a) bone and (b) muscle tissue ........................................................................... 21

Figure 2.2. Isotopic niche measured in Eurasian otter (*Lutra lutra*) bone (green) and muscle (purple) tissue samples ........................................................................................................... 25

Figure 2.3. Relationship between $\delta^{15}$N ratio and abundance of anadromous fish .... 27

Figure 2.4. Relationship between $\delta^{15}$N ratio and distance to the coast ..................... 28

Figure 2.5. Relationship between $\delta^{13}$C ratio and abundance of anadromous fish .... 29

Figure 2.6. Spatial variability in isotopic signatures of Eurasian otters (*Lutra lutra*) sampled across Wales and bordering regions from 1993 – 2007 .............................. 31

Figure 2.7. Isotopic niche measured in Eurasian otter (*Lutra lutra*) muscle tissue samples taken in each season ........................................................................................................... 32

Figure S 2.1. Locations of river catchments (dark grey) in Wales and bordering regions used to match Eurasian otter (*Lutra lutra*) samples to anadromous fish abundance data ........................................................................................................... 39

Figure S 2.2. Anadromous fish abundance in river catchments across Wales and bordering regions from 1993 - 2007 ........................................................................................................... 40

Figure S 2.3. Bayesian standard ellipse area (SEA) sizes and their credible intervals for isotopic signatures of bone and muscle tissue collected from Eurasian otters (*Lutra lutra*) from Mid-England and Wales between 1993 - 2007 ............................................. 61

Figure S 2.4. Bayesian standard ellipse area (SEA) sizes and their credible intervals for isotopic signatures of muscle tissue from male and female Eurasian otters (*Lutra lutra*) in each season, sampled from Mid-England and Wales between 1993 - 200762

Figure 3.1. Spatial distribution of Eurasian otter (*Lutra lutra*) samples ..................... 68

Figure 3.2. Isotopic niche of Eurasian otters (*Lutra lutra*) based on values obtained from multiple whisker subsamples (green) and values obtained from only the basal subsample (purple) ........................................................................................................... 73

Figure 3.3. Variation in isotopic signatures of Eurasian otters (*Lutra lutra*) across England and Wales from 2007 to 2016 ........................................................................................................... 75
Figure 3.4. Variation in $\delta^{15}$N of Eurasian otters (*Lutra lutra*) utilising different aquatic habitat types ........................................................... 77

Figure 3.5. Variation in $\delta^{13}$C of Eurasian otters (*Lutra lutra*) in each season .......... 78

Figure S 3.1. Associations between spatial variables ........................................ 112

Figure S 3.2. Bayesian standard ellipse area (SEA) sizes and their credible intervals for isotopic signatures obtained from Eurasian otters (*Lutra lutra*) across England and Wales from 2007 to 2016 using multiple whisker subsamples and basal whisker subsamples only ........................................................................ 120

Figure S 3.3. Spatial variation in $\delta^{15}$N (a) and $\delta^{13}$C (b) isotopic signatures from basal whisker subsamples of Eurasian otters (*Lutra lutra*) sampled across Wales and bordering regions from 1993 – 2007 .......................................................................................................................... 121

Figure S 3.4. The mean predicted variation in $\delta^{15}$N (a) and $\delta^{13}$C (b) isotopic signatures (i.e. how confident we can be in predictions) of Eurasian otters (*Lutra lutra*) sampled between 2007 and 2016 .................................................................................................................. 122

Figure S 3.5. Variation in $\delta^{15}$N of Eurasian otters (*Lutra lutra*) utilising different aquatic habitat types ........................................................... 123

Figure S 3.6. Variation in $\delta^{13}$C of Eurasian otters (*Lutra lutra*) utilising different aquatic habitat types ........................................................... 124

Figure S 3.7. Variation in $\delta^{13}$C of Eurasian otters (*Lutra lutra*) in each size class .. 125

Figure S 3.8. Variation in $\delta^{15}$N of Eurasian otters (*Lutra lutra*) in each size class .. 126

Figure S 3.9. Within-individual variation (standard deviation) in $\delta^{15}$N of Eurasian otters (*Lutra lutra*) in each size class .......................................................................................................................... 127

Figure 4.1. Eurasian otter (*Lutra lutra*) diet 16S prey counts .................................... 140

Figure 4.2. Eurasian otter (*Lutra lutra*) diet COI prey counts ................................... 142

Figure 4.3. Non-metric multidimensional scaling of relative frequency of occurrence for each taxon following application of different minimum sequence copy thresholds, including different methods and thresholds, to 16S (a) and COI (b) data ............... 149

Figure S 4.1. Eurasian otter (*Lutra lutra*) diet 16S reads ......................................... 182

Figure S 4.2. Eurasian otter (*Lutra lutra*) diet COI reads ....................................... 183

Figure S 4.3. Eurasian otters (*Lutra lutra*) with marine prey present in their diet following application of minimum sequence copy thresholds on 16S data (blue dots) .......................................................................................................................... 184
Figure S 4.4. Eurasian otters (*Lutra lutra*) with marine prey present in their diet following application of minimum sequence copy thresholds on COI data (blue dots) ..........................................................186

Figure 5.1. Summary of the metabarcoding workflow utilised in this study. ...............193

Figure 5.2. Total number of reads remaining in both 16S and COI datasets at each stage of the metabarcoding and data filtering process ..........................................................201

Figure 5.3. Taxon presences in the diet of Eurasian otters (*Lutra lutra*) identified using morphological analysis of prey remains (orange) and DNA metabarcoding (blue) on faecal samples .................................................................................................203

Figure 5.4. Presences of prey groups (grey) and the taxa that contributed to each prey group (orange) in the diet of Eurasian otters (*Lutra lutra*).................................205

Figure 5.5. Model-based unconstrained residual ordination biplot for Eurasian otter (*Lutra lutra*) diet ..................................................................................................................206

Figure 5.6. Frequency of occurrence of prey items in the diet of Eurasian otters (*Lutra lutra*) in each season .................................................................208

Figure 5.7. Frequency of occurrence of prey items in the diet of Eurasian otters (*Lutra lutra*) at different longitudes ............................................................................209

Figure 5.8. Frequency of occurrence of prey items in the diet of Eurasian otters (*Lutra lutra*) at different coastal proximities .................................................................210

Figure S 5.1. Associations between spatial variables .................................................218

Figure S 5.2. Taxon presences in the diet of Eurasian otters (*Lutra lutra*) using DNA metabarcoding on faecal samples .................................................................238

Figure S 5.3. Taxon presences in the diet of Eurasian otters (*Lutra lutra*) using morphological analysis of prey remains and DNA metabarcoding on faecal samples for fish taxa (a) and non-fish taxa (b) .................................................................239

Figure S 5.4. Frequency of occurrence of common rudd (*Scardinius erythrophthalmus*) and common roach (*Rutilus rutilus*) in the diet of Eurasian otters (*Lutra lutra*) in each season. .................................................................241

Figure S 5.5. Frequency of occurrence of European eel (*Anguilla anguilla*) and European bullhead (*Cottus gobio*) in the diet of Eurasian otters (*Lutra lutra*) at different coastal proximities .................................................................242
List of Tables

**Table S 2.1.** Comparison of component models used to make the averaged model for the general linear model (GLM) of muscle $\delta^{15}N$ and general additive models (GAM) of bone and muscle $\delta^{13}C$ in Eurasian otters (*Lutra lutra*) from mid-England, Wales and bordering counties 1993 - 2007 .................................................................59

**Table S 2.2.** Conditional average model outputs for general linear model (GLM) of muscle $\delta^{15}N$ and general additive models (GAM) of bone and muscle $\delta^{13}C$ in Eurasian otters (*Lutra lutra*) from mid-England, Wales and bordering counties 1993 – 2007 .................................................................................................................60

**Table 3.1.** Associated metadata relating to the 300 Eurasian otters (*Lutra lutra*) selected for the study ........................................................................................................................................69

**Table S 3.1.** Conditional average model output for general additive mixed effects model (GAMM) of $\delta^{15}N$ in Eurasian otter (*Lutra lutra*) whisker samples from England and Wales 2007-2016 ........................................................................................................................................113

**Table S 3.2.** Conditional average model output for general additive mixed effects model (GAMM) of $\delta^{13}C$ in Eurasian otter (*Lutra lutra*) whisker samples from England and Wales 2007-2016 ........................................................................................................................................115

**Table S 3.3.** Conditional average model output for general additive model (GAM) of $\delta^{15}N$ in basal segments of Eurasian otter (*Lutra lutra*) whiskers from England and Wales 2007-2016 ........................................................................................................................................116

**Table S 3.4.** Conditional average model output for general additive model (GAM) of $\delta^{13}C$ in basal segments of Eurasian otter (*Lutra lutra*) whiskers from England and Wales 2007-2016 ........................................................................................................................................117

**Table S 3.5.** Conditional average model output for general additive model (GAM) of individual variation in $\delta^{15}N$ from Eurasian otter (*Lutra lutra*) whisker samples taken across England and Wales 2007-2016 ........................................................................................................................................118

**Table S 3.6.** Conditional average model output for general additive model (GAMM) of individual variation in $\delta^{13}C$ from Eurasian otter (*Lutra lutra*) whisker samples taken across England and Wales 2007-2016 ........................................................................................................................................119

**Table 4.1.** Seven methods often applied to eDNA metabarcoding datasets, selected from those identified in the literature (Table S 4.1) .......................................................................................................................................137
Table 4.2. Performance of different minimum sequence copy thresholds on 16S data
...........................................................................................................................................141

Table 4.3. Performance of different minimum sequence copy thresholds on COI data
...........................................................................................................................................143

Table 4.4. Success of different filtering methods in achieving the key objectives of post-bioinformatic data clean-up. Green, orange and red denote positive, neutral and negative outcomes, respectively.................................................................147

Table S 4.1. Application of minimum sequence copy thresholds across 154 studies conducting metabarcoding on eukaryotic DNA for environmental monitoring or dietary analysis.................................................................156

Table S 4.2. Taxa used for primer testing ..............................................................................161

Table 5.1. Taxa identified in Eurasian otter (Lutra lutra) faecal samples from across England and Wales between 2007-2016, along with method of identification, the 'prey group' taxa were clustered into and the reason why taxa were grouped .......196

Table S 5.1. Taxa obtained from morphological analysis of prey remains in Eurasian otter (Lutra lutra) faecal samples .........................................................................................225

Table S 5.2. Taxa obtained from sequencing Eurasian otter (Lutra lutra) faecal samples using 16S primers.......................................................................................................................226

Table S 5.3. Taxa obtained from sequencing Eurasian otter (Lutra lutra) faecal samples using COI primers.......................................................................................................................230
Thesis Summary

Generalist apex predators have broad diets, consuming species from a range of trophic levels and connecting multiple energetic pathways. Investigating the diet of generalist apex predators thus provides a unique opportunity to describe trophic interactions over space and time, and can be used to gain an insight into ecosystem level changes.

Broad spatio-temporal and demographic variation in the diet of the Eurasian otter (*Lutra lutra*), an apex predator of freshwater ecosystems, were investigated using a biobank of tissue samples and data collected from dead otters in England and Wales over 23 years. Molecular methods were utilised to gain a greater insight into otter diet, with stable isotope analysis revealing broad-scale variation in nutrient assimilation, and DNA metabarcoding provided a more precise description of the species consumed by otters.

Isotopic signatures suggested nutrient assimilation by otters was driven by changes in basal nutrient levels and prey availability. Variation in nitrogen isotopes may reflect landscape scale differences in anthropogenic inputs from fertilisers, which were suggested to enrich basal nitrogen signatures within the environment and thus enrich signals throughout the trophic network, with some additional changes driven by differences in the consumption of high trophic level prey by otters. Variation in carbon isotopes of otters reflected changes in the availability of marine-derived nutrients, with otters sampling such nutrients primarily via consumption of marine prey at the coast and anadromous fish inland.

Analysis of DNA metabarcoding data revealed otters consumed a broad range of prey from a variety of habitats. Stocked fish, invasive species and protected species occurred in the diet of otters infrequently, with the exception of the critically endangered European eel (*Anguilla anguilla*), which was one of the most frequently consumed prey items. Dietary variation primarily reflected seasonal changes in prey availability and differences in prey distributions across the landscape pertaining to longitude and coastal proximity. These findings displayed the opportunistic foraging behaviour of otters and their ability to switch predation to more abundant prey.

Prior to ecological analyses of DNA metabarcoding data, a range of minimum sequence copy thresholds were tested to remove erroneous data and provide a more accurate description of the diet of otters. Of these thresholds, combining the removal of
a percentage of reads per sample with removal of the maximum read count in a blank control per taxon performed the best.

Overall the findings of this study demonstrate the high levels of dietary plasticity exhibited by otters. Such adaptability in foraging is suggested to have aided population recovery and distribution expansion by British otter populations and provide resilience to future environmental stressors. The study also highlights how an understanding of the dietary variation exhibited by generalist apex predators can be utilised to investigate trophic interactions and nutrient flows across freshwater ecosystems.
1.1 Freshwater ecosystems

Freshwater ecosystems support a range of environmental processes (e.g. nutrient cycling and carbon storage) and provide essential ecosystem services and socio-economic benefits for humans (e.g. drinking water and recreational activities (Millennium Ecosystem Assessment 2005; Cole et al. 2007; Strayer and Dudgeon 2010; Carpenter et al. 2011). Whilst freshwater ecosystems are some of the most biodiverse habitats on the planet, they are also among the most threatened (Sala et al. 2000; Balian et al. 2007) with more than 25% of freshwater animal species predicted to already be threatened or extinct (Keskin et al. 2016). Species losses can be attributed to a range of anthropogenic factors, including: overexploitation, pollution (e.g. pesticides), modification of river flow (e.g. dams), habitat degradation and introduction of non-native, invasive species (e.g. signal crayfish, *Pacifastacus leniusculus*; Malmqvist and Rundle 2002; Dudgeon et al. 2006; Carpenter et al. 2011; Cazzolla Gatti 2016). Freshwater habitats are particularly vulnerable to anthropogenic activities due to their position in the landscape; not only are freshwater habitats often in basins of valleys, resulting in rivers and lakes receiving sediment, waste and pollutants via run-off, but they are also in close proximity to human settlements (Holt 2000; Sala et al. 2000; Strayer and Dudgeon 2010). Freshwater habitats also contain a much smaller volume of water than marine systems, limiting their ability to dilute pollutants and reduce any impacts (Dudgeon et al. 2006).

Freshwater ecosystems are not isolated systems; freshwater flows through terrestrial environments and into marine systems. Nutrients are able to flow between ecosystems, either directly via water movement transporting compounds or indirectly through organisms moving and excreting bodily fluids or dying (Naiman et al. 2002; Harding et al. 2004; Baxter et al. 2005; Quinn et al. 2009). Areas where one ecosystem transitions into another, such as riparian zones between freshwater and terrestrial systems or estuarine zones between freshwater and marine systems, allow a wide variety of species to come into contact with one another, resulting in a broad range of interactions (Knight et al. 2005; Post et al. 2008; McCoy et al. 2017). The diversity of functions and interactions at areas where ecosystems meet can influence the structure and productivity of each habitat, regardless of whether it is the donor or recipient system, e.g. via energy influxes (Zhang et al. 2003; Baxter et al. 2005; Marczak et al. 2007; Burdon and Harding 2008). This connection between different ecosystems means that anthropogenic influences on one ecosystem can have impacts on another,
such as application of pesticides applied to land being washed into freshwater systems and often affecting fertility or mortality of riverine species (Malmqvist and Rundle 2002; Dudgeon et al. 2006). Protection of freshwater ecosystems thus requires an understanding of the ecology both within freshwater environments, and in adjacent terrestrial ecosystems and marine ecosystems (Dudgeon et al. 2006).

1.2 Food webs and species interactions

Food webs describe the interconnected feeding relationships within ecological communities, from the lowest trophic levels through to top predators (Hall and Raffaelli 1993; Dunne et al. 2002; Bascompte and Melián 2005). Studying food web topologies is a basic requirement for understanding biology and function of ecosystems, bringing together research in biodiversity and metabolic theory to describe which species interact and how this affects movement and storage of nutrients (Elser et al. 2000; Brown and Gillooly 2003; Schmitz et al. 2010). Predator-prey interactions in particular have complex effects that can alter entire ecosystems, with a change in the abundance of one species likely to reflect, and be reflected by, changes in abundance of another (Polis and Winemiller 2013; Berg et al. 2015). Population changes at low trophic levels can affect predators through bottom-up effects, potentially causing alteration in foraging behaviour, reproduction or mortality, which might ultimately impact population size (Pace et al. 1999; Sinclair and Krebs 2002; Elmhagen and Rushton 2007). Changes in the abundance of top predators can also impact species at lower trophic levels through top-down effects, as predators directly influence the density or behaviour of prey species which can in turn indirectly impact other species that interact with these prey species, resulting in trophic cascades (Pace et al. 1999; Shurin et al. 2002; Knight et al. 2005; Dobson et al. 2006).

Predator-prey interactions are influenced by changes in species abundances due to natural (e.g. season or habitat quality) and anthropogenic variation (e.g. introduction/loss of species or habitat fragmentation; Baxter et al. 2005; Schmitz et al. 2010; Birnie-Gauvin et al. 2017), as well as due to changes in a predators foraging behaviour. Foraging theory predicts that predators will choose prey that provide energy gains equal to or greater than that expended on finding, catching and consuming the prey (Pyke et al. 1977). When preferred prey become less available, predators may switch their foraging behaviour and consume alternative prey at a greater rate than predicted by random feeding (Lack 1954; Angelstam et al. 1984), with behavioural switches more common in generalist predators. Dietary variation in predators is able to reflect changes in predator-prey interactions across environmental gradients (Thompson and Townsend 2005; Layer et al. 2010; Pellissier et al. 2018), giving an
insight into the resilience of trophic interactions within an ecosystem (Ives and Carpenter 2007). Apex generalist predators are particularly good indicators for ecosystem health (Dobson et al. 2006; Sergio et al. 2006; Estes et al. 2011; Lemarchand et al. 2011) as their diet responds to and reflects changes in population dynamics at a range of trophic levels (Sergio et al. 2006). Predator diet studies are therefore able to provide information that can guide effective management and conservation for both predator and prey populations (Burgar et al. 2014; Bessey et al. 2019).

1.3 Methods for investigating trophic interactions

1.3.1 Traditional methods

Traditionally, trophic interactions have been studied using direct observation of foraging behaviour (e.g. Margalida et al. 2005; Basha et al. 2009) and morphological analysis of prey remains (e.g. Martins et al. 2011; McCully Phillips et al. 2019). Direct observation can provide useful information for species that are easy to observe (e.g. large terrestrial species), but is less applicable for investigating elusive species (e.g. nocturnal, aquatic or soil dwelling; Darimont and Reimchen 2002; Pompanon et al. 2012) and is biased by disturbance to species and the problem of identification of dietary items at a distance (Kruuk 1995). On the other hand, morphological analysis of prey remains investigates species consumed, allowing dietary choices to be recorded without direct observation. This is achieved through microscopic identification of undigested remains, such as bones, feathers, scales, arthropod exoskeleton or seeds, in stomach contents, pellets or faeces (e.g. Mariano-Jelicich and Favero 2006; Kozlowski et al. 2008; Reys et al. 2009). Stomach contents allow collection of both individual biotic data and samples, but can only be analysed following death of an individual or by invasively inducing regurgitation (e.g. Alonso et al. 2014; Jo et al. 2014). Pellet or faecal collection often lack individual biotic data but can be obtained non-invasively making them more desirable sample types for researching threatened species (Kelly et al. 2012), but associated individual biotic data are often lacking.

Morphological analysis of prey remains has provided information on the diets of a range of species, however, it only provides a snapshot into the diet (Nielsen et al. 2018) and items may be overrepresented or even mis-identified. Species can become overrepresented if their remains are more resistant to digestion (e.g. hard compared to soft bodied prey; Pompanon et al. 2012; Boyer et al. 2015), persist in the gut longer (Carss and Elston 1996; Tollit et al. 2003) or if the species is fully consumed (partial consumption may result in a lack of identifiable remains; Symondson 2002; Adámek et al. 2003). On the other hand, the identification skills of the researcher conducting the
analysis may result in ambiguous or incorrect identifications (Spaulding et al. 2000), especially if remains are morphologically similar (e.g. salmonids; Tollit et al. 2009). These caveats have resulted in the development and utilisation of molecular methods that describe diet over longer periods of time (e.g. stable isotope analysis) or give greater taxonomic detail (e.g. DNA metabarcoding).

### 1.3.2 Stable isotope analysis

Stable isotopes are atoms of the same element that differ in the number of neutrons and do not decay over time. The natural variation in stable isotopes within an ecosystem can be used to infer energy flow through a food web, investigate animal migrations and describe the diet of consumers (Hobson 1999; Rubenstein and Hobson 2004; West et al. 2006; Crawford et al. 2008; Inger and Bearhop 2008). This is achieved by measuring the ratio of heavy to light isotopes (heavy isotopes have more neutrons compared to light) in a sample, compared to a standard, and is reported as delta (δ) units in per mil (‰) (Kelly 2000; Inger and Bearhop 2008; Ben-David and Flaherty 2012). Ratios of stable isotopes vary due to the distribution of inorganic and organic matter, and are further altered by chemical processes during transfer to tissues of organisms (Bedard-Haughn et al. 2003; Kendall et al. 2007; Marshall et al. 2007; Ben-David and Flaherty 2012).

Stable isotopes are incorporated directly from the diet into an animal's tissues, reflecting dietary variation over a range of time periods as nutrient assimilation and cellular turnover differs between tissue types (Tieszen et al. 1983; Hobson and Clark 1992; Dalerum and Angerbjörn 2005). In metabolically active tissues, the rate of cell turnover determines the time period reflected; blood has a high turnover rate representing assimilation from the previous couple of days, the turnover in muscle is slower representing up to a couple of weeks and bone turnover is very slow, typically representing a lifetime (Stenhouse and Baxter 1976; Tieszen et al. 1983; Hobson and Clark 1992). Metabolically inert tissues (such as hair, claws and whiskers) on the other hand remain unchanged following tissue synthesis, and therefore capture isotopic signatures in a chronological order (Darimont and Reimchen 2002; Lewis et al. 2006; Newsome et al. 2009). By sampling a range of tissues from the same individual or conducting serial sampling along the length of a metabolically inert tissue it is possible to show temporal changes in habitat use or diet of organisms (Dalerum and Angerbjörn 2005; Lewis et al. 2006; Phillips and Eldridge 2006; Greer et al. 2015).

The two isotopes most commonly used for ecological research are nitrogen and carbon (Crawford et al. 2008; Inger and Bearhop 2008; Layman et al. 2012). Nitrogen is
reported using atmospheric nitrogen as a standard (Ehleringer and Rundel 1989); atmospheric nitrogen is less enriched in $^{15}$N than nitrogen bound in organic samples, resulting in most organisms possessing a positive $\delta^{15}$N value (Kelly 2000). Nitrogen is primarily used to interpret the trophic position of a consumer in a food web due to the stepwise enrichment in $^{15}$N at each trophic level (Kelly 2000; Crawford et al. 2008; Inger and Bearhop 2008). Preferential excretion of the lighter isotope results in $\delta^{15}$N values increasing by approximately 3.4‰ at each trophic level (Deniro and Epstein 1980; Peterson and Fry 1987; Post 2002), therefore apex predators will have higher $\delta^{15}$N values than meso-predators and herbivores. Nitrogen has also been used as an indicator for anthropogenic inputs, such as agricultural fertilisers and domestic sewage. Inputs from domestic sewage and manure-based fertilisers tend to be more enriched in $^{15}$N than atmospheric nitrogen, giving $\delta^{15}$N values between +2‰ and +30‰, whereas synthetic fertilisers are fixed from the atmosphere, meaning they are less enriched in $^{15}$N with $\delta^{15}$N values between -4‰ and +4‰ (Kendall et al. 2007). Processes such as volatilisation, nitrification and denitrification can further elevate the $\delta^{15}$N values of nitrogen inputs before they enter freshwater ecosystems (Kendall et al. 2007; Hoffman et al. 2012). Thus, anthropogenic inputs can change the basal $\delta^{15}$N value for freshwater food webs, typically enriching the system in $^{15}$N (Bedard-Haughn et al. 2003; Anderson and Cabana 2005; Urton and Hobson 2005; Cole et al. 2006; Hoffman et al. 2012; Baeta et al. 2017).

The standard used for reporting carbon values is a marine fossil limestone formation from South Carolina, PeeDee Belemnite (PBD; Craig 1957), which is enriched in $^{13}$C relative to most organisms and thus typically provides negative $\delta^{13}$C values for tissue samples (O'Leary 1988; Kelly 2000). Unlike $^{15}$N, $^{13}$C does not undergo a considerable enrichment between each trophic level (approximately 0.4‰ - 0.8‰ at each trophic level; Petersen and Fry 1987; Post 2002). Carbon isotopic ratios instead vary depending upon the source of carbon and photosynthetic pathways utilised by primary producers, resulting in a distinct difference in $\delta^{13}$C between food webs based upon C3 plants, C4 plants, crassulacean acid metabolism (CAM) plants or marine primary producers (Kelly 2000; Marshall et al. 2007). In C3 plants, carbon from CO2 is fixed using the RUBISCO enzyme which preferentially assimilates $^{12}$C isotopes and thus results in a low $\delta^{13}$C value (mean of -27‰ but range of -35‰ to -21‰; Petersen and Fry 1987; Kelly 2000; Marshall et al. 2007). C4 plants on the other hand use the enzyme phosphoenolpyruvate (PEP) carboxylase and CAM plants use a third photosynthetic pathway to fix carbon from CO2, both of which have a lower preference for $^{12}$C and thus have a higher $\delta^{13}$C value (mean of -13‰ and range of -14‰ to -10‰;
Petersen and Fry 1987; Kelly 2000; Marshall et al. 2007). Marine primary producers (algae and phytoplankton) also use the RUBISCO pathway, however, the carbon used during photosynthesis is bicarbonate which is more enriched in $^{13}$C and therefore results in a less negative $\delta^{13}C$ value (mean -17‰ range of -24‰ to -3‰; Petersen and Fry 1987; O'Leary 1988; Kelly 2000; Guiry 2019). This makes carbon a good indicator for the carbon resources at the base of the food web and therefore a good indicator for foraging location.
Figure 1.1. Illustration of isotopic signatures in $\delta^{15}N$ and $\delta^{13}C$ from primary producers (plants and zooplankton), herbivores and carnivores. Open circles represent terrestrial/freshwater ecosystems based on C$_3$ food webs, with values adapted from Kristensen et al. (2011) for primary producers and Urton and Hobson (2005) for herbivores and carnivores. Closed circles represent terrestrial/freshwater ecosystems based on C$_4$ food webs, with values adapted from Wang et al. (2010) for primary producers and Codron et al. (2007) for herbivores and carnivores. Grey diamonds represent marine food webs, with values adapted from Madgett et al. (2019). Signatures depicted do not represent definitive isotopic signatures for species within these ecosystems, but rather general predictions for where species may fall within this isotopic niche space. Isotopic signatures of individual species within each ecosystem can vary widely, resulting in overlap in isotopic signatures between each system.
Many ecological studies have combined data from nitrogen and carbon stable isotopes to interpret dietary shifts in a range of predators (e.g. Urton and Hobson 2005; Voigt et al. 2014; Rosenblatt et al. 2015; Jeanniard-Du-Dot et al. 2017). Although this method can help resolve predator diets, it often lacks clarity because of differences in isotopic fractionation and similarities in isotopic signatures (Crawford et al. 2008). Isotopic fractionation differs between and within species due to factors such as nutritional intake and metabolic rate (Mirón et al. 2006; Williams et al. 2007; Gaye-Siessegger et al. 2004; MacAvoy et al. 2006), making it difficult to decipher trophic pathways unless isotopic signatures of both predator and prey populations are measured from the same locality and time period (Crawford et al. 2008). Similarities in isotopic signatures between prey species can further complicate interpretation, by preventing prey species being distinguished based upon stable isotopes (Layman et al. 2012). Further confusion may arise where predator isotopic signatures represent an intermediate value of all the prey consumed, making it possible for individuals with different diets to possess similar isotopic signatures (Middelburg 2014; Hertz et al. 2016). Similarities in isotopic signatures, either between prey or predators, can make it difficult to identify prey species that have contributed isotopic signatures in predator tissues, potentially leading to inaccurate interpretations of the diet (Crawford et al. 2008; Layman et al. 2012). Whilst these caveats limit the use of stable isotope analysis for deciphering specific trophic interactions, this method provides unique insights into nutrient assimilation within predators (Crawford et al. 2008; Inger and Bearhop 2008), and changes in nutrient influxes within the environment (Helfield and Naiman 2001; Kendall et al. 2007), that are not possible through traditional methods.

1.3.3 DNA metabarcoding

DNA metabarcoding provides rapid screening of hundreds (or even thousands) of samples and simultaneous identification of multiple taxa within each sample through the combined use of DNA barcodes and high-throughput sequencing (Taberlet et al. 2018). This method can detect taxa even when there is no visual trace, capturing a greater diversity at a finer taxonomic resolution than traditional morphological methods (Bowser et al. 2013; Roslin and Majaneva 2016; Elbrecht, Vamos, et al. 2017). Metabarcoding is used to identify biodiversity within complex samples that are often comprised of degraded DNA from a mix of organisms; faeces and stomach contents are used to describe a consumers diet (e.g. Moran et al. 2015; McInnes et al. 2017; Aizpurua et al. 2018; Granquist et al. 2018), soil and water samples for environmental monitoring (e.g. Borrell et al. 2017; Fletcher et al. 2017; Oliverio et al. 2018; Treonis et al. 2018) and ancient samples (e.g. bone fragments or permafrost samples) for reconstruction of past environments (e.g. Sonstebo et al. 2010; Grealy et al. 2015).
Due to increases in the availability of reference databases and decreases in sequencing costs, there has been a rapid uptake in the number of studies using DNA metabarcoding (Hawlitschek et al. 2018), with methods being exploited from academic researchers through to natural resource managers (Creer et al. 2016; Deiner et al. 2017).

DNA barcodes are short, variable regions of the genome that allow taxa in unknown samples to be identified by comparing query sequences to known sequences in reference databases (Hebert et al. 2003; Kress et al. 2015). Suitable barcode regions are variable enough to discriminate between taxa, due to interspecific variation being greater than intraspecific, but have conserved flanking regions (primer sites) that allow amplification of multiple taxa (Meyer and Paulay 2005; Vences et al. 2005). Barcoding region and primer choice are key aspects of metabarcoding, determining which taxa are likely to be amplified (Alberdi et al. 2018). Various barcoding regions have been identified for metabarcoding studies - e.g. 18S for microbes (Debroas et al. 2017; Thongsripong et al. 2018), ITS2 for fungi (Rocchi et al. 2017; Banchi et al. 2018) and ITS2 for plants (Dunn et al. 2018; Moorhouse-Gann et al. 2018) - but two of the main barcodes used to distinguish metazoans are COI and 16S rRNA (Clarke et al. 2014; Deagle et al. 2014; Elbrecht et al. 2016). COI is a conserved region with a higher mutation rate than 16S, allowing more closely related species to be distinguished (Andújar, Arribas, Yu, et al. 2018; Braukmann et al. 2019). Both COI and 16S occur in high copy numbers (due to use of mitochondrial copies), but COI contains fewer mutations through insertions and deletions (indels; Hebert et al. 2003; Andújar et al. 2018) and is a protein coding region, allowing pseudogenes to be more easily detected and removed during bioinformatic analyses (Andújar, Arribas, Yu, et al. 2018; Braukmann et al. 2019). COI is often recommended as the primary metazoan barcode of choice though because of the large number of reference sequences available on public databases (Andújar, Arribas, Yu, et al. 2018; Collins et al. 2019); Rennstam Rubmark et al. (2018) found Genbank possessed approximately 2-3 million sequences from 280,000 species for COI, compared to only 380,000 sequences from 90,000 species for 16S. Even so, many researchers advocate the use of 16S over COI because it contains more conserved primer binding sites (Clarke et al. 2014; Deagle et al. 2014; Elbrecht et al. 2016). This allows primers in the 16S region to achieve a similar taxonomic breadth and resolution as COI primers, but with fewer degenerate bases, resulting in fewer primer-template mismatches and less amplification of non-target DNA compared to COI (Siddall et al. 2009; Collins et al. 2019).
Metabarcoding has been used to unravel complex trophic interactions from pollination through to predator foraging (Littlefair et al. 2016; Roslin and Majaneva 2016), providing greater detail into the range of species involved in these links and thus an insight into foraging strategies of individuals (Schwarz et al. 2018), potential human-wildlife conflict (Shehzad et al. 2015), and why populations may be undergoing recovery or decline (Hardy et al. 2017). Metabarcoding is not without its limitations though; for example, metabarcoding gut contents or faecal samples typically only provides a snap-shot into the diet of an individual (approximately 24 – 72 hours in mammals; Deagle et al. 2005; Casper et al. 2007; Thalinger et al. 2016) and cannot distinguish between true consumption, secondary consumption and scavenging, nor between host DNA and cannibalism (Bessey et al. 2019). Methodological biases can be introduced at each step, from extraction and amplification of DNA through to sequencing and bioinformatics (Jusino et al. 2019). Primer biases may cause preferential amplification of particular species (such as DNA from abundant prey or the host itself) which may swamp the other outputs, potentially preventing detection of other species (such as rarely consumed species; Piñol et al. 2014; Elbrecht and Leese 2015; Elbrecht et al. 2017; Alberdi et al. 2018). Quantifiable data therefore cannot be reliably produced, resulting in a reliance on presence-absence data rather than numbers of reads. Inconsistencies in reference databases may produce incorrect taxonomic assignments (Keskin et al. 2016; Rulik et al. 2017; Taberlet et al. 2018), whilst inappropriate bioinformatic thresholds can fail to exclude artefacts or generate false negatives (Deagle et al. 2013; Bessey et al. 2019; Zinger et al. 2019).

Methodological limitations can be overcome to an extent through careful study design (Alberdi et al. 2019); multiple complementary primers can be used to overcome primer biases and provide greater coverage and taxonomic resolution (Hibert et al. 2013; Hawlitschek et al. 2018; Marquina et al. 2019; da Silva et al. 2019), larger sample sizes allow key species to be identified through higher frequency of occurrences and greater detection of mis-assignments, and empirical choice of filtering thresholds can limit the number of false positives and negatives within a study (Clare et al. 2016; Alberdi et al. 2018). With such measures in place, DNA metabarcoding provides one of the most accurate and widely used methods for assessing biodiversity and trophic interactions (Taberlet et al. 2018).

1.4 Eurasian otter: ecology and diet

The Eurasian otter (Lutra lutra) is a semi-aquatic carnivore that is widespread across Europe, Asia and North Africa (Roos et al. 2015). It is an apex predator that inhabits a range of environments, including freshwater rivers, lakes and sea coasts (Jędrzejewska et al. 2001; Parry et al. 2011). Eurasian otters experienced population
declines across their range during the twentieth century due to persecution, habitat
degradation and contamination from organochlorine insecticides and polychlorinated
biphenyls (PCB; Strachan and Jefferies 1996; Roos et al. 2001; Clavero et al. 2010).
Restrictions on use of PCBs from 1986 onwards (Environmental Protection
(Bilcke 2002; Levain et al. 2015), along with legal protection for Eurasian otters
(Wildlife and Countryside Act 1981; Habitat Directive 1992; CITES 2019) have been
key factors allowing otter populations to recover and expand their range (Strachan and
Jefferies 1996; Mason 1998; Roos et al. 2001; Conroy and Chanin 2002). In Great
Britain, surveys have shown substantial recovery and range expansion of otter
populations; in the 1990’s, otters were recorded as present at only 23% of sites in
England (Strachan and Jefferies 1996), 53% of sites in Wales (Andrews et al. 1993)
and 88% of sites in Scotland (Green and Green 1997), however, by 2012 otters had
successfully recolonised all regions of Great Britain with the exception of Kent and East
Sussex (Sainsbury et al. 2019). Whilst IUCN criteria for conservation threats classifies
otters in Great Britain as ‘least concern’ (Mathews et al. 2018), across their whole
range the species is still classified as ‘near threatened’, due to declines or lack of
information in other regions and due to their sensitivity to sudden changes in their
environment (Roos et al. 2015). As British populations undergo range expansion,
individuals may disperse into landscapes that have been altered since otters were last
present, potentially affecting the ecology and behaviour of otters and introducing new
threats to populations (Clavero et al. 2010).

As opportunistic predators, otters consume a wide variety of species and will switch to
alternative prey in response to changing prey abundances (Clavero et al. 2003;
Almeida et al. 2012; Parry et al. 2015). They primarily consume fish; however, they are
also known to take mammals, birds, reptiles, amphibians and invertebrates
(Jędrzejewska et al. 2001; Clavero et al. 2003; Almeida et al. 2012). Dietary variation
occurs over a range of spatio-temporal scales and largely reflects prey availability in
the surrounding environment (Jędrzejewska et al. 2001; Clavero et al. 2003; Lundy and
Montgomery 2010). Differences in habitat and seasonal changes make prey
differentially available in certain areas (e.g. greater opportunities for otters to consume
marine species near coastal habitats) or during certain times of the year (amphibian
breeding aggregations during spring). In more complex or unstable habitats where prey
availability is less consistent, otters are able to broaden their dietary niche and
consume a greater range of alternative prey (Clavero et al. 2004; Clavero et al. 2008;
Ruiz-Olmo and Jiménez 2009). This opportunistic behaviour has likely facilitated the
recovery of otters, allowing them to utilise a variety of prey as they recolonise habitats;
however, prey switches may also result in consumption of prey with less nutritional value or greater contaminant loads, impacting the fitness of individual otters and potentially the persistence of the population (Österblom et al. 2008; Ruiz-Olmo and Jiménez 2009; Lourenço et al. 2011).

Dietary shifts in otters can also impact prey species, with prey populations negatively impacted during periods of increased predation (Latham et al. 2013). This has raised concerns over the impact recovering otter populations may have on threatened prey populations, such as the European eel (*Anguilla anguilla*) which is the favoured prey of British otters (Chanin 1981; Copp and Roche 2003; Miranda et al. 2008), as increased predation rates may result in further population declines. In contrast, increased predation pressures by otters may help with biocontrol of invasive freshwater species as otters adapt to increases in their populations and exploit them as an alternative food source when other prey are restricted. Such dietary shifts have been observed in Mediterranean otters where invasive red swamp crayfish (*Procambarus clarkii*) have become one of the primary prey consumed, particularly during droughts when fish are less available (Adrian and Delibes 1987; Beja 1996; Correia 2001; Barrientos et al. 2014). Another source of alternative prey that otters utilise is stocked fish species (Gutmann Roberts et al. 2017), as these tend to be abundant all year round. Increased predation on these species can deplete stocks, bringing otters into conflict with anglers and aquaculture management and potentially leading to adverse effects on otter populations through human mediated persecution (Kloskowski 2005; Vaclavikova et al. 2011; Poledníková et al. 2013; Grant and Harrington 2015).

Most trophic studies into otter diet have focussed on morphological analysis of prey remains in spraint or stomach contents (e.g. Jędrzejewska et al. 2001; Ruiz-Olmo and Jiménez 2009; Almeida et al. 2012; Moorhouse-Gann et al. 2020), which may miss prey items or neglect broad temporal variation in individuals unless consecutively sampled. The use of molecular methods can overcome such limitations by identifying prey to a greater taxonomic resolution even when prey remains are scarce (e.g. DNA metabarcoding; Bowser et al. 2013; Roslin and Majaneva 2016; Elbrecht, Vamos, et al. 2017) or detecting dietary variation overtime through the analysis of different body tissues (e.g. stable isotope analysis; (Tieszen et al. 1983; Hobson and Clark 1992; Dalerum and Angerbjörn 2005; Nielsen et al. 2018). Molecular analysis of otter diet can therefore not only provide a greater insight into the behaviour and ecology of recovering otters, but also into changes in trophic interactions within freshwater ecosystems (Sergio et al. 2006; Crawford et al. 2008; Estes et al. 2011; Pompanon et al. 2012). The information derived can potentially be used to reduce conflict with
fisheries and help guide conservation of both otters and their prey. There are very few studies utilising molecular methods to analyse otter diet, with none employing stable isotope analysis, one employing DNA barcoding of prey remains (Hong et al. 2019), and five employing DNA metabarcoding (Buglione et al. 2020; Marcolin et al. 2020; Martínez-Abrain et al. 2020; Kumari et al. 2019; Harper et al. 2020). Studies investigating otter diet through DNA analyses have been limited either by small sample size (Hong et al. 2019; Buglione et al. 2020; Kumari et al. 2019; Marcolin et al. 2020; Martínez-Abrain et al. 2020) or utilisation of only one barcoding region (Hong et al. 2019; Buglione et al. 2020; Harper et al. 2020; Marcolin et al. 2020; Martínez-Abrain et al. 2020), potentially missing prey items due to primer bias or poor reference databases (Harper et al. 2020; Marcolin et al. 2020). Dietary analysis of otters can therefore be expanded by utilising both stable isotope analysis and DNA metabarcoding, employing metabarcoding on a greater range of samples and through the use of multiple primer sets in order to identify a greater range of prey items.

1.5 Aims and hypotheses
The aim of this PhD project was to provide a major advance in the understanding of the trophic ecology of the Eurasian otter using molecular techniques; metabarcoding to look at short-term, fine scale identification of prey items (Pompanon et al. 2012; De Barba et al. 2014) and stable isotope analysis to look at long-term variation in nutrient assimilation (Crawford et al. 2008; Layman et al. 2012). Specific aims were to (i) investigate spatial variation in otter foraging ecology as habitats support different species assemblages; (ii) investigate temporal variation in otter foraging ecology as species abundances change seasonally and annually; and (iii) identify dietary differences between demographics (i.e. sex, age-class and body condition). In order to address these ecological questions, this project aims to (iv) develop a suitable metabarcoding approach for the study of otters; and (v) assess the differences between traditional and molecular approaches for describing otter diet; in order to (vi) evaluate both long-term and short-term variation in otter diet.

Chapter two explores the importance of marine derived nutrients to otter diet in Wales. Using stable isotope data collected from otters across Wales and bordering regions of England between 1993 and 2007, we investigate whether marine derived nutrients are associated with coastal feeding or with anadromous fish movement.

Chapter three aims to explore spatio-temporal variation in otter diet across England and Wales between 2007 and 2016 using stable isotope data. Specifically, this chapter aims to explore variation in nutrient assimilation in otters due to allochthonous nutrient
inputs from both terrestrial and marine ecosystems, including anthropogenic inputs and marine derived nutrients. Chapter three also aims to investigate the within-individual variability in diet, across different spatio-temporal scales.

Chapter four details how metabarcoding data were acquired from otter samples and uses the metabarcoding data as a case study in order to evaluate how post-bioinformatic filtering processes affect the final interpretation of the dietary data. Filtering processes tested are outlined and their impacts on otter dietary data are presented.

Chapter five is the final data chapter of the thesis. Here spatio-temporal variation in otter diet across England and Wales between 2007 and 2016 is investigated using DNA metabarcoding, and morphological analysis. Species detected using each method and ecological interpretations of the data from models are presented and compared. More specifically, this chapter considers the impact of otter predation on species of conservation concern, non-native invasive species and fish species prized by anglers.

Finally, Chapter 6 brings together the key findings from the stable isotope, metabarcoding and morphological analyses. This chapter addresses the wider implications of this work for ecosystems and limitations of the study. Remaining knowledge gaps and future research aspects are also suggested.
Chapter Two – There’s something fishy going on: a stable isotope analysis of marine-derived nutrients in the diet of Eurasian otters (Lutra lutra)

2.1 Abstract

Nutrient transfer between terrestrial and marine ecosystems usually refers to movements from inland regions to the sea; however, transfer of marine-derived nutrients (MDN) upstream can also be substantial. Although upstream transfers are well known in North American river systems, little is known about potential transfers in northern European freshwater systems. To evaluate the potential of MDN movements in a European freshwater system, we used the Eurasian otter (Lutra lutra), an apex predator in Northern Europe, as an indicator of MDN availability. We quantified the carbon and nitrogen stable isotope signatures of 174 samples of bone collagen and 222 samples of muscle tissue from Eurasian otters, collected in Wales and England from 1993 to 2007. Overall, isotopic signatures did not reflect a strong marine signature, suggesting MDNs contributed relatively little to the diet of otters. MDNs that were acquired were suggested to be through direct consumption of anadromous fish and marine species, as isotopic signatures were correlated with abundance of anadromous fish, coastal proximity and timing of anadromous fish migrations. Marine isotopic signals were clearer in bone samples compared to muscle tissue, suggesting short term variability in diet may mask longer term MDN inputs.

Keywords: Dietary variation, Eurasian otter, Lutra lutra, marine-derived nutrients, stable isotopes

2.2 Introduction

Nutrients from marine sources (‘marine-derived nutrients’ or MDN) can be transferred upstream and incorporated into freshwater and terrestrial ecosystems via anadromous fish moving inland, and spawning, excreting or decomposing in freshwater streams (Helfield and Naiman 2001; Zhang et al. 2003; Harding et al. 2014). Consumption of marine species by predators (or consumption of riverine fish which have been fed marine subsidies; Bašić et al. 2015; Gutmann Roberts et al. 2017; Nolan et al. 2019), followed by movement of those predators inland, also results in translocation of MDN into freshwater and surrounding terrestrial systems (Hilderbrand et al. 1999; Naiman et al. 2002; Harding et al. 2004). MDN can be assimilated into all levels of freshwater and terrestrial food webs and can result in substantial impacts on the environment, such as
increased algal blooms and productivity in freshwater and terrestrial ecosystems (Helfield and Naiman 2001; Zhang et al. 2003; Harding et al. 2014).

Incorporation of MDN into inland habitats has been well described in North American river systems (Naiman et al. 2002). Large numbers of anadromous Pacific salmon (e.g. *Oncorhynchus* spp.) migrate along these rivers and die on mass following spawning, resulting in the release of MDNs into the surrounding environment (Helfield and Naiman 2001; Gende et al. 2002; Naiman et al. 2002). MDN from Pacific salmon also disperse into neighbouring terrestrial habitats via the movement of piscivorous predators, such as bears, wolves and otters, which transport salmon carcasses on land as well as consume salmon and then excrete in terrestrial habitats (Elliott et al. 1997; Hilderbrand et al. 1999; Naiman et al. 2002; Quinn et al. 2009). Whilst MDNs have been characterised in North America, little is known about their dynamics in Northern European freshwater ecosystems (Walters et al. 2009). Anadromous fish such as Atlantic salmon (*Salmo salar*) and sea trout (*Salmo trutta*) may act as vectors for MDN input into Northern European river systems (Ben-David et al. 1998), but anadromous fish migrations aren’t as numerous in Europe compared to North America. This is due to fewer anadromous species, smaller populations of migrating individuals, and population declines in anadromous species (WWF 2001; Limburg and Waldman 2009). As the Eurasian otter is the dominant mammalian predator of freshwater ecosystems (Krawczyk et al. 2016) across much of its extensive Eurasian range, we hypothesise that otters could play a key role in dispersing MDN into freshwater and terrestrial systems in Northern Europe through consumption of anadromous fish.

Eurasian otters are an apex predator in freshwater habitats but also forage in terrestrial and marine habitats (Jędrzejewska et al. 2001; Parry et al. 2011). They are a long-lived species (up to 10 years) with large ranges (>10 km per night; Chanin 2003). As populations of otters in England and Wales increase and expand their range, otters may begin to utilise more marine foraging opportunities (Strachan and Jefferies 1996; Parry et al. 2011). Otters could therefore be sampling MDN directly, through consumption of marine prey at the coast or anadromous fish inland, and indirectly via consuming prey species lower in the food web that have assimilated MDN into their tissues, such as fish and invertebrates (Naiman et al. 2002; Harding et al. 2004; Vizza et al. 2017). Otters may thus act as a vector for transferring MDN upstream into rivers and neighbouring terrestrial ecosystems.

Stable isotopes measured in animal tissues provide a useful tool for investigating dietary variation (Kelly 2000; Dalerum and Angerbjörn 2005; Crawford et al. 2008). Due
to differing tissue turnover rates, comparisons between tissue sampled at the same time can reveal temporal variation in diet (Bearhop et al. 2003; Dalerum and Angerbjörn 2005), for example muscle tissue can be used to show diet one to two months before sample removal (Hobson and Clark 1992; Maruyama et al. 2001) whereas bone can show diet over the lifetime of some species (Stenhouse and Baxter 1976; Hobson and Clark 1992; Hedges et al. 2007). Carbon and nitrogen isotopes are the most common isotopes used in dietary studies (Kelly 2000; Layman et al. 2012). The levels of $\delta^{13}C$ (the ratio of $^{13}C:^{12}C$ isotopes) in the tissue of an animal can be interpreted to indicate the basal resource of the food web (i.e. the relative contribution made by C$_3$, C$_4$ or marine plants; Kelly 2000; Inger et al. 2006; Post et al. 2007), whereas the levels of $\delta^{15}N$ (the ratio of $^{15}N:^{14}N$ isotopes) can be used to indicate the trophic level at which the consumer is feeding (Deniro and Epstein 1980; Hobson and Clark 1992). Given the greater enrichment of heavier isotopes ($^{13}C$ and $^{15}N$) in marine compared to terrestrial and freshwater food webs (Peterson and Fry 1987; O'Leary 1988; Kelly 2000; Inger et al. 2006), and the maturation of anadromous fish in marine environments (Hendry and Cragg-Hine 2003; Klemetsen et al. 2003; Mark Everard 2013), higher $\delta^{13}C$ signatures are indicative of adult anadromous fish migrating upstream for reproduction (Naiman et al. 2002; Dixon et al. 2012).

Here, we measured the isotopic signature in otter tissues (bone and muscle) to provide evidence for marine derived nutrient contributions to a northern European freshwater system. We hypothesised that (i) isotopic signatures from muscle tissues (representing shorter term dietary variation) would be highly variable, whereas those from bone (representing long term integration) would provide a less variable signature more likely to correlate with putative spatial drivers of enrichment. We hypothesised that (ii) isotopic enrichment of otter samples would be positively correlated with anadromous fish abundance, and (ii) negatively correlated with distance from the coast. Finally, we hypothesised that (iv) isotopic enrichment of otter muscle samples would be highest in the autumn and winter (corresponding with the migration of anadromous fish).

### 2.3 Methods

#### 2.3.1 Sample and data collection

Dead otters were collected from Wales and England as part of a national monitoring programme ([www.otterproject.cf.ac.uk](http://www.otterproject.cf.ac.uk)). During carcass collection, date and location of carcasses (as a grid reference) were recorded. Each otter was classified by season; those found in December, January and February were classified as winter, March, April and May as spring, June, July and August as summer and September, October and November as autumn. Grid references were used for mapping and spatial analysis.
Detailed post mortems were carried out on each otter (data not presented) during which biotic data were recorded, including sex and length (nose to tail) of individuals. Otters were classified into three age groups (juvenile, subadult and adult) based upon weight and signs of sexual maturity (Chadwick 2007). Females weighing less than 2.1 kg were defined as juveniles, whilst heavier females with no signs of reproductive activity (teats not showing, immature uterus) were defined as subadults and those with signs of reproducing at least once were defined as adults. Males weighing less than 3 kg were defined as juveniles, whilst heavier males were defined as subadults if their baculum measured less than 60 mm and adults if the baculum was equal to or greater than this length. During post mortems, bone samples (femur or tibia, n = 174) and two replicate muscle samples (n = 222) were taken from the hind right leg, wrapped in foil and archived at -20 °C. Samples were collected between 1993 and 2007, primarily from Wales and bordering counties, with an additional 10 bone samples from carcasses collected from mid-England (Fig. 2.1) to ensure that otters with no access to marine habitats were represented.

2.3.2 Location data

All mapping and extraction of spatial data was completed in QGIS version 3.4.4 (QGIS Development Team 2019). Each individual otter was assigned to an Environment Agency river catchment based on where the otter carcass was found (Fig. 2.1). To evaluate use of marine prey items, distance from each otter to the coast was calculated along the nearest river rather than as a straight line, since otters travel along river courses more often than across land (Kruuk 2006). To quantify anadromous fish availability, we used Environment Agency fisheries statistics (Environment Agency 1994; 1995; 1996; 1997; 1998; 1999; 2000; 2001; 2002; 2003; 2004; 2005; 2006; 2007; 2008). Angling catch per unit effort of Atlantic salmon and sea trout was calculated for each catchment per year by dividing the number caught by the number of days anglers fished in that catchment; this ranged from 0.018 in the Severn catchment in 1997 to 1.156 in the Glaslyn catchment in 2002 (Fig. S.1 and S 2.2). Angling catch per unit effort was used as a proxy for the anadromous fish abundance per catchment per year and was matched to otters based on the year in which the otter carcass was found. Where relevant fisheries statistics were not available (n = 34 otters) a value of zero was assigned for anadromous fish, because catchments are omitted from Environment Agency data collection only where anadromous fish migrations are absent.
Figure 2.1. Otter sampling locations. Distribution of Eurasian otter (*Lutra lutra*) carcasses collected 1993–2007 and analysed for carbon and nitrogen stable isotopes using (a) bone and (b) muscle tissue. Dark grey shading indicates catchments from which Environment Agency data on anadromous fish species (Atlantic salmon, *Salmo salar*, and sea trout, *Salmo trutta*) were available. Catchment names and anadromous fish abundance data are shown in SI Figures 2.1 and 2.2.
River courses were mapped using the Ordnance Survey Open Rivers map (OS Open Rivers 2017). Distance from the coast was calculated for each individual as shortest distance (km) along a river from the location at which the otter was found to the coastline, because otters tend to travel along water courses rather than across land. First, individuals were assigned to the nearest river, as many otters were found as roadkill and were therefore not necessarily at riverside locations. Distances were then calculated by extracting river lines that represented the shortest route along a river from the otter to the coast, following which the ‘sum line lengths’ function in QGIS was used to calculate the total distance.

### 2.3.3 Body Condition Index

Scaled mass index (SMI) is a more reliable indicator of body condition than other indices and was calculated as follows (Peig and Green 2009; Peig and Green 2010):

\[
SMI = M_i \left( \frac{L_0}{L_i} \right)^{bsMA}
\]

M<sub>i</sub> is the body mass and L<sub>i</sub> is the length measurement of individual <i>i</i>, L<sub>0</sub> is the mean length measurement for the entire study population and bSMA is the scaling exponent. Length was measured from nose to tail-tip to the nearest 5mm. Mean length and the scaling exponent were both calculated from all otter data available as of January 2017 (n = 2477). The scaling exponent is the slope from the standard major axis regression of log transformed values of mass against length. SMI allows the condition of individuals of different sizes to be compared, by predicting the body mass of an individual at a standard body length for the population. Scored SMI was then calculated (Guillemain et al. 2013) as the difference between the SMI of an individual and the average SMI for otters of the same sex and age-class (where age-class was juvenile, sub-adult and adult). This allowed us to assess the body condition of an individual relative to the population.

### 2.3.4 Stable isotope analysis

Stable isotope analysis was undertaken on lipid extracted bone and muscle samples, and on non-lipid extracted muscle samples. Tissues vary in their lipid contents, and as lipids are depleted in ¹³C relative to proteins and carbohydrates, variation in lipid content can affect δ¹³C values (Post et al. 2007). Chemical extraction is therefore used to standardise lipid content; however, this can affect δ¹⁵N values in muscle samples through isotopic fractionation (~ 2.5 ‰; Post et al. 2007), but is unlikely to affect δ¹⁵N values in bone (Medeiros et al. 2015). Lipid extraction was thus conducted on both
bone and muscle samples to account for biases from variable lipid contents between tissues samples, but muscle samples were also processed without lipid extraction to prevent introducing biases to $\delta^{15}N$ analyses.

Sub-samples of cortical bone were taken and cut into small chips. These were demineralised using 0.5M HCl at 4 $^\circ$C for up to four days. HCl was then removed and samples cleaned using demineralised water, and then refrigerated in deionised water. Lipid extraction was carried out on demineralised bone samples, and on one of the two muscle samples, using a Soxhlet extractor. Samples were placed in glass wool in the thimble, and a solvent mixture of chloroform, methanol and water (1:2:0.8) was added and allowed to reflux until the solution ran clear. Samples were then rinsed with deionised water, and all samples (bone, lipid-extracted muscle and non lipid-extracted) were freeze-dried at -60 $^\circ$C and individually homogenised.

Approximately 0.6 mg of each homogenate was weighed into individual tin capsules. Prepared samples were placed under continuous-flow isotope ratio mass spectrometry (CF-IRMS) using a Costech Elemental Analyser (mode ECS 4010) interfaced with a ThermoFinnigan Delta Plus XP mass spectrometer at the NERC Life Sciences Mass Spectrometry Facility at East Kilbride. Ratios of carbon and nitrogen stable isotopes are given as $\delta$-values and expressed in parts per ml ($^{\text{‰}}$), with reference to international standards according to the following equation:

$$\delta X = [(R_{\text{sample}}/ R_{\text{standard}}) - 1] \times 1000$$

Where X represents $^{13}$C or $^{15}$N, $R_{\text{sample}}$ is the corresponding ratio of heavy to light isotopes ($^{13}$C/$^{12}$C or $^{15}$N/$^{14}$N) in the sample, and $R_{\text{standard}}$ is the ratio of the international references, PeeDee Belemnite for $\delta^{13}$C and atmospheric nitrogen for $\delta^{15}$N.

### 2.3.5 Statistical analysis

In order to explore the associations between stable isotope ratios, and biotic and abiotic drivers, we used generalised linear models (GLM) or generalised additive models (GAM) in R [version 3.3.3] and R Studio [version 1.0.136] (R Core Team 2019; scripts available in SI. 2). Four models were developed, with (1) bone $\delta^{15}$N, (2) bone $\delta^{13}$C, (3) muscle $\delta^{15}$N and (4) muscle $\delta^{13}$C as the response variables. For muscle analyses, $\delta^{15}$N values were obtained from non-lipid extracted muscle samples, whereas $\delta^{13}$C values were obtained from lipid-extracted muscle samples (Post et al. 2007). All models included the following fixed variables: sex (male/female), individual length.
(mm), distance from the coast (km), and anadromous fish abundance per catchment per year, along with two-way interactions between sex and length, length and anadromous fish abundance and sex and anadromous fish abundance. Models using muscle data as the response variable also included the fixed variable season (spring, summer, autumn or winter), along with two-way interactions between season and sex, and season and anadromous fish abundance per catchment. Season was included for these models because muscle reflects a shorter time period than bone and can therefore reflect seasonal variation (Hobson and Clark 1992; Maruyama et al. 2001). Year was not included because the anadromous fish variable already incorporated change with time.

Preliminary analyses were used to compare GLM and GAM versions of the same models, with GAMs built using the R package ‘mgcv’ (Wood 2011) and a penalised, thin-plate smoothing spline fitted to distance from the coast. All models were assigned a Gaussian distribution with an identity link function with the exception of model one (bone $\delta^{15}$N) which was assigned a Gamma distribution and log link function (based on evaluation of model fit and AIC). The estimated degrees of freedom (edf) value was used to determine whether there was evidence for non-linearity. Model assumptions were plotted, allowing outliers that were having a large influence on the data to be identified and removed (e.g. individual points largely skewing QQ plots or points with a Cook's distance >1). Optimum GAM / GLM versions of each model were compared using ANOVA and AIC values. Comparisons showed that GAMs were better at explaining the data in models 1, 2 and 4 whereas for model 3 (muscle $\delta^{15}$N) a GLM was selected. A linear regression was also conducted to test for significant differences for each isotope between the different sample types.

Model selection was achieved by carrying out multi-model inference and model averaging on the global model using the R-package ‘MuMIn’ (Barton 2019). Where more than one ‘top model’ was found (i.e. there was a delta AIC of less than two between multiple candidate models) conditional average values were used to infer significant factors and interactions (models 2, 3 and 4). Where only one top model was found (i.e. only one model had a delta AIC <2) no model averaging was carried out, and significance was inferred directly from the model summary (model 1). Where not all fixed variables in the averaged model were significant, only variables with a p-value of less than 0.05 were reported in full, but all are listed in SI: 3: Tables 1 and 2 (all variables included in the top model for model 1 were significant). Significant factors and interactions that included categorical variables were plotted using the R-package ‘Stable Isotope Bayesian Ellipses in R’ (SIBER; Jackson et al. 2011) to produce 95%
confidence ellipses for isotopic data. Graphs produced through SIBER allowed any differences or overlap in isotopic niches between groups of animals to be visualised. Spatial variation in isotopes was also plotted using the multilevel b-spline function in QGIS.

### 2.4 Results

Muscle and bone $\delta^{15}N$ values were not significantly different from one another and showed a similar range of isotopic values, whereas for $\delta^{13}C$ values there was a significant difference between samples types (LM: $t = -13.81$, $p = <0.001$) with bone showing a greater enrichment in $^{13}C$ (Fig. 2.2). Overall variability was less in muscle tissue than bone, for both $\delta^{15}N$ and $\delta^{13}C$ (bone $\delta^{15}N$ mean = 14.67‰ ±2.43‰; muscle $\delta^{15}N$ mean 14.30‰ ±2.05‰; bone $\delta^{13}C$ mean 21.41‰ ±2.35‰, muscle $\delta^{13}C$ mean 24.51‰ ±2.18‰; Fig. S 2.3).

![Figure 2.2](image_url)

**Figure 2.2.** Isotopic niche measured in Eurasian otter (*Lutra lutra*) bone (green) and muscle (purple) tissue samples. Outer ellipses show 95% standard ellipse areas; inner ellipses show the 95% confidence intervals around the bivariate means. Samples were taken from across Mid-England, Wales and bordering counties from 1993 – 2007. The grey dashed line indicates an approximate $\delta^{13}C$ threshold between consumption of marine and freshwater prey, with isotopic signatures above the line more likely to indicate marine prey and signatures below the line more likely to indicate freshwater prey. The value is based on isotopic evidence from Madgett et al. 2019 (muscle and liver samples collected from marine invertebrates and vertebrates in Scotland; suggested $\delta^{13}C$ threshold of -19‰). A similar threshold is also suggested by Guiry 2019 (fish bone samples recovered during archaeological palaeodietary studies, from various locations; suggested $\delta^{13}C$ threshold of -17‰).
2.4.1 Variation in $\delta^{15}N$

Anadromous fish abundance was negatively associated with $\delta^{15}N$ in both bone and muscle tissue (model 1: Bone N, GAM: $t = -3.139$, $p = 0.002$; model 2: Muscle N, GAM: $z = 2.320$, $p = 0.02$; Fig. 2.3). The nature of these associations did not differ with interacting terms (i.e. sex, body length or season). Distance from the coast was also negatively associated with $\delta^{15}N$ in bone and muscle tissue, but where muscle tissue showed a linear pattern of decline, bone showed no detectable decline in $\delta^{15}N$ until approximately 120 km from the coast (model 1, GAM: edf = 2.161, $F = 3.337$, $p = 0.02$; model 3, GLM: $z = 3.910$, $p = <0.001$; Fig. 2.4). Although otter sex and body length were included in some of the component models used in model averaging (Tables S 2.1 and 2.2) neither variable showed a significant association with $\delta^{15}N$, in either bone or muscle tissue. Season was not found to be important in any of the component models for variation in muscle $\delta^{15}N$. 
Figure 2.3. Relationship between δ^{15}N ratio and abundance of anadromous fish. Anadromous fish abundance is represented by the angling catch per unit effort (number of fish caught divided by the number of days anglers fished) of Atlantic salmon and sea trout per catchment per year. Points shown represent isotopic values obtained from Eurasian otter (Lutra lutra) bone and muscle samples. Solid lines indicate the results predicted from models (general additive models for bone samples and general linear models for muscle samples) that included all the factors important in the final averaged models. Dashed lines indicate the upper and lower confidence intervals for predicted regressions.
Figure 2.4. Relationship between $\delta^{15}$N ratio and distance to the coast. Points shown represent isotopic values obtained from Eurasian otter (*Lutra lutra*) bone and muscle samples. Solid lines indicate the results predicted from models (general additive models for bone samples and general linear models for muscle samples) that included all the factors important in the final averaged models. Dashed lines indicate the upper and lower confidence intervals for predicted regressions.
2.4.2 Variation in $\delta^{13}C$

Bone showed a positive association between $\delta^{13}C$ and anadromous fish abundance (model 2, GAM: $z = 2.871, p = 0.004$), particularly for males (Fig. 2.5), whereas muscle tissue showed no such association. Muscle $\delta^{13}C$ signatures varied seasonally, in a sex dependent manner (Fig. 2.6; Table S 2.1), such that females showed greater enrichment in $^{13}$C than males in spring, whereas the reverse was true in winter (GAM: $z = 2.591, p = 0.01$). Body length and distance from the coast were important to some component models, but did not show any significant associations with $\delta^{13}C$ in either bone or muscle tissue (Tables S 2.1 and 2.2).

![Graph showing relationship between $\delta^{13}C$ ratio and abundance of anadromous fish.](image)

Figure 2.5. Relationship between $\delta^{13}C$ ratio and abundance of anadromous fish. Anadromous fish abundance is represented by the angling catch per unit effort (number of fish caught divided by the number of days anglers fished) of Atlantic salmon and sea trout per catchment per year. Red lines and points show female Eurasian otter (*Lutra lutra*) bone samples and blue shows male bone samples. Solid lines indicate the results predicted from general additive models that included all the factors important in the final averaged models. Dashed lines indicate the upper and lower confidence intervals for predicted regressions.
2.4.3 Nitrogen and carbon isotopic signatures

Considering both $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ together, it is apparent that isotopic signatures vary considerably across the landscape (Fig. 2.6). It is also apparent that isotopic niche size (standard ellipse area, SEA) was larger for males than females in spring, whereas in all other seasons SEA was similar between sexes (Fig. 2.7; Fig. S 2.4).
Figure 2.6. Spatial variability in isotopic signatures of Eurasian otters (Lutra lutra) sampled across Wales and bordering regions from 1993 – 2007. Landscape scale variation was interpolated from isotopic signatures obtained from otter bone and muscle tissue. Lighter colours represent more enrichment in the heavier isotopes and black circles show the locations of otters.
Figure 2.7. Isotopic niche measured in Eurasian otter (Lutra lutra) muscle tissue samples taken in each season. Outer ellipses show 95% standard ellipse areas; inner ellipses show the 95% confidence intervals around the bivariate means. Blue ellipses and points represent male otters and red represent females.
2.5 Discussion

Previous studies have shown otters have highly plastic diets which vary across different localities and habitats (e.g. Clavero et al. 2003; Almeida et al. 2012; Krawczyk et al. 2016). This is reflective of the opportunistic foraging behaviour of otters; otters will consume prey that are abundant and easy to catch, switching to alternative prey when preferred prey are less readily available (Jędrzejewska et al. 2001; Smiroldo et al. 2009). Our findings are consistent with these studies, with the broad range of δ¹⁵N and δ¹³C values observed indicating inter-individual variation in predation of species from different trophic levels and basal resources respectively (Kelly 2000; Matthews and Mazumder 2004; Newsome et al. 2009; Nolan et al. 2019).

In comparison to other studies investigating MDN acquisition by terrestrial/freshwater predators (e.g. bears, wolves and northern pike; Hilderbrand et al. 1999; Quinn et al. 2009; Levi et al. 2012; Koshino et al. 2013; Samways et al. 2018; Nolan et al. 2019), otters were on average more depleted in ¹³C, with isotopic values less reflective of a marine signal (typically reported as a δ¹³C value between -19‰ to -7‰; Fig. 2.2; Kelly 2000; Guiry 2019; Madgett et al. 2019), although no such depletion was observed for ¹⁵N. As carbon is a better indicator for basal resources (Kelly 2000; Marshall et al. 2007), signatures suggest that otters in Wales predominantly consumed freshwater prey, with marine prey utilised infrequently. Similarly, Darimont and Reimchen (2002) suggested that consumption of marine prey by wolves in British Columbia was limited, based on isotopic signatures being more reflective of terrestrial nutrients. Although we suggest that assimilation of MDNs across the otter population was low, variation in ¹³C and ¹⁵N enrichment between individuals indicated differential assimilation of MDNs, reflecting opportunistic foraging behaviours with differing access to marine resources (a similar finding to isotopic studies from other opportunistic consumers, e.g. Naiman et al. 2002; Inger et al. 2006; Dixon et al. 2012; Bašić et al. 2015; Gutmann Roberts et al. 2017; Nolan et al. 2019).

2.5.1 Distance from the coast

We expected enrichment in heavy isotopes to decline with distance from the coast as there would be fewer opportunities for otters to forage on marine prey. Although enrichment in ¹⁵N followed this pattern, no such association was observed with ¹³C. In most individuals, both near the coast and further inland, carbon signatures acquired from bone and muscle tissue did not indicate a strong marine signal (i.e. δ¹³C value between -19‰ to -7‰; Kelly 2000; Guiry 2019; Madgett et al. 2019), suggesting otters at all distances from the coast primarily consumed freshwater prey, with infrequent
consumption of marine prey resulting in only minor $^{13}$C enrichment in average carbon signatures of otters. Acquisition of MDNs via predation on marine prey in coastal habitats (Peterson and Fry 1987; O'Leary 1988; Kelly 2000; Inger et al. 2006) may also be balanced further inland by predation on anadromous fish (Naiman et al. 2002; Dixon et al. 2012) or stocked fish fed marine subsidies (Bašić et al. 2015; Gutmann Roberts et al. 2017; Nolan et al. 2019), thus producing similar average isotopic signatures in otters at all distances from the coast (Middelburg 2014; Hertz et al. 2016).

The depletion in $^{15}$N with distance from the coast seems unlikely to reflect a decline in marine derived nutrient inputs, given the lack of supporting evidence from carbon. Instead, we suggest that variation in nitrogen signatures might indicate changes in the trophic level of prey consumed (Deniro and Epstein 1980; Hobson and Clark 1992). Species richness typically increases with river discharge as there is an increase in ecosystem area (McGarvey and Hughes 2008; Iwasaki et al. 2012), so downstream regions are more likely to support a greater range of prey and include more species from higher trophic levels (Holt et al. 1999; Srivastava et al. 2008). Another explanation might be that anthropogenic inputs (such as fertilisers) washed into riverine systems through surface run-off and groundwater (Holt 2000; Hoffman et al. 2012; Mahl et al. 2015) are driving downstream enrichment in $^{15}$N, becoming incorporated into aquatic food webs and impacting nitrogen isotopic composition of various species (Lake et al. 2001; Anderson and Cabana 2005; Baeta et al. 2017; Lee et al. 2018). Whilst muscle showed a gradual decline in $^{15}$N enrichment, in bone this decline was not observed until after around 125 km, before which there was very little change in nitrogen enrichment. Otter home ranges can vary from 10 km to 40 km (Erlinge 1967; Chanin 2013; International Otter Survival Fund 2020) and the location of their range may change over their lifetime. Nitrogen signatures indicated by bone may thus reflect shifts in otter ranges to downstream, more species rich regions over broader (lifetime) temporal periods.

The enrichment in heavier isotopes ($^{13}$C and $^{15}$N) in some coastal otters indicated possible exploitation of marine prey, whereas comparative depletion in others suggested some coastal individuals only consumed freshwater/terrestrial prey, with intermediate signatures reflecting utilisation of both food sources. Similar findings have been observed in other apex predators, such as wolves and alligators, and have been attributed to the availability of multiple trophic networks (e.g. marine and terrestrial; Darimont et al. 2009; Rosenblatt et al. 2015). Coastal habitats in Wales are diverse, allowing otters to exploit a range of prey items. The variability in isotopic signatures in coastal regions may thus reflect the opportunistic foraging behaviour of otters and their
ability to exploit many different prey items in more diverse habitats. However, only a small proportion of otters had isotopic signatures that fell within the typical isotopic range for marine sources (Kelly 2000; Guiry 2019; Madgett et al. 2019), suggesting that even at the coast, otters in Wales primarily consume freshwater prey.

### 2.5.2 Anadromous fish abundance

We expected otters to be more enriched with heavier isotopes in areas with greater anadromous fish abundance due to the transport of MDNs into freshwater by anadromous fish as they migrate upstream (Naiman et al. 2002; Dixon et al. 2012). Both Atlantic salmon and sea trout adults migrate upstream from October to February in the UK, releasing MDNs into freshwater systems as they reproduce, excrete and die in rivers (Helfield and Naiman 2001; Gende et al. 2002; Naiman et al. 2002). Juveniles will initially possess a marine signature derived from their yolk sac, which will gradually become a freshwater signature as they switch to freshwater prey. Carbon isotopic signatures obtained from bone samples supported our hypothesis, revealing otters in areas with greater anadromous fish abundances were more enriched in $^{13}$C and suggesting that otters, particularly males, acquired MDNs via anadromous fish. This relationship was not observed in muscle tissue, likely because, in many cases, the shorter time signal indicated by muscle did not coincide with the seasonal influx of MDNs. This suggests that the signal is predominantly acquired via direct (and therefore seasonal) consumption of MDNs from anadromous fish rather than indirectly via freshwater prey in food webs enriched by MDNs. As most otters did not have isotopic signatures reflective of a typical marine signature (Kelly 2000, Guiry 2019; Madgett et al. 2019), it is unlikely that consumption of anadromous fish contributed a large proportion of the diet; instead signatures indicated that most otters predominantly consumed prey from freshwater sources.

Our hypothesis predicted that otter in areas with more anadromous fish would be enriched in $^{15}$N, however, both bone and muscle showed the inverse, with the greatest enrichment observed towards the east of Wales and England. Land use changes from west to east may be responsible for this, with a greater proportion of land used for arable and horticultural farming in the east (Morton et al. 2011). Associated change in anthropogenic inputs, such as fertilisers, are likely to contribute to increased $^{15}$N enrichment of aquatic systems. It should also be noted that anadromous fish abundances were based on angling data and thus relied on reports from fishermen, which could confound the results through abundance inaccuracies.

### 2.5.3 Seasonal variation
There is some evidence for niche separation between males and females in spring and to a lesser extent in winter. Males appear to exploit a broader range of species from different basal resources than females in spring, with a relative depletion in $^{13}$C. In winter, although the overall range is similar, males exhibit enrichment in $^{13}$C relative to females. Such differences suggest dietary preferences, or foraging strategies, differ between the sexes alongside seasonal variation in prey availability; for example, isotopic differences in winter may reflect greater consumption of migrating anadromous fish by males compared to females. Species specific identification of prey is required to explore this further.

### 2.5.4 Long and short-term isotopic signatures

We anticipated greater variability in muscle tissue compared to bone due to the shorter term signal, with such variation becoming smoothed over longer time scales (Hobson and Clark 1992; Dalerum and Angerbjörn 2005; Guiry et al. 2016). However, across the dataset we observed just as much variation in bone as muscle tissue. This suggests that individuals are dietary specialists, maintaining differential feeding habits over the long term, and resulting in high population level variation. These findings are consistent with other studies, revealing generalist predators may consist of individual specialists (Matthews and Mazumder 2004; Newsome et al. 2015; Robertson et al. 2015; Rosenblatt et al. 2015; Svanbäck et al. 2015). Dietary variation within otter populations is thus suggested to be driven by differences between individuals, rather than individual variation over time.

Bone samples were enriched in $^{13}$C compared to muscle tissue, but no differences were observed for nitrogen signatures. Previous studies have found bone tends to be enriched in $^{13}$C and depleted in $^{15}$N relative to muscle tissue (Sholto-douglas et al. 1991; Kelly 2000; Syväranta et al. 2008; Jansen et al. 2012), potentially indicating that differences observed here are due to tissue-dependent isotopic fractionation. However, differences between tissue types may also indicate dietary shifts across different timescales (Bearhop et al. 2003; Dalerum and Angerbjörn 2005; Jansen et al. 2012). This may explain the lack of difference in nitrogen signatures between the sample types; if otters consume more MDNs over their lifetime compared to their more recent diet, bone samples will become enriched in heavier isotopes, increasing the difference in carbon signatures between the sample types, but resulting in similar values for nitrogen signatures. The lack of a typical marine signature (Kelly 2000; Guiry 2019; Madgett et al. 2019) in either sample type suggests only occasional consumption of marine-derived prey by most individuals, with average isotopic signatures being more reflective of freshwater prey, with enrichment in heavier isotopes from marine species.
diluted by greater consumption of freshwater species less enriched in $^{15}$N and $^{13}$C. This may occur due to transient foraging opportunities, with MDN availability being influenced by seasonal changes in diadromous fish abundance, anglers feeding fish marine subsidies (Bašić et al. 2015; Gutmann Roberts et al. 2017; Nolan et al. 2019) and range changes affecting an otter’s proximity to the coast. These findings thus suggest that over long term, otters may opportunistically consume MDNs, but such nutrients are unlikely to contribute a large proportion to otter diet.

2.5.5 Limitations

Similarities in isotopic signatures between species can make interpretation of isotopic data difficult; this can occur due to different prey having indistinguishable isotopic signatures (Crawford et al. 2008; Layman et al. 2012), or through consumption of a variety of prey producing similar intermediate isotopic signatures in predators (Middelburg 2014; Hertz et al. 2016). Freshwater and marine species are often reported as having distinguishable isotopic signatures, with marine species more enriched in heavier isotopes (freshwater: $\delta^{13}$C -35‰ to -21‰ and $\delta^{15}$N 7‰ to 14‰; marine: $\delta^{13}$C -19‰ to -7‰ and $\delta^{15}$N 15‰ to 24‰; Kelly 2000; Guiry 2019; Madgett et al. 2019), however, freshwater species have been shown to express a greater range of isotopic values ($\delta^{13}$C of -36‰ to -3‰ and $\delta^{15}$N 5‰ to 25‰; Guiry 2019), resulting in overlap with marine signals. Such overlap in isotopic signatures may therefore lead to uncertainty when interpreting specific nutrient inputs from different basal resources, although marine species are on average more enriched in $^{15}$N and $^{13}$C (Petersen and Fry 1987; O'Leary 1988; Kelly 2000), providing some reassurance that enriched signatures are more likely to reflect consumption of these prey. Additionally, interpretation of isotopic data may have been influenced by differences in spatial distribution of bone and muscle tissue samples (Fig. 2.1), potentially skewing spatial conclusions and comparisons between long and short timescales. Whilst these caveats may limit the use of stable isotope analysis for deciphering specific trophic interactions, isotopic data still provided a unique insight into nutrient changes in otter diet and potential assimilation of MDNs.

2.5.6 Conclusions

Isotopic evidence suggests that although MDNs are consumed by otters, they only constitute a small proportion of the overall diet. These findings thus suggest that compared to North American systems with abundant salmonid runs, MDNs contribute relatively little to the nutrient composition of freshwater ecosystems in Wales and bordering counties, where anadromous fish are less abundant. This is likely to be
broadly representative of other northern European freshwaters, due to a similar lack of large populations of anadromous fish species. The broad range of isotopic values in both bone and muscle tissue suggests a broad diversity in diet across the otter population, incorporating long term individual specialisms, as well as opportunistic foraging reflecting variable prey availability. Due to the limitations here in the ability to decipher MDN contribution between marine species, anadromous species or fish fed marine fishmeal bait, we recommend further analysis into particular species being consumed by Eurasian otters, e.g. through methods such as DNA metabarcoding.

2.6 Acknowledgements
This work was funded by KESS II (Knowledge Economy Skills Scholarship) and the Wildlife Trust of South and West Wales who partnered this project. Additional funding was provided by NERC Life Sciences Mass Spectrometry facilities. Sample collection was conducted by Cardiff University Otter Project employees and placement students during post mortems of otter carcasses. Processing of samples for stable isotope analysis was conducted by Rona McGill at the NERC Life Sciences Mass Spectrometry facilities in East Kilbride.
2.7 Supplementary information

2.7.1 Anadromous fish catchments and abundance

Figure S 2.2. Anadromous fish abundance in river catchments across Wales and bordering regions from 1993 - 2007. Anadromous fish abundance is represented by the angling catch per unit effort (number of fish caught divided by the number of days anglers fished) of Atlantic salmon and sea trout per catchment per year. Information is only provided for catchments which Eurasian otters were sampled from. Catchments without relevant fishery statistics were clustered into ‘other’, along with areas in Wales that were not defined by a catchment. Areas in England that were not defined by a catchment were clustered together into ‘England’.


2.7.2 Scripts for analysing data in R

R Code used for analysing stable isotope data acquired from Eurasian otter bone and muscle tissue. Code was run using R [version 3.6.0] and R studio [version 1.2.1335] (R Core Team 2019) and converted into document format using R markdown (Xie et al. 2018; Allaire et al. 2020). Executable code is presented in grey boxes.

Load packages

```r
library(mgcv)
library(MuMIn)
library(arm)
library(rsq)
library(car)
library(ggplot2)
options(na.action = "na.pass")
```

Plot data to visualise changes in anadromous fish abundance per catchment per year

```r
cpueGraph <- read.csv("Fish abundance for graph.csv", header = T)
summary(cpueGraph)

ggplot(cpueGraph, aes(Year, Fish.catch)) +
  geom_line(colour = "grey60") +
  geom_point(colour = "black") +
  facet_wrap(~ Catchment, nrow = 12, ncol = 3) +
  theme_bw() + labs(y = "Anadromous fish catch per unit effort", cex=2) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        axis.text=element_text(size=10),
        axis.title=element_text(size=14,face="bold"))
```

Compare isotopic signatures obtained from Bone and Muscle tissue

Open data and run linear models, one to compare nitrogen in each sample type and one for carbon.

```r
BVM <- read.csv("SIBERBoneVsMuscle.csv")
summary(BVM)

Nitrogen.SampleType <- lm(BVM$iso2~BVM$group)
summary(Nitrogen.SampleType)
```
Carbon.SampleType <- lm(BVM$iso1 ~ BVM$group)
summary(carbon.SampleType)

Analyse isotopic variation in bone data with spatial, temporal and demographic variation
Open data, remove NA's, create a variable for distance in Km and check distribution of isotopic data.
Bone <- read.csv("Otter SIA Bone.csv", header = T)
levels(Bone$Sex)[1] <- NA
Bone1 <- na.omit(Bone)
Bone1$Dist <- Bone1$Distance.along.river/1000
summary(Bone1)

hist(Bone1$nitrogen)
hist(Bone1$carbon)

Run models using nitrogen as response variable
Run generalised additive model (GAM) and check model assumptions
BoneN1 <- gam(nitrogen ~ Sex + Length.total.mm. + Sex:Length.total.mm. + s(Dist, fx = F, k = -1) + Fish.Catch + Sex:Fish.Catch + Length.total.mm.:Fish.Catch, data = Bone1, family = Gamma(link = log))

summary.gam(BoneN1)
plot.gam(BoneN1)
par(mfrow=c(2,2))
gam.check(BoneN1)
par(mfrow=c(1,1))
sresidBN <- (BoneN1$residuals - mean(BoneN1$residuals))/sd(BoneN1$residuals)
qqp(sresidBN)
anova(BoneN1)
AIC(BoneN1)

Run generalised linear model (GLM) and check model assumptions
BoneN2 <- glm(nitrogen ~ Sex + Length.total.mm. + Sex:Length.total.mm. + Dist + Fish.Catch + Sex:Fish.Catch + Length.total.mm.:Fish.Catch, data = Bone1, family = Gamma(link = log))

summary(BoneN2)
```r
par(mfrow = c(2,2))
plot(BoneN2)
par(mfrow = c(1,1))
sresidBN2 <- (BoneN2$residuals - mean(BoneN2$residuals))/sd(BoneN2$residuals)
qqp(sresidBN2)

Compare GAM and GLM
anova(BoneN2, BoneN1, test = "Chisq")
AIC(BoneN2, BoneN1)

GAM found to be better, therefore use this model to identify significant associations
Conduct model averaging
BoneN1.set <- dredge(BoneN1)
BoneN1.avg <- model.avg(BoneN1.set, subset = delta < 2, fit = TRUE)

There was only one model (nitrogen ~ Fish.Catch + s(Dist, fx = F, k = -1) + 1),
therefore extract the top model and interpret results from this
BoneN1.avg <- get.models(BoneN1.set, subset = delta < 2, fit = TRUE)
BoneN1.avg

topNmodel <- gam(nitrogen ~ Fish.Catch + s(Dist, fx = F, k = -1) + 1, +
               data = Bone1, family = Gamma(link = log))
summary(topNmodel)

Run models using carbon as the response variable
Run GAM and check model assumptions
BoneC1 <- gam(carbon ~ Sex + Length.total.mm. + Sex:Length.total.mm. +
               + s(Dist, fx = F, k = -1) + Fish.Catch + Sex:Fish.Catch +
               Length.total.mm.:Fish.Catch, data = Bone1,
               family = gaussian(link = identity))
summary.gam(BoneC1)
plot.gam(BoneC1)
par(mfrow=c(2,2))
gam.check(BoneC1)
par(mfrow=c(1,1))
sresidBC <- (BoneC1$residuals - mean(BoneC1$residuals))/sd(BoneC1$residuals)
qqp(sresidBC)
```
Model checks revealed there were two outliers, therefore remove these data points and re-run the model.

```r
Bone1C <- Bone1[-c(8, 29),]
BoneC1 <- gam(carbon ~ Sex + Length.total.mm. + Sex:Length.total.mm. + 
              s(Dist, fx=F, k=-1) + Fish.Catch + Sex:Fish.Catch + 
              Length.total.mm.:Fish.Catch, data = Bone1C,
              family = gaussian(link = identity))
```

```r
summary.gam(BoneC1)
plot.gam(BoneC1)
par(mfrow=c(2,2))
gam.check(BoneC1)
par(mfrow=c(1,1))
anova(BoneC1)
AIC(BoneC1)
```

Run GLM and check model assumptions.

```r
BoneC2 <- glm(carbon ~ Sex + Length.total.mm. + Sex:Length.total.mm. + 
              Dist + Fish.Catch + Sex:Fish.Catch + Length.total.mm.:Fish.Catch, 
              data = Bone1)
```

```r
summary(BoneC2)
par(mfrow=c(2,2))
plot(BoneC2)
par(mfrow=c(1,1))
sresidBC2 <- (BoneC2$residuals - mean(BoneC2$residuals))/sd(BoneC2$residuals)
qqp(sresidBC2)
```

GLM also identified two outliers, both of which were the same as identified in the GAM. Therefore, re-run the model without these outliers.

```r
BoneC2 <- glm(carbon ~ Sex + Length.total.mm. + Sex:Length.total.mm. + Dist + 
              Fish.Catch + Sex:Fish.Catch + Length.total.mm.:Fish.Catch, 
              data = Bone1C)
```

```r
summary(BoneC2)
par(mfrow=c(2,2))
plot(BoneC2)
par(mfrow=c(1,1))
```
\[
\text{sresidBC2} \leftarrow (\text{BoneC2}\$\text{residuals} - \text{mean}(\text{BoneC2}\$\text{residuals})) / \text{sd}(\text{BoneC2}\$\text{residuals}) \\
\text{qqp}(\text{sresidBC2})
\]

Compare GAM and GLM

\text{anova}(\text{BoneC2}, \text{BoneC1}, \text{test} = "\text{Chisq}")

\text{AIC}(\text{BoneC2}, \text{BoneC1})

GAM found to be better, therefore use this model to identify significant associations

Conduct model averaging and use summary to identify significant associations

\text{BoneC1.set} \leftarrow \text{dredge}(\text{BoneC1})

\text{BoneC1.avg} \leftarrow \text{model.avg}(\text{BoneC1.set}, \text{subset} = \text{delta} < 2, \text{fit} = \text{TRUE})

\text{summary}(\text{BoneC1.avg})

Analyse variation in muscle data with spatial, temporal and demographic variation

Open data, reformat season from a numeric into a factor, remove NA’s, create a variable for distance in Km and check distribution of isotopic data

\text{Muscle} \leftarrow \text{read.csv("Otter SIA Muscle MAC MBN.csv", header = T)}

\text{Muscle}\$\text{Season} \leftarrow \text{as.factor}(\text{Muscle}\$\text{Season})

\text{Muscle1} \leftarrow \text{na.omit}(\text{Muscle})

\text{Muscle1}\$\text{Dist} \leftarrow \text{Muscle1}\$\text{Distance.along.river}/1000

\text{summary}(\text{Muscle1})

\text{hist}(\text{Muscle1}\$\text{nitrogen})

\text{hist}(\text{Muscle1}\$\text{carbon})

Run models using nitrogen as the response variable

Run GAM and check model assumptions

\text{MuscleN1} \leftarrow \text{gam}(\text{nitrogen} \sim \text{Sex} + \text{Length.total.mm.} + \text{Sex:Length.total.mm.} +
\hspace{1cm} + \text{s(}\text{Dist, fx} = \text{F, k} = -1\text{)} + \text{Fish.Catch} + \text{Sex:Fish.Catch} +
\hspace{1cm} + \text{Length.total.mm.:Fish.Catch} + \text{s(}\text{Dist, fx} = \text{F, k} = -1\text{, by=Season}) +
\hspace{1cm} + \text{Season} + \text{Sex:Fish.Catch} + \text{Length.total.mm.:Season} + \text{Fish.Catch:Season},
\hspace{1cm} \text{data} = \text{Muscle1})

\text{summary.gam}(\text{MuscleN1})

Output of GAM revealed distance was linear, therefore run as a GLM instead and check model assumptions
```r
MuscleN2 <- glm(nitrogen ~ Sex + Length.total.mm. + Sex:Length.total.mm. +
                Dist + Fish.Catch + Sex:Fish.Catch +
                Length.total.mm.:Fish.Catch + Dist:Season +
                Season + Sex:Season + Length.total.mm.:Season + Fish.Catch:Season,
                data = Muscle1)

summary(MuscleN2)
par(mfrow = c(2,2))
plot(MuscleN2)
par(mfrow = c(1,1))
sresidMN <- (MuscleN2$residuals - mean(MuscleN2$residuals)) / sd(MuscleN2$residuals)
qqp(sresidMN)

Conduct model averaging
MuscleN2.set <- dredge(MuscleN2)
MuscleN2.avg <- model.avg(MuscleN2.set, subset = delta < 2, fit = TRUE)
summary(MuscleN2.avg)

Run models using carbon as the response variable
Run GAM and check model assumptions
MuscleC1 <- gam(carbon ~ Sex + Length.total.mm. + Sex:Length.total.mm. +
                + s(Dist, fx = F, k = -1) + Fish.Catch + Sex:Fish.Catch +
                Length.total.mm.:Fish.Catch + Dist:Season +
                Season + Sex:Season + Length.total.mm.:Season + Fish.Catch:Season,
                data = Muscle1)

summary.gam(MuscleC1)
plot.gam(MuscleC1)
par(mfrow=c(2,2))
gam.check(MuscleC1)
par(mfrow=c(1,1))
sresidMC <- (MuscleC1$residuals - mean(MuscleC1$residuals)) / sd(MuscleC1$residuals)
qqp(sresidMC)
anova(MuscleC1)
AIC(MuscleC1)
```
Run GLM and check model assumptions

```r
MuscleC2 <- glm(carbon ~ Sex + Length.total.mm. + Sex:Length.total.mm. + Dist + Fish.Catch + Sex:Fish.Catch + Length.total.mm.:Fish.Catch + Dist:Season + Season + Sex:Season + Length.total.mm.:Season + Fish.Catch:Season, data = Muscle1)
```

```r
summary(MuscleC2)
rsq(MuscleC2)
par(mfrow = c(2,2))
plot(MuscleC2)
par(mfrow = c(1,1))
sresidMC2 <- (MuscleC2$residuals - mean(MuscleC2$residuals)) / sd(MuscleC2$residuals)
qqp(sresidMC2)
```

Compare GAM and GLM

```r
anova(MuscleC2, MuscleC1, test = "Chisq")
AIC(MuscleC2, MuscleC1)
```

GAM found to be better, therefore use this model to identify significant associations

Conduct model averaging and use summary to identify significant associations

```r
MuscleC1.set <- dredge(MuscleC1)
MuscleC1.avg <- model.avg(MuscleC1.set, subset = delta < 2, fit = TRUE)
summary(MuscleC1.avg)
```

Season was significant therefore relevel and rerun the model to find out which levels were significantly different from each other

```r
levels(Muscle1$Season)
Muscle1$Season <- relevel(Muscle1$Season, ref = "2")
update(MuscleC1, . ~.)
MuscleC1.set <- dredge(MuscleC1)
MuscleC1.avg <- model.avg(MuscleC1.set, subset = delta < 2, fit = TRUE)
summary(MuscleC1.avg)
```

```r
Muscle1$Season <- relevel(Muscle1$Season, ref = "3")
update(MuscleC1, . ~.)
MuscleC1.set <- dredge(MuscleC1)
```
MuscleC1.avg <- `model.avg`(MuscleC1.set, subset = delta < 2, fit = TRUE)
summary(MuscleC1.avg)

Return data to the original order
Muscle1$Season <- `relevel`(Muscle1$Season, ref = "2")
Muscle1$Season <- `relevel`(Muscle1$Season, ref = "1")
levels(Muscle1$Season)

Plot Significant Terms
Start by making a dummy dataset including only the variables in the final averaged (or top) model, make predictions using dummy data and the final averaged (or top) model and then plot data with predicted trends

Nitrogen bone and muscle data plotted against anadromous fish abundance
Change graphical output so to plot two graphs side by side
par(mfrow=c(1,2))

First plot data/predictions from the Bone model
gampdatBNFish <- `expand.grid`(Fish.Catch = seq(min(Bone1$Fish.Catch), max(Bone1$Fish.Catch), length.out = 1000),
Dist=mean(Bone1$Dist, na.rm= T))
gampredBNFish <- `predict`(topNmodel, newdata= gampdatBNFish, na.rm= T,
type="response", se.fit=TRUE)
predframegamBNFish <- `data.frame`(gampdatBNFish, preds= gampredBNFish$fit,
se=gampredBNFish$se.fit)
plot(nitrogen ~ Fish.Catch, data = Bone1, bty = 'L', cex.axis=1.5, cex.lab=1.75,
cex.main=1.75, ylab = "", xlab = 'Anadromous fish abundance',
main = "Bone samples", ylim = c(9,23), xlim = c(0,1.2))
title(ylab = expression({delta}^{15}N~"\mu"2030'), line = 2.4, cex.lab = 1.75)
lines(preds ~ Fish.Catch, data = predframegamBNFish, lwd=2)
lines(preds + se ~ Fish.Catch, data = predframegamBNFish, lty = 2, lwd=2, col = "red")
lines(preds - se ~ Fish.Catch, data = predframegamBNFish, lty = 2, lwd=2, col = "red")

Second plot data/predictions from the Muscle model
gampdatMNFish <- `expand.grid`(Fish.Catch = seq(min(Muscle1$Fish.Catch),
max(Muscle1$Fish.Catch), length.out = 1000),
Dist=mean(Muscle1$Dist, na.rm= T),
Length.total.mm.=mean(Muscle1$Length.total.mm., na.rm= T),
Sex = c("F")

gampredMNFish <- predict(MuscleN2.avg, newdata = gampdatMNFish, na.rm = T, type = "response", se.fit = TRUE)
predframegamMNFish <- data.frame(gampdatMNFish, preds = gampredMNFish$fit, se = gampredMNFish$se.fit)

plot(nitrogen ~ Fish.Catch, data = Muscle1, bty = 'L', cex.axis = 1.5, cex.lab = 1.75, cex.main = 1.75, ylab = "", xlab = 'Anadromous fish abundance', main = "Muscle samples", ylim = c(9, 23), xlim = c(0, 1.2))
title(ylab = expression(delta^15*N), line = 2.4, cex.lab = 1.75)
lines(preds ~ Fish.Catch, data = predframegamMNFish, lwd = 2)
lines(preds + se ~ Fish.Catch, data = predframegamMNFish, lty = 2, lwd = 2, col = "red")
lines(preds - se ~ Fish.Catch, data = predframegamMNFish, lty = 2, lwd = 2, col = "red")

Reset graphical parameters
par(mfrow = c(1, 1))

Nitrogen bone and muscle data plotted against distance to the coast
Change graphical output so to plot two graphs side by side
par(mfrow = c(1, 2))

First plot data/predictions from the Bone model
gampdatBNDist <- expand.grid(Dist = seq(min(Bone1$Dist), max(Bone1$Dist), length.out = 1000), Fish.Catch = mean(Bone1$Fish.Catch, na.rm = T))
gampredBNDist <- predict(topNmodel, newdata = gampdatBNDist, na.rm = T, type = "response", se.fit = TRUE)
predframegamBNDist <- data.frame(gampdatBNDist, preds = gampredBNDist$fit, se = gampredBNDist$se.fit)

plot(nitrogen ~ Dist, data = Bone1, bty = 'L', cex.axis = 1.5, cex.lab = 1.75, cex.main = 1.75, ylim = c(10, 23), ylab = "", xlab = 'Distance from the coast (Km)', main = "Bone samples")
title(ylab = expression((delta)^15*N), line = 2.4, cex.lab = 1.75)
lines(preds ~ Dist, data = predframegamBNDist, lwd = 2)
lines(preds + se ~ Dist, data = predframegamBNDist, lty = 2, lwd = 2, col = "red")
lines(preds - se ~ Dist, data = predframegamBNDist, lty = 2, lwd = 2, col = "red")
Second plot data/predictions from the Muscle model

```r
gampdatMNDist <- expand.grid(Dist = seq(min(Muscle1$Dist), max(Muscle1$Dist), length.out = 1000),
                              Fish.Catch=mean(Muscle1$Fish.Catch, na.rm=T),
                              Length.total.mm.=mean(Muscle1$Length.total.mm., na.rm=T),
                              Sex=c("F"))

gampredMNDist <- predict(MuscleN2.avg, newdata = gampdatMNDist, na.rm=T,
                          type="response", se.fit=TRUE)
predframegamMNDist <- data.frame(gampdatMNDist, preds=gampredMNDist$fit, se=gampredMNDist$se.fit)

plot(nitrogen ~ Dist, data = Muscle1, bty = 'L', cex.axis=1.5, cex.lab=1.75,
     cex.main=1.75, ylim=c(10,23), ylab = '', xlab = 'Distance from the coast (Km)',
     main = "Muscle samples")

title(ylab = expression((delta)^15*N\sim\nu2030'), line = 2.4, cex.lab = 1.75)

lines(preds ~ Dist, data = predframegamMNDist, lwd=2)
lines(preds + se ~ Dist, data = predframegamMNDist, lty =2, lwd=2, col = "red")
lines(preds - se ~ Dist, data = predframegamMNDist, lty =2, lwd=2, col = "red")

Reset graphical parameters
par(mfrow=c(1,1))
```

Carbon bone data plotted to show interaction between anadromous fish abundance and sex

Plot data and colour points red for females and blue for males

```r
plot(carbon ~ Fish.Catch, data = Bone1, bty = 'L', cex.axis=1.25, cex.lab=1.5, ylab = '',
     xlab = 'Anadromous fish abundance')

title(ylab = expression((delta)^13*C\sim\nu2030'), line = 2.5, cex.lab = 1.5)

points(carbon ~ Fish.Catch, data = Bone1, subset = (Sex == "F"), col = "red")
points(carbon ~ Fish.Catch, data = Bone1, subset = (Sex == "M"), col = "blue")
```

Make predictions using females as the base level for the model and add predicted trend lines for female models

```r
gampdatBCFish <- expand.grid(Fish.Catch = seq(min(Bone1$Fish.Catch),
                                              max(Bone1$Fish.Catch), length.out = 1000),
                              Dist=mean(Bone1$Dist, na.rm=T),
                              Length.total.mm.=mean(Bone1$Length.total.mm., na.rm=T),
```
Make predictions with males as the base level for the model and add predicted trend lines for male otters

gampdatBCFish <- \texttt{expand.grid}(\texttt{Fish.Catch = seq}(\texttt{min(Bone1$Fish.Catch)}, \texttt{max(Bone1$Fish.Catch)}, \texttt{length.out = 1000}),
\texttt{Dist=mean}(\texttt{Bone1$Dist, na.rm=T}),
\texttt{Length.total.mm.=mean}(\texttt{Bone1$Length.total.mm., na.rm=T}),
\texttt{Sex=c("M")})

gampredBCFish <- \texttt{predict}(\texttt{BoneC1.avg, newdata= gampdatBCFish, na.rm=T},
\texttt{type="response"}, \texttt{se.fit=TRUE})
predframegamBCFish <- \texttt{data.frame}(gampdatBCFish, \texttt{preds=gampredBCFish$fit},
\texttt{se=gampredBCFish$se.fit})
\texttt{lines(preds ~ Fish.Catch, data = predframegamBCFish, col = "blue", lwd = 1.5)}
\texttt{lines(preds + se ~ Fish.Catch, data = predframegamBCFish, lty = 2, col = "blue")}
\texttt{lines(preds - se ~ Fish.Catch, data = predframegamBCFish, lty = 2, col = "blue")}

Use SIBER package (Jackson et al. 2011) to plot ellipses that represent isotopic niches for significant categorical variables

Set the parameters for plotting the data

\texttt{set.seed(1)}
\texttt{library(SIBER)}
\texttt{palette(c("green", "purple"))}
\texttt{community.hulls.args <- list(col = 1, lty = 1, lwd = 1)}
\texttt{group.ellipses.args  <- list(n = 100, p.interval = 0.95, lty = 1, lwd = 2)}
\texttt{group.hull.args      <- list(lty = 2, col = "grey20")}
Plot isotopic niches depicted by each sample type
Read in the data, make a SIBER object and then plot the data (red = muscle tissue, blue = bone). Add predicted ellipses that encompass approximately 95% of the data (large ellipses) and ellipses that depict the 95% confidence interval around the bivariate means (small ellipses). Also, add a dashed line to denote typical threshold between freshwater and marine carbon isotopic signatures

```r
BoneVsMuscle <- read.csv("SIBERBoneVsMuscle.csv", header=T)
siber.BoneVsMuscle <- createSiberObject(BoneVsMuscle)

plotSiberObject(siber.BoneVsMuscle,
    ax.pad = 2,
    hulls = F, community.hulls.args,
    ellipses = T, group.ellipses.args,
    group.hulls = F, group.hull.args,
    bty = "L",
    iso.order = c(1, 2),
    xlab = "",
    ylab = "",
    cex.axis = 1.5)

title(ylab = expression({delta}^{15}N~\u2030), line = 2.4, cex.lab = 1.75)
title(xlab = expression({delta}^{13}C~\u2030), line = 3, cex.lab = 1.75)
plotGroupEllipses(siber.BoneVsMuscle, n = 100, p.interval = 0.95, ci.mean = T,
    lty = 1, lwd = 2)
abline(v=-18, col="grey20", lty=2, lwd=2)
```

Calculate summary statistics for each group (sample type): TA, SEA, SEAc and range of isotopic values

```r
group.ML <- groupMetricsML(siber.BoneVsMuscle)
print(group.ML)
summary(BoneVsMuscle[BoneVsMuscle$group == 1, ])
summary(BoneVsMuscle[BoneVsMuscle$group == 2, ])
```

Fit Bayesian model to the data. First set the options for running jags

```r
library(rjags)
parms <- list()
parms$n.iter <- 2*10^5  # number of iterations to run the model for
```
parms$n.burnin <- 1 * 10^3  # discard the first set of values
parms$n.thin <- 10        # thin the posterior by this many
parms$n.chains <- 2      # run this many chains

Define the priors

priors <- list()
priors$R <- 1 * diag(2)
priors$k <- 2
priors$tau.mu <- 1.0E-3

Fit ellipses using the priors and use these to calculate SEA.B for each group. SEA.B is
then plotted and red x's are added to denote the maximum likelihood estimates of
SEA-c to the Bayesian estimates

ellipses.posterior <- siberMVN(siber.BoneVsMuscle, parms, priors)
SEA.B <- siberEllipses(ellipses.posterior)

siberDensityPlot(SEA.B, xticklabels = c("", ""),
                 xlab = "",
                 ylab = "",
                 bty = "L",
                 las = 1,
                 main = "",
                 cex.axis=1.5)

points(1:ncol(SEA.B), group.ML[3,], col="red", pch="x", lwd = 2)
title(main = "SIBER ellipses on each group", cex.main = 1.5)
title(ylab = expression("Standard Ellipse Area " (^u2030^2)), line = 2.3,
     cex.lab = 1.5)
title(xlab = "Sample Type", line = 2.5, cex.lab = 1.5)
axis(1, at=c(1,2), labels = c("Bone", "Muscle"), las=1, cex.axis=1.5)

Calculate credible intervals and modes

cr.p <- c(0.95, 0.99)
SEA.B.credibles <- lapply(
  as.data.frame(SEA.B),
  function(x,...){tmp<-hdrcde::hdr(x)$hdr},
  prob = cr.p)
SEA.B.credibles
SEA.B.modes <- lapply(as.data.frame(SEA.B), function(x,...){tmp< -hdrcde::hdr(x)$mode}, prob = cr.p, all.modes=T)

Carbon muscle data plotted to show interaction between season and sex
Read in data and make a SIBER object
MuscleSexSeason <- read.csv("SIBERMuscleSexSeason.csv", header=T)
siber.MuscleSexSeason <- createSiberObject(MuscleSexSeason)

Plot the data (red = female, blue = male) and then add the 95% confidence interval around the bivariate means
plotSiberObject(siber.MuscleSexSeason,
               ax.pad = 2,
hulls = F, community.hulls.args,
ellipses = T, group.ellipses.args,
group.hulls = F, group.hull.args,
bty = "L",
iso.order = c(1, 2),
xlab = expression({delta}^13*C~\u2030),
ylab = expression({delta}^15*N~\u2030)
)

plotGroupEllipses(siber.MuscleSexSeason, n = 100, p.interval = 0.95, ci.mean = T,
                   lty = 1, lwd = 2)

Calculate summary statistics for each group (sex): TA, SEA, SEAc and range of isotopic values

Calculate Layman metrics on each of the communities (season).

summary(MuscleSexSeason[MuscleSexSeason$gender == 1, ])
summary(MuscleSexSeason[MuscleSexSeason$gender == 2, ])

Community.ML <- communityMetricsML(siber.MuscleSexSeason)
print(community.ML)
Fit bayesian model to the data: Ellipses fit using the priors and use these to calculate SEA.B for each group. SEA.B is then plotted and red x’s are added to denote the maximum likelihood estimates of SEA-c to the Bayesian estimates.

```r
ellipses.posterior <- siberMVN(siber.MuscleSexSeason, parms, priors)
SEA.B <- siberEllipses(ellipses.posterior)
```

```r
siberDensityPlot(SEA.B,
    xticklabels = c("", "", "", "", "", "", "", ""),
    xlab = "", ylab = "",
    bty = "L",
    las = 1,
    main = "",
    ylim = c(0,30),
    cex.axis=1.5
  )

points(1:ncol(SEA.B), group.ML[3,], col="red", pch="x", lwd = 2)
title(main = "SIBER ellipses on each group", cex.main = 1.5)
title(ylab = expression("Standard Ellipse Area " (^2)), line = 2.3,
    cex.lab = 1.5)
axis(1, at=seq(1,8,1), labels = c("F.Sp", "M.Sp", "F.Su", "M.Su", "F.A", "M.A","M.W",
    "F.W"), las=1, cex.axis=1.5)
```

Calculate credible intervals and modes

```r
cr.p <- c(0.95, 0.99)
SEA.B.credibles <- lapply(
    as.data.frame(SEA.B),
    function(x,...){tmp<-hdrcde::hdr(x)$hdr},
    prob = cr.p)
SEA.B.credibles
```

```r
SEA.B.modes <- lapply(
    as.data.frame(SEA.B),
    function(x,...){tmp<-hdrcde::hdr(x)$mode},
    prob = cr.p, all.modes=T)
SEA.B.modes
```

Plot season:sex interaction as four graphs (one for each season) in one window.
First set the arguments to be passed to each of the plotting functions

```r
par(mfrow=c(2,2))
palette(c("blue", "red"))
```

Plot data for Spring

```r
MuscleSexSpring <- read.csv("SIBERMuscleSexSpring.csv", header=T)
siber.MuscleSexSpring <- createSiberObject(MuscleSexSpring)

plotSiberObject(siber.MuscleSexSpring,
    ax.pad = 2,
    Hulls = F, community.hulls.args,
    ellipses = T, group.ellipses.args,
    group.hulls = F, group.hull.args,
    bty = "L",
    iso.order = c(1, 2),
    xlab = "",
    ylab = "",
    cex.axis = 1.5,
    x.limits = c(-33,-17),
    y.limits = c(8,21)
)

title(main = "Spring", cex.main=2)
title(ylab = expression((delta)^{15}N~\text{"\u2030"}), line = 2.4, cex.lab = 1.75)
title(xlab = expression((delta)^{13}C~\text{"\u2030"}), line = 3, cex.lab = 1.75)
plotGroupEllipses(siber.MuscleSexSpring, n = 100, p.interval = 0.95, ci.mean = T, lty = 1, lwd = 2)
```

Plot data for Summer

```r
MuscleSexSummer <- read.csv("SIBERMuscleSexSummer.csv", header=T)
siber.MuscleSexSummer <- createSiberObject(MuscleSexSummer)

plotSiberObject(siber.MuscleSexSummer,
    ax.pad = 3,
    Hulls = F, community.hulls.args,
    ellipses = T, group.ellipses.args,
    group.hulls = F, group.hull.args,
    bty = "L",
    iso.order = c(1, 2),
)```
Plot data for Autumn

MuscleSexAutumn <- read.csv("SIBERMuscleSexAutumn.csv", header=T)
siber.MuscleSexAutumn <- createSiberObject(MuscleSexAutumn)

plotSiberObject(siber.MuscleSexAutumn,
ax.pad = 2,
hulls = F, community.hulls.args,
ellipses = T, group.ellipses.args,
group.hulls = F, group.hull.args,
bty = "L",
iso.order = c(1, 2),
xlab = "",
ylab = "",
cex.axis = 1.5,
x.limits = c(-33,-17),
y.limits = c(8,21))

Plot data for Winter

MuscleSexWinter <- read.csv("SIBERMuscleSexWinter.csv", header=T)
siber.MuscleSexWinter <- createSiberObject(MuscleSexWinter)
plotSiberObject(siber.MuscleSexWinter, 
   ax.pad = 2, 
   hulls = F, community.hulls.args, 
   ellipses = T, group.ellipses.args, 
   group.hulls = F, group.hull.args, 
   bty = "L", 
   iso.order = c(1, 2), 
   xlab = "", 
   ylab = "", 
   cex.axis = 1.5, 
   cex.main = 10, 
   x.limits = c(-33,-17), 
   y.limits = c(8,21),
)

title(main = "Winter", cex.main=1.75)
title(ylab = expression({delta}^15*N~\u2030), line = 2.4, cex.lab = 1.75)
title(xlab = expression({delta}^13*C~\u2030), line = 3, cex.lab = 1.75)
plotGroupEllipses(siber.MuscleSexWinter, n = 100, p.interval = 0.95, ci.mean = T, 
   lty = 1, lwd = 2)

Reset graphical parameters
par(mfrow=c(1,1))

2.7.3  Model averaging tables
Table S 2.1. Comparison of component models used to make the averaged model for the general linear model (GLM) of muscle δ\(^{15}\)N and general additive models (GAM) of bone and muscle δ\(^{13}\)C in Eurasian otters (*Lutra lutra*) from mid-England, Wales and bordering counties 1993 - 2007. The degrees of freedom (df), log-likelihood (Log(L)), Akaike Information Criterion for small sample size (AICc), delta AICc and the AICc weight are given for each of the component models. Delta AICc shows the difference between the AICc for a particular model and the lowest AICc amongst all the models. AICc weights show the likelihood of each model; as the weight increases, the model becomes more likely.

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Explanatory Variable</th>
<th>Fixed Variables</th>
<th>df</th>
<th>logLik</th>
<th>AICc</th>
<th>delta AICc</th>
<th>AICc weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bone δ(^{13})C</td>
<td>Anadromous fish abundance + s(Distance to the coast, fx=F, k=-1) + Sex + Anadromous fish abundance:Sex</td>
<td>7.48</td>
<td>-315.40</td>
<td>646.59</td>
<td>0.00</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Bone δ(^{13})C</td>
<td>Anadromous fish abundance + s(Distance to the coast, fx=F, k=-1) + Body length (mm) + Sex + Anadromous fish abundance:Sex</td>
<td>8.47</td>
<td>-315.26</td>
<td>648.55</td>
<td>1.96</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>δ(^{15})N</td>
<td>Distance to the coast + Anadromous fish abundance + Sex</td>
<td>5.00</td>
<td>-419.74</td>
<td>849.79</td>
<td>0.00</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>δ(^{15})N</td>
<td>Distance to the coast + Anadromous fish abundance</td>
<td>4.00</td>
<td>-420.85</td>
<td>849.90</td>
<td>0.11</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>δ(^{15})N</td>
<td>Distance to the coast + Anadromous fish abundance + Body length (mm)</td>
<td>5.00</td>
<td>-420.38</td>
<td>851.06</td>
<td>1.28</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>δ(^{15})N</td>
<td>Distance to the coast + Anadromous fish abundance + Sex + Anadromous fish abundance:Sex</td>
<td>6.00</td>
<td>-419.67</td>
<td>851.77</td>
<td>1.99</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance + Body length (mm) + s(Distance to the coast, fx=F, k=-1) + Season</td>
<td>11.02</td>
<td>-400.44</td>
<td>824.31</td>
<td>0.00</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance:Body length (mm)</td>
<td>11.87</td>
<td>-399.54</td>
<td>824.44</td>
<td>0.12</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Season:Sex</td>
<td>14.69</td>
<td>-396.38</td>
<td>824.63</td>
<td>0.32</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance:Body length (mm):Sex + Season:Sex</td>
<td>15.79</td>
<td>-395.08</td>
<td>824.63</td>
<td>0.32</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance + Body length (mm) + s(Distance to the coast, fx=F, k=-1) + Season + Sex + Anadromous fish abundance:Sex + Body length (mm):Sex + Season:Sex</td>
<td>16.81</td>
<td>-393.94</td>
<td>824.77</td>
<td>0.46</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance:Body length (mm) + s(Distance to the coast, fx=F, k=-1) + Season + Sex + Anadromous fish abundance:Body length (mm):Sex</td>
<td>13.88</td>
<td>-397.44</td>
<td>824.87</td>
<td>0.55</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance + Body length (mm) + s(Distance to the coast, fx=F, k=-1) + Season + Sex + Anadromous fish abundance:Sex + Body length (mm):Sex</td>
<td>13.98</td>
<td>-397.51</td>
<td>825.22</td>
<td>0.91</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance + Body length (mm) + s(Distance to the coast, fx=F, k=-1) + Season + Sex + Anadromous fish abundance:Body length (mm):Sex + Season:Sex</td>
<td>16.68</td>
<td>-394.35</td>
<td>825.27</td>
<td>0.96</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance:Sex + Season:Sex</td>
<td>15.71</td>
<td>-395.53</td>
<td>825.31</td>
<td>1.00</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>length (mm):Sex</td>
<td>13.00</td>
<td>-398.97</td>
<td>825.89</td>
<td>1.57</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance + Body length (mm) + s(Distance to the coast, fx=F, k=-1) + Season + Sex + Anadromous fish abundance:Body length (mm) + Season:Sex</td>
<td>15.58</td>
<td>-396.08</td>
<td>826.12</td>
<td>1.80</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Muscle δ(^{13})C</td>
<td>Anadromous fish abundance + Body length (mm) + s(Distance to the coast, fx=F, k=-1) + Season + Sex + Anadromous fish abundance:Body length (mm):Sex + Anadromous fish abundance:Sex + Body length (mm):Sex</td>
<td>14.89</td>
<td>-396.89</td>
<td>826.12</td>
<td>1.80</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>
Table S 2.2. Conditional average model outputs for general linear model (GLM) of muscle δ¹⁵N and general additive models (GAM) of bone and muscle δ¹³C in Eurasian otters (*Lutra lutra*) from mid-England, Wales and bordering counties 1993 – 2007. Averaged models were created from general additive models (GAM) with Akaike’s Information Criterion (AIC) values of less than two. The explanatory variable was the linear corrected stable isotope δ¹⁵N or δ¹³C in muscle samples from otters. Fixed variables preceded by an ‘s’ represent a variable with a smoothing function with the number following representing the knot number. Effect size (estimate), standard error (std. error), adjusted standard error (adjusted SE), z value and p value (Pr(>|z|)) are given for each fixed variable in the final averaged model.

| Sample Type | Explanatory Variable | Fixed Variable | Estimate | Std. Error | Adjusted SE | z value | Pr(>|z|) |
|-------------|----------------------|----------------|----------|------------|-------------|---------|----------|
| Bone        | Anadromous fish abundance | 1.91 | 0.92 | 0.93 | 2.06 | 0.04 |
|             | Male (compared to female) | -0.22 | 0.43 | 0.44 | 0.50 | 0.61 |
|             | Anadromous fish abundance:Sex | 3.77 | 1.36 | 1.37 | 2.76 | 0.01 |
|             | (Distance to the coast).1 | -0.03 | 0.47 | 0.48 | 0.06 | 0.96 |
|             | (Distance to the coast).2 | 0.16 | 1.21 | 1.22 | 0.13 | 0.90 |
|             | (Distance to the coast).3 | 0.02 | 0.26 | 0.26 | 0.08 | 0.94 |
|             | (Distance to the coast).4 | -0.47 | 0.74 | 0.74 | 0.64 | 0.52 |
|             | (Distance to the coast).5 | 0.16 | 0.27 | 0.27 | 0.60 | 0.55 |
|             | (Distance to the coast).6 | -0.43 | 0.58 | 0.58 | 0.74 | 0.46 |
|             | (Distance to the coast).7 | 0.05 | 0.10 | 0.10 | 0.52 | 0.60 |
|             | (Distance to the coast).8 | -2.11 | 1.60 | 1.61 | 1.31 | 0.19 |
|             | (Distance to the coast).9 | -0.10 | 0.68 | 0.69 | 0.14 | 0.89 |
|             | Body length (mm) | 0.00 | 0.00 | 0.00 | 0.49 | 0.63 |
|             | Anadromous fish abundance:Sex | -4.3 | 1.17 | 1.18 | 0.37 | 0.72 |
|             | Bone δ¹³C | Distance to the coast | -0.01 | 0.00 | 0.00 | 3.91 | 0.00 |
|             | Anadromous fish abundance | -1.60 | 0.68 | 0.69 | 2.32 | 0.02 |
|             | (compared to female) | 0.44 | 0.32 | 0.32 | 1.39 | 0.16 |
|             | Anadromous fish abundance:Sex | -0.43 | 1.17 | 1.18 | 0.37 | 0.72 |
| Muscle      | Anadromous fish abundance | 0.12 | 5.12 | 5.14 | 0.02 | 0.98 |
|             | Body length (mm) | 0.00 | 0.00 | 0.00 | 0.25 | 0.80 |
|             | Summer (compared to Spring) | 0.31 | 0.67 | 0.67 | 0.46 | 0.65 |
|             | Autumn (compared to Spring) | 0.20 | 0.55 | 0.55 | 0.36 | 0.72 |
|             | Winter (compared to Spring) | 0.56 | 0.67 | 0.67 | 0.83 | 0.41 |
|             | Autumn (compared to Summer) | 0.51 | 0.57 | 0.58 | 0.88 | 0.38 |
|             | Winter (compared to Summer) | 0.86 | 0.55 | 0.56 | 1.55 | 0.12 |
|             | Winter (compared to Autumn) | 0.36 | 0.48 | 0.49 | 0.73 | 0.46 |
|             | (Distance to the coast).1 | 0.0 | 0.62 | 0.62 | 0.58 | 0.57 |
|             | (Distance to the coast).2 | -2.70 | 2.23 | 2.24 | 1.21 | 0.23 |
|             | (Distance to the coast).3 | 0.23 | 0.65 | 0.66 | 0.35 | 0.72 |
|             | (Distance to the coast).4 | 1.05 | 1.34 | 1.35 | 0.78 | 0.44 |
|             | (Distance to the coast).5 | 0.31 | 0.36 | 0.36 | 0.34 | 0.74 |
|             | (Distance to the coast).6 | 1.06 | 1.15 | 1.15 | 0.92 | 0.36 |
|             | (Distance to the coast).7 | -0.49 | 0.46 | 0.46 | 1.05 | 0.30 |
|             | (Distance to the coast).8 | -4.04 | 2.92 | 2.93 | 1.38 | 0.17 |
|             | (Distance to the coast).9 | -0.20 | 1.13 | 1.14 | 0.18 | 0.86 |
|             | Anadromous fish abundance:Body length (mm) | 0.01 | 0.01 | 0.01 | 1.35 | 0.18 |
|             | (compared to female) | -4.35 | 3.15 | 3.16 | 1.37 | 0.17 |
|             | Summer:Male (compared to Spring:Female) | 1.23 | 0.87 | 0.88 | 1.52 | 0.13 |
|             | Autumn:Male (compared to Spring:Female) | 0.13 | 0.71 | 0.72 | 1.44 | 0.15 |
|             | Winter:Male (compared to Spring:Female) | 1.85 | 0.69 | 0.69 | 2.67 | 0.01 |
|             | Autumn:Male (compared to Summer:Female) | 0.10 | 0.88 | 0.88 | 0.34 | 0.74 |
|             | Winter:Male (compared to Summer:Female) | 0.52 | 0.84 | 0.84 | 0.61 | 0.54 |
|             | Winter:Male (compared to Winter:Female) | 0.82 | 0.68 | 0.69 | 1.19 | 0.23 |
|             | Body length (mm):Sex | 0.00 | 0.00 | 0.00 | 1.68 | 0.09 |
|             | Anadromous fish abundance:Sex | 1.52 | 1.14 | 1.15 | 1.33 | 0.19 |
2.7.4 *SIBER* Bayesian standard ellipse area plots

![Bayesian standard ellipse area plots](image)

**Figure S 2.3.** Bayesian standard ellipse area (SEA) sizes and their credible intervals for isotopic signatures of bone and muscle tissue collected from Eurasian otters (*Lutra lutra*) from Mid-England and Wales between 1993 - 2007. Black circles represent the SEA mode, red crosses represent the sample size corrected SEA (SEA$_c$) and boxes indicate 50%, 75% and 95% credible intervals from inner to outer.
Figure S 2.4. Bayesian standard ellipse area (SEA) sizes and their credible intervals for isotopic signatures of muscle tissue from male and female Eurasian otters (*Lutra lutra*) in each season, sampled from Mid-England and Wales between 1993 - 2007. Black circles represent the SEA mode, red crosses represent the sample size corrected SEA (SEA$_c$) and boxes indicate 50%, 75% and 95% credible intervals from inner to outer. Females are represented by F and males by M. Spring is represented by Sp, summer by Su, autumn by A and winter by W.
Chapter Three – Investigating nutrient fluxes in freshwater ecosystems through isotopic variation in Eurasian otters (*Lutra lutra*) and the implications for otter diet.

### 3.1 Abstract

Freshwater ecosystems are susceptible to changing nutrient flows from both autochthonous and allochthonous sources. Nutrient influxes can directly influence primary producers in freshwater habitats and indirectly impact species of higher trophic levels, as nutrients are transferred through the trophic network. To evaluate the influence of different nutrient sources on trophic networks in British freshwater ecosystems, we used the Eurasian otter (*Lutra lutra*), an apex predator in European freshwater habitats, as an indicator of nutrient availability. Nitrogen and carbon stable isotopes were quantified from 300 whisker samples collected from otters across England and Wales between 2007 and 2016. Multiple subsamples were taken from each whisker to assess temporal variation in isotopic composition within individuals. Our findings suggested that the isotopic niche of otters is driven by availability of prey and changes in basal isotopic signatures across the landscape, with nitrogen signatures reflecting spatial variation in fertiliser inputs and availability of high trophic level prey, whilst carbon signatures reflected the availability of marine derived nutrients via marine species and anadromous fish. We thus concluded that allochthonous inputs have variable impacts on freshwater ecosystems in Britain, with freshwater habitats in close proximity to coastal habitats and those in West Britain more likely to be influenced by marine derived nutrients, whilst those in East Britain are more likely to be influenced by anthropogenic inputs from fertilisers.

*Keywords*: Arable and horticultural inputs, dietary variation, Eurasian otter, *Lutra lutra*, marine-derived nutrients, stable isotopes

### 3.2 Introduction

Freshwater ecosystems are important pathways for nutrient flow, connecting terrestrial and marine habitats and allowing transfer of subsidies (e.g. water, biological material and nutrients) between the three environments (Polis *et al.* 2004; Richardson and Sato 2015). Nutrient flows have considerable implications for the dynamics of recipient...
systems, affecting productivity, diversity and community structure (Schindler et al. 2005; Holt 2008; Yang et al. 2017). Transfer of nutrients through the food web can result in effects at all trophic levels within the recipient system (Leroux and Loreau 2012; Yang et al. 2017; Lee et al. 2018; Samways et al. 2018). Freshwaters bodies not only transfer nutrients between terrestrial and marine ecosystems, but also receive nutrients from both (Polis et al. 2004) making freshwater species susceptible to changes in nutrients from multiple sources.

Freshwater productivity is governed by both autochthonous (derived from within the system, e.g. biomass from aquatic primary producers including macrophytes, phytoplankton and algae) and allochthonous organic matter (derived from outside the system, e.g. terrestrial detritus; Finlay and Kendall 2007; Marchese et al. 2014). Allochthonous subsidies can enter freshwaters through organisms decaying or excreting in freshwaters (Helfield and Naiman 2001; Zhang et al. 2003; Harding et al. 2014), dissolved material in run-off and groundwater, or via broken down particulate material (such as leaf litter; Richardson and Sato 2015). Allochthonous subsidies (including anthropogenic inputs such as wastewater and agricultural chemicals; Holt 2000; Bedard-Haughn et al. 2003; Anderson and Cabana 2005; O’Brien and Wehr 2010) contribute to the energy budget of freshwater systems and subsidise nutrient poor freshwater habitats (Finlay and Kendall 2007; Marcarelli et al. 2011). Whilst these subsidies can increase productivity in freshwater, they can also have detrimental impacts if nutrient inputs are too high, e.g. via eutrophication (Khan and Ansari 2005). Rivers play a key role in transporting nutrients from terrestrial and freshwater habitats to marine habitats, while also providing a route for marine nutrients to move upstream, where they may impact landlocked ecosystems (Flecker et al. 2010). Upstream movement of nutrients primarily occurs through the movement of animals, specifically diadromous fish and their piscivorous predators (Helfield and Naiman 2001; Schindler et al. 2003). Diadromous fish are those that spend part of their life cycle at sea and part in freshwater. Anadromous fish (e.g. Atlantic salmon, *Salmo salar*) spend most of their development and adult lives feeding at sea before returning to freshwater systems to breed, whereas catadromous fish (e.g. European eel, *Anguilla anguilla*) breed at sea and return to freshwater systems to mature and live as adults. Movement of diadromous individuals from marine to freshwater can result in nutrients being transferred via excrement, gametes or carcasses (Schindler et al. 2003; Samways et al. 2015). For freshwater systems, this can be a valuable source of nutrient rich material, increasing production within the ecosystem (Harding et al. 2014; Samways et al. 2018). Marine nutrients can also reach freshwater and terrestrial systems through predators hunting marine prey and either discarding the carcass or excreting in

Nutrient inputs have direct impacts on lower trophic levels by driving an increase in primary production, altering abundances of species within an ecosystem. This can indirectly affect predators by changing prey availability. Together, these bottom-up effects are capable of impacting the health of predators or causing a change in their foraging behaviour (Pace *et al.* 1999; Sinclair and Krebs 2002; Elmhagen and Rushton 2007). Dietary variation in top predators may therefore be indicative of impacts felt by the whole food web.

Eurasian otters (*Lutra lutra*, hereafter referred to as ‘otters’) are semi-aquatic carnivores that forage in a range of habitats, from rivers and streams to sea coasts (Jędrzejewska *et al.* 2001; Parry *et al.* 2011). Since the 1970’s, otters have been undergoing population recovery, following declines in the 1950’s, allowing them to recolonise parts of their former range (Roos *et al.* 2001; Conroy and Chanin 2002, Sainsbury *et al.* 2019). They exhibit opportunistic foraging behaviour, feeding on a wide range of species (Clavero *et al.* 2003; Almeida *et al.* 2012; Parry *et al.* 2015). Otters primarily consume freshwater fish; however, they also take other freshwater prey (such as amphibians and water fowl), marine prey (such as salmon, eels and flatfish) and terrestrial prey (such as rabbits and rats; Jędrzejewska *et al.* 2001; Clavero *et al.* 2003; Almeida *et al.* 2012). As apex predators of freshwater ecosystems, otters are susceptible to changes lower in the food web and can act as indicators for nutrient inputs, and as vectors of nutrients between systems (Ruiz-Olmo *et al.* 1998; Sergio *et al.* 2006; Estes *et al.* 2011; Lemarchand *et al.* 2011).

Stable isotopes measured in the tissues of an organism reflect the nutrients assimilated by that organism (Tieszen *et al.* 1983; Hobson and Clark 1992; Dalerum and Angerbjörn 2005) and are therefore good indicators of nutrient flow through a food web. Isotopic signatures in metabolically active tissues (e.g. bone and muscle) are constantly changing, with the rate of cellular turnover determining the period of time each tissue reflects (Tieszen *et al.* 1983; Hobson and Clark 1992; Dalerum and Angerbjörn 2005). Metabolically inert tissues (e.g. whiskers) retain the signature from when they were first synthesised and capture nutrient signatures in chronological order (Darimont and Reimchen 2002; Lewis *et al.* 2006; Newsome *et al.* 2009). This provides the opportunity to use serial sampling along the length of a metabolically inert tissue to compare within-individual isotopic signatures (and thus nutrient assimilation) between time points.
The ratio of heavy to light stable isotopes varies due to a range of biogeochemical processes (such as photosynthesis and nitrogen fixation; Bedard-Haughn et al. 2003; Kendall et al. 2007; Marshall et al. 2007; Ben-David and Flaherty 2012). The most commonly used isotopes in ecological studies are nitrogen and carbon. Typically, nitrogen isotopic signatures (represented by the δ¹⁵N, the ratio of ¹⁵N:¹⁴N isotopes) are used to reflect which trophic level a consumer is feeding at (Kelly 2000; Crawford et al. 2008; Inger and Bearhop 2008). On consumption, ¹⁵N is more likely to be incorporated into consumer tissues than ¹⁴N, resulting in an increase in δ¹⁵N of ~ 3 ‰ for each trophic level (Deniro and Epstein 1980; Peterson and Fry 1987; Post 2002). Nitrogen isotopic signatures have also been used as an indicator for anthropogenic inputs into freshwater ecosystems, as manure-based fertilisers and domestic sewage are typically more enriched in ¹⁵N than nitrogen fixed from the atmosphere, therefore resulting in higher δ¹⁵N values in basal nitrogen signatures of food webs (Bedard-Haughn et al. 2003; Anderson and Cabana 2005; Urton and Hobson 2005; Cole et al. 2006; Hoffman et al. 2012; Baeta et al. 2017). Carbon isotopic signatures (represented by the δ¹³C, the ratio of ¹³C:¹²C isotopes) are typically used to reflect the basal resource for the trophic pathway, with differences between marine, freshwater and terrestrial primary producers. Carbon isotopic signatures in streams have been used to distinguish between allochthonous terrestrial inputs and autochthonous carbon sources (Finlay 2001; Ishikawa et al. 2012), but overlap in signatures between terrestrial and freshwater sources can make these systems difficult to distinguish. Marine ecosystems are more enriched in ¹³C than either terrestrial or freshwater ecosystems (Peterson and Fry 1987; O’Leary 1988; Kelly 2000), giving a more distinctive isotopic signature. Stable isotopes can therefore be used when investigating nutrient flow from terrestrial and marine into freshwater ecosystems (e.g Naiman et al. 2002; Loomer et al. 2014; Vizza et al. 2017; Lee et al. 2018).

Most studies of nutrient inputs into freshwater ecosystems focus on primary consumers (e.g. Finlay 2001; Bannon and Roman 2008; Ishikawa et al. 2012; Harding et al. 2014; Samways et al. 2015; Montaña and Schalk 2018); the implications for top predators are infrequently reported (although see; MacAvoy et al. 2000; Lake et al. 2001; Darimont and Reimchen 2002; Nolan et al. 2019). Here we aim to investigate nutrient inputs from terrestrial and marine ecosystems into freshwater ecosystems, as detected in an apex predator. To do this we used stable isotopes to investigate changes in nitrogen and carbon signatures in the otter. Nitrogen was used to detect differences in trophic level consumption and anthropogenic input, whereas carbon was used to detect changes in basal resource use and marine input. We tested the following hypotheses: (i) otters in
lowland regions of south-east England would be enriched in $^{15}$N due to greater proportions of arable and horticultural land use, compared to areas of high altitude in Wales, south-west and north England, (ii) otters in coastal areas would be enriched in $^{13}$C compared to inland areas due to consumption of marine prey, (iii) there would be greater variability in carbon isotopic signatures in otters in coastal areas compared to inland areas due to the availability of both marine and freshwater prey, (iv) otters in urban areas would be enriched in $^{15}$N due to anthropogenic inputs and (v) isotopic signatures of otters would vary temporally due to changes in prey availability and climatic changes affecting run-off of nutrients from land.

3.3 Methods

3.3.1 Sample and data collection

All samples analysed were from otters in the Cardiff University Otter Project collection, a national monitoring programme for dead otters sampled from across Great Britain (https://www.cardiff.ac.uk/otter-project). Most otters collected were killed by road traffic accidents, with a minority dying through drowning, being shot, starvation, or disease. Information on date (year and month) and location (as grid reference) of carcass collection were recorded at the site of collection. Grid references were used to plot data for spatial analysis. Detailed post mortems were performed for each carcass (data not presented) during which biotic data (e.g. sex and size of individual) was recorded and biological samples, including multiple whiskers per otter, were collected and stored at -20 °C. Otters (n = 300) were selected to reflect a range of biotic and spatio-temporal scales (Fig. 3.1 and Table 3.1). Otters from Scotland were omitted due to small sample size and limited spread (most records were from the western Isles of Scotland).

Following post mortems, scaled mass index (SMI) was calculated for each individual otter (Peig and Green 2009; Peig and Green 2010; see Chapter Two for details) and converted into scored SMI (Guillemain et al. 2013). Otters were also classified into size categories based upon their total length (nose to tail tip) using the ‘bins’ function in R (OneR v2.2 package; von Jouanne-Diedrich 2017), which applies a clustering method using Jenks natural breaks optimisation. Male and female otters were clustered separately into small (males <104.6cm, females <936cm long), medium (males between 104.6cm and 113.1cm, females between 936cm and 103.1cm), and large (males > 113.1cm, females > 103.1cm).
Figure 3.1. Spatial distribution of Eurasian otter (*Lutra lutra*) samples. Blue dots represent male otters whilst red represent females.
Table 3.1. Associated metadata relating to the 300 Eurasian otters (*Lutra lutra*) selected for the study.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Female</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>161</td>
</tr>
<tr>
<td>Season</td>
<td>Spring</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Summer</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Autumn</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Winter</td>
<td>80</td>
</tr>
<tr>
<td>Year</td>
<td>2007 - 2016</td>
<td>~30 per year</td>
</tr>
<tr>
<td>Body condition using scored scaled mass index</td>
<td>Range from -3.1 to +3.16</td>
<td>300</td>
</tr>
<tr>
<td>Size</td>
<td>Small</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>120</td>
</tr>
</tbody>
</table>

3.3.2 Spatial data

Spatial data describing proximity to the coast, land use, altitude, slope and primary water habitat were collated using QGIS version 3.4.4 (QGIS Development Team 2019). Distance from the coast was calculated as shortest distance (km) along a river from the location at which the otter was found to the low tide point of the mouth of the river (hereafter referred to as ‘river distance’), using the package RivEX (Hornby 2020), because otters tend to travel along water courses rather than across land. As most otters were found as roadkill, and not all were adjacent to rivers, each otter was first assigned to the nearest river. Otters found more than 1000m from a river were individually checked, and if there was more than one river along which they might have travelled, then river distance was calculated for all rivers and an average distance used. All otters within 1000m of the coast were given a distance of zero if they were closer to the coastline than a river.

To assess whether land use was influencing dietary variation of otters, the proportion of different land uses in the area where an otter carcass was found was calculated. Land use was mapped using the UK land cover map from 2007 (Morton *et al.* 2011) with 25 m resolution. Otter locations were mapped as points, and circular areas of 20 km diameter (hereafter referred to as ‘buffers’) mapped around each. Otter range size is variable, but estimates suggest a primarily linear range along rivers of up to 20 or even 40 km (Chanin 2013; International Otter Survival Fund 2020). Buffer zones 20 km in diameter were chosen to encompass both the most likely area of freshwater habitat, and the adjacent terrestrial land forming the river catchment. We chose to focus on the proportions of three main land use types that were hypothesised to influence nutrient inputs in freshwaters: arable and horticulture, improved grassland, and urban
(comprised of both urban and suburban). The proportion of each land use was extracted for all buffer areas. Altitude and slope values were mapped using a European Digital Elevation Model (EU-DEM) map (European Environment Agency 2011) and the average value was calculated within all buffer areas.

Otters in England and Wales typically feed in freshwater river systems but will opportunistically feed in lakes or at the coast if these habitats are within range (Jędrzejewska et al. 2001; Clavero et al. 2004; Parry et al. 2011). Available prey differ between lakes, coasts and river systems, as well as between different parts of the river network (e.g. tributary, main river channel). To assess whether water habitat type influenced dietary variation, we designated each otter to one of the following: transitional water (coastal and estuarine), lake, main river channel, or tributary. Otter locations were mapped, and those within 5 km of transitional water, or a lake (based on Water Framework Directive 2000/60/EC designations mapped using GIS shapefiles provided by Natural Resources Wales and Environment Agency) were assigned accordingly. Otters found more than 5 km from transitional or lake habitat were assumed to be feeding primarily in the river network. The RivEX network map (Hornby 2020) was used to map all rivers, and individuals were further categorised according to whether their assumed habitat was primarily main river or tributary. To do this, the total length of main river channels and tributaries were calculated within each 20 km buffer. The length of main channels was weighted 10 times greater to account for the greater cross section of a main channel compared to tributaries (Benda et al. 2004) since waterways with greater areas are assumed to support more prey (Samarasin et al. 2014). The sum of weighted main river lengths and tributary lengths was calculated, and if more than 50 percent of each buffer was weighted main river channel, the otter was assigned to the main river channel, otherwise it was assigned to tributary.

Following spatial data collation, variables were checked for correlation using the R package ‘corrplot’ (Wei and Simko 2017). Land use categories ‘arable and horticulture’, and ‘improved grassland’, as well as altitude and slope were all found to be highly correlated with longitude (Fig. S 3.1). This showed that from west to east, altitude, slope and percentage of improved grassland all decreased whilst percentage of arable and horticulture land use increased. We were therefore unable to separately evaluate the association between these variables and isotopic signatures, and instead used longitude as a proxy.
3.3.3 **Stable isotope analysis**

One whisker per otter was selected at random and cleaned with distilled water. Starting at the base, whiskers were cut into ~1mm sections and placed into a tin capsule until the subsample weighed ~0.7mg (mean = 0.712mg, SE Mean = 0.024). The tin capsule was then closed for analysis and the process was repeated until all the whisker had been used. Across all individuals, between two and 16 subsamples were produced per whisker. Labelling ensured that subsample location could be identified, from the basal segments (providing the most recent isotopic signature) to the tip (the oldest signal). Preliminary analysis was carried out on a subset of 50 whiskers to explore within-whisker variation. Mean isotopic value and isotopic variation per whisker was calculated for each whisker twice: once using all subsamples of a whisker and once only using every other subsample. Simple linear models showed a very strong correlation between the two datasets (Mean $\delta^{15}\text{N}$: $t = 99.136$, $p = <0.001$, $R^2 = 0.995$; Variance in $\delta^{15}\text{N}$: $t = 34.257$, $p = <0.001$, $R^2 = 0.958$; Mean $\delta^{13}\text{C}$: $t = 103.004$, $p = <0.001$, $R^2 = 0.995$; Variance in $\delta^{13}\text{C}$: $t = 20.016$, $p = <0.001$, $R^2 = 0.885$). Therefore, in order to save analytical time, subsequent analyses used every other subsample for each whisker (starting with the base subsample), with the exception of whiskers with less than five subsamples, for which all subsamples were analysed.

All whisker subsamples analysed were placed under continuous-flow isotope ratio mass spectrometry using a Sercon Integra 2 Elemental Analyser (Crewe, UK) at the University of Exeter for nitrogen and carbon analysis. Ratios of carbon and nitrogen stable isotopes are given as $\delta$-values and expressed in parts per mil (‰), with reference to international standards according to the following equation:

$$\delta X = [(R_{\text{sample}} / R_{\text{standard}}) - 1] \times 1000$$

Where $X$ represents $^{13}\text{C}$ or $^{15}\text{N}$, $R_{\text{sample}}$ is the corresponding ratio of heavy to light isotopes ($^{13}\text{C} / ^{12}\text{C}$ or $^{15}\text{N} / ^{14}\text{N}$), and $R_{\text{standard}}$ is the ratio of the international references, PeeDee Belemnite for $\delta^{13}\text{C}$ and atmospheric nitrogen for $\delta^{15}\text{N}$.

3.3.4 **Statistical analysis**

Generalised additive mixed effects models (GAMM) and generalised additive models (GAM) were used to explore associations between stable isotope ratios, biotic and abiotic drivers using R [version 3.6.0] and R studio [version 1.2.1335] (R Core Team 2019; scripts available in SI. 2). In all models the dependent variable was the isotope ratio, either $\delta^{15}\text{N}$ or $\delta^{13}\text{C}$. Models one and two used multiple samples per individual and individual identity was fitted as a random term in a mixed model (GAMM) to control for
repeated measures between individuals, using the ‘uGamm’ function in the R package ‘MuMIn’ (Barton 2019). Multiple samples included either all whisker subsamples (where n segments < 5), or alternate subsamples (where n ≥ 5). Models three and four described isotope ratios in a single, basal, subsample from each individual, and models five and six described within whisker variation in isotope ratios for each individual. All models were built using the R package ‘mgcv’ (Wood 2011) to run GAMs.

All models included the following fixed variables: sex (male/female), size of otter (small/medium/large), scored SMI, year of carcass collection, water habitat (transitional/lake/main channel/tributary), proportion of urban land use, distance from the coast along a river (km), latitude and longitude, along with two-way interactions latitude:longitude, sex:size, sex:water habitat, size:water habitat, sex:distance from the coast and size:distance from the coast. Models three and four (using data from basal segments only) also included fixed variables season (spring/summer/autumn/winter) and mean monthly rainfall (mm), both assigned based upon the month of death for each otter, along with two-way interactions between sex and season and size and season. Season was excluded from models one, two, five and six because for these models signals were measured from the whole whisker, therefore isotopic signature integration was long term, making seasonal interpretations ambiguous.

Models were checked for non-linearity and only terms and interactions with estimated degrees of freedom (edf) over two were kept as non-linear in the final model. A thin-plate smoothing spline (Wood 2003) was applied to percentage of urban land use in models one to four (this term was linear in models five and six) and a tensor product spline (Wood et al. 2013) was applied to the interaction between longitude and latitude in all six models. Models one to four were assigned a Gaussian distribution with an identity link function, whilst models five and six were assigned a gamma distribution with an inverse link function. Model selection was achieved by carrying out multi-model inference and model averaging on the global model using the R package ‘MuMIn’ (Barton 2019). The dredge function was used to create all possible model combinations, using the variables in the global model, and then rank these by Akaike’s information criterion (AIC) values. Models between which there was a delta AIC of less than two were used to create the average model and conditional average values were used to infer significant factors and interactions.

Isotopic niche area was plotted using the R-package ‘Stable Isotope Bayesian Ellipses in R’ (SIBER; Jackson et al. 2011) to produce 95% confidence ellipses for isotopic data. Graphs produced through SIBER allowed any differences or overlap in isotopic
niches to be visualised. Significant interactions between latitude and longitude were plotted using the R-package ‘IsoriX’ (Courtiol and Rousset 2017; Courtiol et al. 2019) to produce isoscapes (isotope landscapes) and show variation in each isotope over the study area.

3.4 Results

The range of isotopic signatures identified across multiple whisker subsamples ($\delta^{15}N$ mean = $15.44\%_{oo} \pm 2.98\%_{oo}$, $\delta^{13}C$ mean = $-23.73\%_{oo} \pm 2.63\%_{oo}$) was similar to that identified across only basal subsamples ($\delta^{15}N$ mean = $15.29\%_{oo} \pm 2.96\%_{oo}$, $\delta^{13}C$ mean = $-23.77\%_{oo} \pm 2.59\%_{oo}$; Fig. 3.2; Fig S 3.2). Within individual variation, taken as the standard deviation across subsamples of a whisker, ranged from 0.02‰ to 3.94‰ for $\delta^{15}N$ (mean = 0.83‰ ±0.65) and 0.02‰ to 5.11‰ for $\delta^{13}C$ (mean = 0.79‰ ±0.63).

Figure 3.2. Isotopic niche of Eurasian otters (Lutra lutra) based on values obtained from multiple whisker subsamples (green) and values obtained from only the basal subsample (purple). Dashed lines show convex hulls. Outer ellipses show 95% standard ellipse areas; inner ellipses show the 95% confidence intervals around the bivariate means. Samples were taken from across England and Wales from 2007 – 2016.
All six starting models produced more than one top model and underwent model averaging. However, not all variables important to the final averaged model were significant; here we report only variables that were identified as both important in one or more top models, and significant in the averaged model (with a p-value of <0.05). Full details of model outputs are provided in Tables S3.1 to 3.6.

### 3.4.1 Spatial variation

The interaction between latitude and longitude was significant in all averaged models (Tables S3.1 to S3.6). Data showed the highest δ^{15}N values (i.e. more enriched in ^{15}N) and lowest δ^{13}C values (i.e. more depleted in ^{13}C) in South-East England (primarily around Cambridgeshire [52.5° N, 0° E]) compared to regions in Wales, South-West and North England, with the highest enrichment in ^{13}C found in coastal regions of Wales and South-West England (Figures 3.3; data presented are from multiple whisker subsamples; basal samples alone showed the same pattern, Fig. S3.3). Within individual variability in δ^{15}N was greatest in South-East England (primarily around Cambridgeshire) and lowest in Wales and South-West England (primarily around the Bristol channel) whereas individual variation in δ^{13}C was greatest in areas near the coast and lowest in central England (Fig. 3.3).
Figure 3.3. Variation in isotopic signatures of Eurasian otters (Lutra lutra) across England and Wales from 2007 to 2016. Upper plots show the average isotope values, where lighter colours represent more enrichment in the respective isotopes. Lower plots show the variability in isotope values within individuals, where lighter colours represent greater variation in the isotope values (i.e. more individual variation). The mean predicted variation in nitrogen and carbon isotopes (i.e. how confident we can be in predictions) are presented in supplementary information (Fig. S 3.4). Orange dots are locations of individual otters. Isotopic data are from multiple whisker subsamples obtained from an individual otter.
Nitrogen ratios ($\delta^{15}$N) differed significantly between aquatic habitat types, revealing sex dependent differences. Females had higher $\delta^{15}$N values than males in all aquatic habitats, except tributaries, where males had higher $\delta^{15}$N values; results showed the same pattern whether basal segments only, or multiple segments, were analysed (Fig. 3.4; data presented are from multiple whisker subsamples; basal samples alone showed the same pattern, Fig. S3.5; statistical details Tables S3.1 and S3.3). Nitrogen ($\delta^{15}$N) signatures of females were highest in individuals found near coastal/estuarine regions and lowest in those associated with tributaries, whereas $\delta^{15}$N signatures of males were highest in individuals found near coastal/estuarine regions and lowest in those associated with lakes or main river channels. Data describing carbon ratios ($\delta^{13}$C) showed no significant differences between aquatic habitat types when multiple subsamples were analysed, but differences between aquatic habitats were apparent when using only basal subsamples (Table S3.4). For basal segments, otters near coastal/estuarine areas had higher $\delta^{13}$C values (i.e. more enriched in $^{13}$C) compared to those associated with the river network (Fig. S3.6). There were no significant differences in isotopic signatures between otters found near coastal/estuarine regions and those found near lakes, or between otters found near lakes and those associated with main river channels or tributaries. No significant association was found between aquatic habitat type and individual variation in $\delta^{15}$N or $\delta^{13}$C. There was no evidence for any association with distance from the coast or proportion of urban land use across all six models.
Variation in $\delta^{15}N$ of Eurasian otters ($Lutra lutra$) utilising different aquatic habitat types. Symbols show the predicted mean (red = female, blue = male) plus or minus standard errors. Isotopic data were acquired from multiple whisker subsamples obtained from otters across England and Wales from 2007 to 2016.

### 3.4.2 Temporal Variation

Females had lower $\delta^{13}C$ values than males in all seasons, with both sexes showing highest $\delta^{13}C$ values in summer and autumn compared to winter and spring (Fig. 3.5; statistical detail Table S 3.4). There was no evidence for seasonality in $\delta^{15}N$ signatures. $\delta^{15}N$ was positively associated with increasing rainfall (Table S 3.3), but no such trend was observed for $\delta^{13}C$. There was no evidence for long term temporal change in $\delta^{15}N$ or $\delta^{13}C$ over the 10-year period (for multiple whisker subsamples, basal subsamples or individual variation).
3.4.3 Biotic Variation

Males showed both greater $\delta^{13}C$ values on average, and greater within individual variability in $\delta^{13}C$, than females (for statistical details see Tables S 3.2, S 3.4 and S 3.6). For multiple whisker segments, $\delta^{13}C$ enrichment in males was independent of size, but for basal segments large and medium males were enriched compared to females, whereas for small otters no difference was seen (Fig. S 3.7). Large otters showed greater $\delta^{15}N$ values on average, and lower individual variability in $\delta^{15}N$, than medium and small otters, although greater $\delta^{15}N$ enrichment was only observed using basal whisker subsamples (Figs. S 3.8 and S 3.9; for statistical details see Tables S 3.1, S 3.3 and S 3.5).

3.5 Discussion

The isotopic niche of otters in England and Wales varied widely, suggesting a highly variable diet across the population. This parallels other large generalist predators which take prey items from different isotopic pools (e.g. sea otters; Newsome et al. 2009, Eurasian badgers; Robertson et al. 2015, American alligators; Rosenblatt et al.
2015, American mink; Bodey et al. 2010, double crested cormorants; Doucette et al. 2011). By conducting serial sampling across a whisker, we were able to investigate changes in diet within an individual over different temporal scales (Darimont and Reimchen 2002; Bearhop et al. 2004; Dalerum and Angerbjörn 2005; Newsome et al. 2009); base segments represent the most recent diet, with subsequent segments reflecting dietary composition further in the past. Such analyses showed isotopic signatures of otters varied spatially, temporally and between individuals from different demographic groups, reflecting the differing abilities of individual otters to catch and consume a variety of prey from different habitats. The observed variation also reflected changes in basal isotopic signatures within the environment, allowing us to infer how allochthonous inputs, from anthropogenic and marine sources, into freshwaters can affect nutrient flow through trophic pathways and result in changes to the isotopic signature of a top predator.

3.5.1 Whisker subsamples

Otters were found to occupy a similar isotopic niche over long- and short-term timescales, shown by the similarity between isotopic signatures averaged across multiple whisker segments (i.e. representing a broad time span), to those measured from only basal segments (i.e. representing recent acquisition). These findings suggest that individuals maintain feeding habits over time, with differences between individuals being the primary driver of the large isotopic variation observed across the population, rather than variable feeding habits exhibited within individuals. This finding was supported by the low levels of within-individual variation in isotopic signatures identified by comparing different whisker segments from an individual, suggesting that, on average, individuals did not vary their diets considerably over time. Therefore, whilst the broad range of isotopic signatures indicates otters are generalist predators, the lack of variation over time suggests populations may consist of individual specialists, a similar observation to that of other generalist predators (Matthews and Mazumder 2004; Newsome et al. 2015; Robertson et al. 2015; Rosenblatt et al. 2015; Svanbäck et al. 2015).

3.5.2 Spatial variation

Otters showed the greatest $^{15}$N enrichment in the East of England around Cambridgeshire, an area with high proportions of arable and horticultural land use (Morton et al. 2011; Department for Environment Food and Rural Affairs 2020). We suggest that observed enrichment in $^{15}$N is likely driven by fertiliser run-off from arable and horticultural practices. Fertiliser inputs into freshwater systems can enrich nitrogen signatures of primary producers in $^{15}$N and subsequently enrich the trophic network as
a whole (Lake et al. 2001; Anderson and Cabana 2005; Hoffman et al. 2012; Lee et al. 2018), thus driving higher nitrogen signatures in apex predators, such as otters, that forage in these areas. Differences in species assemblages may also contribute to spatial variation in nitrogen; freshwater habitats at higher altitudes (e.g. in the north and west of the UK) tend to occupy a smaller ecosystem area, and therefore support a narrower range of aquatic species (McGarvey and Hughes 2008; Iwasaki et al. 2012) with fewer species from higher trophic levels (Holt et al. 1999; Srivastava et al. 2008). The lower availability of high trophic level prey in higher altitudinal regions may thus result in otters consuming fewer of these prey, contributing to the lower enrichment in $^{15}$N observed in otters from such areas.

In addition to higher overall levels of $^{15}$N enrichment, East England also evinced the highest level of variability in nitrogen within individuals. Previous studies have found that the application of fertilisers, and further modifications, can result in broad ranges of nitrogen isotopic values in a system (Vander Zanden and Rasmussen 1999; Anderson and Cabana 2005; Loomer et al. 2014), which may then result in broad differences in nitrogen signatures across prey items. If generalist species like otters sample a range of these prey items, large variations in the individual nitrogen signatures over time will be observed. However, this variation could also be a result of otters in this region sampling prey from a broader range of trophic levels (Kelly 2000; Crawford et al. 2008; Inger and Bearhop 2008). We cannot disentangle changes in nitrogen signatures due to fertiliser inputs from those due to consumption of prey from different trophic levels, therefore we suggest that both may be contributing to the landscape variation in nitrogen observed here.

Nitrogen signatures of otters were also expected to differ due to urban land use as anthropogenic inputs, such as sewage and wastewater, can enrich basal nitrogen signatures in $^{15}$N (Lake et al. 2001; Anderson and Cabana 2005; Cole et al. 2006; Kendall et al. 2007; Baeta et al. 2017). No such association was observed with either average nitrogen isotopic values nor individual variability, suggesting that inputs from urbanisation are not having a strong impact on otter diet (or the underlying trophic pathway). However, because we used the proportion of urban land use in our models rather more direct measures of anthropogenic inputs, we cannot rule out possible associations being undetected.

Otters in South-West England, North England and Wales displayed greater $^{13}$C enrichment than conspecifics in the midlands and South-East England. In England and Wales, anadromous fish are restricted to South-West England, North England and
Wales (Parrish et al. 1998; Environment Agency 2017), therefore we suggest that landscape scale variation in carbon signatures may reflect availability of MDNs in river catchments from anadromous fish. Acquisition of MDNs from anadromous fish has led to carbon signatures enriched in $^{13}$C in a range of predators, including bears, wolves and northern pike (Hilderbrand et al. 1999; Darimont and Reimchen 2002; Quinn et al. 2009; Levi et al. 2012; Samways et al. 2018; Nolan et al. 2019). Although catadromous fish may also act as a source of MDNs, where anadromous fish retain their marine signatures during upstream migrations (Limburg and Waldman 2009; Flecker et al. 2010), catadromous lose theirs and therefore contribute fewer MDNs to catchments. Catadromous fish are also more widespread than anadromous fish, yet enriched carbon signals did not match catadromous fish distributions (NBN atlas 2020c). We therefore concluded that MDNs in rivers were primarily driven by anadromous fish, although, overall the carbon signatures gave only a weak indication of MDNs.

At a finer scale, comparison by habitat type showed enrichment of both $^{13}$C and $^{15}$N in coastal habitats (compared with lake, main river, or tributaries), suggesting acquisition of MDN through marine prey consumption (differences between the sexes with respect to habitat are discussed below). Despite the significant enrichment in coastal habitat, we found no significant association between $^{13}$C enrichment and distance from the coast. Previous dietary studies suggest British otters utilise marine resources to differing extents, with otters in the Scottish Isles consuming more marine prey than those in coastal regions of England and Wales (e.g. Kruuk and Moorhouse 1990; Parry et al. 2011; Moorhouse-Gann et al. 2020). We suggest that use of marine prey by coastal otters in England and Wales varies considerably both between regions and individuals. Very few individuals showed a clear marine signal, but otters in coastal regions of South-West England and Wales had high levels of within-individual variation in carbon, suggesting switching between marine and freshwater prey.

### 3.5.3 Temporal Variation

Across Britain, many species are experiencing changes to their abundances and distributions (Hayhow et al. 2019), including species predated by otters. Whilst some prey species are undergoing declines (e.g. the European eel, Anguilla anguilla; Bark et al. 2007; Aprahamian and Walker 2008; ICES 2019), others are expanding their range and increasing their populations (e.g. invasive signal crayfish, Pacifastacus leniusculus; Sibley et al. 2002; Holdich et al. 2014). As opportunistic predators, otters will consume prey that are more abundant, and although we expected changes to the isotopic niche over time, no such trend was observed. Other studies have suggested long-term changes in otter prey (e.g. Kruuk 2014, Moorhouse-Gann et al 2020), thus
the absence of isotopic differences here may reflect either switches between prey species with similar signatures, or consumption of different prey resulting in similar average isotopic signatures in otters. Higher resolution (species specific) information is required to identify whether long-term changes in the freshwater community are reflected in otter diet.

Seasonal variation was only evident for carbon isotopes, with both males and females displaying carbon signatures enriched in $^{13}$C in autumn compared to spring. This suggests a change in basal resources between these seasons, and may reflect an autumn peak in availability of anadromous fish as adults return to freshwater to breed (Everard 2013). Contrastingly, nitrogen did not vary between the seasons but did become more enriched in $^{15}$N during periods of high rainfall. Although it is possible that additional turbidity releases nutrients from river banks, we suggest that variation in nitrogen primarily reflects run-off washing fertilisers into freshwaters (Holt 2000; Seitzinger et al. 2005; Tsoi et al. 2011; Mahl et al. 2015), thus enriching the system in $^{15}$N.

3.5.4 Biotic Variation

Our findings provided evidence for differential use of prey by males and females. Differences in nitrogen enrichment suggested females utilise prey from higher trophic levels than males in all aquatic habitats except tributaries. Greater $^{13}$C enrichment and within-individual variability in males suggested that males not only exploit different basal resources, but also a broader range of sources than females. Sex differences may reflect differing prey availability due to larger male ranges, or differing abilities to capture prey due to larger male body size. Size was also important, with no differentiation between the sexes among the smallest size class, but at larger size classes males exploited a different range of basal resources than females. The isotopic similarity of small males to females potentially reflects the reliance of younger otters on prey caught by their mother (Polotti 1995; Kruuk 2006). We found greater $^{15}$N enrichment in large otters, but lower individual variation. These findings may reflect the greater ability of larger otters to catch and handle prey from higher trophic levels (such as salmonids), allowing them to exhibit greater prey preferences than smaller otters, which may rely more on opportunistic predation events. Size class differences in carbon and nitrogen signatures were only observed at shorter timescales (i.e. from basal whisker segments) as might be expected due to changing body size over time.

Surprisingly, we did not observe any isotopic difference between otters of different body conditions, suggesting that the isotopic niche occupied is not reflected in the
condition of the individual. However, this may be misleading because otters with a consistent, isotopically 'average' diet, would exhibit the same signature as otters that consume prey across a wide range of different signatures (Middelburg 2014; Hertz et al. 2016). More information on specific species consumed by otters is therefore required to fully explore the association between dietary variation and body condition of otters.

### 3.5.5 Caveats

Due to the broad spatio-temporal scales investigated in this study, it was impractical for us to directly measure basal ecosystem isotopic signatures across the study area. However, by controlling statistically for the strong landscape scale spatial variation in carbon and nitrogen, we have been able to reveal associations with allochthonous inputs, such as anthropogenic inputs and marine derived nutrients. These associations have been detected in an apex predator, but reflect changes to basal isotopic signatures (Lee et al. 2018; Montaña and Schalk 2018; Samways et al. 2018), therefore reflecting changes in isotopic signatures across multiple trophic levels.

There is no available evidence describing the growth rate of otter whiskers, and rates obtained from other mustelids reveal large differences between species: stoats (0.6mm per day; Spurr 2002), badgers (0.43mm per day; Robertson et al. 2013) and sea otters (0.21mm per day; Tyrrell et al. 2013). Therefore, a specific time frame could not be applied to the variation observed in this study. However, by carrying out separate analyses using just the base of the whisker, we were able to show how the average isotopic signature of otters varies over short time frames compared to the average over a longer period (shown by analysis of multiple whisker segments).

### 3.5.6 Conclusions

Here we have shown how variation in the isotopic signature of a top predator can reflect basal isotopic changes in the surrounding environment. Arable and horticultural practices appear to have a strong influence on the nitrogen signature of otters, with fertiliser application elevating nitrogen levels throughout the food web. Variation in carbon signatures showed otters assimilated MDNs from coastal and diadromous species, although this appears to only contribute small proportions to the diet of otters. Differences in the trophic level and basal resources consumed, over varying spatial, temporal and biotic scales, revealed the large plasticity in the diet of otters. This opportunistic foraging behaviour is likely to give otters greater resilience to future environmental changes, such as through climate change, and has potentially helped them recolonise habitats during their population recovery. We recommend further
analyses focus on gaining greater taxonomic resolution into the diet of otters (e.g. through methods such as DNA metabarcoding) in order to assess whether variation observed here is being driven by changes in the prey species consumed.

3.6 Acknowledgements
This work was funded by KESS II (Knowledge Economy Skills Scholarship) and the Wildlife Trust of South and West Wales who partnered this project. Sample collection was conducted by Cardiff University Otter Project employees and placement students during post mortems of otter carcasses. We would like to extend thanks to Alexandra Nederbragtin for providing access to weighing facilities in Cardiff University School of Earth and Ocean Sciences. Thank you to Christopher Mitchell at the Environment and Sustainability Institute, University of Exeter for running samples though the mass spectrometer. Thank you to Jez Smith for guidance on statistical analyses.
3.7 Supplementary Information

3.7.1 Scripts for analysing data in R

R Code used for analysing stable isotope data acquired from Eurasian otter whiskers. Code was run using R [version 3.6.0] and R studio [version 1.2.1335] (R Core Team 2019) and converted into document format using R markdown (Xie et al. 2018; Allaire et al. 2020). Executable code is presented in grey boxes.

Load packages

```r
library("Matrix")
library("lme4")
library("MASS")
library("car")
library("MuMIn")
library("dplyr")
library("mgcv")
library("Rcpp")
library("arm")
library("rsq")
library("gamm4")
library("ggplot2")
library("gridExtra")
library("OneR")
library("ggpubr")
library("tidyverse")
library("corrplot")
```

Analyses using multiple whisker segments

Load in data and check distributions of isotopic variables (response variables).

```r
Whiskers <- read.csv("Combined data for SIA Tidy2.csv", header = T)
summary(Whiskers)
str(Whiskers)
hist(Whiskers$X15N)
hist(Whiskers$X13C)
qqp(Whiskers$X15N, "norm")
qqp(Whiskers$X13C, "norm")
```

Reformat data
Make a variable called ‘size’ that accounts for males being larger than females. First check the difference in size of otters between males and females.

```r
ggplot(Whiskers2, aes(x = Length, group = Sex, fill = Sex)) + geom_histogram(position = "dodge") + theme_bw()
tapply(Whiskers2$Length, Whiskers2$Sex, summary)
```

Make an object with data just for males, split into 3 bins based upon otter length and then define each classification.

```r
Males <- filter(Whiskers, Sex == "Male")
Males <- na.omit(Males)
summary(Males)

Males$Size <- bin(Males$Length, nbins = 3, labels = c("small", "medium", "large"), method = "clusters")
plot(Males$Length ~ Males$Size)
tapply(Males$Length, Males$Size, summary)
```

Make an object with data just for females, split into 3 bins based upon otter length and then define each classification.

```r
Females <- filter(Whiskers, Sex == "Female")
Females <- na.omit(Females)
summary(Females)

Females$Size <- bin(Females$Length, nbins = 3, labels = c("small", "medium", "large"), method = "clusters")
plot(Females$Length ~ Females$Size)
tapply(Females$Length, Females$Size, summary)
```

Add size classifications to the full dataset and remove NA's.

```r
Whiskers <- Whiskers %>%
  mutate(Size = case_when(
    Sex == "Female" & Length < 936 ~ "Small",
    Sex == "Female" & Length > 936 & Length < 1031 ~ "Medium",
    Sex == "Female" & Length > 1031 ~ "Large",
    Sex == "Male" & Length < 1046 ~ "Small",
    Sex == "Male" & Length > 1046 & Length < 1131 ~ "Medium",
    Sex == "Male" & Length > 1131 ~ "Large",
  ))

Whiskers$Size <- as.factor(Whiskers$Size)
summary(Whiskers)
```
str(Whiskers)
Whiskers2 <- na.omit(Whiskers)
summary(Whiskers2)
str(Whiskers2)
Adjust continuous variables so scales are more similar.

head(Whiskers2)
Whiskers2$KmRiverDist <- Whiskers2$RiverDistToCoast/1000
Whiskers2$LengthCM <- Whiskers2$Length/10
Whiskers2$Year2 <- Whiskers2$Year - 2006

Check correlation between continuous variables
Correlation plot:
colnames(Whiskers2)
corrcheck <- Whiskers2[,c(10,13,16,17,19:21,23:26)]
str(corrcheck)
cor1 <- cor(corrcheck)
corrplot(cor1, type = "upper", order = "hclust")
Make correlation plot with only the variables that are highly correlated.
colnames(corrcheck)
correlation <- corrcheck[,c(3:6,10)]
cor2 <- cor(correlation)
corrplot(cor2, type = "upper", order = "hclust")

PCA:
PCAcheck <- prcomp(corrcheck, scale=TRUE)
summary(PCAcheck)
plot(PCAcheck)
PCAccheck
biplot(PCAcheck)
Make PCA with only the variables that are highly correlated:
PCAcorr <- prcomp(correlation, scale = T)
summary(PCAcorr)
plot(PCAcorr)
PCAcorr
biplot(PCAcorr)
These plots show agriculture, longitude, improved grassland, slope and altitude are correlated. Agriculture and longitude were positively correlated whilst both of these variables were negatively correlated with improved grassland, altitude and slope. It was therefore decided to use longitude only in models to represent all these variables, as
agriculture, improved grassland, altitude and slope are landscape variables that vary across a longitudinal gradient.

**Nitrogen model**

Run generalised additive mixed effects model (GAMM) and check model assumptions

```r
FinalNitMumin <- uGamm(X15N ~ Sex + Size + Sex:Size + Scored.SMI + Year2 +
                        WaterClass + Sex:WaterClass + Size:WaterClass + s(Urban) +
t2(lat,long) + s(lat) + s(long) + KmRiverDist + KmRiverDist:Sex +
                        KmRiverDist:Size, random = ~ (1|UWC), family =
                        gaussian(link=identity), REML = T, data=Whiskers2)
```

```r
summary(FinalNitMumin$gam)
anova(FinalNitMumin$gam)
summary(FinalNitMumin$mer)
plot(FinalNitMumin$gam)
```

```r
fitted.gamN <- FinalNitMumin$gam$fitted.values
residuals1N <- resid(FinalNitMumin$gam, type = "pearson")
residuals2N <- resid(FinalNitMumin$mer, type = "pearson")
```

```r
plot(fitted.gamN)
plot(residuals1N)
plot(residuals2N)
sresidN <- (residuals2N - mean(residuals2N))/sd(residuals2N)
hist(sresidN)
plot(sresidN ~ fitted.gamN)
AIC(FinalNitMumin$mer)
```

Conduct model averaging

```r
test <- dredge(FinalNitMumin)
test.set <- model.avg(test, subset = delta < 2, fit = TRUE)
summary(test.set)
importance(test.set)
```

Significant terms included the interaction between latitude and longitude, the interaction between sex and water class. Final averaged model also included the interaction between size and water class, therefore relevel both water class and size in order to decipher potential significance between levels.

Relevel and then rerun model averaging to check for differences between levels of size and water class (base levels of global model = Large and Coast/Estuary). First relevel to make Large and main river channel the base variables for size and water class respectively.
Whiskers2$WaterClass <- \texttt{relevel(Whiskers2$WaterClass, ref = "Main")}
\texttt{update}(FinalNitMumin, . ~.)
test <- \texttt{dredge(FinalNitMumin)}
test.setmain <- \texttt{model.avg(test, subset = delta < 2, fit = TRUE)}
\texttt{summary}(test.setmain)
\texttt{importance}(test.setmain)

Relevel to Large:Tributary

Whiskers2$WaterClass <- \texttt{relevel(Whiskers2$WaterClass, ref = "Tributary")}
\texttt{update}(FinalNitMumin, . ~.)
test <- \texttt{dredge(FinalNitMumin)}
test.setLake1 <- \texttt{model.avg(test, subset = delta < 2, fit = TRUE)}
\texttt{summary}(test.setLake1)
\texttt{importance}(test.setLake1)

Relevel to Medium:Coast/Estuary

Whiskers2$Size <- \texttt{relevel(Whiskers2$Size, ref = "Medium")}
Whiskers2$WaterClass <- \texttt{relevel(Whiskers2$WaterClass, ref = "Coast/Estuary")}
\texttt{update}(FinalNitMumin, . ~.)
test <- \texttt{dredge(FinalNitMumin)}
test.setmed <- \texttt{model.avg(test, subset = delta < 2, fit = TRUE)}
\texttt{summary}(test.setmed)
\texttt{importance}(test.setmed)

Relevel to Medium:Lake

Whiskers2$WaterClass <- \texttt{relevel(Whiskers2$WaterClass, ref = "Lake")}
\texttt{update}(FinalNitMumin, . ~.)
test <- \texttt{dredge(FinalNitMumin)}
test.setlake <- \texttt{model.avg(test, subset = delta < 2, fit = TRUE)}
\texttt{summary}(test.setlake)
\texttt{importance}(test.setlake)

Relevel to Medium:Main

Whiskers2$WaterClass <- \texttt{relevel(Whiskers2$WaterClass, ref = "Main")}
\texttt{update}(FinalNitMumin, . ~.)
test <- \texttt{dredge(FinalNitMumin)}
test.setmain2 <- \texttt{model.avg(test, subset = delta < 2, fit = TRUE)}
\texttt{summary}(test.setmain2)
\texttt{importance}(test.setmain2)

Return levels back to the same base levels as the global model.
Carbon model
Run GAMM and check model assumptions

```r
Carbon model
Run GAMM and check model assumptions

FinalCarbonMumin <- uGamm(X13C ~ Sex + Size + Size:Sex + Scored.SMI + Year2 +
WaterClass + Sex:WaterClass + Size:WaterClass + s(Urban) +
t2(lat,long) + s(lat) + s(long) + KmRiverDist +
KmRiverDist:Sex + KmRiverDist:Size, random = ~ (1|UWC),
family = gaussian(link=identity), REML = T, data=Whiskers2)
```

```
summary(FinalCarbonMumin$gam)
anova(FinalCarbonMumin$gam)
summary(FinalCarbonMumin$mer)
```
Base <- `filter`(Whiskers, segment == "A")
Base2 <- `na.omit`(Base)

`summary`(Base2)
`str`(Base2)

Scale variables and change month into season (1 - 4 where 1 = spring and 4 = winter).

Base2$KmRiverDist <- Base2$RiverDistToCoast/1000
Base2$LengthCM <- Base2$Length/10
Base2$Year2 <- Base2$Year - 2006

`library`(plyr)

Base2$Season <- `as.factor`(mapvalues(Base2$Month, from = c("1", "2", "3", "4", "5", "6", "7", "8", "9", "10", "11", "12"),
                                           to = c("4", "4", "1", "1", "2", "2", "3", "3", "3", "4")))

Make 'size' category.

Base <- Base %>%
    `mutate`(Size = case_when(
                  Sex == "Female" & Length < 936 ~ "Small",
                  Sex == "Female" & Length > 936 & Length < 1031 ~ "Medium",
                  Sex == "Female" & Length > 1031 ~ "Large",
                  Sex == "Male" & Length < 1046 ~ "Small",
                  Sex == "Male" & Length > 1046 & Length < 1131 ~ "Medium",
                  Sex == "Male" & Length > 1131 ~ "Large",
                  TRUE ~ "X"
                ),
    Size <- `as.factor`(Base$Size)

`summary`(Base)
`str`(Base)

Save the new dataset as a `.csv` file.

`write.csv`(Base2, "One value whiskers.csv")

Check distributions of isotopic variables (response variables).

`hist`(Base2$X15N)
`hist`(Base2$X13C)
`qqp`(Base2$X15N, "norm")
`qqp`(Base2$X13C, "norm")

*Nitrogen model*

Run generalised additive model (GAM) and check model assumptions.

BaseNitrogen <- `gam`(X15N ~ Sex + Size + Size:Sex + Scored.SMI + Year2 + Season +
                        Sex:Season + Size:Season + WaterClass + Rainfall + Sex:WaterClass +...)
Size:WaterClass + s(Urban) + te(lat,long) + long + s(lat) + KmRiverDist + KmRiverDist:Sex + KmRiverDist:Size, family = \textit{gaussian} (link = identity), data=Base2)

\begin{verbatim}
summary.gam(BaseNitrogen)
plot.gam(BaseNitrogen)
par(mfrow=c(2,2))
gam.check(BaseNitrogen)
par(mfrow=c(1,1))
sresidBN <- (BaseNitrogen$residuals - mean(BaseNitrogen$residuals))/sd(BaseNitrogen$residuals)
plot(sresidBN ~ BaseNitrogen$fitted.values, pch = 20, cex = 2, cex.lab = 1.5)
qqp(sresidBN)
AIC(BaseNitrogen)
\end{verbatim}

Conduct model averaging.

\begin{verbatim}
BaseNitrogen.set <- dredge(BaseNitrogen)
BaseNitrogen.avg <- model.avg(BaseNitrogen.set, subset = delta < 2, fit = TRUE)
summary(BaseNitrogen.avg)
importance(BaseNitrogen.avg)
\end{verbatim}

Significant terms included the interaction between latitude and longitude, interaction between sex and waterclass, size and rainfall. No other interactions or multi-level terms were significant in the final averaged model, therefore only need relevel size and water class.

Relevel to 'medium' and then rerun model averaging to check for differences between levels of size, then return to the same base levels as in the global model.

\begin{verbatim}
Base2$Size < - relevel(Base2$Size, ref = "Medium")
levels(Base2$Size)
update(BaseNitrogen, . ~.)
BaseNitrogen.set <- dredge(BaseNitrogen)
BaseNitrogen.avg <- model.avg(BaseNitrogen.set, subset = delta < 2, fit = TRUE)
summary(BaseNitrogen.avg)
importance(BaseNitrogen.avg)
Base2$Size < - relevel(Base2$Size, ref = "Large")
\end{verbatim}

Relevel and then rerun model averaging to check differences between levels of water class. First relevel to 'Lake'.

\begin{verbatim}
Base2$WaterClass < - relevel(Base2$WaterClass, ref = "Lake")
update(BaseNitrogen, . ~.)
BaseNitrogen.set <- dredge(BaseNitrogen)
\end{verbatim}
BaseNitrogen.avg <- model.avg(BaseNitrogen.set, subset = delta < 2, fit = TRUE)
summary(BaseNitrogen.avg)
importance(BaseNitrogen.avg)

Then relevel to main river channel

Base2$WaterClass <- relevel(Base2$WaterClass, ref = "Main")
update(BaseNitrogen, ~ .)
BaseNitrogen.set <- dredge(BaseNitrogen)
BaseNitrogen.avg <- model.avg(BaseNitrogen.set, subset = delta < 2, fit = TRUE)
summary(BaseNitrogen.avg)
importance(BaseNitrogen.avg)

Return levels back to the same base levels as the global model

Base2$WaterClass <- relevel(Base2$WaterClass, ref = "Lake")
Base2$WaterClass <- relevel(Base2$WaterClass, ref = "Coast/Estuary")

Carbon model
Run GAM and check model assumptions.

BaseCarbon <- gam(X13CBase ~ Sex + Size + Size:Sex + Scored.SMI + Year2 + Season +
    Sex:Season + Size:Season + WaterClass + Rainfall + Sex:WaterClass +
    Size:WaterClass + s(Urban) + te(lat,long) + lat + long + KmRiverDist
    + KmRiverDist:Sex + KmRiverDist:Size, family=gaussian(link=identity),
    data=Base2)

summary.gam(BaseCarbon)
plot.gam(BaseCarbon)
par(mfrow=c(2,2))
gam.check(BaseCarbon)
par(mfrow=c(1,1))
sresidBC <- (BaseCarbon$residuals - mean(BaseCarbon$residuals))/sd(BaseCarbon$residuals)
plot(sresidBC ~ BaseCarbon$fitted.values, pch = 20, cex = 2, cex.lab = 1.5)
qqp(sresidBC)
AIC(BaseCarbon)

Conduct model averaging.

BaseCarbon.set <- dredge(BaseCarbon)
BaseCarbon.avg <- model.avg(BaseCarbon.set, subset = delta < 2, fit = TRUE)
summary(BaseCarbon.avg)
importance(BaseCarbon.avg)
Significant terms included the interaction between latitude and longitude, interaction between size and sex, interaction between season and sex, and water class. No other interactions or multi-level terms were significant in the final averaged model, therefore relevel by size, season and water class.

Relevel to ‘medium’ and rerun model averaging to check differences between levels of size, then return to global model base levels.

```r
Base2$Size <- relevel(Base2$Size, ref = "Medium")
update(BaseCarbon, ~.)
BaseCarbon.set <- dredge(BaseCarbon)
BaseCarbon.avg <- model.avg(BaseCarbon.set, subset = delta < 2, fit = TRUE)
summary(BaseCarbon.avg)
importance(BaseCarbon.avg)
```

Relevel and then rerun model averaging to check differences between levels of water class. First relevel to ‘Lake’.

```r
Base2$WaterClass <- relevel(Base2$WaterClass, ref = "Lake")
update(BaseCarbon, ~.)
BaseCarbon.set <- dredge(BaseCarbon)
BaseCarbon.avg <- model.avg(BaseCarbon.set, subset = delta < 2, fit = TRUE)
summary(BaseCarbon.avg)
importance(BaseCarbon.avg)
```

Then relevel to main river channel, before returning levels to the global model base levels.

```r
Base2$WaterClass <- relevel(Base2$WaterClass, ref = "Main")
update(BaseCarbon, ~.)
BaseCarbon.set <- dredge(BaseCarbon)
BaseCarbon.avg <- model.avg(BaseCarbon.set, subset = delta < 2, fit = TRUE)
summary(BaseCarbon.avg)
importance(BaseCarbon.avg)
```

Relevel and then rerun model averaging to check differences between levels of season. First relevel to ‘2’ (summer).

```r
Base2$Season <- relevel(Base2$Season, ref = "2")
update(BaseCarbon, ~.)
BaseCarbon.set <- dredge(BaseCarbon)
BaseCarbon.avg <- model.avg(BaseCarbon.set, subset = delta < 2, fit = TRUE)
```
summary(BaseCarbon.avg)
importance(BaseCarbon.avg)

Then relevel to ‘3’ (Autumn) before returning levels back to global model base levels.
Base2$Season <- relevel(Base2$Season, ref = "3")
update(BaseCarbon, ~.)
BaseCarbon.set <- dredge(BaseCarbon)
BaseCarbon.avg <- model.avg(BaseCarbon.set, subset = delta < 2, fit = TRUE)
summary(BaseCarbon.avg)
importance(BaseCarbon.avg)
Base2$Season <- relevel(Base2$Season, ref = "2")
Base2$Season <- relevel(Base2$Season, ref = "1")

Analyses for individual variation
Assessed using the standard deviation in isotopic values between whisker segments for each individual. First load in and check data
diff <- read.csv("One value whiskers.csv", header = T)
summary(diff)
str(diff)
hist(diff$N_Whisker_StDev)
hist(diff$C_Whisker_StDev)
qqp(diff$N_Whisker_StDev)
qqp(diff$C_Whisker_StDev)

Nitrogen model
Run GAM and check model assumptions.
NStDev <- gam(N_Whisker_StDev ~ Sex + Size + Size:Sex + Scored.SMI + Year2 +
   WaterClass + Sex:WaterClass + Size:WaterClass + Urban + te(lat,long) +
   lat + long + KmRiverDist + KmRiverDist:Sex + KmRiverDist:Size,
   family = Gamma, data=diff)

summary.gam(NStDev)
plot.gam(NStDev)
par(mfrow=c(2,2))
gam.check(NStDev)
par(mfrow=c(1,1))
sresidNStDev <- (NStDev$residuals - mean(NStDev$residuals))/sd(NStDev$residuals)
plot(sresidNStDev ~ NStDev$fitted.values, pch = 20, cex = 2, cex.lab = 1.5)
qqp(sresidNStDev)
AIC(NStDev)

Conduct model averaging.
StDevNitrogen.set <- dredge(NStDev)
StDevNitrogen.avg <- model.avg(StDevNitrogen.set, subset = delta < 2, fit = TRUE)
summary(StDevNitrogen.avg)
importance(StDevNitrogen.avg)

Significant terms included the interaction between latitude and longitude, and size. No other interactions or multi-level terms were significant in the final averaged model, therefore only relevel by size.

Relevel to ‘medium’ and then rerun model averaging to check for differences between levels of size, then return to the same base levels as in the global model.

diff$Size <- relevel(diff$Size, ref = "Medium")
update(NStDev, . ~ .)
StDevNitrogen.set <- dredge(NStDev)
StDevNitrogen.avg <- model.avg(StDevNitrogen.set, subset = delta < 2, fit = TRUE)
summary(StDevNitrogen.avg)
importance(StDevNitrogen.avg)
diff$Size <- relevel(diff$Size, ref = "Large")

Carbon model

Run GAM and check model assumptions.

CStDev <- gam(C_Whisker_StDev ~ Sex + Size + Sex:Sex + Scored.SMI + Year2 + WaterClass + Sex:WaterClass + Size:WaterClass + Urban + te(lat,long) + lat + long + KmRiverDist + KmRiverDist:Sex + KmRiverDist:Size,
family = Gamma(link = inverse), data=diff)

summary.gam(CStDev)
plot.gam(CStDev)
par(mfrow=c(2,2))
gam.check(CStDev)
par(mfrow=c(1,1))
sresidCStDev <- (CStDev$residuals - mean(CStDev$residuals))/sd(CStDev$residuals)
plot(sresidCStDev ~ CStDev$fitted.values, pch = 20, cex = 2, cex.lab = 1.5)
qqp(sresidCStDev)
AIC(CStDev)

Conduct model averaging.
StDevCarbon.set <- `dredge`(CStDev)
StDevCarbon.avg <- `model.avg`(StDevCarbon.set, subset = delta < 2, fit = TRUE)
`summary`(StDevCarbon.avg)
`importance`(StDevCarbon.avg)

Significant terms included the interaction between latitude and longitude, and sex. No other interactions or multi-level terms were significant in the final averaged model, therefore no need to relevel.

**Plot significant spatial variables**

*Interaction between latitude and longitude using multiple whisker segments*

The interaction between latitude and longitude was significant in all six models. Data from multiple whisker segments and individual variation was plotted using ‘isoriX’ package to build isoscapes. Basal segment variation was plotted in QGIS using multi-level b-spline interpolation and showed a similar pattern to plots using multiple whisker segments.

Load isoriX package
```
`install.packages`("IsoriX", dependencies = TRUE)
`library`(IsoriX)
```

Make isoscapes for Nitrogen isotopes:

Create a dataframe only including the necessary data and reformat.
```
`colnames`(Whiskers2)
columns_to_keep <- Whiskers2[c(2,4,25,16,12,11,4)]
columns_to_keep$UWC <- `as.factor`(columns_to_keep$UWC)
`summary`(columns_to_keep)
`colnames`(columns_to_keep) <- c("source_ID", "lat", "long", "elev", "year", "month", "source_value")
```

Produce an overall average of isotopic values per individual otter save the data as a `.csv`.
```
Nagg <- `prepsources`(data = columns_to_keep,
                     long_min = -6, long_max = 2,
                     lat_min = 50, lat_max = 56)
`write.csv`(Nagg, "N isoscape data.csv")
```

Fit the model to make predictions from, check the assumptions and save the fitted model.
```
NaggFit <- `isofit`(data=Nagg, mean_model_fix = list(elev=T, lat_abs=T, long=T))
`plot`(NaggFit)
NaggFit

\texttt{AIC(NaggFit$mean\_fit)}

\texttt{save(NaggFit, file = "NaggFit.rda", compress = "xz")}

Prepare the raster.

\texttt{getelev()}

\texttt{library("rgdal" )}

\texttt{ElevWorld <- raster("gmted2010\_30mn.tif")}

ElevWorld

Crop the raster.

\texttt{ElevN <- prepraster(raster = ElevWorld, isofit = NaggFit, aggregation\_factor = 1)}

Check what the structural raster looks like.

\texttt{levelplot(ElevN,}

\hspace{1cm} margin = FALSE,

\hspace{1cm} main = "Structural raster") +

\hspace{1cm} layer(sp.polygons(CountryBorders)) +

\hspace{1cm} layer(sp.polygons(OceanMask, fill = "cyan"))

Take the structural raster and fitted model as arguments to predict isoscapes.

\texttt{NIsoscape <- isoscape(raster = ElevN, isofit = NaggFit)}

NIsoscape$isoscapes

Make plots, save plots and remove white lines. First plot the predictions of the average isotopic value in space. Second plot the residual variation in isotopic values per otter in order to quantify temporal variation (shows individual variation). Third plot the predicted variation in isotopic values per otter in order to quantify the uncertainty in plotted predictions.

\texttt{library("Cairo")}

\texttt{CairoPNG(filename = "MyNisoscape.png",}

\hspace{1cm} height = 1080,

\hspace{1cm} width = 1920,

\hspace{1cm} res = 250)

\texttt{plot(NIsoscape,}

\hspace{1cm} sources = list(pch = 19, col = "darkorange2", cex = 0.2),

\hspace{1cm} borders = list(col = "white"),

\hspace{1cm} mask = list(fill = "darkgrey"),

\hspace{1cm} palette = list(step = 1, fn = NULL, range = \texttt{c(4,22))},

\hspace{1cm} y_title = list(which = F, title = ""))

\texttt{dev.off()}

\texttt{CairoPNG(filename = "MyNresidVarisoscape.png",}

\hspace{1cm} ...
Make isoscapes for Carbon isotopes:
Create a dataframe only including the necessary data and reformat.

```
colnames(Whiskers2)
columns_to_keep2 <- Whiskers2[,c(2,25,26,14,11,10,5)]
columns_to_keep2$UWC <- as.factor(columns_to_keep2$UWC)
summary(columns_to_keep2)
colnames(columns_to_keep2) <- c("source_ID", "lat", "long", "elev", "year", "month", "source_value")
```

Produce an overall average of isotopic values per individual otter save the data as a `.csv`.

```
Cagg <- prepsources(data = columns_to_keep2,
                    long_min = -6, long_max = 2,
                    lat_min = 50, lat_max = 56)
write.csv(Cagg, "C isoscape data.csv")
```
```r
CaggFit <- isofit(data=Cagg, mean_model_fix = list(elev=T, lat_abs=T, long=T))
plot(CaggFit)
CaggFit
AIC(CaggFit$mean_fit)
save(CaggFit, file = "CaggFit.rda", compress = "xz")#

Crop the raster.
ElevC <- prepraster(raster = ElevWorld, isofit = CaggFit, aggregation_factor = 1)
Check what the structural raster looks like.
levelplot(ElevC, margin = FALSE, main = "Structural raster") +
layer(sp.polygons(CountryBorders)) +
layer(sp.polygons(OceanMask, fill = "cyan"))

Take the structural raster and fitted model as arguments to predict isoscapes.
CIsoscape <- isoscape(raster = ElevC, isofit = CaggFit)
CIsoscape$isoscapes

Make plots, save plots and remove white lines. First plot the predictions of the average isotopic value in space. Second plot the residual variation in isotopic values per otter in order to quantify temporal variation (shows individual variation). Third plot the predicted variation in isotopic values per otter in order to quantify the uncertainty in plotted predictions.

CairoPNG(filename = "MyCisoscape.png",
          height = 1080,
          width = 1920,
          res = 250)
plot(CIsoscape,
sources = list(pch = 19, col = "darkorange4", cex = 0.2),
borders = list(col = "white"),
mask = list(fill = "darkgrey"),
palette = list(step = 1, fn = NULL, range = c(-27,-17)),
y_title = list(which = F, title = ""))
dev.off()

CairoPNG(filename = "MyCresidVarisoscape.png",
          height = 1080,
          width = 1920,
          res = 250)
plot(CIsoscape, which = "mean_residVar",
```
Interaction between latitude and longitude using basal whisker segments

To map significant interactions between latitude and longitude for basal whisker segment models, first predict isotopic values based on final averaged models and then save these predictions as a ‘.csv’ so they can be plotted and interpolated in QGIS.

Plot data for Nitrogen isotopes:

Create model including only the variables that were important in the final averaged model.

```r
latlongN <- gam(X15N ~ Sex + Size + Year2 + WaterClass + Rainfall +
                 Sex:WaterClass + s(Urban) + te(lat,long) + long +
                 s(lat) + KmRiverDist + KmRiverDist:Sex, data=Base2)
```

Create a dummy dataset and then make predictions using dummy data and the model created above.

```r
pdat.latlongN <- expand.grid(long = seq(min(Base2$long), max(Base2$long),
                                    length.out = 100),
                           lat = seq(min(Base2$lat), max(Base2$lat), length.out = 100),
                           Sex = c("Female"),
                           Size = c("Large"),
                           Year2 = mean(Base2$Year2, na.rm=T),
sources = list (pch = 19, col = "darkorange2", cex = 0.2),
borders = list(col = "white"),
mask = list(fill = "darkgrey"),
palette = list(fn = NULL),
y_title = list(which = F, title = ""))
device.off()
CairoPNG(filename = "MyCpredVarisoscape.png",
         height = 1080,
         width = 1920,
         res = 250)
plot(CIsoscape, which = "mean_predVar",
     sources = list (pch = 19, col = "darkorange2", cex = 0.2),
borders = list(col = "white"),
mask = list(fill = "darkgrey"),
palette = list(fn = NULL),
y_title = list(which = F, title = ""))
device.off()
```

Interaction between latitude and longitude using basal whisker segments

To map significant interactions between latitude and longitude for basal whisker segment models, first predict isotopic values based on final averaged models and then save these predictions as a ‘.csv’ so they can be plotted and interpolated in QGIS.

Plot data for Nitrogen isotopes:

Create model including only the variables that were important in the final averaged model.

```r
latlongN <- gam(X15N ~ Sex + Size + Year2 + WaterClass + Rainfall +
                 Sex:WaterClass + s(Urban) + te(lat,long) + long +
                 s(lat) + KmRiverDist + KmRiverDist:Sex, data=Base2)
```

Create a dummy dataset and then make predictions using dummy data and the model created above.

```r
pdat.latlongN <- expand.grid(long = seq(min(Base2$long), max(Base2$long),
                                    length.out = 100),
                           lat = seq(min(Base2$lat), max(Base2$lat), length.out = 100),
                           Sex = c("Female"),
                           Size = c("Large"),
                           Year2 = mean(Base2$Year2, na.rm=T),
sources = list (pch = 19, col = "darkorange2", cex = 0.2),
borders = list(col = "white"),
mask = list(fill = "darkgrey"),
palette = list(fn = NULL),
y_title = list(which = F, title = ""))
device.off()
CairoPNG(filename = "MyCpredVarisoscape.png",
         height = 1080,
         width = 1920,
         res = 250)
plot(CIsoscape, which = "mean_predVar",
     sources = list (pch = 19, col = "darkorange2", cex = 0.2),
borders = list(col = "white"),
mask = list(fill = "darkgrey"),
palette = list(fn = NULL),
y_title = list(which = F, title = ""))
device.off()
```
Rainfall = \texttt{mean(Base2$Rainfall, na.rm=T)},
WaterClass = \texttt{c("Main")},
KmRiverDist = \texttt{mean(Base2$KmRiverDist, na.rm=T)},
Urban = \texttt{mean(Base2$Urban, na.rm=T)}

pred.latlongN <- \texttt{predict(latlongN, newdata=pdat.latlongN, na.rm=T, type="response", se.fit=TRUE)}
predframe.latlongN <- \texttt{data.frame(pdat.latlongN, preds=pred.latlongN$fit, se=pred.latlongN$se.fit)}
\texttt{write.csv(predframe.latlongN, "predictedN.csv")}

Plot data for Carbon isotopes:
Create model including only the variables that were important in the final averaged model.

lat.longC <- \texttt{gam(X13C \sim \text{Sex} + \text{Size} + \text{Size:Sex} + \text{Scored.SMI} + \text{Season} + \text{WaterClass} + \text{Sex:WaterClass} + \text{s(Urban)} + \text{te(lat,long)} + \text{s(lat)} + \text{s(long)} + \text{KmRiverDist} + \text{Rainfall}, data=Base2)}

Create a dummy dataset and then make predictions using dummy data and the model created above.

pdat.latlongC <- \texttt{expand.grid(long = \texttt{seq(min(Base2$long), max(Base2$long), length.out = 100)},
lat = \texttt{seq(min(Base2$lat), max(Base2$lat), length.out = 100)},
Size = \texttt{c("Large")},
Scored.SMI = \texttt{mean(Base2$Scored.SMI, na.rm=T)},
Season=\texttt{c("1")},
Rainfall = \texttt{mean(Base2$Rainfall, na.rm=T)},
WaterClass = \texttt{levels(Base2$WaterClass)},
Sex = \texttt{c("Female")},
KmRiverDist = \texttt{mean(Base2$KmRiverDist, na.rm=T)},
Urban = \texttt{mean(Base2$Urban, na.rm=T)}
)

pred.latlongC <- \texttt{predict(lat.longC, newdata=pdat.latlongC, na.rm=T, type="response", se.fit=TRUE)}
predframe.latlongC <- \texttt{data.frame(pdat.latlongC, preds=pred.latlongC$fit, se=pred.latlongC$se.fit)}
Nitrogen interaction between sex and aquatic habitat

Plot data for multiple whisker segments:

Create model including only the variables that were important in the final averaged model.

```r
SexWCFullN <- uGamm(X15N ~ Size + Sex + WaterClass + Sex:WaterClass +
                      Size:WaterClass +
                      s(Urban) + t2(lat,long) + s(lat) + s(long),
                      random = ~ (1) | UWC),
                      family = gaussian(link = identity),
                      REML = T, data = Whiskers2)
```

Create a dummy dataset and then make predictions using dummy data and the model created above.

```r
pdat.SexWaterFullN <- expand.grid(WaterClass = levels(Base2$WaterClass),
                                   Sex = levels(Base2$Sex),
                                   long = mean(Base2$long, na.rm = T),
                                   Size = c("Large"),
                                   lat = mean(Base2$lat, na.rm = T),
                                   Urban = mean(Base2$Urban, na.rm = T)
)

pred.SexWaterFullN <- predict(SexWCFullN, newdata = pdat.SexWaterFullN, na.rm = T,
                               type = "response", se.fit = TRUE)
predframe.SexWaterFullN <- data.frame(pdat.SexWaterFullN, preds = pred.SexWaterFullN$fit,
                                        se = pred.SexWaterFullN$se.fit)
```

Plot data/predictions.

AquaticHab <- c("Coast/Estuary", "Lake", "Main River", "Tributary")

```r
ggplot(predframe.SexWaterFullN, aes(x = WaterClass, y = preds, fill = Sex)) +
  geom_errorbar(aes(ymin = preds - se, ymax = preds + se),
                width = .2, position = position_dodge(.3)) +
  geom_point(size = 5, shape = 21, position = position_dodge(.3)) +
  theme_classic() +
  labs(y = expression({delta}^15*N \sim \nu2030'), x = "Aquatic Habitat") +
  theme(axis.title.x = element_text(size = 18), axis.title.y = element_text(size = 18),
        axis.text.x = element_text(size = 14, colour = "black"),
        axis.text.y = element_text(size = 14, colour = "black"),
        axis.ticks.length = unit(.3, "cm"),
```

```
se = pred.latlongC$se.fit
write.csv(predframe.latlongC, "predictedC.csv")
```
plot.margin = unit(c(0.5,0.5,1,1),"cm"),
legend.position = "none") +
scale_y_continuous(limits = c(15.5, 19.5), breaks = scales::pretty_breaks(n = 10)) +
scale_x_discrete(labels= AquaticHab) +
scale_color_manual(values = c("blue", "red"))

Plot data for basal whisker segments:
Create model including only the variables that were important in the final averaged model.
SexWCBaseN <- gam(X15N ~ Sex + Size + Year2 + WaterClass + Rainfall +
Sex:WaterClass +
s(Urban) + te(lat,long) + long + s(lat) + KmRiverDist +
KmRiverDist:Sex, data=Base2)

Create a dummy dataset and then make predictions using dummy data and the model created above.
pdat.SexWaterN <- expand.grid(WaterClass = levels(Base2$WaterClass),
                           Sex = levels(Base2$Sex),
                           Rainfall = mean(Base2$Rainfall, na.rm=T),
                           Year2 = mean(Base2$Year2, na.rm=T),
                           long = mean(Base2$long, na.rm=T),
                           Size = c("Large"),
                           lat = mean(Base2$lat, na.rm=T),
                           KmRiverDist = mean(Base2$KmRiverDist, na.rm=T),
                           Urban = mean(Base2$Urban, na.rm=T)
)
pred.SexWaterN <- predict(RainfallN, newdata=pdat.SexWaterN, na.rm=T,
type="response",
  se.fit=TRUE)
predframe.SexWaterN <- data.frame(pdat.SexWaterN, preds=pred.SexWaterN$fit,
  se = pred.SexWaterN$se.fit)

Plot data/predictions.
ggplot(predframe.SexWaterN, aes(x=WaterClass, y=preds, fill=Sex)) +
  geom_errorbar(aes(ymin=preds - se, ymax=preds + se),
                width=.2, position=position_dodge(.3)) +
  geom_point(size = 5, shape = 21, position = position_dodge(.3)) +
  theme_classic() +
Association between carbon and aquatic habitats (multiple whisker segments model)

AquaticHab <- c("Coast/Estuary", "Lake", "Main River", "Tributary")
ggplot(Base2, aes(x=WaterClass, y=X13C)) +
  geom_boxplot() +
  theme_classic() +
  labs(y=expression((delta)^13*N~"\u2030"), x = "Aquatic Habitat") +
  theme(axis.title.x = element_text(size=18, vjust=-0.5),
        axis.title.y = element_text(size=18),
        axis.text.x = element_text(size=14, colour = "black"),
        axis.text.y = element_text(size=14, colour = "black"),
        axis.ticks.length=unit(.3, "cm"),
        plot.margin=unit(c(0.5,0.5,1,1),"cm"),
        legend.position = "none") +
  scale_x_discrete(labels= AquaticHab)

Temporal variables

Association between nitrogen and rainfall (basal segments model)

Create model including only the variables that were important in the averaged model

RainfallN <- gam(X15NBase ~ Sex + Size + Year2 + WaterClass + Rainfall +
                  Sex:WaterClass + s(Urban) + te(lat,long) + long + s(lat) +
                  KmRiverDist + KmRiverDist:Sex, data=Base2)

Create a dummy dataset and then make predictions using dummy data and the model created above.

pdat.RainfallN <- expand.grid(Rainfall = seq(min(Base2$Rainfall), max(Base2$Rainfall),
                   length.out = 1000),
               Sex = c("Female"),
Size = c("Large"),
Year2 = mean(Base2$Year2, na.rm=T),
long = mean(Base2$long, na.rm=T),
WaterClass = c("Main"),
lat = mean(Base2$lat, na.rm=T),
KmRiverDist = mean(Base2$KmRiverDist, na.rm=T),
Urban = mean(Base2$Urban, na.rm=T)
)

pred.RainfallN <- predict(RainfallN, newdata=pdat.RainfallN, na.rm=T,
type="response",
  se.fit=TRUE)
predframe.RainfallN <- data.frame(pdat.RainfallN, preds=pred.RainfallN$fit,
  se = pred.RainfallN$se.fit) Plot data/predictions.
plot(X15N ~Rainfall, data = Base2, bty = 'L',
  ylab = expression({delta}^{15}N~\'\u2030' ))
lines(preds ~ Rainfall, data = predframe.RainfallN)
lines(preds + se ~ Rainfall, data = predframe.RainfallN, lty = 2, col = "red")
lines(preds - se ~ Rainfall, data = predframe.RainfallN, lty = 2, col = "red")

Carbon interaction between season and sex (basal segments model)
Create model including only the variables that were important in the final averaged model.
WaterC <- gam(X13CBase ~ Sex + Size + Sex:Size:Sex + Scored.SMI + Season +
WaterClass +
  Sex:WaterClass + s(Urban) + te(lat,long) + s(lat) + s(long) +
  KmRiverDist + Rainfall, data=Base2)
Create a dummy dataset and then make predictions using dummy data and the model created above.
pdat.SexSeasonC <- expand.grid(Sex = levels(Base2$Sex),
  Season = levels(Base2$Season),
  Size = c("Large"),
  Scored.SMI = mean(Base2$Scored.SMI, na.rm=T),
  Rainfall = mean(Base2$Rainfall, na.rm=T),
  long = mean(Base2$long, na.rm=T),
  WaterClass = c("Coast/Estuary"),
  lat = mean(Base2$lat, na.rm=T),
KmRiverDist = mean(Base2$KmRiverDist, na.rm=T),
Urban = mean(Base2$Urban, na.rm=T)
)

pred.SexSeasonC <- predict(WaterC, newdata=pdat.SexSeasonC, na.rm=T,
type="response",
    se.fit=TRUE)
predframe.SexSeasonC <- data.frame(pdat.SexSeasonC,
preds=pred.SexSeasonC$fit,
    se = pred.SexSeasonC$se.fit)

Plot data/predictions.
Season <- c("Spring", "Summer", "Autumn", "Winter")
ggplot(predframe.SexSeasonC, aes(x = Season, y = preds, fill = Sex)) +
    geom_errorbar(aes(ymin=preds - se, ymax=preds + se),
        width=.2, position=position_dodge(.3)) +
    geom_point(size = 5, shape = 21, position = position_dodge(.3)) +
    theme_classic() +
labs(y=expression((delta)^13*C~\'‰'), x = "Season") +
theme(axis.title.x = element_text(size=18), axis.title.y = element_text(size=18),
    axis.text.x = element_text(size=14, colour = "black"),
    axis.text.y = element_text(size=14, colour = "black"),
    axis.ticks.length=unit(.3, "cm"),
    plot.margin=unit(c(0.5,0.5,1,1),"cm"),
    legend.position = "none") +
scale_y_continuous(breaks = scales::pretty_breaks(n = 10)) +
scale_x_discrete(labels= Season) +
scale_color_manual(values = c("blue", "red"))

Biotic variables

Association between carbon and sex (multiple whisker segments model)
plot(X13C ~ Sex, data = Whiskers2, bty = 'L', ylab =
expression((delta)^13*C~\'‰'),
    type = "n")

Carbon interaction between size and sex (basal segments model)
Create model including only the variables that were important in the averaged model.
WaterC <- gam(X13CBase ~ Sex + Size + Size:Sex + Scored.SMI + Season +
WaterClass +

107
Sex: WaterClass + s(Urban) + te(lat,long) + s(lat) + s(long) + KmRiverDist + Rainfall, data=Base2)

Create a dummy dataset and then make predictions using this dataset and the model created above.

```r
pdat.SizeC <- expand.grid(Size = levels(Base2$Size),
  WaterClass = c("Coast/Estuary"),
  Scored.SMI = mean(Base2$Scored.SMI, na.rm=T),
  Season=c("1"),
  Rainfall = mean(Base2$Rainfall, na.rm=T),
  long = mean(Base2$long, na.rm=T),
  Sex = levels(Base2$Sex),
  lat = mean(Base2$lat, na.rm=T),
  KmRiverDist = mean(Base2$KmRiverDist, na.rm=T),
  Urban = mean(Base2$Urban, na.rm=T)
)

pred.SizeC <- predict(WaterC, newdata=pdat.SizeC, na.rm=T, type="response", se.fit=TRUE)
predframe.SizeC <- data.frame(pdat.SizeC, preds=pred.SizeC$fit, se = pred.SizeC$se.fit)
```

Plot data/predictions.

```r
ggplot(predframe.SizeC, aes(x = Size, y = preds, fill = Sex)) +
  geom_errorbar(aes(ymin=preds - se, ymax=preds + se),
               width=.2, position=position_dodge(.3)) +
  geom_point(size = 5, shape = 21, position = position_dodge(.3)) +
  theme_classic() +
  labs(y=expression((delta)^13*C~"\'u2030"'), x = "Size Class") +
  theme(axis.title.x = element_text(size=18), axis.title.y = element_text(size=18),
        axis.text.x = element_text(size=14, colour = "black"),
        axis.text.y = element_text(size=14, colour = "black"),
        axis.ticks.length=unit(.3, "cm"),
        plot.margin=unit(c(0.5,0.5,1,1),"cm"),
        legend.position = "none") +
  scale_y_continuous(breaks = scales::pretty_breaks(n = 10)) +
  scale_color_manual(values = c("blue", "red"))
```

Association between carbon and sex (individual variation model)

```r
plot(C_Whisker_StDev ~ Sex, data = diff, bty = 'L')
```
Association between nitrogen and size (multiple whisker segments model)

```r
plot(X15NBase ~ Size, data = Base2, bty = 'L',
     ylab = expression(\(\delta^{15}\!\text{N} \sim \mu\text{u2030}\))
```

Association between nitrogen and size (individual variation model)

```r
plot(N_Whisker_StDev ~ Size, data = diff, bty = 'L')
```

Use SIBER to plot isotopic variation observed in models using multiple whisker segments and those using only basal segments

Set the parameters for plotting the data

```r
set.seed(1)
library(SIBER)
palette(c("purple", "green"))
community.hulls.args <- list(col = 1, lty = 1, lwd = 1)
group.ellipses.args <- list(n = 100, p.interval = 0.95, lty = 1, lwd = 2)
group.hull.args <- list(lty = 2, col = "grey20")
```

Read in the data, make a SIBER object and then plot the data (red = multiple whisker segments, blue = basal segments). Add convex hulls (grey dashed lines), predicted ellipses that encompass approximately 95% of the data (large ellipses) and ellipses that depict the 95% confidence interval around the bivariate means (small ellipses).

```r
Segment <- read.csv("BaseVsFullSIBER.csv", header=T)
siber.Segment <- createSiberObject(Segment)
plotSiberObject(siber.Segment,
               ax.pad = 2,
               hulls = F, community.hulls.args,
               ellipses = T, group.ellipses.args,
               group.hulls = T, group.hull.args,
               bty = "L",
               iso.order = c(1, 2),
               xlab = "",
               ylab = "",
               cex.axis = 1.5)
```

```r
title(ylab = expression(\(\delta^{15}\!\text{N} \sim \mu\text{u2030}\)), line = 2.4, cex.lab = 1.75)
```
Calculate summary statistics for each group (sample type): TA, SEA, SEAc and range of isotopic values.

```r
group.ML <- groupMetricsML(siber.Segment)
print(group.ML)
summary(Segment[Segment$group == 1,])
summary(Segment[Segment$group == 2,])
```

Fit bayesian model to the data. First set the options for running jags.

```r
library(rjags)
parms <- list()
parms$n.iter <- 2 * 10^5
parms$n.burnin <- 1 * 10^3
parms$n.thin <- 10
parms$n.chains <- 2
Define the priors.
priors <- list()
priors$R <- 1 * diag(2)
priors$k <- 2
priors$tau.mu <- 1.0E-3
```

Fit ellipses using the priors and use these to calculate SEA.B for each group. SEA.B is then plotted and red x’s are added to denote the maximum likelihood estimates of SEA-c to the Bayesian estimates.

```r
ellipses.posterior <- siberMVN(siber.Segment, parms, priors)
SEA.B <- siberEllipses(ellipses.posterior)

siberDensityPlot(SEA.B, xticklabels = c("", ""),
                 xlab = "",
                 ylab = "",
                 bty = "L",
                 las = 1,
                 main = "",
                 cex.axis = 1.5)

points(1:ncol(SEA.B), group.ML[3,], col="red", pch = "x", lwd = 2)
```
Calculate credible intervals and modes.

cr.p <- c(0.95, 0.99)
SEA.B.credibles <- \texttt{lapply}(\texttt{as.data.frame}(SEA.B), \texttt{function}(x,...){tmp<-hdrcde::hdr(x)$hdr}, \texttt{prob = cr.p})

SEA.B.credibles

SEA.B.modes <- \texttt{lapply}(\texttt{as.data.frame}(SEA.B), \texttt{function}(x,...){tmp<-hdrcde::hdr(x)$mode}, \texttt{prob = cr.p, all.modes=T})

SEA.B.modes
3.7.2 Correlation between landscape variables

Figure S 3.1. Associations between spatial variables. Plot shows the correlation using a Pearson’s correlation coefficient. Spatial variables represent the landscape within a 20 km buffer of each individual otter: ‘Agriculture’ describes the proportion of arable and horticultural land use within each buffer, ‘Longitude’ is the position an otter was found at (center of the buffer), ‘Imp.Grassland’ describes the proportion of land use assigned to improved grassland and ‘Altitude’ and ‘Slope’ are the mean values within each buffer. Blue circles represent positive correlations whilst red circles represent negative correlations. Darker colours and larger circles represent stronger correlations.
3.7.3 Model averaging tables

Table S 3.1. Conditional average model output for general additive mixed effects model (GAMM) of δ¹⁵N in Eurasian otter (Lutra lutra) whisker samples from England and Wales 2007-2016. Averaged models were created from models with Akaike’s Information Criterion (AIC) values of less than two. Fixed variables preceded by an ‘s’ or ‘t2’ represent a variable with a smoothing function with the number following representing the knot number. Fixed variables labelled ‘small’, ‘medium’, ‘large’ represent otter size. Effect size (estimate), standard error (std. error), adjusted standard error (adjusted SE), z value and p value (Pr(>|z|)) are given for each fixed variable in the final averaged model.

| Variable                        | Relative Importance | Level                                             | Estimate | Std. Error | Adjusted SE | z value | Pr(>|z|) |
|---------------------------------|---------------------|---------------------------------------------------|----------|------------|-------------|---------|---------|
| Water habitat type              | 0.72                | Lake (compared to coast/testuary)                 | -0.85    | 0.45       | 0.45        | 1.92    | 0.055   |
|                                 |                     | Main River Channel (compared to coast/testuary)   | -0.68    | 0.49       | 0.49        | 1.38    | 0.167   |
|                                 |                     | Tributary (compared to coast/testuary)            | -0.94    | 0.51       | 0.51        | 1.85    | 0.064   |
|                                 |                     | Main River Channel (compared to lake)              | 0.13     | 0.41       | 0.41        | 0.33    | 0.740   |
|                                 |                     | Tributary (compared to lake)                      | -0.09    | 0.49       | 0.49        | 0.17    | 0.862   |
|                                 |                     | Tributary (compared to main river channel)        | -0.35    | 0.57       | 0.57        | 0.61    | 0.545   |
| Proportion of urban land use    | 1                   | s(Urban Land Use).1                               | 0.50     | 0.32       | 0.32        | 1.56    | 0.120   |
|                                 |                     | s(Urban Land Use).2                               | -0.19    | 1.29       | 1.29        | 1.02    | 0.311   |
|                                 |                     | s(Urban Land Use).3                               | 0.02     | 0.15       | 0.15        | 0.19    | 0.062   |
|                                 |                     | s(Urban Land Use).4                               | 0.08     | 0.84       | 0.84        | 0.09    | 0.927   |
|                                 |                     | s(Urban Land Use).5                               | 0.02     | 0.21       | 0.22        | 0.11    | 0.914   |
|                                 |                     | s(Urban Land Use).6                               | 0.07     | 0.77       | 0.77        | 0.09    | 0.930   |
| Latitude.Longitude              | 1                   | 2(Latitude,Longitude).1                           | 2.22     | 0.81       | 0.81        | 2.74    | 0.006   |
|                                 |                     | 2(Latitude,Longitude).2                           | -0.43    | 0.52       | 0.52        | 0.83    | 0.405   |
|                                 |                     | 2(Latitude,Longitude).3                           | -0.68    | 0.87       | 0.87        | 0.79    | 0.430   |
|                                 |                     | 2(Latitude,Longitude).4                           | -0.18    | 0.66       | 0.66        | 0.28    | 0.781   |
|                                 |                     | 2(Latitude,Longitude).5                           | 0.11     | 0.33       | 0.33        | 0.33    | 0.739   |
|                                 |                     | 2(Latitude,Longitude).6                           | 0.72     | 0.49       | 0.49        | 1.46    | 0.145   |
|                                 |                     | 2(Latitude,Longitude).7                           | -1.21    | 0.33       | 0.33        | 3.66    | <0.001  |
|                                 |                     | 2(Latitude,Longitude).8                           | 0.05     | 0.37       | 0.37        | 0.13    | 0.897   |
|                                 |                     | 2(Latitude,Longitude).9                           | -2.68    | 0.74       | 0.74        | 3.64    | <0.001  |
|                                 |                     | 2(Latitude,Longitude).10                          | -0.06    | 0.52       | 0.52        | 0.16    | 0.876   |
|                                 |                     | 2(Latitude,Longitude).11                          | -0.91    | 0.63       | 0.63        | 1.44    | 0.149   |
|                                 |                     | 2(Latitude,Longitude).12                          | -0.85    | 0.96       | 0.96        | 0.89    | 0.375   |
|                                 |                     | 2(Latitude,Longitude).13                          | 2.23     | 0.68       | 0.68        | 3.27    | 0.001   |
|                                 |                     | 2(Latitude,Longitude).14                          | -0.50    | 0.51       | 0.51        | 0.99    | 0.322   |
|                                 |                     | 2(Latitude,Longitude).15                          | 0.10     | 0.35       | 0.35        | 0.28    | 0.783   |
|                                 |                     | 2(Latitude,Longitude).16                          | 0.14     | 0.36       | 0.36        | 0.40    | 0.693   |
|                                 |                     | 2(Latitude,Longitude).17                          | 0.09     | 0.32       | 0.32        | 0.28    | 0.777   |
|                                 |                     | 2(Latitude,Longitude).18                          | 0.02     | 0.34       | 0.34        | 0.07    | 0.946   |
|                                 |                     | 2(Latitude,Longitude).19                          | 0.13     | 0.60       | 0.60        | 0.22    | 0.828   |
|                                 |                     | 2(Latitude,Longitude).20                          | 0.19     | 0.33       | 0.33        | 0.58    | 0.561   |
|                                 |                     | 2(Latitude,Longitude).21                          | -0.11    | 0.75       | 0.75        | 0.14    | 0.888   |
|                                 |                     | 2(Latitude,Longitude).22                          | -0.02    | 0.59       | 0.59        | 0.04    | 0.972   |
|                                 |                     | 2(Latitude,Longitude).23                          | 0.87     | 0.79       | 0.79        | 1.11    | 0.268   |
|                                 |                     | 2(Latitude,Longitude).24                          | -1.99    | 0.96       | 0.96        | 2.08    | 0.037   |
| Sex                             | 0.5                 | Male (compared to female)                         | -0.41    | 0.37       | 0.37        | 1.11    | 0.268   |
| Size                            | 0.54                | Medium (compared to large)                        | -0.58    | 0.32       | 0.32        | 1.81    | 0.070   |
|                                 |                     | Small (compared to large)                         | -0.11    | 0.39       | 0.39        | 0.27    | 0.784   |
|                                 |                     | Small (compared to medium)                        | 0.47     | 0.42       | 0.42        | 1.12    | 0.263   |
| Sex:Water habitat type          | 0.28                | Male:Lake (compared to female:coast/testuary)    | -0.44    | 0.68       | 0.68        | 0.65    | 0.519   |
|                                 |                     | Male:Main River Channel (compared to female:coast/testuary) | -0.26 | 0.69 | 0.69 | 0.38 | 0.707 |
|                                 |                     | Male:Tributary (compared to female:coast/testuary) | 0.94 | 0.62 | 0.62 | 1.51 | 0.132 |
|                                 |                     | Male:Main River Channel (compared to female:lake) | 0.18 | 0.76 | 0.76 | 0.24 | 0.810 |
|                                 |                     | Male:Tributary (compared to female:lake)          | 1.36 | 0.69 | 0.69 | 1.99 | 0.046 |
|                                 |                     | Male:Tributary (compared to female:main river channel) | 0.33 | 0.65 | 0.65 | 0.51 | 0.613 |
| Variable                                      | Relative Importance | Level                                      | Estimate | Std. Error | Adjusted SE | z value | Pr(>|z|) |
|-----------------------------------------------|---------------------|--------------------------------------------|----------|------------|-------------|---------|---------|
| Medium.Lake (Compared to large:coast/estuary)| 0.09                | 0.77                                      | 0.78     | 0.12       | 0.905       |         |         |
| Small.Lake ( compared to large:coast/estuary)| 0.07                | 0.95                                      | 0.95     | 0.07       | 0.941       |         |         |
| Medium:Main River Channel (compared to large:coast/estuary) | 0.26                | 0.78                                      | 0.78     | 0.34       | 0.735       |         |         |
| Small:Main River Channel (compared to large:coast/estuary) | 0.92                | 1.00                                      | 1.00     | 0.92       | 0.356       |         |         |
| Medium:TrIBUTARY (compared to large:coast/estuary) | -0.76               | 0.69                                      | 0.69     | 1.10       | 0.272       |         |         |
| Small:TrIBUTARY (compared to large:coast/estuary) | 0.66                | 0.83                                      | 0.83     | 0.83       | 0.409       |         |         |
| Medium:Lake (Compared to large:tributary)    | 0.76                | 0.69                                      | 0.69     | 1.10       | 0.273       |         |         |
| Small:Lake (compared to large:tributary)     | -0.68               | 0.83                                      | 0.83     | 0.82       | 0.411       |         |         |
| Medium:Main River Channel (compared to large:tributary) | 0.85                | 0.80                                      | 0.80     | 1.07       | 0.285       |         |         |
| Small:Main River Channel (compared to large:tributary) | -0.61               | 0.92                                      | 0.92     | 0.66       | 0.509       |         |         |
| Medium:Coast:Estuary (compared to large:tributary) | 1.03                | 0.80                                      | 0.81     | 1.28       | 0.201       |         |         |
| Small:Coast:Estuary (compared to large:tributary) | 0.24                | 1.00                                      | 1.00     | 0.24       | 0.808       |         |         |
| Medium:Lake (Compared to large:main river channel) | -0.01               | 0.22                                      | 0.22     | 0.05       | 0.961       |         |         |
| Small:Lake (compared to large:main river channel) | -0.05               | 0.34                                      | 0.34     | 0.15       | 0.877       |         |         |
| Medium:Coast:Estuary (compared to large:main river channel) | -0.02               | 0.20                                      | 0.20     | 0.08       | 0.936       |         |         |
| Small:Coast:Estuary (compared to large:main river channel) | -0.06               | 0.33                                      | 0.33     | 0.17       | 0.864       |         |         |
| Medium:TrIBUTARY (compared to large:main river channel) | -0.06               | 0.32                                      | 0.32     | 0.20       | 0.842       |         |         |
| Small:TrIBUTARY (compared to large:main river channel) | -0.01               | 0.26                                      | 0.26     | 0.06       | 0.954       |         |         |
| Large:Lake (compared to medium:coast/estuary) | -0.09               | 0.77                                      | 0.78     | 0.12       | 0.905       |         |         |
| Small:Lake (compared to medium:coast/estuary) | -0.02               | 0.93                                      | 0.93     | 0.02       | 0.981       |         |         |
| Large:Main River Channel (compared to medium:coast/estuary) | -0.26               | 0.78                                      | 0.78     | 0.34       | 0.736       |         |         |
| Small:Main River Channel (compared to medium:coast/estuary) | 0.66                | 0.97                                      | 0.97     | 0.68       | 0.496       |         |         |
| Large:TrIBUTARY (compared to medium:coast/estuary) | 0.76                | 0.69                                      | 0.69     | 1.10       | 0.272       |         |         |
| Small:TrIBUTARY (compared to medium:coast/estuary) | 1.44                | 0.87                                      | 0.87     | 1.66       | 0.098       |         |         |
| Large:Coast:Estuary (compared to medium:lake)  | 0.09                | 0.77                                      | 0.78     | 0.12       | 0.905       |         |         |
| Small:Coast:Estuary (compared to medium:lake)  | 0.02                | 0.93                                      | 0.93     | 0.02       | 0.981       |         |         |
| Large:Main River Channel (compared to medium:lake) | -0.17               | 0.87                                      | 0.87     | 0.20       | 0.843       |         |         |
| Small:Main River Channel (compared to medium:lake) | 0.68                | 1.03                                      | 1.03     | 0.66       | 0.507       |         |         |
| Large:TrIBUTARY (compared to medium:lake)     | 0.85                | 0.80                                      | 0.80     | 1.07       | 0.284       |         |         |
| Small:TrIBUTARY (compared to medium:lake)     | 1.47                | 0.94                                      | 0.95     | 1.55       | 0.121       |         |         |
| Large:Lake (compared to medium:main river channel) | 0.17                | 0.87                                      | 0.87     | 0.20       | 0.843       |         |         |
| Small:Lake (compared to medium:main river channel) | -0.68               | 1.03                                      | 1.03     | 0.66       | 0.508       |         |         |
| Large:Coast:Estuary (compared to medium:main river channel) | 0.26                | 0.78                                      | 0.78     | 0.34       | 0.735       |         |         |
| Small:Coast:Estuary (compared to medium:main river channel) | -0.66               | 0.97                                      | 0.97     | 0.68       | 0.497       |         |         |
| Large:TrIBUTARY (compared to medium:main river channel) | 1.03                | 0.80                                      | 0.80     | 1.27       | 0.203       |         |         |
| Small:TrIBUTARY (compared to medium:main river channel) | 0.79                | 0.96                                      | 0.96     | 0.82       | 0.414       |         |         |
Table S 3.2. Conditional average model output for general additive mixed effects model (GAMM) of δ¹³C in Eurasian otter (*Lutra lutra*) whisker samples from England and Wales 2007-2016. Averaged models were created from models with Akaike’s Information Criterion (AIC) values of less than two. Fixed variables preceded by an ‘s’ or ‘t2’ represent a variable with a smoothing function with the number following representing the knot number. Effect size (estimate), standard error (std. error), adjusted standard error (adjusted SE), z value and p value (Pr(>|z|)) are given for each fixed variable in the final averaged model.

| Variable     | Relative Importance | Level                     | Estimate | Std. Error | Adjusted SE | z value | Pr(>|z|) |
|--------------|---------------------|---------------------------|----------|------------|-------------|---------|----------|
| Sex          | 1                   | Male (compared to female) | 0.53     | 0.23       | 0.23        | 2.26    | 0.024    |
| Latitude     | 0.59                | s(Latitude).1             | 0.32     | 0.53       | 0.53        | 0.61    | 0.543    |
|              |                     | s(Latitude).2             | 0.56     | 1.17       | 1.17        | 0.48    | 0.634    |
|              |                     | s(Latitude).3             | 0.31     | 0.33       | 0.33        | 0.94    | 0.349    |
|              |                     | s(Latitude).4             | -0.03    | 0.57       | 0.57        | 0.06    | 0.952    |
|              |                     | s(Latitude).5             | 0.02     | 0.18       | 0.18        | 0.09    | 0.933    |
|              |                     | s(Latitude).6             | -0.01    | 0.54       | 0.54        | 0.01    | 0.989    |
|              |                     | s(Latitude).7             | 0.02     | 0.16       | 0.16        | 0.13    | 0.895    |
|              |                     | s(Latitude).8             | -0.32    | 2.00       | 2.00        | 0.16    | 0.874    |
|              |                     | s(Latitude).9             | 0.74     | 0.76       | 0.76        | 0.97    | 0.333    |
| Longitude    | 1                   | s(Longitude).1            | -0.18    | 1.36       | 1.37        | 0.13    | 0.897    |
|              |                     | s(Longitude).2            | -1.40    | 2.92       | 2.92        | 0.48    | 0.632    |
|              |                     | s(Longitude).3            | -1.59    | 0.65       | 0.66        | 2.43    | 0.015    |
|              |                     | s(Longitude).4            | 2.15     | 2.03       | 2.03        | 1.06    | 0.290    |
|              |                     | s(Longitude).5            | -0.89    | 0.67       | 0.67        | 1.32    | 0.185    |
|              |                     | s(Longitude).6            | -2.53    | 1.82       | 1.82        | 1.39    | 0.165    |
|              |                     | s(Longitude).7            | -0.72    | 0.64       | 0.64        | 1.13    | 0.257    |
|              |                     | s(Longitude).8            | -5.24    | 5.32       | 5.33        | 0.98    | 0.325    |
|              |                     | s(Longitude).9            | -1.37    | 2.29       | 2.29        | 0.60    | 0.551    |
| Latitude:Longitude | 0.41 | t2(Latitude:Longitude).1 | -0.15    | 0.24       | 0.24        | 0.64    | 0.525    |
|              |                     | t2(Latitude:Longitude).2 | 0.05     | 0.21       | 0.21        | 0.25    | 0.806    |
|              |                     | t2(Latitude:Longitude).3 | -0.01    | 0.24       | 0.24        | 0.03    | 0.978    |
|              |                     | t2(Latitude:Longitude).4 | -0.05    | 0.22       | 0.22        | 0.21    | 0.836    |
|              |                     | t2(Latitude:Longitude).5 | -0.20    | 0.17       | 0.17        | 1.23    | 0.219    |
|              |                     | t2(Latitude:Longitude).6 | -0.10    | 0.22       | 0.22        | 0.48    | 0.632    |
|              |                     | t2(Latitude:Longitude).7 | 0.22     | 0.15       | 0.15        | 1.47    | 0.143    |
|              |                     | t2(Latitude:Longitude).8 | -0.25    | 0.11       | 0.11        | 2.33    | 0.020    |
|              |                     | t2(Latitude:Longitude).9 | 0.10     | 0.38       | 0.38        | 0.25    | 0.802    |
|              |                     | t2(Latitude:Longitude).10| -0.01    | 0.38       | 0.38        | 0.04    | 0.970    |
|              |                     | t2(Latitude:Longitude).11| 0.01     | 0.40       | 0.40        | 0.02    | 0.985    |
|              |                     | t2(Latitude:Longitude).12| 0.11     | 0.35       | 0.35        | 0.33    | 0.744    |
|              |                     | t2(Latitude:Longitude).13| 0.27     | 0.31       | 0.31        | 0.88    | 0.379    |
|              |                     | t2(Latitude:Longitude).14| -0.35    | 0.21       | 0.21        | 1.68    | 0.093    |
|              |                     | t2(Latitude:Longitude).15| 0.00     | 0.01       | 0.01        | 0.05    | 0.957    |
|              |                     | t2(Latitude:Longitude).16| 0.00     | 0.01       | 0.01        | 0.05    | 0.957    |
|              |                     | t2(Latitude:Longitude).17| 0.00     | 0.00       | 0.00        | 0.00    | 0.957    |
|              |                     | t2(Latitude:Longitude).18| 0.00     | 0.00       | 0.00        | 0.00    | 0.957    |
|              |                     | t2(Latitude:Longitude).19| 0.01     | 0.12       | 0.12        | 0.05    | 0.957    |
|              |                     | t2(Latitude:Longitude).20| 0.00     | 0.03       | 0.03        | 0.05    | 0.957    |
|              |                     | t2(Latitude:Longitude).21| 0.14     | 0.36       | 0.36        | 0.40    | 0.692    |
|              |                     | t2(Latitude:Longitude).22| 0.04     | 0.37       | 0.37        | 0.12    | 0.903    |
|              |                     | t2(Latitude:Longitude).23| -0.18    | 0.49       | 0.49        | 0.37    | 0.711    |
Table S 3.3. Conditional average model output for general additive model (GAM) of δ^{15}N in basal segments of Eurasian otter (Lutra lutra) whiskers from England and Wales 2007-2016. Averaged models were created from models with Akaike’s Information Criterion (AIC) values of less than two. Fixed variables preceded by an ‘s’ or ‘te’ represent a variable with a smoothing function with the number following representing the knot number. Fixed variables labelled ‘small’, ‘medium’, ‘large’ represent otter size. Effect size (estimate), standard error (std. error), adjusted standard error (adjusted SE), z value and p value (Pr(>|z|)) are given for each fixed variable in the final averaged model.

| Variable | Relative Importance | Level | Estimate | Std. Error | Adjusted SE | z value | Pr(>|z|) |
|----------|---------------------|-------|----------|------------|------------|----------|---------|
| Distance from the coast (Km) | 1 | - | -0.01 | 0.00 | 0.00 | 1.64 | 0.100 |
| Longitude | 0.5 | - | -0.33 | 0.46 | 0.46 | 0.72 | 0.471 |
| Rainfall (mm) | 1 | - | 0.01 | 0.00 | 0.00 | 2.74 | 0.006 |
| Sex | 0.9 | Male (compared to female) | -0.34 | 0.42 | 0.43 | 0.80 | 0.422 |
| Water habitat type | 0.9 | Lake (compared to coastline) | -0.91 | 0.48 | 0.48 | 1.89 | 0.059 |
| | | Main River Channel (compared to coastline) | -0.83 | 0.55 | 0.55 | 1.51 | 0.132 |
| | | Tributary (compared to coastline) | -0.83 | 0.69 | 0.69 | 1.21 | 0.228 |
| | | Main River Channel (compared to lake) | 0.09 | 0.55 | 0.56 | 0.16 | 0.876 |
| | | Tributary (compared to lake) | 0.08 | 0.72 | 0.72 | 0.11 | 0.910 |
| | | Tributary (compared to main river channel) | 0.00 | 0.68 | 0.68 | 0.01 | 0.994 |
| Sex:Distance from the coast | 0.59 | Male:Distance from the coast (Km) (compared to female) | -0.01 | 0.00 | 0.00 | 1.91 | 0.056 |
| Sex:Water habitat type | 0.5 | Male:Main Water Habitat Channel (compared to female) | -0.08 | 0.77 | 0.77 | 0.10 | 0.921 |
| Latitude | 0.54 | s(Latitude).1 | 0.38 | 1.07 | 1.08 | 0.35 | 0.726 |
| | | s(Latitude).2 | -0.32 | 1.81 | 1.82 | 0.17 | 0.862 |
| | | s(Latitude).3 | -0.39 | 0.53 | 0.53 | 0.73 | 0.486 |
| | | s(Latitude).4 | 0.02 | 0.93 | 0.93 | 0.02 | 0.982 |
| | | s(Latitude).5 | 0.00 | 0.27 | 0.27 | 0.02 | 0.986 |
| | | s(Latitude).6 | 0.03 | 0.90 | 0.90 | 0.03 | 0.974 |
| | | s(Latitude).7 | 0.01 | 0.26 | 0.26 | 0.04 | 0.972 |
| | | s(Latitude).8 | 0.28 | 3.64 | 3.66 | 0.08 | 0.939 |
| | | s(Latitude).9 | 0.05 | 1.56 | 1.56 | 0.03 | 0.974 |
| Proportion of urban land use | 1 | s(Urban Land Use).1 | 0.14 | 0.20 | 0.20 | 0.68 | 0.485 |
| | | s(Urban Land Use).2 | -0.11 | 0.65 | 0.66 | 0.16 | 0.870 |
| | | s(Urban Land Use).3 | 0.00 | 0.09 | 0.09 | 0.05 | 0.957 |
| | | s(Urban Land Use).4 | 0.12 | 0.42 | 0.43 | 0.28 | 0.779 |
| | | s(Urban Land Use).5 | 0.03 | 0.09 | 0.09 | 0.31 | 0.759 |
| | | s(Urban Land Use).6 | 0.12 | 0.39 | 0.40 | 0.31 | 0.754 |
| | | s(Urban Land Use).7 | 0.01 | 0.04 | 0.04 | 0.14 | 0.889 |
| | | s(Urban Land Use).8 | 0.62 | 1.50 | 1.51 | 0.41 | 0.679 |
| | | s(Urban Land Use).9 | 0.46 | 0.35 | 0.35 | 1.31 | 0.192 |
| Latitude: Longitude | 1 | s(Latitude:Longitude).1 | -1.16 | 2.23 | 2.24 | 0.52 | 0.603 |
| | | s(Latitude:Longitude).2 | -1.34 | 2.65 | 2.66 | 0.51 | 0.613 |
| | | s(Latitude:Longitude).3 | -8.83 | 3.58 | 3.60 | 2.46 | 0.014 |
| | | s(Latitude:Longitude).4 | 8.40 | 6.85 | 6.88 | 1.22 | 0.222 |
| | | s(Latitude:Longitude).5 | 0.30 | 2.08 | 2.09 | 0.15 | 0.882 |
| | | s(Latitude:Longitude).6 | -0.94 | 0.72 | 0.72 | 1.31 | 0.192 |
| | | s(Latitude:Longitude).7 | 0.48 | 0.78 | 0.78 | 0.62 | 0.537 |
| | | s(Latitude:Longitude).8 | -0.65 | 1.28 | 1.28 | 0.51 | 0.613 |
| | | s(Latitude:Longitude).9 | 4.62 | 2.79 | 2.80 | 1.65 | 0.099 |
| | | s(Latitude:Longitude).10 | 1.16 | 1.89 | 1.90 | 0.61 | 0.541 |
| | | s(Latitude:Longitude).11 | -2.46 | 0.93 | 0.94 | 2.63 | 0.009 |
| | | s(Latitude:Longitude).12 | -0.27 | 0.68 | 0.68 | 0.39 | 0.693 |
| | | s(Latitude:Longitude).13 | 5.32 | 1.54 | 1.55 | 3.44 | 0.001 |
| | | s(Latitude:Longitude).14 | 1.27 | 2.06 | 2.06 | 0.62 | 0.537 |
| | | s(Latitude:Longitude).15 | 1.94 | 2.51 | 2.52 | 0.77 | 0.441 |
| | | s(Latitude:Longitude).16 | -0.51 | 1.80 | 1.80 | 0.29 | 0.776 |
| | | s(Latitude:Longitude).17 | 0.71 | 1.21 | 1.22 | 0.59 | 0.559 |
| | | s(Latitude:Longitude).18 | 4.37 | 1.91 | 1.91 | 2.28 | 0.023 |
| | | s(Latitude:Longitude).19 | -1.71 | 2.09 | 2.10 | 0.81 | 0.416 |
| | | s(Latitude:Longitude).20 | 1.76 | 11.03 | 11.08 | 0.16 | 0.873 |
| | | s(Latitude:Longitude).21 | -2.99 | 2.61 | 2.62 | 1.15 | 0.252 |
| | | s(Latitude:Longitude).22 | -0.80 | 2.30 | 2.30 | 0.35 | 0.729 |
| | | s(Latitude:Longitude).23 | -7.71 | 9.08 | 9.11 | 0.85 | 0.397 |
| | | s(Latitude:Longitude).24 | -13.70 | 10.33 | 10.38 | 1.32 | 0.187 |
| Size | 0.42 | Large (compared to medium) | 0.63 | 0.30 | 0.31 | 2.05 | 0.040 |
| | | Small (compared to medium) | 0.42 | 0.36 | 0.36 | 1.14 | 0.253 |
| | | Small (compared to large) | -0.21 | 0.36 | 0.36 | 0.59 | 0.554 |
| Year | 0.14 | - | -0.04 | 0.05 | 0.05 | 0.89 | 0.372 |
Table S 3.4. Conditional average model output for general additive model (GAM) of ∂13C in
basal segments of Eurasian otter (Lutra lutra) whiskers from England and Wales 2007-2016.
Averaged models were created from models with Akaike’s Information Criterion (AIC) values of
less than two. Fixed variables preceded by an ‘s’ or ‘te’ represent a variable with a smoothing
function with the number following representing the knot number. Fixed variables labelled
‘small’, ‘medium’, ‘large’ represent otter size. Effect size (estimate), standard error (std. error),
adjusted standard error (adjusted SE), z value and p value (Pr(> z )) are given for each fixed
variable in the final averaged model.
Variable
Distance from the coast
(Km)
Latitude
Longitude
Scored Body Condition

Relative
Importance

Level

0.67

-

0.5
0.5
0.83

Summer (compared to spring)
Autumn (compared to spring)
Winter (compared to spring)
Autumn (compared to summer)
Winter (compared to summer)
Winter (compared to Autumn)
Male (compared to female)
Medium (compared to large)
Small (compared to large)
Small (compared to medium)
Lake (compared to coast/estuary)
Main River Channel (compared to coast/estuary)
Tributary (compared to coast/estuary)
Main River Channel (compared to lake)
Tributary (compared to lake)
Tributary (compared to main river channel)
Male:Summer (compared to female:spring)
Male:Autumn (compared to female:spring)
Male:Winter (compared to female:spring)
Male:Autumn (compared to female:summer)
Male:Winter (compared to female:summer)
Male:Winter (compared to female:Autumn)
s(Urban Land Use).1
s(Urban Land Use).2
s(Urban Land Use).3
s(Urban Land Use).4
s(Urban Land Use).5
s(Urban Land Use).6
s(Urban Land Use).7
s(Urban Land Use).8
s(Urban Land Use).9
te(Latitude,Longitude).1
te(Latitude,Longitude).2
te(Latitude,Longitude).3
te(Latitude,Longitude).4
te(Latitude,Longitude).5
te(Latitude,Longitude).6
te(Latitude,Longitude).7
te(Latitude,Longitude).8
te(Latitude,Longitude).9
te(Latitude,Longitude).10
te(Latitude,Longitude).11
te(Latitude,Longitude).12
te(Latitude,Longitude).13
te(Latitude,Longitude).14
te(Latitude,Longitude).15
te(Latitude,Longitude).16
te(Latitude,Longitude).17
te(Latitude,Longitude).18
te(Latitude,Longitude).19
te(Latitude,Longitude).20
te(Latitude,Longitude).21
te(Latitude,Longitude).22
te(Latitude,Longitude).23
te(Latitude,Longitude).24
Male:Medium (compared to female:large)
Male:Small (compared to female:large)
Male:Large (compared to female:medium)
Male:Small (compared to female:medium)
-

Season

0.81

Sex

1

Size

0.96

Water habitat type

1

Sex:Season

0.64

Proportion of Urban land
use

1

Latitude: Longitude

1

Sex:Size

0.42

Rainfall (mm)

0.32

Estimate
0.00
-0.45
-0.28
0.30
0.20
-0.23
-0.59
-0.43
-0.79
-0.36
0.15
0.17
-0.39
-0.56
-0.39
-1.19
-1.10
-0.80
-0.72
0.08
0.76
1.96
1.02
1.20
0.25
-0.94
-0.06
0.39
-0.03
-0.36
-0.11
-0.41
-0.01
-2.03
0.16
0.10
-1.20
-1.38
2.47
1.38
0.08
-1.24
-1.51
0.91
1.89
0.47
-0.48
-1.61
-0.75
2.09
0.59
-0.11
-0.90
-1.93
2.19
1.80
2.12
2.95
-4.19
0.15
-1.25
-0.15
-1.40
0.00

Std. Adjusted
z
Pr(> z )
Error
SE
value
0.00
0.02
0.25
0.16
0.55
0.64
0.51
0.62
0.52
0.53
0.67
0.35
0.60
0.57
0.37
0.41
0.40
0.41
0.38
0.38
0.75
0.70
0.69
0.75
0.74
0.69
0.27
0.97
0.13
0.63
0.14
0.58
0.07
2.16
0.48
0.88
1.05
1.39
3.34
1.16
0.47
0.46
0.68
1.77
1.25
0.40
0.44
0.86
1.21
1.50
0.65
0.52
0.95
1.33
3.75
1.22
1.02
1.41
5.14
0.57
0.70
0.57
0.69
0.00

0.00
0.02
0.25
0.16
0.55
0.64
0.51
0.62
0.52
0.53
0.67
0.36
0.60
0.57
0.38
0.41
0.40
0.41
0.38
0.38
0.76
0.70
0.70
0.75
0.74
0.69
0.27
0.98
0.13
0.63
0.14
0.58
0.07
2.17
0.48
0.88
1.05
1.39
3.35
1.17
0.47
0.46
0.68
1.78
1.26
0.40
0.44
0.86
1.21
1.50
0.66
0.53
0.95
1.33
3.77
1.22
1.02
1.41
5.16
0.57
0.70
0.57
0.70
0.00

1.57
26.11
1.16
1.91
0.37
0.36
1.16
0.70
1.51
0.69
0.23
0.48
0.65
0.97
1.03
2.90
2.75
1.96
1.86
0.21
1.01
2.79
1.46
1.59
0.34
1.37
0.24
0.40
0.26
0.57
0.79
0.70
0.20
0.94
0.33
0.12
1.14
0.99
0.74
1.19
0.18
2.68
2.23
0.51
1.50
1.19
1.08
1.86
0.62
1.39
0.90
0.22
0.95
1.45
0.58
1.47
2.08
2.10
0.81
0.27
1.78
0.27
2.01
1.34

0.117
< 0.001
0.247
0.056
0.713
0.720
0.245
0.487
0.130
0.494
0.820
0.631
0.518
0.331
0.301
0.004
0.006
0.051
0.062
0.830
0.312
0.005
0.143
0.112
0.731
0.171
0.810
0.691
0.799
0.569
0.428
0.482
0.841
0.349
0.743
0.906
0.255
0.321
0.461
0.236
0.860
0.007
0.026
0.609
0.133
0.236
0.283
0.062
0.538
0.164
0.370
0.828
0.343
0.148
0.560
0.141
0.037
0.036
0.416
0.789
0.076
0.789
0.044
0.179

117


Table S 3.5. Conditional average model output for general additive model (GAM) of individual variation in δ15N from Eurasian otter (*Lutra lutra*) whisker samples taken across England and Wales 2007-2016. Individual variation was defined by the standard deviation δ15N for each vibrissae sample. Averaged models were created from models with Akaike’s Information Criterion (AIC) values of less than two. Fixed variables preceded by an ‘s’ or ‘te’ represent a variable with a smoothing function with the number following representing the knot number. Fixed variables labelled ‘small’, ‘medium’, ‘large’ represent otter size. Effect size (estimate), standard error (std. error), adjusted standard error (adjusted SE), z value and p value (Pr(>|z|)) are given for each fixed variable in the final averaged model.

| Variable                              | Relative Importance | Level                              | Estimate  | Std. Error | Adjusted SE | z value | Pr(>|z|) |
|---------------------------------------|--------------------|------------------------------------|-----------|------------|-------------|---------|---------|
| Longitude                             | 0.78               |                                    | -0.02     | 0.17       | 0.17        | 0.09    | 0.929   |
| Size                                  | 0.97               | Medium (compared to large)         | -0.28     | 0.12       | 0.12        | 2.21    | 0.027   |
|                                       |                    | Small (compared to large)          | -0.33     | 0.15       | 0.15        | 2.20    | 0.028   |
|                                       |                    | Small (compared to medium)         | -0.05     | 0.13       | 0.13        | 0.37    | 0.708   |
| Distance from the coast (Km)          | 0.7                | -                                  | 0.00      | 0.00       | 0.00        | 1.56    | 0.119   |
| Scored Body Condition                 | 0.35               | -                                  | 0.08      | 0.06       | 0.06        | 1.28    | 0.200   |
| Latitude:Longitude                    | 0.45               | te(Latitude,Longitude).1           | -0.09     | 0.56       | 0.56        | 0.17    | 0.869   |
|                                       |                    | te(Latitude,Longitude).2           | 0.40      | 0.67       | 0.67        | 0.59    | 0.554   |
|                                       |                    | te(Latitude,Longitude).3           | 0.78      | 1.17       | 1.18        | 0.67    | 0.506   |
|                                       |                    | te(Latitude,Longitude).4           | 2.01      | 2.87       | 2.88        | 0.70    | 0.485   |
|                                       |                    | te(Latitude,Longitude).5           | -0.13     | 0.94       | 0.94        | 0.14    | 0.892   |
|                                       |                    | te(Latitude,Longitude).6           | 0.37      | 0.24       | 0.24        | 1.51    | 0.132   |
|                                       |                    | te(Latitude,Longitude).7           | -0.02     | 0.26       | 0.26        | 0.06    | 0.952   |
|                                       |                    | te(Latitude,Longitude).8           | -0.15     | 0.49       | 0.50        | 0.31    | 0.755   |
|                                       |                    | te(Latitude,Longitude).9           | 0.99      | 1.04       | 1.05        | 0.95    | 0.343   |
|                                       |                    | te(Latitude,Longitude).10          | 0.12      | 0.86       | 0.86        | 0.14    | 0.890   |
|                                       |                    | te(Latitude,Longitude).11          | 0.38      | 0.27       | 0.27        | 1.41    | 0.160   |
|                                       |                    | te(Latitude,Longitude).12          | 0.83      | 0.32       | 0.32        | 2.58    | 0.010   |
|                                       |                    | te(Latitude,Longitude).13          | -0.29     | 0.54       | 0.54        | 0.54    | 0.588   |
|                                       |                    | te(Latitude,Longitude).14          | -0.15     | 0.90       | 0.91        | 0.17    | 0.868   |
|                                       |                    | te(Latitude,Longitude).15          | 0.99      | 1.07       | 1.08        | 0.92    | 0.358   |
|                                       |                    | te(Latitude,Longitude).16          | -0.04     | 0.26       | 0.26        | 0.17    | 0.869   |
|                                       |                    | te(Latitude,Longitude).17          | 0.12      | 0.24       | 0.24        | 0.49    | 0.627   |
|                                       |                    | te(Latitude,Longitude).18          | -0.10     | 0.54       | 0.54        | 0.18    | 0.857   |
|                                       |                    | te(Latitude,Longitude).19          | -1.46     | 0.88       | 0.88        | 1.66    | 0.096   |
|                                       |                    | te(Latitude,Longitude).20          | 1.76      | 3.66       | 3.67        | 0.48    | 0.632   |
|                                       |                    | te(Latitude,Longitude).21          | 0.51      | 0.74       | 0.74        | 0.70    | 0.487   |
|                                       |                    | te(Latitude,Longitude).22          | 0.47      | 0.42       | 0.43        | 1.10    | 0.272   |
|                                       |                    | te(Latitude,Longitude).23          | -0.86     | 0.70       | 0.71        | 1.22    | 0.223   |
|                                       |                    | te(Latitude,Longitude).24          | -5.88     | 3.80       | 3.81        | 1.54    | 0.123   |
| Latitude                              | 0.25               | -                                  | 0.03      | 0.03       | 0.03        | 1.05    | 0.296   |
| Sex                                   | 0.24               | Male (compared to female)          | -0.09     | 0.11       | 0.11        | 0.87    | 0.387   |
| Proportion of Urban land use           | 0.09               | -                                  | 0.01      | 0.01       | 0.01        | 0.56    | 0.575   |
| Year                                  | 0.03               | -                                  | -0.01     | 0.02       | 0.02        | 0.37    | 0.714   |
Table S 3.6. Conditional average model output for general additive model (GAMM) of individual variation in δ¹³C from Eurasian otter (*Lutra lutra*) whisker samples taken across England and Wales 2007-2016. Individual variation was defined by the standard deviation in δ¹³C for each vibrissae sample. Averaged models were created from models with Akaike’s Information Criterion (AIC) values of less than two. Fixed variables preceded by an ‘s’ or ‘te’ represent a variable with a smoothing function with the number following representing the knot number. Effect size (estimate), standard error (std. error), adjusted standard error (adjusted SE), z value and p value (Pr(>|z|)) are given for each fixed variable in the final averaged model.

| Variable                      | Relative Importance | Level                      | Estimate | Std. Error | Adjusted SE | z value | Pr(>|z|) |
|-------------------------------|---------------------|----------------------------|----------|------------|-------------|---------|---------|
| Longitude                     | 0.5                 | -                          | -0.03    | 0.21       | 0.21        | 0.14    | 0.888   |
| Scored Body Condition         | 0.78                | -                          | -0.10    | 0.06       | 0.06        | 1.62    | 0.104   |
| Sex                           | 1                   | Male (compared to female)  | -0.27    | 0.11       | 0.11        | 2.49    | 0.013   |
| Latitude: Longitude           | 1                   | te(Latitude,Longitude).1   | -0.21    | 0.47       | 0.47        | 0.45    | 0.653   |
|                               |                     | te(Latitude,Longitude).2   | -0.01    | 0.62       | 0.62        | 0.02    | 0.982   |
|                               |                     | te(Latitude,Longitude).3   | 0.10     | 1.05       | 1.06        | 0.09    | 0.927   |
|                               |                     | te(Latitude,Longitude).4   | 0.83     | 2.50       | 2.51        | 0.33    | 0.741   |
|                               |                     | te(Latitude,Longitude).5   | -0.62    | 0.78       | 0.79        | 0.79    | 0.430   |
|                               |                     | te(Latitude,Longitude).6   | -0.16    | 0.20       | 0.20        | 0.79    | 0.427   |
|                               |                     | te(Latitude,Longitude).7   | 0.23     | 0.24       | 0.24        | 0.96    | 0.336   |
|                               |                     | te(Latitude,Longitude).8   | -0.41    | 0.44       | 0.44        | 0.94    | 0.350   |
|                               |                     | te(Latitude,Longitude).9   | 0.46     | 0.91       | 0.92        | 0.50    | 0.617   |
|                               |                     | te(Latitude,Longitude).10  | -0.31    | 0.68       | 0.68        | 0.46    | 0.647   |
|                               |                     | te(Latitude,Longitude).11  | 0.03     | 0.24       | 0.24        | 0.11    | 0.910   |
|                               |                     | te(Latitude,Longitude).12  | 0.66     | 0.28       | 0.28        | 2.34    | 0.019   |
|                               |                     | te(Latitude,Longitude).13  | 0.83     | 0.46       | 0.46        | 1.79    | 0.074   |
|                               |                     | te(Latitude,Longitude).14  | -0.26    | 0.75       | 0.75        | 0.35    | 0.724   |
|                               |                     | te(Latitude,Longitude).15  | 0.74     | 0.98       | 0.98        | 0.75    | 0.451   |
|                               |                     | te(Latitude,Longitude).16  | -0.11    | 0.24       | 0.24        | 0.44    | 0.659   |
|                               |                     | te(Latitude,Longitude).17  | 0.03     | 0.20       | 0.20        | 0.15    | 0.882   |
|                               |                     | te(Latitude,Longitude).18  | 0.32     | 0.44       | 0.44        | 0.72    | 0.473   |
|                               |                     | te(Latitude,Longitude).19  | -0.57    | 0.78       | 0.78        | 0.73    | 0.469   |
|                               |                     | te(Latitude,Longitude).20  | 2.51     | 3.16       | 3.17        | 0.79    | 0.430   |
|                               |                     | te(Latitude,Longitude).21  | 0.47     | 0.69       | 0.70        | 0.67    | 0.501   |
|                               |                     | te(Latitude,Longitude).22  | -0.10    | 0.38       | 0.38        | 0.26    | 0.796   |
|                               |                     | te(Latitude,Longitude).23  | -0.39    | 0.62       | 0.62        | 0.62    | 0.533   |
|                               |                     | te(Latitude,Longitude).24  | -1.17    | 3.46       | 3.48        | 0.34    | 0.737   |
| Latitude                      | 0.5                 | -                          | 0.03     | 0.01       | 0.01        | 4.27    | < 0.001 |
| Year                          | 0.17                | Year                       | -0.02    | 0.02       | 0.02        | 0.97    | 0.332   |
| Distance from the coast (Km)  | 0.15                | Distance from the coast (Km)| 0.00   | 0.00       | 0.00        | 0.81    | 0.419   |
| Proportion of Urban land use  | 0.15                | Urban Land Use             | 0.01     | 0.01       | 0.01        | 0.73    | 0.465   |
3.7.4 Additional figures representing stable isotope variation in whisker segments of Eurasian otters (Lutra lutra)

**Figure S 3.2.** Bayesian standard ellipse area (SEA) sizes and their credible intervals for isotopic signatures obtained from Eurasian otters (*Lutra lutra*) across England and Wales from 2007 to 2016 using multiple whisker subsamples and basal whisker subsamples only. Black circles represent the SEA mode, red crosses represent the sample size corrected SEA (SEAc) and boxes indicate 50%, 75% and 95% credible intervals from inner to outer.
Figure S 3.3. Spatial variation in $\delta^{15}$N (a) and $\delta^{13}$C (b) isotopic signatures from basal whisker subsamples of Eurasian otters (*Lutra lutra*) sampled across Wales and bordering regions from 1993 – 2007. Landscape scale variation was interpolated from predicted isotopic signatures obtained averaged statistical models. Lighter colours represent more enrichment in the respective isotopes and orange circles show the locations of otters.
Figure S 3.4. The mean predicted variation in δ¹⁵N (a) and δ¹³C (b) isotopic signatures (i.e. how confident we can be in predictions) of Eurasian otters (*Lutra lutra*) sampled between 2007 and 2016. Darker colours represent greater confidence in the predicted values. Orange dots are locations of individual otters.
Figure S 3.5. Variation in $\delta^{15}$N of Eurasian otters (*Lutra lutra*) utilising different aquatic habitat types. Symbols show the predicted mean (red = female, blue = male) plus or minus standard errors. Isotopic data were acquired from basal whisker subsamples of otters across England and Wales from 2007 to 2016.
Figure S 3.6. Variation in $\delta^{13}C$ of Eurasian otters (*Lutra lutra*) utilising different aquatic habitat types. Isotopic data were acquired from basal whisker subsamples of otters across England and Wales from 2007 to 2016.
Figure S 3.7. Variation in $\delta^{13}C$ of Eurasian otters (*Lutra lutra*) in each size class. Symbols show the predicted mean (red = female, blue = male) plus or minus standard errors. Isotopic data were acquired from basal whisker subsamples obtained from otters across England and Wales from 2007 to 2016.
Figure S 3.8. Variation in $\delta^{15}$N of Eurasian otters ($Lutra lutra$) in each size class. Isotopic data were acquired from basal whisker subsamples obtained from otters across England and Wales from 2007 to 2016.
Figure S 3.9. Within-individual variation (standard deviation) in δ¹⁵N of Eurasian otters (Lutra lutra) in each size class. Isotopic data were acquired from multiple whisker subsamples obtained from otters across England and Wales from 2007 to 2016.
Chapter Four – True or false: identifying artefacts in metabarcoding data and methods to reduce their prevalence

4.1 Abstract
Metabarcoding provides a powerful tool for investigating biodiversity and trophic interactions, but the high sensitivity of this methodology makes it vulnerable to errors, resulting in artefacts in the final data. More recently, metabarcoding studies have been utilising minimum sequence copy thresholds (MSCTs) to remove artefacts that remain in datasets; however, there is no consensus on best practice for the use of MSCTs. To mitigate erroneous reporting of results and inconsistencies, this study discusses and provides guidance for best-practice filtering of metabarcoding data for the ascertainment of conservative and accurate data. The most common MSCTs identified in the literature were applied to example datasets of Eurasian otter (*Lutra lutra*) dietary samples sequenced using different barcoding regions. Changing both the method and threshold value considerably impacted the outcome of datasets. Of the MSCTs tested, it was concluded that the optimal method combined a sample-based threshold with removal of maximum taxon contamination, providing stringent filtering of artefacts whilst retaining target data. Choice of threshold value differed between datasets due to variation in artefact abundance and sequencing depth, thus studies should employ controls (mock communities, negatives and unused MID-tag combinations) to select appropriate threshold values for each individual study.

*Keywords*: Artefacts, contamination, eDNA, environmental DNA, faecal analysis, metabarcoding

4.2 Introduction
Metabarcoding provides a powerful tool for ecological studies of biodiversity and trophic interactions (Deiner *et al.* 2017; Taberlet *et al.* 2018). By combining high throughput sequencing (HTS) with DNA barcoding, large volumes of high-resolution data can be generated from many samples simultaneously (Taberlet *et al.* 2018). As an accurate means of detecting and identifying not just common species, but also cryptic and rare species, metabarcoding has in many cases superseded traditional methods such as morphological analysis of prey remains in gut contents and faeces, and direct observation (Bowser *et al.* 2013; Roslin and Majaneva 2016; Elbrecht, Vamos, *et al.*
The high sensitivity of metabarcoding does, however, render it vulnerable to errors (Alberdi et al. 2018; Jusino et al. 2019), with the practices employed for processing and analysing samples varying from study to study. Differences in how samples are treated can lead to different results and conclusions being made from the same samples (Alberdi et al. 2018; Alberdi et al. 2019). Better guidelines for how best to process samples and analyse the data are thus needed as the use of metabarcoding increases.

Errors can lead to artefacts (i.e. false positives) being introduced at each stage of the metabarcoding process, from sample collection through to bioinformatic analysis (Alberdi et al. 2019; Jusino et al. 2019). These can occur through contamination from environmental or lab sources (Leonard et al. 2007; Siddall et al. 2009; Czurda et al. 2016), tag-jumping and sample mis-assignment (transfer of sample-specific tags between samples; Schnell et al. 2015) or PCR and sequencing errors (chimeras or mis-identified sequences; Shin et al. 2014; Bjørnsgaard Aas et al. 2017). Artefacts may also be produced through inconsistencies in reference databases (such as GenBank and BOLD; Valentini et al. 2009), resulting in sequences being assigned to the wrong taxon (Keskin et al. 2016; Rulik et al. 2017; Taberlet et al. 2018). Many of these artefacts can be limited through careful sample handling and laboratory procedures (e.g. pre- and post-PCR workstations; King et al. 2008; Murray et al. 2015) or the use of bioinformatics software to detect and remove erroneous sequences (e.g. UNOISE; Edgar 2016). However, it is likely that some artefacts will remain in the data regardless of precautionary steps taken (Nakagawa et al. 2018; Weyrich et al. 2019), thus inflating species richness within samples (Schnell et al. 2015; Clare et al. 2016; Zinger et al. 2019) and leading to an incorrect interpretation of the data.

One of the methods commonly applied to reduce the prevalence of artefacts is to use minimum sequence copy thresholds (MSCTs; e.g. Hänfling et al. 2016). Application of MSCTs remains ambiguous and non-standardised though, with many studies employing entirely distinct methodologies and thresholds (e.g. Gebremedhin et al. 2016; Guardiola et al. 2016; McInnes et al. 2017). A review of the relevant literature - 154 papers conducting metabarcoding on eukaryotic DNA for environmental monitoring or dietary analysis (Table S 4.1) - revealed a large proportion of studies do not employ MSCTs (29%) or only use basic methods (e.g. 32% removed sequences with low total read counts). Many studies utilising MSCTs employ one threshold across the dataset (18%), generally discarding sequences with an abundance of less than 10 reads per sample (10%) in order to remove PCR errors, which usually occur in low frequencies (Brown et al. 2015; Leray and Knowlton 2017). This is a largely arbitrary value though,
as increases in sequencing depth mean erroneous sequences can occur in abundances greater than 10 (De Barba et al. 2014; Elbrecht and Leese 2015). To circumvent this issue, relative thresholds may be applied to remove all low abundance sequences, regardless of sequencing depth (Elbrecht and Leese 2017; Zinger et al. 2019). Such thresholds are being utilised more frequently and may be based upon total read abundance (2%), read abundance per sample (18%) or read abundance per taxon (9%). Each of these methods is likely to remove artefacts to a different extent, introducing large differences in final datasets as a consequence.

The choice of threshold must be carefully considered because it can considerably impact the data; low thresholds will be unsuccessful at removing artefacts, whereas high thresholds may remove too much data and result in false negatives (Hänfling et al. 2016). This is especially true for dietary studies contending with amplification of consumer DNA, as general primers that amplify the consumer will result in a lower proportion of each sample being assigned to food item DNA, whereas specific primers that avoid amplifying the consumer may reduce amplification of some food items over others due to primer bias (Piñol et al. 2014). This variation increases the risk of target sequences being excluded if inappropriate filtering thresholds are selected. In such cases, experimental controls are valuable components for empirically assessing thresholds, providing an estimate for the proportion of artefacts within a dataset (Taberlet et al. 2018; Alberdi et al. 2019). Mock communities - mixtures of DNA at known concentrations from a selection of different species - are especially useful as metabarcoding positive controls due to their comparability to eDNA samples (Taberlet et al. 2018). Ideally, these are composed of DNA from species that do not occur in the study system, allowing errors from cross-contamination or tag-jumping between mock communities and eDNA to be identified more easily (e.g. De Barba et al. 2014; Hänfling et al. 2016). Mock communities can also show when a threshold has been set too high, removing target sequences from mock communities and therefore potentially removing target sequences from eDNA samples (e.g. De Barba et al. 2014; Hänfling et al. 2016).

Despite many eDNA studies utilising MSCTs, there is as yet no consensus on whether they are necessary, and if so, how best to employ them. In this chapter, we trial the most commonly used practices for removing artefacts from eDNA metabarcoding using an example dataset of Eurasian otter (Lutra lutra) dietary DNA extracted from faecal samples. Samples were processed alongside experimental controls, allowing the practicality of controls for selecting filtering thresholds to be assessed. Using these data, we aim to illustrate the optimal solutions for different data types and provide a
basis for the standardisation of this process in metabarcoding. We hypothesised that; (i) data with MSCTs applied would still contain artefacts; (ii) artefact removal would differ depending upon the method of MSCT application, with MSCTs removing artefacts from different sources (e.g. artefacts in blanks vs. those in mock communities), producing different results; (iii) using different MSCTs in combination would remove more artefacts than MSCTs applied on their own; (iv) low filtering thresholds would fail to remove many artefacts; (v) high thresholds would remove too much data, resulting in the loss of target sequences and hence trophic relationships; (vi) experimental controls would greatly benefit the choice of filtering method and threshold through identification of known target sequences and artefacts.

4.3 Methods

4.3.1 Otter faecal samples and data collection
Samples were obtained from the Cardiff University Otter Project collection, a biobank of tissues taken from dead otters during post-mortems, as described in Chapter three. Faecal samples were collected from the rectum during post-mortem examination, wrapped in foil and stored at -20 °C. One faecal sample per otter was thawed and transferred to a sterilised zip lock bag to be homogenised by hand before a sub-sample of ~200mg was taken and stored in a 2ml eppendorf tube at -80 °C.

4.3.2 DNA metabarcoding

4.3.2.1 DNA Extraction
DNA extraction was carried out using the QIAamp DNA mini stool kit following the manufacturer’s instructions (Qiagen, Hilden, Germany), with the exception that only half the recommended volume of buffers and InhibitEX tablets were used (Deagle et al. 2005). Faecal DNA extractions took place in a laminar flow hood to reduce the chances of contamination. DNA was also extracted from muscle tissue of otter and a range of potential prey species (Table S 4.2) using the DNeasy blood and tissue kit following the manufacturer’s instructions (Qiagen). Extraction negatives (containing no faeces or tissue) were included alongside each group of extractions to check for contamination (King et al. 2008). Following DNA extractions, DNA and extraction negatives were stored at -20 °C.

4.3.2.2 Primer selection and optimisation
We chose to use two primer sets from different barcoding regions in order to broaden the range of taxa amplified and overcome biases of each region (Deagle et al. 2014; Aizpurua et al. 2018; da Silva et al. 2019): the 16S barcoding region was selected to
amplify vertebrate DNA whilst COI was chosen to amplify invertebrate DNA. All primers considered underwent *in silico* testing using ecoPCR (Boyer *et al.* 2015) and were further tested in vitro. Temperature gradient polymerase chain reactions (PCRs) were performed to ascertain the optimal annealing temperature which would amplify most of the prey taxa but reduce the amount of predator DNA amplified. PCRs were run using 5 µl reaction volumes under the following conditions: 2.5 µl multiplex (Qiagen), 1.75 µl RNA/DNA free water, 0.05 µl BSA (0.05 ug/ml), 0.1 µl of each primer at 10 µm concentration and 0.5 µl of template DNA, with an initial denaturation at 95 °C for 15 minutes, followed by 35 cycles of 94 °C for 30 seconds, annealing temperature (40 °C – 60 °C, temperature gradient dependent) for 90 seconds, 72 °C for 90 seconds and then a final extension at 72 °C for 10 minutes before being held at 15 °C. Negative (RNA/DNA free water) and positive (predator and prey DNA) controls were included in each PCR. PCR products were run through a 2% agarose gel stained with SYBR®Safe (ThermoFisher Scientific, Paisley, UK) and if a band was visible then DNA was considered to have been successfully amplified.

A preliminary experiment (outlined in SI: 4.6.2) was carried out using a subset of the successful DNA extractions to test whether extraction and amplification methods were appropriate. Sequencing showed the extraction and amplification method was suitable, as well as confirming that the barcoding regions were successfully amplifying desired target species whilst reducing amplification of predator DNA. Following this experiment, further modifications were made to the primers of choice in order to increase taxonomic breadth. The 16S primer set that performed best was FN2199 (5’-yayaagcaggagaacctc-3’) and R8B7 (5’-ttatccctrggtarcthgg-3’) (modified from Deagle *et al.* 2009), which targeted a 225-267 bp amplicon (including primers). For COI, the primer set included Mod_mCOIintF (5’-ggwacwggwtgaacwgtwtaycc-3’) (modified from Leray *et al.* 2013) and HCO-2198 (5’-taaacttcagggtaccaaaaaatca-3’) (Folmer *et al.* 1994), which targeted a 365 bp amplicon (including primers). These primer sets were found to amplify desired taxa (Table S 4.2), with an optimal annealing temperature of 60 °C for 16S and 57 °C for COI. COI primers were also found to amplify a range of vertebrate taxa (as well as invertebrates) but did not cover the same range as the 16S primers, further justifying the use of both primer sets.

4.3.2.3 DNA Amplification

DNA extracted from faecal samples was tested for successful amplification using both primer sets, using the optimal PCR conditions identified during the optimisation process described above. Faecal samples that showed successful DNA amplification (visible band on agarose gel) underwent amplification using NGS-grade primers with
multiplex identifier sequences (or molecular identification tags; MIDs) to enable individual identification (of otters) from pooled data during bioinformatics analysis. A selection of 16 samples were amplified twice but with different MID-tag combinations in order to assess PCR errors, contamination and MID-tag primer biases during bioinformatics analysis.

MID-tagged primers were designed by adding a unique 10 bp sequence onto the 5'-end of each primer. By choosing 38 different 10 bp sequences (Eurofins, Luxembourg), we designed 19 forward primers and 19 reverse primers for each primer set, giving 361 unique combinations of forward and reverse primer pairs. MID-tagged primer combinations were tested in silico for primer-dimer and self-dimer using ThermoFisher Multiple Primer Analyser (Thermo Fisher Scientific 2020), only primer combinations with no primer-dimer or self-dimer were used for subsequent analyses. Temperature gradient PCRs were performed on a variety of MID-tagged primer combinations with a range of different annealing temperatures to decipher the optimal annealing temperature that would amplify prey DNA from tissue and within faecal samples, but reduce the amount of predator amplified, across a subset of the MID-tagged primers. The optimal annealing temperature was found to be 57 °C for 16S MID-tagged primers and 54 °C for COI. A 25µl reaction volume was used for MID-tag PCRs under the following conditions; 12.5 µl multiplex (Qiagen), 6.75 µl RNA/DNA free water, 0.25 µl BSA (0.05ug/ml), 0.5 µl of forward primer at 10 µm concentration, 2.5 µl of reverse primer at 2 µm and 2.5 µl of template DNA using the same PCR cycle as described above but with the optimal temperature for the MID-tagged primers used. Results from PCRs using MID-tagged primers were visualised using 3.5 µl of product on a QIAxcel (Qiagen), which gives concentrations of amplified DNA along with band presence and length of amplicon.

4.3.2.4 Mock Communities
To make mock communities suitable for this study, we used marine species that have not previously been detected in the diet of otters: 16S mock communities included pollock (*Pollachius pollachius*), blonde ray (*Raja brackyura*), smooth hound (*Mustelus asterias*), john dory (*Zeus faber*) and European sprat (*Sprattus sprattus*), whilst COI included smooth hound, john dory, common dragonet (*Callionymus lyra*), compass jellyfish (*Chrysaora hysoscella*) and sea sponge (*Suberites sp.*). First, DNA was extracted from tissue samples from a range of marine species following the protocol used for prey tissue samples. Extracted DNA was tested for successful amplification (following the optimal protocol identified during primer selection and optimisation) using both primer sets and species that were amplified by both were then used to create the
mock communities. DNA was quantified for each extraction of the chosen species using a Qubit (ThermoFisher) and diluted to 4 ng/µl. Extracted DNA was then combined into six mixtures; one mixture included equimolar concentrations of each species (5 µl of each species) while each of the other five were dominated by one species (added to the mixture at a greater volume [15 µl]) with the remaining four species equimolar. The equimolar mock community was split into three aliquots, whilst mock communities dominated by one species were split into two aliquots. Each mock community aliquot was then amplified using different MID-tag primers and quantified using a QIAxcel as described above.

4.3.2.5 High-Throughput Sequencing (HTS)
Samples successfully amplified during PCRs using MID-tagged primers (along with necessary negative controls and mock communities) were pooled to create equimolar mixtures. Mixtures consisted of 13 mock community samples per barcoding region, 16 negative controls per barcoding region, 268 faecal samples (plus 17 repeats) for 16S and 261 faecal samples (plus 16 repeats) for COI. Faecal samples and mock communities were pooled relative to their concentration depicted by the Qiaxcel, whilst negatives were pooled using the average concentration for reactions within their specific PCR run. Samples were first pooled by PCR run, giving six pools for 16S and seven for COI tagged samples. These pools were cleaned and concentrated using a left-side SPRI bead size selection, following the manufacturer’s protocol (Beckman Coulter, Brea, USA) with a SPRI bead ratio of 1.2x for 16S pools and 0.9x for COI pools (ratios determined by amplicon size). Pools were quantified using Qubit dsDNA high sensitivity assay kit (Thermo Fisher Scientific, Waltham, U.S.A), quality checked using TapeStation 2200 (Agilent, Santa Clara, USA) and combined to create one final equimolar mixture per barcoding region. Final pools underwent library preparation following the protocol outlined in the NEXTflex Rapid DNA-seq kit (Biox Scientific, Austin, USA); protocol steps included end repair and adenylation, ligation of Illumina adapters, bead size selection based on average amplicon length (insert peak size of 250bp for 16S and 365bp for COI) and PCR amplification (98 °C for two minutes, six cycles of 98 °C for 30 seconds, 65 °C for 30 seconds and 72 °C for 60 seconds, and then 72 °C for four minutes). Libraries then underwent high-throughput sequencing on an Illumina MiSeq at Cardiff University Genomics Hub. Sequencing two libraries of different amplicon sizes on the same sequencing run can introduce biases due to preferential amplification of smaller amplicons; because 16S (225 - 267 bp) and COI (365 bp) libraries differed in amplicon size, the two libraries were sequenced on separate runs using V2 sequencing chip with 2x250bp paired-end reads (expected capacity 12 - 15 million reads; Illumina 2020).
4.3.3 Sequence analysis

Bioinformatics analyses were carried out using a custom pipeline (scripts available in SI: 4.6.3). First, sequences were checked for truncation of MID-tags to assess whether an adequate number of sequences could be related back to a particular sample using the MID-tags. This was done by filtering sequence files (R1 and R2) from the Miseq into sequences with the forward primer and sequences with the reverse primer. Following this, sequences were separated into files for sequences which started with exactly 10bp before the primer and those that did not, allowing us to quantify how many sequences had a complete unique identifier at the beginning of their sequence. The highest proportion of truncated sequences was obtained from 16S R2 sequences with reverse primers, which showed nine percent of sequences had truncated MIDs and would not be assigned to a sample identifier. This was deemed acceptable and we continued with the pipeline using the original files.

FastP (Chen et al. 2018) was used to check the quality of reads, discard poor quality reads (<Q30, <125bp long or too many unqualified bases, denoted by ‘N’) and merge read pairs from Miseq files (R1 and R2). Merged reads were then assigned a sample ID based on the combination of MID-tags at the beginning and end of the sequence using the ‘trim.seqs’ function from Mothur v.1.39.5 (Schloss et al. 2009), this also removed the MID-tag and primer sequences from the reads. Using the files created by Mothur, reads were demultiplexed to obtain one file per sample ID. Read headers were modified for each file to include the sample ID and reads were then concatenated back into one file. This allowed each sequence to be identified to a particular sample simply by reading the sequence header. We then used the commands ‘fastx_uniques’, ‘unoise3’ and ‘otutab’ in Usearch (v. 11) to denoise (remove PCR and sequencing errors), cluster sequences into zero-radius operational taxonomic units (zOTUs) and create an OTU table (Edgar 2016; Edgar 2020). Taxonomic assignment for each zOTU was obtained using the ‘blastn’ command in BLAST+, using a threshold of 97% similarity and e-value of 0.00001, against a downloaded database of DNA barcoding sequences submitted to online databases (e.g. Genbank; National Center for Biotechnology Information 2008; Camacho et al. 2009).

Before assigning taxonomic identities to each zOTU, BLAST results were filtered using the ‘dplyr’ package in R [version 3.6.0] using R Studio [version 1.2.1335] (R Core Team 2019). This was used to retain only accession codes with the top BITscore for each zOTU. These data were then input into MEGAN [version 6.12.3] (Huson et al. 2007; Huson et al. 2016) to assign taxon names to each zOTU. As erroneous entries on
online databases can prevent species-level assignments, zOTUs for which the top BLAST hit was resolved higher than species-level were manually checked and assigned the appropriate taxon. Taxonomic identity for each zOTU was added to the OTU table produced by Usearch and reads were aggregated by taxonomic identity for each sample in R using the ‘aggregate’ function with a sum base function. Sequences were identified to taxon rather than using OTU to overcome the issue of OTUs over-splitting taxonomic groups and allow for ecological interpretations to be made (such as identifying marine taxa in non-coastal otters). Using taxon identifications also allowed artefacts and contaminants to be more clearly identified during clean-up protocols.

4.3.4 Minimum Sequence Copy Thresholds (MSCTs)

A preliminary analysis found that applying MSCTs to total read counts for a sequence resulted in low abundance reads persisting in the data once the read counts for sequences were split across samples. We therefore concluded this was an ineffective filtering method on its own and instead focussed on filters applied to read counts assigned to a sample and taxon. We tested seven MSCT (Table 4.1) methodologies and their effects on both the 16S and COI datasets. Methods were enacted in excel using ‘IF’ formulae. The methods employed were the most common identified from the literature (Table S 4.1):
Table 4.1. Seven methods often applied to eDNA metabarcoding datasets, selected from those identified in the literature (Table S 4.1). The ‘method name’, herein used to refer to these methods, is given alongside the description (how the methods are executed) and the aim of each. The range of thresholds tested were chosen based upon artefacts identified in control samples, starting with a low threshold and increasing the value until most of the identifiable artefacts were removed.

<table>
<thead>
<tr>
<th>Method Name</th>
<th>Method Description</th>
<th>Method Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. No filter</td>
<td>No taxon or sample filtering.</td>
<td>No clean-up/maximum preservation of data.</td>
</tr>
<tr>
<td>2. Singletons</td>
<td>Remove any read counts of one.</td>
<td>Remove extremely low frequency artefacts (e.g. sequencing artefacts).</td>
</tr>
<tr>
<td>3. &lt; 10</td>
<td>Remove any read counts of less than 10.</td>
<td>Remove low frequency artefacts (e.g. sequencing artefacts, PCR contamination)</td>
</tr>
<tr>
<td>4. Max contamination</td>
<td>Remove any read counts within each taxon that are lower than the highest read count within a blank (negative control or unused MID-tag combination) for that taxon.</td>
<td>Remove contamination detected by the blanks (e.g. extraction/PCR contamination, tag-jumping).</td>
</tr>
<tr>
<td>5. Total %</td>
<td>Remove read counts less than a proportion of the total read count for the dataset read count. Thresholds tested ranged from 0.001% to 0.02% for 16S and 0.0003% to 0.003% for COI.</td>
<td>Remove low frequency artefacts (e.g. sequencing artefacts, PCR contamination)</td>
</tr>
<tr>
<td>6. Sample %</td>
<td>Remove any read counts within a sample that are less than a proportion of the total read count for that sample. Thresholds tested ranged from 0.3% to 8% for 16S and 0.05% to 0.5% for COI.</td>
<td>Remove sample contamination (e.g. environmental, extraction or PCR contamination)</td>
</tr>
<tr>
<td>7. Taxon %</td>
<td>Remove read counts with an abundance less than a proportion of the total read count for that OTU. Thresholds tested ranged from 0.1% to 3% for both 16S and COI.</td>
<td>Remove cross contamination (e.g. cross contamination, tag-jumping)</td>
</tr>
</tbody>
</table>
If the read count (i.e. number of reads per sample per taxon) did not pass the designated threshold then the read count was set to zero, rather than subtracting the threshold from all read counts, to prevent changing the read counts of the remaining reads. For each of the proportion methods (5-7, Table 4.1), a variety of thresholds were tested (Table 1) to explore how choice of threshold can affect data output. The range of thresholds tested were chosen based upon artefacts identified in control samples; we started with a low threshold and increased the value until most of the identifiable artefacts were removed. We also explored the effectiveness of using different MSCTs in combination; this involved combining ‘Max Contamination’ with each proportional threshold method (5-7), and ‘Sample %’ with ‘Taxon %’.

Basic statistics were calculated in order to assess the effectiveness of each filtering method; total read count was used to assess loss of reads across the whole dataset, the presence of singleton reads was used to assess removal of PCR and sequencing errors, reads in blanks (negative controls and unused MID-tags) were used to assess levels of contamination and tag-jumping, and mock communities were used to assess presence of false positives within samples. Artefacts could also be identified through taxa incorrectly occurring in samples, such as taxa from eDNA samples in control samples, marine taxa in otters that did not have access to marine habitats, and mock community taxa in negative controls, unused MID tags or faecal samples.

To visualise the results of each filtering method, tables of reads were converted into heat charts using ‘ggplot2’ (Wickham 2016) in R and R Studio (scripts available in SI: 4.6.4). Frequency of occurrence for each taxon across all MID-tag combinations was also calculated for each filtering method and used to create heat charts. Relative frequencies were calculated by dividing frequency of occurrence by the total number of MID-tag combinations (361 for both barcoding regions); these values then underwent non-metric multidimensional scaling (NMDS) to visualise dissimilarity between the taxa present following application of each MSCT. This was conducted using the ‘metaMDS’ function in the ‘vegan’ package (Oksanen et al. 2013) with two dimensions (stress <0.1) and a Bray-Curtis dissimilarity calculation (Bray and Curtis 1957; scripts available in Supplementary Information 3). Ellipses were created using the ‘ordiellipse’ function with the default ‘sd’ setting (standard deviation).

4.4 Results
Sequencing of the 16S library yielded 17.6 million paired-end reads, which decreased to 11.7 million following bioinformatic analysis. Sequencing of the COI library yielded
13.7 million paired end reads, which decreased to 7.9 million following bioinformatic analysis.

**4.4.1 No filter (‘No Filter’)**

The highest read counts and occurrence of artefacts were observed in data with no MSCT applied, both for COI and 16S. This was evident through the high numbers of reads in blanks (Tables 4.2 and 4.3; Figs. S 4.1 and S 4.2), false positives in mock communities (Figs. 4.1 and 4.2), mock community taxa present in blanks and eDNA samples, taxa from eDNA samples occurring in control samples, and marine taxa occurring in faecal samples from otters with no access to marine habitats (Tables 4.2 and 4.3; Figs. S 4.3 and S 4.4). Artefacts appeared to be much more prevalent for taxa with high read counts (e.g. mock community taxa, taxa commonly consumed by otters and taxa identified as otter). Many low abundance reads, including singletons, were also observed in the unfiltered data, and whilst this suggests that rare species will be detected, it also suggests that the data contains many sequencing errors.
Figure 4.1. Eurasian otter (*Lutra lutra*) diet 16S prey counts.

Number of presences of each taxon following application of each minimum sequence copy threshold (low count = yellow, high count = purple). ‘Low’, ‘Mid’ and ‘High’ depict the value utilised for proportional thresholds (‘Total %’, ‘Sample %’ and ‘Taxon %’), with ‘Opt + MC’ denoting the ‘optimal’ threshold combined with ‘Max Contamination’ methods.
Table 4.2. Performance of different minimum sequence copy thresholds on 16S data. ‘Low’, ‘Mid’ and ‘High’ depict the value utilised for proportional thresholds (‘Total %’, ‘Sample %’ and ‘Taxon %’), with ‘Opt + MC’ denoting the ‘optimal’ threshold combined with ‘Max Contamination’ methods. Expected presences of marine taxa (~) were defined by the number of Eurasian otters (*Lutra lutra*) displaying reads for each marine taxon that were located along the coast or near an estuary.

<table>
<thead>
<tr>
<th>Minimum Sequence Copy Threshold</th>
<th>Total</th>
<th>Singletons</th>
<th>Blanks</th>
<th>Mock Communities</th>
<th>Marine taxa presences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Summed read count</td>
<td>Taxa</td>
<td>Number of presences</td>
<td>Summed read count</td>
<td>Average read count</td>
</tr>
<tr>
<td>No filter</td>
<td>11723871</td>
<td>105</td>
<td>2767</td>
<td>117460</td>
<td>1864</td>
</tr>
<tr>
<td>Singletons</td>
<td>11721104</td>
<td>105</td>
<td>0</td>
<td>117032</td>
<td>1858</td>
</tr>
<tr>
<td>&lt; 10</td>
<td>11705943</td>
<td>99</td>
<td>0</td>
<td>114675</td>
<td>1820</td>
</tr>
<tr>
<td>Maximum Contamination</td>
<td>10938496</td>
<td>102</td>
<td>63</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Total %</td>
<td>11534535</td>
<td>71</td>
<td>0</td>
<td>96869</td>
<td>1538</td>
</tr>
<tr>
<td>Mid Total %</td>
<td>11349821</td>
<td>60</td>
<td>0</td>
<td>78023</td>
<td>1238</td>
</tr>
<tr>
<td>High Total %</td>
<td>10733900</td>
<td>46</td>
<td>0</td>
<td>35916</td>
<td>570</td>
</tr>
<tr>
<td>Opt Total % + MC</td>
<td>10874148</td>
<td>63</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Sample %</td>
<td>11559268</td>
<td>89</td>
<td>218</td>
<td>116737</td>
<td>1853</td>
</tr>
<tr>
<td>Mid Sample %</td>
<td>11478669</td>
<td>68</td>
<td>0</td>
<td>113804</td>
<td>1806</td>
</tr>
<tr>
<td>High Sample %</td>
<td>10631707</td>
<td>46</td>
<td>0</td>
<td>86797</td>
<td>1378</td>
</tr>
<tr>
<td>Opt Sample % + MC</td>
<td>10875890</td>
<td>65</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Taxon %</td>
<td>11031736</td>
<td>105</td>
<td>742</td>
<td>45985</td>
<td>730</td>
</tr>
<tr>
<td>Mid Taxon %</td>
<td>8669244</td>
<td>105</td>
<td>267</td>
<td>30812</td>
<td>489</td>
</tr>
<tr>
<td>High Taxon %</td>
<td>3660086</td>
<td>104</td>
<td>25</td>
<td>30645</td>
<td>486</td>
</tr>
<tr>
<td>Opt Taxon % + MC</td>
<td>8589029</td>
<td>102</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Sample % + Mid Taxon %</td>
<td>10187214</td>
<td>72</td>
<td>2</td>
<td>30851</td>
<td>490</td>
</tr>
<tr>
<td>Mid Sample % + Low Taxon %</td>
<td>10959369</td>
<td>68</td>
<td>0</td>
<td>44471</td>
<td>706</td>
</tr>
<tr>
<td>Mid Sample % + Mid Taxon %</td>
<td>10177475</td>
<td>67</td>
<td>0</td>
<td>30434</td>
<td>483</td>
</tr>
<tr>
<td>Mid Sample % + High Taxon %</td>
<td>8647191</td>
<td>67</td>
<td>0</td>
<td>29865</td>
<td>474</td>
</tr>
<tr>
<td>High Sample % + Mid Taxon %</td>
<td>10155032</td>
<td>60</td>
<td>0</td>
<td>29866</td>
<td>474</td>
</tr>
</tbody>
</table>
Figure 4.2. Eurasian otter (*Lutra lutra*) diet COI prey counts. Number of presences of each taxon following application of each minimum sequence copy threshold (low count = yellow, high count = purple). 'Low', 'Mid' and 'High' depict the value utilised for proportional thresholds (‘Total %’, ‘Sample %’ and ‘Taxon %’), with ‘Opt + MC’ denoting the ‘optimal’ threshold combined with ‘Max Contamination’ methods.
Table 4.3. Performance of different minimum sequence copy thresholds on COI data. ‘Low’, ‘Mid’ and ‘High’ depict the value utilised for proportional thresholds (‘Total %’, ‘Sample %’ and ‘Taxon %’), with ‘Opt + MC’ denoting the ‘optimal’ threshold combined with ‘Max Contamination’ methods. Expected presences of marine taxa (–) were defined by the number of Eurasian otters (*Lutra lutra*) displaying reads for each marine taxon that were located along the coast or near an estuary.

<table>
<thead>
<tr>
<th>Minimum Sequence Copy Threshold</th>
<th>Total</th>
<th>Singletons</th>
<th>Blanks</th>
<th>Mock Communities</th>
<th>Marine taxa presences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Summed read count</td>
<td>Taxa</td>
<td>Number of presences</td>
<td>Summed read count</td>
<td>Average read count</td>
</tr>
<tr>
<td>No filter</td>
<td>7893115</td>
<td>232</td>
<td>2060</td>
<td>216089</td>
<td>3044</td>
</tr>
<tr>
<td>Singletons</td>
<td>7891055</td>
<td>232</td>
<td>0</td>
<td>215573</td>
<td>3036</td>
</tr>
<tr>
<td>&lt; 10</td>
<td>7881635</td>
<td>225</td>
<td>0</td>
<td>213925</td>
<td>3013</td>
</tr>
<tr>
<td>Maximum Contamination</td>
<td>5309935</td>
<td>228</td>
<td>187</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Total %</td>
<td>7869536</td>
<td>204</td>
<td>0</td>
<td>211968</td>
<td>2985</td>
</tr>
<tr>
<td>Mid Total %</td>
<td>7840726</td>
<td>161</td>
<td>0</td>
<td>208134</td>
<td>2931</td>
</tr>
<tr>
<td>High Total %</td>
<td>7783200</td>
<td>95</td>
<td>0</td>
<td>203183</td>
<td>2862</td>
</tr>
<tr>
<td>Opt Total % + MC</td>
<td>5296463</td>
<td>160</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Sample %</td>
<td>7880755</td>
<td>222</td>
<td>287</td>
<td>215704</td>
<td>3038</td>
</tr>
<tr>
<td>Mid Sample %</td>
<td>7834579</td>
<td>147</td>
<td>8</td>
<td>214039</td>
<td>3015</td>
</tr>
<tr>
<td>High Sample %</td>
<td>7815331</td>
<td>119</td>
<td>0</td>
<td>213254</td>
<td>3004</td>
</tr>
<tr>
<td>Opt Sample % + MC</td>
<td>5290528</td>
<td>142</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Taxon %</td>
<td>7572455</td>
<td>232</td>
<td>799</td>
<td>70721</td>
<td>996</td>
</tr>
<tr>
<td>Mid Taxon %</td>
<td>1952228</td>
<td>232</td>
<td>263</td>
<td>1065</td>
<td>15</td>
</tr>
<tr>
<td>High Taxon %</td>
<td>1450771</td>
<td>230</td>
<td>26</td>
<td>901</td>
<td>13</td>
</tr>
<tr>
<td>Opt Taxon % + MC</td>
<td>1949510</td>
<td>223</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Low Sample % + Mid Taxon %</td>
<td>1948979</td>
<td>206</td>
<td>12</td>
<td>1013</td>
<td>14</td>
</tr>
<tr>
<td>Mid Sample % + Low Taxon %</td>
<td>3870133</td>
<td>147</td>
<td>1</td>
<td>1031</td>
<td>15</td>
</tr>
<tr>
<td>Mid Sample % + Mid Taxon %</td>
<td>1940988</td>
<td>146</td>
<td>1</td>
<td>929</td>
<td>13</td>
</tr>
<tr>
<td>Mid Sample % + High Taxon %</td>
<td>1650235</td>
<td>144</td>
<td>1</td>
<td>929</td>
<td>13</td>
</tr>
<tr>
<td>High Sample % + Mid Taxon %</td>
<td>1934367</td>
<td>117</td>
<td>0</td>
<td>889</td>
<td>13</td>
</tr>
</tbody>
</table>
4.4.2 Remove singleton reads (‘Singletons’)
Removing singleton reads resulted in data very similar to that of unfiltered data for both barcodes, with only a low number of artefacts removed (Figs. 4.1 and 4.2; Tables 4.2 and 4.3; Figs S 4.1 - 4.4).

4.4.3 Remove read counts less than 10 (‘< 10’)
Removing reads with an abundance less than 10 reduced the occurrence of artefacts in blanks, mock communities and the presence of mock community taxa in other samples. Artefacts persisted in all controls and samples though, producing data very similar to unfiltered data (Figs. 4.1 and 4.2; Tables 4.2 and 4.3; Figs. S 4.1 - 4.4).

4.4.4 Remove maximum taxon contamination (‘Max Contamination’)
Removing reads less than or equal to the maximum read count in blanks per taxon removed no reads from some taxa and high values from others (16S: minimum read removal = 0, maximum = 8757, average = 394; COI: minimum = 0, maximum 23413, average = 117). Taxa experiencing high levels of read removal tended to come from those with high total read counts. This cleared all blanks of reads (Tables 4.2 and 4.3; Figs. S 4.1 and S 4.2), all mock community taxa from eDNA samples and taxa with high read abundances in eDNA samples from controls (Figs. 4.1 and 4.2). False positives were still present in mock communities though (Figs. 4.1 and 4.2), as were singleton reads. This method also cleared out some marine taxa from inland otters but some still remained (Tables 4.2 and 4.3; Figs. S 4.3 and S 4.4).

4.4.5 Proportion of total read count (‘Total %’)
This method removed artefacts present in blanks (Tables 4.2 and 4.3; Figs. S 4.1 and 4.2), false positives in mock communities (Figs. 4.1 and 4.2) and marine taxa from inland otters (Tables 4.2 and 4.3; Figs. S 4.3 and S 4.4). Mock community taxa were cleared from blanks and eDNA samples to an extent, but some were still present even at high thresholds (Tables 4.2 and 4.3). Taxa from eDNA samples with high read abundances were not filtered very well though, with many occurring in controls even at high thresholds. The lowest thresholds tested only filtered out a proportion of the artefacts, whilst the highest thresholds filtered out all false positives within mock communities and almost all reads in blanks (Figs. 4.1 and 4.2; Tables 4.2 and 4.3), however, the latter also removed target reads shown by the loss of mock community taxa within mock communities. A lower threshold was therefore necessary to give a balance between false positives and false negatives; the optimal threshold for achieving such a balance was 0.003% for 16S and 0.0008% for COI, removing reads with abundances less than 352 for 16S and 79 for COI.
4.4.6 Proportion of read count per sample (‘Sample %’)

This method removed false positives from mock communities (Figs. 4.1 and 4.2) and marine taxa from inland otters (Tables 4.2 and 4.3; Figs. S.4.3 and S.4.4). Low abundance taxa (e.g. foreign taxa occurring through sequencing errors) were less prevalent (Figs. 4.1 and 4.2), as were singletons. Taxa with high read abundances (e.g. mock community taxa and common taxa in eDNA samples) and reads present in blanks were only filtered to an extent (Tables 4.2 and 4.3; Figs. S.4.1 and S.4.2), resulting in artefacts from both being prevalent in filtered data regardless of the threshold utilised. This method removed fewer reads from samples with low total read counts, therefore samples were more likely to still contain artefacts. The highest thresholds were needed to remove all false positives from mock communities (Figs. 4.1 and 4.2). A much higher threshold was required for 16S due to the greater relative read counts of taxa compared to COI. The high thresholds required to clear mock communities of false positives in both 16S and COI datasets also removed some target reads (shown by the loss of mock community taxa), thus lower thresholds were selected to give a balance between false positives and false negatives; the optimal threshold for achieving such a balance was 1% for 16S and 0.3% for COI, removing reads to a varying degree (16S: minimum read removal for a sample = 0, maximum = 8757, average = 394, COI: minimum = 0, maximum 23413, average = 117).

4.4.7 Proportion of read count per taxon (‘Taxon %’)

This method filtered out reads in blanks (Tables 4.2 and 4.3; Figs. S.4.1 and S.4.2), as well as artefacts from taxa with high read abundances, clearing most of these from the data sets. A large proportion of reads were removed using this method (Tables 4.2 and 4.3; Figs. S.4.1 and S.4.2), especially from taxa with high total read counts (Figs. 4.1 and 4.2). Taxa with low read counts had fewer reads removed, resulting in these taxa containing more artefacts, shown by the prevalence of singleton reads and taxa identified as PCR or sequencing errors (e.g. foreign taxa; Figs. 4.1 and 4.2). This method was not good at removing false positives within eDNA samples, with false positives prevalent in mock communities regardless of the threshold utilised, and marine taxa were only removed from inland otters when using a high threshold (Tables 4.2 and 4.3; Figs. S.4.3 and S.4.4). With low thresholds applied, many more artefacts were observed in blanks, but a threshold of 3% cleared most of these artefacts from the datasets. The highest thresholds removed a high proportion of reads, therefore lower thresholds were selected to give a balance between clearing out artefacts and not losing too many reads; the optimal threshold for achieving such a balance was 0.5% for 16S and 0.8% for COI, removing reads to a varying degree (16S: minimum
read removal for a taxon  = 0, maximum = 26039, average = 553, COI: minimum = 0, maximum 2040, average = 49).

4.4.8 Combining methods

Many of the thresholds tested for MSCTs based on a proportion of read counts (‘Total %’, ‘Sample %’ and ‘Taxon %’) did not clear all artefacts, particularly regarding clearance of blanks. Proportional methods were thus combined with ‘Max Contamination’ to overcome this issue. ‘Sample %’ thresholds were also combined with ‘Taxon %’ thresholds because of their complementary removal of artefacts. Combining methods removed more artefacts than just using one filter. ‘Total %’ thresholds or ‘Sample %’ thresholds combined with ‘Max Contamination’ left very few artefacts in the data. These methods were found to complement one another, with proportional thresholds clearing most false positives from mock communities (Tables 4.2 and 4.3; Figs. 4.1 and 4.2) and marine taxa from inland otters (Figs. S 4.3 and S 4.4), whilst the contamination threshold cleared reads in blanks and artefacts from taxa with high read counts (e.g. mock community taxa in non-mock community samples and eDNA taxa in controls; Tables 4.2 and 4.3; Figs. S 4.1 and S 4.2). These combinations also cleared singletons and taxa identified as PCR or sequencing errors (Tables 4.2 and 4.3; Figs. 4.1 and 4.2). Combining methods sometimes allowed lower thresholds to be used concurrently for optimal results, but in other cases did not change the thresholds required (16S: optimal ‘total %’ = 0.002%, optimal ‘sample %’ = 0.5%; COI: optimal ‘total %’ = 0.0008%, optimal ‘sample %’ = 0.2%).

‘Taxon %’ thresholds combined with ‘Max Contamination’ still contained many artefacts; all reads in blanks and singletons were removed, but false positives were still present in mock communities as were marine taxa in inland otters (although in lower abundances compared to either filter alone; Tables 4.2 and 4.3; Figs. S 4.3 and S 4.4). Combining ‘Taxon %’ thresholds with ‘Sample %’ thresholds removed more artefacts and performed similarly to MSCTs combining ‘Sample %’ thresholds with ‘Max Contamination’. Combining these methods cleared the majority of reads from blanks, all singleton reads, artefacts from taxa with high read counts and most false positives in mock communities (Tables 4.2 and 4.3; Figs. 4.1 and 4.2); however, it also removed a lot of reads across the dataset, there were still artefacts present in the controls and marine taxa were still present in inland otters. The optimal combination of thresholds changed between datasets (16S: ‘sample %’ = 0.5% with ‘taxon %’ = 0.3%; COI: ‘sample %’ = 0.2% with ‘taxon %’ = 0.3%). Lowering the sample threshold introduced more false positives into the data, whilst increasing the threshold removed target reads. Lowering the taxon threshold retained more reads in blanks and artefacts from
taxa with high read counts, whilst increasing the taxon threshold decreased the total read count, resulting in loss of target reads.

**Table 4.4.** Success of different filtering methods in achieving the key objectives of post-bioinformatic data clean-up. Green, orange and red denote positive, neutral and negative outcomes, respectively. ‘Low’, ‘Mid’ and ‘High’ depict the value utilised for proportional thresholds (‘Total %’, ‘Sample %’ and ‘Taxon %’), with ‘Opt + MC’ denoting the ‘optimal’ threshold combined with ‘Max Contamination’ (Max Contam) methods.
4.4.9 Repeats
Repeated eDNA samples and mock communities were inconsistent across filtered and unfiltered data. Taxa observed at a high read count in one repeat were more likely to also occur in the other repeat, however, taxa with low read abundances often only occurred in one of the two repeats.

4.4.10 NMDS analysis
Choice of MSCT method had major effects on the final composition of the data, as shown by the NMDS analysis (Fig. 4.3). Application of ‘No Filter’, ‘Singletons’ and ‘< 10’ MSCTs were found to be similar, with ‘< 10’ showing greater similarities to MSCTs based on ‘Total %’ and ‘Sample %’ in COI compared to 16S data. ‘Sample %’ and ‘Total %’ thresholds were most similar to each other and gave distinct results from those of taxon MSCTs (‘Taxon %’ and ‘Max Contamination’). By combining taxon MSCTs with either ‘Sample %’ or ‘Total %’ thresholds, an intermediate result was obtained. All combinations of taxon filters with ‘Sample %’ or ‘Total %’ performed similarly to one another in the COI data; however, in the 16S data those that combined ‘Sample %’ or ‘Total %’ with ‘Maximum Contamination’ were more dissimilar to taxon methods than combinations between ‘Sample %’ and ‘Taxon %’.
Figure 4.3. Non-metric multidimensional scaling of relative frequency of occurrence for each taxon following application of different minimum sequence copy thresholds, including different methods and thresholds, to 16S (a) and COI (b) data. Ellipse colours denote each method with No Filter, Singletons, <10 and Maximum Contamination not having ellipses given the lack of modifiable threshold.
4.5 Discussion

Here we have illustrated the efficacy of different filtering methods and thresholds for removal of artefacts from metabarcoding data. For optimisation of thresholds, previous studies have disproportionately emphasised the importance of mock communities (e.g. Elbrecht and Leese 2017; Jusino et al. 2019); however, since the biases affecting true unknown mixtures of eDNA are almost impossible to experimentally replicate (Alberdi et al. 2018), mock communities cannot solely be used to adequately filter data. By sequencing and analysing mock communities, blank samples and eDNA together, we were able to assess fully which filters and thresholds were optimal in cleaning metabarcoding data of this nature.

4.5.1 Identifying artefacts

Despite all appropriate precautionary steps being taken to reduce contamination (e.g. screening negative controls, pre- and post-PCR workstations), and bioinformatic programmes used to remove erroneous sequences, artefacts were still observed in the unfiltered data. Such contamination is, however, largely unavoidable when using a method so broad-spectrum and sensitive (Alberdi et al. 2018; Jusino et al. 2019). Artefacts primarily manifested as unexpected reads in control samples, but also as erroneous taxa and mis-assigned reads. Erroneous taxa, usually existing in low read counts in the unfiltered data (De Barba et al. 2014; Ficetola et al. 2015) are, in this case, taxa produced through PCR or sequencing errors that are ecologically highly unlikely to appear in their respective samples (e.g. foreign species), thus rendering them easy to identify and eliminate. Mis-assigned reads were more difficult to identify but a proportion were able to be detected through the presence of marine taxa in landlocked sites (Figs. S 4.3 and S 4.4). In such cases, reads were assumed to be derived from eDNA samples collected from otters inhabiting marine environments, producing artefacts during HTS in other eDNA samples through cross-contamination, tag-jumping or mis-assignment (Schnell et al. 2015; Alberdi et al. 2019). Mis-assigned reads from samples that could feasibly contain the same taxa (e.g. faecal samples from two freshwater otters) are likely to remain undetected. We also aimed to use repeated samples to detect artefacts, but inconsistencies in taxa present between replicates meant this was not possible, a similar finding to other metabarcoding studies utilising repeated samples (e.g. Alberdi et al. 2018; Rennstam Rubbmark et al. 2019; Zizka et al. 2019). In this study, replicates were amplified using different MID-tag combinations, therefore mismatches between MID-tags and target DNA may have affected primer binding efficiencies and resulted in differential amplification of taxa during PCRs (O’Donnell et al. 2016; Mata et al. 2019).
Theoretically, blanks should contain no reads, and mock communities should contain reads only from selected taxa, with these taxa only occurring within mock communities. In reality, controls are rarely found to be free of all contamination following metabarcoding. Detection of artefacts is, however, then facilitated through the presence of these unexpected reads. Such reads in negative controls may occur due to low levels of contamination (e.g. from reagents or samples; Leonard et al. 2007; Czurda et al. 2016; Alberdi et al. 2019) that went undetected during screening of samples and may be present throughout only a few, or potentially all samples. Reads present in blanks may also occur due to tag-jumping or mis-assignment (Schnell et al. 2015), which are primarily identifiable through unused MID-tag combinations. These artefacts are hard to detect without blanks because they are frequently assigned to taxa that legitimately occur in high read abundances across many samples (Jensen et al. 2015; Carew et al. 2018), such as mock community taxa and common taxa in eDNA samples (e.g. commonly consumed taxa or the consumer itself). Further artefacts were detected through the presence of mock community taxa in eDNA samples and common eDNA taxa in mock communities; these were concluded to be primarily due to tag-jumping or mis-assignment rather than sample cross-contamination because eDNA and mock community samples were processed separately. Unexpected reads in mock communities also allowed low abundance artefacts from contaminants and PCR or sequencing errors to be identified, which may also have occurred across the eDNA samples. Control samples showed that artefacts were prevalent throughout the unfiltered data, with those identified through blanks increasing the frequency of occurrence of taxa, those identified through mock communities inflating sample diversity and both contributing to higher total read counts and, ultimately, false positives.

The composition of the mock communities is of great importance to the process of identifying artefacts. If the mock communities are comprised of species that may feasibly occur in the eDNA samples taken from the focal study system, the utility of these controls would be reduced. Although our mock communities were composed of marine taxa unlikely to be consumed by otters, high read counts were observed in our COI mock communities for brill (*Scophthalmus rhombus*), a taxon known to be consumed by otters but not included in mock community mixtures. The marine samples from which DNA was extracted were collected as part of a larger marine surveying initiative and, whilst care was taken by the practitioners responsible for the collection, cross-contamination by other species is possible. Since this taxon could legitimately occur in mock communities and eDNA samples, false presences are harder to confirm,
but its marine origin meant that in areas lacking access to marine prey by otters, reads could still be identified as artefacts.

**4.5.2 Performance of Minimum Sequence Copy Thresholds (MSCTs)**

Artefacts were removed to varying extents depending upon the filtering method and threshold utilised. Basic MSCTs commonly utilised in the literature, such as removing singletons (e.g. Oliverio *et al.* 2018) or reads with an abundance less than 10 (e.g. Gebremedhin *et al.* 2016), removed very few artefacts. This was due to the high sequencing depth achieved in the study, increasing the chance of artefacts occurring in read abundances greater than 10 (De Barba *et al.* 2014; Elbrecht and Leese 2015). MSCTs removing reads with an abundance below a proportion of the total read count performed better, reducing the abundance of all detectable artefacts; however, applying one threshold across all read counts potentially indiscriminately removes target reads with low abundances and retains abundant artefacts. This bias can be overcome by using MSCTs based on sample read counts, as the read count will inevitably vary between samples despite best efforts to facilitate consistent sample read depths (Deagle *et al.* 2019). Such MSCTs perform similarly to total read thresholds but account for this variation in sequencing depth between samples. Sample MSCTs efficaciously removed artefacts from within samples, with lowered levels of cross-contamination (shown through removal of false positives of marine taxa) and erroneous taxa, but did not clear artefacts from blanks nor abundant taxa.

MSCTs that removed reads based on the amount present in blanks (‘Maximum Contamination’) or a proportion of each total taxon read count (‘Taxon %’) removed artefacts from blanks and abundant taxa, but not from mock communities or erroneous taxa (e.g. false marine taxa presences). Of these two methods, removal of maximum taxon contamination was more suitable as it removed all artefacts from negative controls and taxa with high read counts without removing too many reads overall. To achieve the same result using thresholds based on taxon read counts resulted in much greater read losses, increasing the likelihood of false negatives. Proportional taxon thresholds also showed a strong bias towards removing reads from abundant taxa, helping remove artefacts produced through tag-jumping, but would potentially produce false negatives if taxa legitimately occur in many samples; this was observed in our 16S data through the loss of many of predator reads within faecal samples. Comparing proportional taxon thresholds to others that cleared out similar amounts of artefacts revealed that proportional taxon thresholds produced the highest loss of reads, thus making this method more likely to lead to false negatives. Removal of maximum taxon contamination is logically superior given that the taxa for which the greatest number of
reads will be removed will be based on those that are verifiably contaminating the blanks. Care must, however, be taken to ensure that the protocols followed to produce the blanks are sufficiently stringent, but are not unnecessarily conservative (e.g. negative control volumes included being based on the average volume pooled per plate, vs. the maximum volume pooled per plate), since this will cause this filtering method to produce many false negatives through overly strict removal of data.

4.5.3 Combining MSCTs

Combining different MSCTs improved the performance of all filters, leading to a greater reduction in artefact presence. The weakest combination used proportional taxon thresholds with removal of maximum taxon contamination (‘Taxon %’ with ‘Max Contamination’); these analogous methods removed artefacts in similar ways (i.e. removal based on reads present across taxa, rather than across samples), with neither sufficiently mitigating artefacts within samples. Artefacts persisting in blanks following application of total read count thresholds were removed by combining this method with removal of maximum taxon contamination; however, this combination may introduce biases by not accounting for read depth variations between samples. Taxon-based thresholds were complementary to sample-based thresholds, with one removing artefacts identified through blanks and abundant taxa and the other removing artefacts within samples, including false marine presences and erroneous taxa. Combining sample-based thresholds with removal of taxon contamination performed better than combinations with proportional taxon thresholds, as a greater proportion of artefacts were removed with a lower total read loss, reducing the likelihood of false negatives. Due to its consistently improved performance over other MSCTs across both metabarcoding datasets, we can conclude that combining a sample-based threshold with removal of maximum taxon contamination is the optimal method for stringent filtering of metabarcoding data whilst retaining target data.

4.5.4 Choosing an appropriate threshold

In metabarcoding studies, removal of false positives tends to be prioritised over false negatives due to the assumption that reads prove taxon presence whilst a lack of reads does not prove absence because false negatives can occur due to experimental biases (e.g. sampling or primer bias; Oehm et al. 2011; Piñol et al. 2015). A trade-off exists whereby removal of false positives leads to an increase in false negatives (Zepeda-Mendoza et al. 2016; Alberdi et al. 2019), observed here when utilising high thresholds which removed many artefacts but also removed target reads, biasing results to taxa with high read abundance. By relaxing thresholds and allowing some artefacts to persist, a balance can be achieved where a high proportion of false positives are
removed whilst retaining only very few false negatives that are easily disregarded (Clare et al. 2016; Hänfling et al. 2016; Zizka et al. 2019), thus better reflecting the diversity within samples. The threshold at which this balance is achieved varies between studies depending upon the sequencing depth and breadth of taxa; in our study, 16S had a higher total read count, but fewer taxa as it only amplified vertebrates (whereas COI amplified vertebrates, invertebrates, fungi and bacteria). This resulted in the presence of artefacts in higher abundances, therefore requiring a higher threshold. Appropriate thresholds should be chosen based upon artefact removal from control samples. The aim of the study should, however, also be considered as many MSCTs may remove rare taxa along with artefacts due to their low read abundance (Clare et al. 2016; Hänfling et al. 2016; Zizka et al. 2019). Studies concerning commonly-detected taxa can employ more stringent filters that remove more artefacts at the expense of losing rare taxa that may not be of interest anyway, whereas studies concerning rare taxa should consider employing lower thresholds whilst accepting that artefacts will be prevalent.

In this study, we chose to assess the effectiveness of different thresholds using taxa read counts as well as occurrences (count data converted to presence or absence). Occurrence data is often assumed to be a conservative method of assessing metabarcoding data, as recovery biases (e.g. primer bias, starting amount of DNA) have a lower impact on such data (Deagle et al. 2019). Although occurrence data can inflate the importance of taxa that occur at low read counts (e.g. rare taxa or taxa consumed in small amounts; Deagle et al. 2019), and therefore also artefacts, we found it provided a simple and concise method for assessing artefact prevalence. Other methods, such as relative read abundance (RRA), may provide an alternative method for assessing abundance of artefacts and their impact on metabarcoding datasets by considering the proportion of reads each taxon contributes to a sample’s total read count (this is analogous to the ‘Sample %’ MSCT). However, conversion of reads to RRA can produce misleading results due to biases such as differential digestion rates or primer amplifications (Pompanon et al. 2012; Clare 2014; Piñol et al. 2014; Thomas et al. 2014; Elbrecht and Leese 2015; Elbrecht et al. 2017; Alberdi et al. 2018), whilst the loss of read count data can potentially obscure interpretations of overall data loss. For these reasons we chose not to convert read count data into RRA but instead use raw read counts to assess the use of different MSCTs, thus allowing both artefact abundance and overall loss of reads to be assessed and directly compared. Future developments may make RRA a useful tool for artefact detection and removal though, allowing identification of artefacts that are having a proportionally large impact on metabarcoding data.
4.5.5 Conclusions

Here we have shown artefacts persist in metabarcoding data even following stringent lab and bioinformatic procedures. Whilst this study concerns just two metabarcoding datasets, the findings are consistent with the existing literature, confirming a broader relevance of the suggested thresholds. Although artefacts often occur in low abundances, they can create a disproportionate representation of biodiversity and produce misleading results, highlighting the need for read count filters. MSCTs removed artefacts to differing extents, but combining sample-based thresholds with removal of maximum taxon contamination provided an optimal outcome. Whilst the optimal method was the same for both datasets, thresholds applied differed due to variation in sequencing depth and differential taxon amplification. The choice of threshold must thus depend on the individual study, taking into consideration the sequencing depth, breadth of taxa amplified, artefact abundance and the fundamental question under investigation. Control samples were crucial in assessing filters and selecting appropriate thresholds, providing a means for assessing removal of artefacts and target reads. We recommend that future metabarcoding studies include mock communities and blanks in every study, and, if possible, identify taxa detected within eDNA samples that can be used to identify artefacts in the resultant metabarcoding data (e.g. marine taxa in inland samples). Given the broad variation in MSCTs applied to metabarcoding studies, inconsistent results between studies are inevitable to mitigate erroneous reporting of results and inconsistencies, effective guidance for best-practice filtering of metabarcoding data for the ascertainment of conservative and accurate data should be followed.

4.6 Acknowledgements

This work was funded by KESS II (Knowledge Economy Skills Scholarship) and the Wildlife Trust of South and West Wales who partnered this project. Sample collection was conducted by Cardiff University Otter Project employees and placement students during post mortems of otter carcasses. Samples for primer testing were provided by Manuel Nicolaus from CEFAS (marine species), Shaun Leonard at Wild Trout Trust (freshwater fish), Francois Edward and Monika Juergens at CEH (freshwater invertebrates) and Derek Gow Consultancy (mammals). Thank you to Helen Hipperson and the NBAF team at University of Sheffield for their initial training in bioinformatic analyses. Thank you to Angela Marchbank of the Cardiff University Genomics Hub for training and assistance with high-throughput sequencing of samples.
## Supplementary Information

### 4.7.1 Literature review of the use of minimum sequence copy thresholds

Table S 4.1. Application of minimum sequence copy thresholds across 154 studies conducting metabarcoding on eukaryotic DNA for environmental monitoring or dietary analysis.

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Number of studies</th>
<th>Studies</th>
</tr>
</thead>
</table>
### Table S.4.1. (continued)

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Number of studies</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singletons</strong></td>
<td>After assigning reads to OTU and sample, remove any read counts with an abundance of one</td>
<td>9 (6%)</td>
<td>Albaina <em>et al.</em> 2016; Guardiola <em>et al.</em> 2016; Andruszkiewicz <em>et al.</em> 2017; Kaunisto <em>et al.</em> 2017; Vamos <em>et al.</em> 2017; Alberdi <em>et al.</em> 2018; Theissinger <em>et al.</em> 2018; Stat <em>et al.</em> 2019</td>
</tr>
<tr>
<td><strong>Sequences with low abundance across a dataset</strong></td>
<td>Following clustering of sequences based on similarities, remove any clusters that have a total less than a particular value. Values vary from 3 to 1000, with most (15/25) utilising 10 as the threshold.</td>
<td>25 (16%)</td>
<td>Hibert <em>et al.</em> 2013; Guardiola <em>et al.</em> 2015; Kartzinel <em>et al.</em> 2015; Miya <em>et al.</em> 2015; Murray <em>et al.</em> 2015; Hänfling <em>et al.</em> 2016; Shaw <em>et al.</em> 2016; Srivathsan <em>et al.</em> 2016; Valentini <em>et al.</em> 2016; Arulandhu <em>et al.</em> 2017; Emilson <em>et al.</em> 2017; Fletcher <em>et al.</em> 2017; Lobo <em>et al.</em> 2017; Pochon <em>et al.</em> 2017; Rocchi <em>et al.</em> 2017; Sato <em>et al.</em> 2017; Smart <em>et al.</em> 2017; Arrizabalaga-Escudero <em>et al.</em> 2018; Bylemans <em>et al.</em> 2018; Carew <em>et al.</em> 2018; Frontalini <em>et al.</em> 2018; Galan <em>et al.</em> 2018; Harper <em>et al.</em> 2018; Hawlitschek <em>et al.</em> 2018; Nakagawa <em>et al.</em> 2018; Toju and Baba 2018; Treonis <em>et al.</em> 2018; Wangensteen <em>et al.</em> 2018; Milhau <em>et al.</em> 2019; Nichols <em>et al.</em> 2019; Rytkönen <em>et al.</em> 2019</td>
</tr>
<tr>
<td></td>
<td>After assigning reads to OTU and sample, remove any read counts with an abundance less than a particular value. Values vary from three to 139, with most (16/27) utilising 10 as the threshold.</td>
<td>27 (18%)</td>
<td>Cowart <em>et al.</em> 2015; Civade <em>et al.</em> 2016; Galal-Khallaf <em>et al.</em> 2016; Gebremedhin <em>et al.</em> 2016; Li <em>et al.</em> 2016; Miller <em>et al.</em> 2016; Pomon <em>et al.</em> 2016; Lopes <em>et al.</em> 2017; Prosser and Hebert 2017; Rodgers <em>et al.</em> 2017; Alberdi <em>et al.</em> 2018; Ailsos <em>et al.</em> 2018; Anslan <em>et al.</em> 2018; Divoll <em>et al.</em> 2018; Dunn <em>et al.</em> 2018; Elbrecht <em>et al.</em> 2018; Greiman <em>et al.</em> 2018; Pont <em>et al.</em> 2018; Theissinger <em>et al.</em> 2018; Bessey <em>et al.</em> 2019; Kitson <em>et al.</em> 2019; Pont <em>et al.</em> 2019; Siegenthaler <em>et al.</em> 2019</td>
</tr>
<tr>
<td></td>
<td>Remove reads with an abundance less than a proportion of the total read count for all reads assigned to an OTU and sample</td>
<td>3 (2%)</td>
<td>Grealy <em>et al.</em> 2015; Klymus, Richter, <em>et al.</em> 2017; Braukmann <em>et al.</em> 2019</td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Number of studies</td>
<td>Studies</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Removing contamination</td>
<td>Remove highest read number within a negative for all samples in that OTU</td>
<td>10 (6%)</td>
<td>Guardiola et al. 2015; Andruszkiewicz et al. 2017; Bell, Loeffler, et al. 2017; Bell, Fowler, et al. 2017; Alsos et al. 2018; Dunn et al. 2018; Galan et al. 2018; Greiman et al. 2018; Grey et al. 2018; Bell et al. 2019</td>
</tr>
<tr>
<td>Proportion of OTU read count</td>
<td>Remove reads with an abundance less than a proportion of an OTUs total read count for all samples in an OTU</td>
<td>14 (9%)</td>
<td>Guardiola et al. 2015; Civade et al. 2016; Guardiola et al. 2016; Pornment et al. 2016; Valentini et al. 2016; Lopes et al. 2017; Galan et al. 2018; Pont et al. 2018; Wangensteen et al. 2018; Milhau et al. 2019; Pont et al. 2019</td>
</tr>
<tr>
<td>Samples with low read count</td>
<td>Remove samples with a low total read count</td>
<td>9 (6%)</td>
<td>Jarman et al. 2013; Kartzinel et al. 2015; Cordier et al. 2017; Krehenwinkel et al. 2017; Laroche et al. 2017; McInnes, Alderman, Lea, et al. 2017; McInnes, Alderman, Deagle, et al. 2017; Granquist et al. 2018; da Silva et al. 2019</td>
</tr>
<tr>
<td>Other OTU Method</td>
<td>Remove OTUs present in negative controls, Remove OTUs only present in one sample, Remove OTUs with an abundance below a certain value</td>
<td>11 (7%)</td>
<td>Albaina et al. 2016; Guardiola et al. 2016; Yoon et al. 2016; Andruszkiewicz et al. 2017; Yamamoto et al. 2017; Aizpurua et al. 2018; Deagle et al. 2018; Nakagawa et al. 2018; Wangensteen et al. 2018; Milhau et al. 2019</td>
</tr>
</tbody>
</table>
4.7.2 Preliminary metabarcoding experiment

4.7.2.1 DNA Extraction and Amplification

A preliminary experiment was performed on 60 faecal samples taken from otter carcasses collected in 2015 and 2016. Faecal samples underwent DNA extraction and PCR with non-MID-tagged primers as described in the main text, with the exception that recommended volumes of buffers and InhibitEX tablets were used. A different 16S primer set was used for this preliminary study which was a fish specific primer: FN1824 (5’ – agacccdtctagcctwag – 3’) and R5D4 (5’ – ttatcctrggtarctgyg – 3’) (modified from Deagle et al. 2009). Primers were tested on a range of species prior to use (Table S 4.2) MID-tagged primer PCRs were conducted the same way as described in the main text with the exception of 10ul reaction volumes were used; 5ul multiplex, 0.1ul BSA (0.05ug/ml), 0.2ul (10um) forward primer, 1ul (2um) reverse primer, 1.7ul RNase free water an 2ul DNA. MID-tag primer PCRs used an annealing temperature of 63 °C for 16S primers and 57 °C for COI primers.

4.7.2.2 High-Throughput Sequencing

Following MID-tag PCRs, samples were pooled based upon gel brightness; bright bands were considered 3x more concentrated than very faint bands with mid-brightness band in between the two concentrations, therefore 6ul of product from samples with faint bands (and negative controls) was added to the pool, 4 ul from mid-brightness and 2 ul from bright bands. Samples were first pooled by PCR run, giving two pools per barcoding region, following which each pool was cleaned and concentrated using a left-side SPRI bead size selection, following the manufacturer’s protocol with a SPRI bead ratio of 1.2x for both 16S and COI pools (Beckman Coulter, Brea, USA). Pools were quantified using Qubit dsDNA high sensitivity assay kit (Thermo Fisher Scientific, Waltham, U.S.A), quality checked using TapeStation 2200 (Agilent, Santa Clara, USA) and combined to create one equimolar mixture per barcoding region. These equimolar mixtures underwent library preparation following the protocol outlined in the NEXTflex Rapid DNA-seq kit (Bioo Scientific, Austin, USA), with different adapters used for each barcoding region, following which libraries were quantified using Qubit dsDNA high sensitivity assay kit. The 16S library was diluted to a similar concentration as COI and the two libraries were combined in order to run them both on the same sequencing run. As smaller amplicon libraries are preferentially sequenced, a greater volume of the the larger amplicon library is required when combining; 16S had a smaller amplicon (243 bp including primers) than COI (363 bp including primers), therefore the two libraries were combined using 4.4 µl of 16S and 5.6 µl of COI. Samples were then sequenced using a nano sequencing chip with 2x250
bp paired-end reads (expected capacity of one million reads; Illumina 2020) on an Illumina MiSeq at Cardiff University Genomics Hub.
### Table S 4.2. Taxa used for primer testing.

DNA was obtained from muscle tissue collected from identified samples of each of the taxa. The ‘COI’ column shows DNA amplified by the primer set Mod_mCOIintF (5’ - ggwacgwgwtacwgtwttcycc -3’) (modified from Leray et al. 2013) and HCO-2198 (5’- taactcagggtgacacaaaaatatca -3’) (Folmer et al. 1994). The ‘General vertebrate 16S’ column shows DNA amplified by the primer set FN2199 (5’- yayaagcagagaacct -3’) and R8B7 (5’- ttatccctrggtarctg -3’) (modified from Deagle et al. 2009). The ‘Fish specific 16S’ column shows DNA amplified by the primer set FN1824 (5’ – agacctdtragctwag – 3’) and R5D4 (5’ – ttatccctrggtarctgyggt – 3’) (modified from Deagle et al. 2009). Amplification of DNA was assessed through gel electrophoresis; ‘No’ states that not band was observed on the gel and thus DNA was not amplified, ‘Yes’ states a band was observed and DNA was amplified, and ‘Faint’ states only a faint band was observed and therefore DNA was only amplified in low levels.

<table>
<thead>
<tr>
<th>Group</th>
<th>Latin name</th>
<th>Common name</th>
<th>COI</th>
<th>General vertebrate 16S</th>
<th>Fish specific 16S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salmo trutta</td>
<td>Brown trout</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Thymallus thymallus</td>
<td>Grayling</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Tinca tinca</td>
<td>Tench</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Cyprinus carpio</td>
<td>Common carp</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Gymnocephalus cernua</td>
<td>Ruffe</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Gasterosteus aculeatus</td>
<td>Three-spined stickleback</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Phoxinus phoxinus</td>
<td>Minnow</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Lampetra fluviatilis</td>
<td>River lamprey</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Perca fluviatilis</td>
<td>Perch</td>
<td>Faint</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Carassius carassius</td>
<td>Crucian carp</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Rutilus rutilus</td>
<td>Roach</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Scardinius erythrophthalmus</td>
<td>Rudd</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Squalius cephalus</td>
<td>Chub</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Cottus gobio</td>
<td>Bullhead</td>
<td>Faint</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Barbus barbus</td>
<td>Barbel</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Esox lucius</td>
<td>Northern pike</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Onchorhynchus mykiss</td>
<td>Rainbow Trout</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Zeus faber</td>
<td>John Dory</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Pollachius pollachius</td>
<td>Pollock</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Raja brackyura</td>
<td>Blonde ray</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Mustelus asterias</td>
<td>Smooth hound</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Catilomynus lyra</td>
<td>Common dragonet</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Amphians</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bufo bufo</td>
<td>Common toad</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Triturus cristatus</td>
<td>Great crested newt</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Reptiles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anguis fragilis</td>
<td>Slow worm</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Birds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turdus merula</td>
<td>Blackbird</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Caprimulgus europaeus</td>
<td>Nightjar</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Crex crex</td>
<td>Corncrake</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Egretta garzetta</td>
<td>Little egret</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Aythya fuligula</td>
<td>Tufted duck</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Aythya ferina</td>
<td>Common pochard</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Oceanodroma monteiro</td>
<td>Monteiro's storm petrel</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Oceanodroma castro</td>
<td>Band-rumped storm petrel</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Hydrobates pelagicus</td>
<td>European storm petrel</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Bulweria bulwerii</td>
<td>Bulwer's petrel</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Streptopelia turtur</td>
<td>Turtle dove</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Sylvia atricapilla</td>
<td>Eurasian blackcap</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Ficedula hypoleuca</td>
<td>Pied flycatcher</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Sylvia borin</td>
<td>Garden warbler</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Alcedo atthis</td>
<td>European Kinfisher</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Carduelis carduelis</td>
<td>Goldfinch</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Erithacus rubecula</td>
<td>European robin</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Fringilla coelebs</td>
<td>Common chaffinch</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Passer domesticus</td>
<td>House sparrow</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Phylloscopus collybita</td>
<td>Chiffchaff</td>
<td>Faint</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
<tr>
<td>Phylloscopus trochilus</td>
<td>Willow warbler</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td></td>
</tr>
</tbody>
</table>
### Table S 4.2. (continued)

<table>
<thead>
<tr>
<th>Group</th>
<th>Latin name</th>
<th>Common name</th>
<th>COI</th>
<th>General vertebrate 16S</th>
<th>Fish specific 16S</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mammals</strong></td>
<td>Lutra lutra</td>
<td>Eurasian otter</td>
<td>Faint</td>
<td>Faint</td>
<td>Faint</td>
</tr>
<tr>
<td></td>
<td>Chiroptera</td>
<td>Bat</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Sorex minutus</td>
<td>Pygmy shrew</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Neomys fodiens</td>
<td>Water shrew</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Sorex araneus</td>
<td>Common shrew</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Micromys minutus</td>
<td>Harvest mouse</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Talpa europaea</td>
<td>Mole</td>
<td>No</td>
<td>No</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Arvicola amphibius</td>
<td>Water vole</td>
<td>Faint</td>
<td>Yes</td>
<td>Not tested</td>
</tr>
<tr>
<td><strong>Invertebrates</strong></td>
<td>Odonata</td>
<td>Damselfly</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Lumbricidae</td>
<td>Worm</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Pisidium tenuilineatum</td>
<td>Pea mussel</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Decapoda</td>
<td>Crab</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Pacifastacus leniusculus</td>
<td>Signal Crayfish</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Mytilus edulis</td>
<td>Mussel</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Gammarus pulex</td>
<td>Freshwater shrimp</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Carabidae</td>
<td>Ground beetle</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Penaeus esculentus</td>
<td>Shrimp</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Neuroptera</td>
<td>Lacewing</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Arion ater</td>
<td>Black slug</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Dreissena polymorpha</td>
<td>Zebra mussel</td>
<td>Faint</td>
<td>No</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Unio pictorum</td>
<td>Painter's mussel</td>
<td>Yes</td>
<td>No</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Anodonata anatina</td>
<td>Duck mussel</td>
<td>Yes</td>
<td>No</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Unio tumidus</td>
<td>Swollen river mussel</td>
<td>Yes</td>
<td>No</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Patella vulgaris</td>
<td>Limpet</td>
<td>Yes</td>
<td>No</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Mytilus edulis</td>
<td>Blue mussel</td>
<td>Yes</td>
<td>No</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Actiniaia</td>
<td>Sea anemone</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Gastropoda</td>
<td>Sea slug</td>
<td>No</td>
<td>No</td>
<td>Not tested</td>
</tr>
<tr>
<td></td>
<td>Gastropoda</td>
<td>Sea snail</td>
<td>Yes</td>
<td>No</td>
<td>Not tested</td>
</tr>
</tbody>
</table>
4.7.3 Bioinformatics scripts

Codes 1–8 were run using bash code on Cardiff University High Performance Computers, whilst codes 9–10 were run in R [version 3.6.0] and R studio [version 1.2.1335] (R Core Team 2019). In codes 1–8, lines starting with ‘#SBATCH’ are instructions for the server (e.g. what to call your job and how much memory to use), lines starting with ‘##’ or ‘# ’ are comments which are not read and lines with no # are the command to run the code. All code was run once using 16S data and then again for COI data (filenames changed). Executable code is presented in grey boxes. Code was converted into document format using R Markdown (Xie et al. 2018; Allaire et al. 2020).

1) Check for truncated MID-tags

Extract and quantify reads with exactly 10bp (the MID-tag) before your primer. This was carried out in the command line rather than using Perl and Shell scripts. Primers must be written without degenerate bases e.g. Y becomes [CT]. This was carried out for both read1 and read2 files, acquired from the Illumina MiSeq, using the forward primer sequence and then the reverse primer sequence.

First, extract all reads in the read1 file with the forward primer:

```bash
grep "^[CT][AT][AG][AC][GT]AAGACGAGAAAGACCCT" R01.fastq > R01_For_all.txt
```

Next extract all the reads in the new file with exactly 10 characters before the primer and then all the reads in the new file without exactly 10bp before the primer:

```bash
grep "^[CT][AT][AG][AC][GT]AAGACGAGAAAGACCCT" R01_For_all.txt > R01_For_10.txt
grep --v "^[CT][AT][AG][AC][GT]AAGACGAGAAAGACCCT" R01_For_all.txt > R01_For_trn.txt
```

Finally, count the number of reads in each of the files made and use these values to quantify truncation by the percentage of reads in the ‘trn’ file compared to the ‘all’ file.

```bash
wc -l R01_For_ *
```

2) Run FastP on files obtained from Illumina MiSeq

```bash
#!/bin/bash
#SBATCH --partition=mammoth       # the requested queue
#SBATCH --nodes=1              # number of nodes to use
#SBATCH --tasks-per-node=1     #
#SBATCH --cpus-per-task=1      #
#SBATCH --mem-per-cpu=1000000   # in megabytes, unless unit explicitly stated
#SBATCH --error=FastP.err      # redirect stderr to this file
#SBATCH --output=FastP.out     # redirect stdout to this file
#SBATCH --job-name=FastP      # name your job
```
## we will do FastQC quality check, merge the paired end reads and trim the sequences
## in one go using FastP to get the complete amplicon sequence
/mnt/data/GROUP-sabwocs/c1638428/scripts/LornaPipeline/fastp -i
  16S_S1_L001_R1.fastq -l
  16S_S1_L001_R2.fastq -l 125 -m --discard_unmerged -o merged.fastq

## next convert the fastq file to fasta format
module load fastx_toolkit/0.0.14
fastq_to_fasta -i merged.fastq -Q 33 -o merged.fa

3) Run Mothur and assign sequences to sample IDs
#!/bin/bash
#SBATCH --partition=mammoth       # the requested queue
#SBATCH --nodes=1              # number of nodes to use
#SBATCH --tasks-per-node=1     #
#SBATCH --cpus-per-task=1       #
#SBATCH --mem-per-cpu=1000000    # in megabytes, unless unit explicitly stated
#SBATCH --error=Mothur.err      # redirect stderr to this file
#SBATCH --output=Mothur.out     # redirect stdout to this file
#SBATCH --job-name=Mothur       # name your job

## we will identify the sequences that match the oligos used, allowing for 1 mismatch.

#Run Mothur
module load mothur/1.39.5
mothur "#trim.seqs(fasta=merged.fa, oligos=16S_Oligos.txt, checkorient=t, pdiffs=1)"

#split .groups file into A and B
grep 'a$' merged.groups > mergedA.groups
grep 'b$' merged.groups > mergedB.groups

#remove 'a' and 'b' labels
sed -i 's/a//g' mergedA.groups
sed -i 's/b//g' mergedB.groups

4) Demultiplex merged files to produce one file per sample ID.
First the Perl script is presented which is used to produce a new fasta file per sample ID, then the shell script is presented which submits the code to the server.

Perl script:

```perl
#!/usr/bin/perl
unless ($#ARGV == 0)
{
    print "Usage: 3_Demultiplex.pl FastaList_16S.txt";

die;
}

open (INLIST, "<$ARGV[0]"") || die;

# replace 'XXX' with your username, and if you want to put the output into another
# directory you can add that to the 'outdir' path here
$indir = "/mnt/data/GROUP -sabwocs/c1638428/NewPipeTest/Deplex#2";
$outdir = "/mnt/data/GROUP -sabwocs/c1638428/NewPipeTest/Deplex#2";

# Loops through the list of your samples ('SampleList') and performs the commands for
# each one
while (<INLIST>) {
    $lib = $_
    chomp($lib);

    # A shortcut to read or write a file for each of your samples, each file having the
    # same extension
    $readidsa = $lib . "_a_ids.txt";
    $readidsb = $lib . "_b_ids.txt";
    $readidsab = $lib . "_ab_ids.txt"

    $fa1 = $lib . ".fa";
    $fa2 = $lib . ".fasta"

    # split fasta read IDs into files grouped by sample ID. Replace 'XX' with the name of
    # you '.groups' file (output from mothur)
    system("grep -w $lib $indir/mergedA.groups | awk '{print \$1}' > $outdir/$readidsa" );
    system("grep -w $lib $indir/mergedB.groups | awk '{print \$1}' > $outdir/$readidsb" );
```

# combine the list of sequence names for 'a' and 'b' matches
system("cat $outdir/$readidsa $outdir/$readidsb >> $outdir/$readidsab");

# split the trimmed fasta file into reads specific to each sample. Replace 'XX' with
# the name of your trimmed fasta file (output from mothur)
my $command1 = 'perl -ne' . ".if(/^>( \S+)/){$c=$i{$1}}$c?print:chomp;$i{$_}=1 if" . "." . $outdir/$readidsab $indir/merged.trim.fasta > $outdir/$fa1";

system($command1);

system("awk '{print \$1}' $indir/$fa1 > $indir/$fa2");
}
exit;

Shell script:
#!/bin/bash
#SBATCH -- partition=mammoth       # the requested queue
#SBATCH -- nodes=1              # number of nodes to use
#SBATCH -- tasks-per-node=1     #
#SBATCH -- cpus-per-task=1      #
#SBATCH -- mem-per-cpu=1000000   # in megabytes, unless unit explicitly stated
#SBATCH -- error=Demultiplex.err # redirect stderr to this file
#SBATCH -- output=Demultiplex.out    # redirect stdout to this file
#SBATCH -- job-name=Demultiplex   # name your job

perl 3_Demultiplex.pl FastaList_16S.txt

5) Edit headers so each sequence begins with an identifier of the sample it was assigned to.

First the Perl script is presented (creates one fasta file per sample ID) and then the
Shell script (submits the code to the server).
Perl script:
#!/usr/bin/perl
unless ($#ARGV == 0) {
    print "Usage: 4_Edit_Headers.pl FastaList_16S.txt"
    die;
}
open (INLIST, "<$ARGV[0]" ) || die;

$indir = "/mnt/data/GROUP-sabwocs/c1638428/NewPipeTest/Deplex#2/FastaFiles";
$outdir = "/mnt/data/GROUP-sabwocs/c1638428/NewPipeTest/Deplex#2/FastaFiles";

while (<INLIST>) {
    $lib = ";
    chomp ($lib);

    $fa1 = $lib . ".fasta";
    $fa2 = $lib . ".edit.fasta";

    system( qq(sed "s/^/>$lib;/g" "$indir/$fa1" > "$indir/$fa2"));
}
exit;

Shell script:
#!/bin/bash
#SBATCH -- partition=mammoth       # the requested queue
#SBATCH -- nodes=1              # number of nodes to use
#SBATCH -- tasks-per-node=1     #
#SBATCH -- cpus-per-task=1      #
#SBATCH -- mem-per-cpu=1000000   # in megabytes, unless unit explicitly stated
#SBATCH -- error=headers.err     # redirect stderr to this file
#SBATCH -- output=headers.out    # redirect stdout to this file
#SBATCH -- job-name=headers     # name your job

perl 4_Edit_Headers.pl FastaList_16S.txt

6) Merge sample ID files back into one file.
This is achieved using the command line rather than Perl and Shell scripts
   cat *edit.fasta > All16S.fasta

7) Run Unoise 3 to denoise and cluster sequences into zOTUs and OTUs.
This also produces read count tables (number of reads for each zOTU/OTU in each sample).
   #!/bin/bash
  #SBATCH -- partition=mammoth       # the requested queue
#SBATCH --nodes=1          # number of nodes to use
#SBATCH --tasks-per-node=1   #
#SBATCH --cpus-per-task=1    #
#SBATCH --mem-per-cpu=1000000 # in megabytes, unless unit explicitly stated
#SBATCH --error=Unoise3.err  # redirect stderr to this file
#SBATCH --output=Unoise3.out # redirect stdout to this file
#SBATCH --job-name=Unoise3  # name your job

# removes identical replicates from the fasta input,
# output for next step = SampleName_rc_uniques.fasta
/mnt/data/GROUP-sabwocs/c1638428/scripts/usearch_11 -fastx_uniques All16S.fasta
   -fastaout Unique.fasta -sizeout -strand both -relabel Uniq -threads 4

# sort by size
/mnt/data/GROUP-sabwocs/c1638428/scripts/usearch_11 -sortbysize Unique.fasta
   -fastaout Sorted.fasta

# Cluster OTUs
/mnt/data/GROUP-sabwocs/c1638428/scripts/usearch_11 -cluster_otus Sorted.fasta -otus OTU.fasta -relabel Otu

# denoise and cluster using unoise3 to make zOTUs
/mnt/data/GROUP-sabwocs/c1638428/scripts/usearch_11 -unoise3 Sorted.fasta
   -zotus zOTU.fasta

# make list of zOTU's and the number of sequences per zOTU (size)
/mnt/data/GROUP-sabwocs/c1638428/scripts/usearch_11 -otutab All16S.fasta
   -zotus zOTU.fasta -otutabout zOTUtable_16S.txt -strand both -threads 4

# make list of OTU's and the number of sequences per OTU (size)
/mnt/data/GROUP-sabwocs/c1638428/scripts/usearch_11 -otutab All16S.fasta
   -otus OTU.fasta -otutabout OTUtable_16S.txt -strand both -threads 4

8) Use blast to identify zOTUs and OTUs
#!/bin/bash
#SBATCH --partition=mammoth     # the requested queue
#SBATCH --nodes=1                # number of nodes to use
#SBATCH --tasks-per-node=1      #
#SBATCH --cpus-per-task=4  
#SBATCH --mem-per-cpu=1000000  
#SBATCH --error=blast.err  
#SBATCH --output=blast.out  
#SBATCH --job-name=blast

# blast the clusters from usearch
module load blast/2.7.1
export BLASTDB=/mnt/data/GROUP-sabwocs/c1638428/scripts/BLAST-DB

blastn -query zOTU.fasta -db nt -num_threads 4 -evalue 0.00001 -perc_identity 97
-outfmt 6 -out zOTU.txt

blastn -query OTU.fasta -db nt -num_threads 4 -evalue 0.00001 -perc_identity 97
-outfmt 6 -out OTU_blastOutput.txt

9) Filter blast sequences
Keep only the taxonomic identifiers (i.e. Accession codes) with the top BIT score per zOTU.

blast <- read.table("zOTU.txt")
summary(blast)
library(dplyr)
blast_filter <- blast %>%
  group_by(V1) %>%
  filter(V12 == max(V12))
write.table(blast_filter, "16S_zOTU_TopHit_blastOutput.txt")

10) Cluster read counts per sample based on taxon name.
Taxonomic names are acquired by running blast data through MEGAN [version 6.12.3] (Huson et al. 2016) and then these names are matched to read count tables using zOTU/OTU identifiers (carried out using =VLOOKUP() function in Excel).

To_Agg <- read.csv("zOTUTable_withTaxonNames.csv", header = T)
Agg <- aggregate(~Taxon, data=COI_to_Agg, sum)
write.table(Agg, "zOTUTable_Aggregated.csv")

4.7.4 Scripts for visualising metabarcoding data in R following application of minimum sequence copy thresholds
R Code used to create graphs and compare metabarcoding data following application of different minimum sequence copy thresholds. Code was run using R [version 3.6.0] and R studio [version 1.2.1335] (R Core Team 2019) and converted into document format using R markdown (Xie et al. 2018; Allaire et al. 2020). Executable code is presented in grey boxes.

**Load libraries**

```r
library(ggplot2)
library(plyr)
library(reshape2)
library(viridis)
library(scales)
library(dplyr)
library(RColorBrewer)
```

**Generating Heat Charts for 16S and COI datasets with different Minimum Sequence Copy Thresholds (MSCT) applied**

In order to visualise data following application of different MSCTs (different filters, proportional thresholds and combinations of filters), functions were built to process all datasets in one go. These included:

1) A function to rename taxa with labels that were too long in the 16S datasets

```r
RenameTaxon <- function(x){x$Taxon <- revalue(x$Taxon, c("(Cyprinus carpio 'mirror' x Cyprinus carpio 'singuonensis') x Carassius auratus red var ")="Carp Hybrid", "Chelon labrosus or Mugil cephalus or Liza aurata or Oedalechilus labeo"="Mullet", "Gadus morhua or Pollachius pollachiuis"="G.morhua/P.pollachiuis", "Leuciscus idus or Leuciscus leuciscus"="L.idus/L.leuciscus", "Merlangius merlangus or Gadus morhua"="M.merlangus/G.morhua", "Squalius cephalus or Leuciscus leuciscus or Rutilus rutilus"="S.cephaalus/L.leuciscus or R.rutilus", "Scardinius erythrophtalmus or Rutilus rutilus"="S.erythrophtalmus/R.rutilus");x}
```

2) A function to reformat an object created from a list of files into one dataframe

```r
meltingHTS <- function(x){ddply(melt(x), .(variable), transform, rescale = rescale(value))}
```

3) A function to create heat charts using ggplot. Read count limits and breaks were set so all graphs were created on the same scale and thus were comparable.


**Run code for 16S datasets**

Make a list of all `.csv` files in the current directory.

```r
my_16S_files <- list.files(pattern = "\.csv$")
```

Use this list to read in the csv files (a list of dataframes).

```r
my_16S_data <- lapply(my_16S_files, read.csv)
```

Change the names of the dataframes in the list so they don't have `.csv` at the end.

```r
names(my_16S_data) <- gsub("\\.csv$, "", my_16S_files)
```

Optional checks.

```r
str(my_16S_data)
levels(my_16S_data$'01) No Clean-up`$Taxon)
```

Use the RenameTaxon function created above to rename the really long taxons to shorter names.

```r
my_16S_data2 <- lapply(my_16S_data, RenameTaxon)
```

Check the names have changed.

```r
levels(my_16S_data2$'01) No Clean-up`$Taxon)

summary(my_16S_data2)
```

Use the meltingHTS function created above to get the data in the correct format.

```r
melt_16S_Data <- lapply(my_16S_data2, meltingHTS)
```

Use the ggHeatChart function made above to make an object of heat charts that are all made to the same standards.

```r
plots_16S <- lapply(melt_16S_Data, ggHeatChart)
```

Add the dataframe name to each plot as the title.

```r
plots_16S2 <- lapply(seq_along(plots_16S), function(i) {
  plots_16S[[i]] + ggtitle(names(plots_16S)[i])
})
```

Put these plots into a PDF.
Run code for COI datasets

Make a list of all `.csv` files in the current directory.
my_COI_files <- list.files(pattern = "\csv$")

Use this list to read in the csv files (a list of dataframes).
my_COI_data <- lapply(my_COI_files, read.csv)

Change the names of the dataframes in the list so they don't have `.csv` at the end.
names(my_COI_data) <- gsub("\csv", "", my_COI_files)

Optional checks.
str(my_COI_data)
levels(my_COI_data$'01) Original $'Taxa)

Use the meltingHTS function created above to get the data in the correct format.
melt_COI_Data <- lapply(my_COI_data, meltingHTS)

Use the ggHeatChart function made above to make an object of heat charts that are all
made to the same standards.
plots_COI <- lapply(melt_COI_Data, ggHeatChart)

Add the dataframe name to each plot as the title.
plots_COI2 <- lapply(seq_along(plots_COI), function(i) {
    plots_COI[[i]] + ggtitle(names(plots_COI)[i])
})

Put these plots into a PDF.

Create summary heat charts from data using a selection of MSCT

16S data showing presences of each taxon per filtering method

Read in the data and reformat it so that when the graph is made filtering method one is
row one.

Presences16S2 <- read.csv("16S methods presences Condensed.csv", header = T)

summary(Presences16S2)

Presences16S2$Method = with(Presences16S2, factor(Method, levels =
rev(levels(Method))))
meltedPresences16S2 <- melt(Presences16S2)
Create a heat chart using ggplot adjusting the scales to cover all MID-tag combinations utilised in the study. Change axis labels, add a legend and fix the ratio.

Presenceplot <- ggplot(meltedPresences16S2, aes(variable, Method)) +
geom_tile(aes(fill = value), colour = "white") +
scale_fill_gradientn(na.value = "white",
  colours=viridis(361, direction = -1),
  values=rescale(c(1,10,20,30,40,50,100,200,300)),
  breaks = c(10,50,100,200,300,361),
  limits=c(1,361),
  oob = scales::squish) +
coord_fixed(ratio = 4, xlim = NULL, ylim = NULL, expand = TRUE) +
theme(axis.title.y=element_blank(),
  axis.title.x=element_blank(),
  axis.text.x=element_blank(),
  axis.ticks.x=element_blank(),
  axis.text.y = element_text(size = 10),
  legend.position="bottom",
  legend.text = element_text(size = 10),
  legend.key.width = unit(4, "cm"),
  legend.text.align = 0,
  legend.title = element_blank())+
guides(fill = guide_colourbar(ticks = FALSE))

Presenceplot

Create a row that depicts which column was represented by each taxonomic classification. Read in the data and then create the row using ggplot

Taxa <- read.csv("Taxa errors.csv", header = T)
summary(Taxa)
meltedTaxa <- melt(Taxa[1:2,])
taxaplot <- ggplot(meltedTaxa, aes(variable, Order)) +
geom_tile(aes(fill = value), colour = "grey") +
scale_fill_gradientn(na.value = "white", colours=brewer.pal(8, "Dark2"),
  breaks = c(1,2,3,4,5,6,7,8),
  labels = c("Mocks","Amphibians","Birds","Fish","Mammals",
             "Otter","NA","Errors"),
  guide = "legend") +
coord_fixed(ratio = 4, xlim = NULL, ylim = NULL, expand = TRUE) +
theme(axis.title.y=element_blank(),
taxaplot

16S data showing the total read count for each sample per filtering method

Read in the data and reformat it so that when the graph is made filtering method one is row one.

```r
SampleReads16S2 <- read.csv("16S reads per sample condensed.csv", header = T)
summary(SampleReads16S2)
```

Create a heat chart using ggplot adjusting the scales to cover all values of read counts in the study. Change axis labels, add a legend and fix the ratio.

```r
meltedSampleReads16S2 <- melt(SampleReads16S2)
Samplesplot <- ggplot(meltedSampleReads16S2, aes(variable, Method)) +
  geom_tile(aes(fill = value), colour = "white") +
  scale_fill_gradientn(na.value = "white",
                      colours=viridis(5, direction = -1),
                      breaks = c(1000,10000,20000,30000,40000,50000,60000),
                      limits=c(10,60000),
                      oob = scales::squish) +
  coord_fixed(ratio = 10, xlim = NULL, ylim = NULL, expand = TRUE) +
  theme(axis.title.y=element_blank(),
        axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank(),
        axis.text.y = element_text(size = 10),
        legend.position="bottom",
        legend.text = element_text(size = 10),
        legend.key.width = unit(4, "cm"),
        legend.text.align = 0,
        legend.title = element_blank()) +
```
**Sample plot**

Create a row that depicts which column was represented by each MID-tag combination and which category that combination was assigned to (mock community, negative control, unused MID-tag or eDNA sample). Read in the data and then create the row using ggplot.

```r
Samples <- read.csv("Samples List.csv", header = T)
summary(Samples)
meltedSamples <- melt(Samples)

SamplesList <- ggplot(meltedSamples, aes(variable, Method)) +
  geom_tile(aes(fill = value), colour = "grey") +
  scale_fill_gradientn(na.value = "white",
    colours = brewer.pal(4, "Spectral"),
    breaks = c(1, 2, 3, 4),
    labels = c("Mocks", "Unused MID-tag Combinations", "Negatives",
      "eDNA Samples"),
    guide = "legend") +
  coord_fixed(ratio = 10, xlim = NULL, ylim = NULL, expand = TRUE) +
  theme(axis.title.y = element_blank(),
    axis.title.x = element_blank(),
    axis.text.x = element_blank(),
    axis.ticks.x = element_blank(),
    axis.text.y = element_text(size = 1),
    legend.position = "top",
    legend.title = element_blank(),
    legend.key.size = unit(8, "mm"),
    legend.text = element_text(size = 20))

SamplesList
```

**COI data showing presences of each taxon per filtering method**

Read in the data and reformat it so that when the graph is made filtering method one is row one.

```r
PresencesCOI2 <- read.csv("COI Presences per Taxa Condensed.csv", header = T)
summary(PresencesCOI2)

PresencesCOI2$Methods <- with(PresencesCOI2, 
  factor(Methods, levels = rev(levels(Methods))))
meltedPresencesCOI2 <- melt(PresencesCOI2)
```
Create a heat chart using ggplot adjusting the scales to cover all MID-tag combinations utilised in the study. Change axis labels, add a legend and fix the ratio.

```r
PresenceplotCOI <- ggplot(meltedPresencesCOI2, aes(variable, Methods)) + 
  geom_tile(aes(fill = value), colour = "white") + 
  scale_fill_gradientn(na.value = "white", 
                        colours = viridis(361, direction = -1), 
                        values = rescale(c(1,10,20,30,40,50,100,200,300)), 
                        breaks = c(10,50,100,200,300,361), 
                        limits = c(1,361), 
                        oob = scales::squish) + 
  coord_fixed(ratio = 8, xlim = NULL, ylim = NULL, expand = TRUE) + 
  theme(axis.title.y = element_blank(), 
        axis.title.x = element_blank(), 
        axis.text.x = element_blank(), 
        axis.ticks.x = element_blank(), 
        axis.text.y = element_text(size = 10), 
        legend.position = "bottom", 
        legend.text = element_text(size = 10), 
        legend.key.width = unit(4, "cm"), 
        legend.text.align = 0, 
        legend.title = element_blank()) + 
  guides(fill = guide_colourbar(ticks = FALSE))
```

Create a row that depicts which column was represented by each taxonomic classification. Read in the data and then create the row using ggplot.

```r
COITaxa <- read.csv("COITaxaErrors2.csv", header = T) 
summary(COITaxa) 
meltedCOITaxa <- melt(COITaxa) 

library(RColorBrewer) 
display.brewer.all() 
brewer.pal(n=11, "Dark2") 
COItaxaplot <- ggplot(meltedCOITaxa, aes(variable, Group)) + 
  geom_tile(aes(fill = value), colour = "grey") + 
  scale_fill_gradientn(na.value = "white", 
                       colours = brewer.pal(7, "Greys")) + 
  coord_fixed(ratio = 8, xlim = NULL, ylim = NULL, expand = TRUE) + 
  theme(axis.title.y = element_blank(), 
```
COI data showing the total read count for each sample per filtering method

Read in the data and reformat it so that when the graph is made filtering method one is row one.

```r
SampleReadsCOI2 <- read.csv("COI Reads per Sample Condensed.csv", header = T)
summary(SampleReadsCOI2)
SampleReadsCOI2$Methods = with(SampleReadsCOI2,
  factor(Methods, levels = rev(levels(Methods))))
meltedSampleReadsCOI2 <- melt(SampleReadsCOI2)
```

Create a heat chart using ggplot adjusting the scales to cover all values of read counts in the study. Change axis labels, add a legend and fix the ratio.

```r
SamplesplotCOI <- ggplot(meltedSampleReadsCOI2, aes(variable, Methods)) +
  geom_tile(aes(fill = value), colour = "white") +
  scale_fill_gradientn(na.value = "white",
    colours = viridis(5, direction = -1),
    breaks = c(1000, 10000, 20000, 30000, 40000, 50000, 60000),
    limits = c(10, 60000),
    oob = scales::squish) +
  coord_fixed(ratio = 10, xlim = NULL, ylim = NULL, expand = TRUE) +
  theme(axis.title.y = element_blank(),
    axis.title.x = element_blank(),
    axis.text.x = element_blank(),
    axis.ticks.x = element_blank(),
    axis.text.y = element_text(size = 10),
    legend.position = "bottom",
    legend.text = element_text(size = 10),
    legend.key.width = unit(4, "cm"),
    legend.text.align = 0,
    legend.title = element_blank()) +
```
Create a row that depicts which column was represented by each MID-tag combination and which category that combination was assigned to (mock community, negative control, unused MID-tag or eDNA sample). Read in the data and then create the row using ggplot

```r
COISamples <- read.csv("COISamples List.csv", header = T)
summary(COISamples)
meltedCOISamples <- melt(COISamples)

COISamplesList <- ggplot(meltedCOISamples, aes(variable, Methods)) +
geom_tile(aes(fill = value), colour = "grey") +
scale_fill_gradientn(na.value = "white",
  colours = brewer.pal(4, "Spectral"),
  breaks = c(1, 2, 3, 4),
  labels = c("Mocks", "Unused MID-tag Combinations", "Negatives",
    "eDNA Samples"),
  guide = "legend") +
coord_fixed(ratio = 10, xlim = NULL, ylim = NULL, expand = TRUE) +
theme(axis.title.y = element_blank(),
  axis.title.x = element_blank(),
  axis.text.x = element_blank(),
  axis.ticks.x = element_blank(),
  axis.text.y = element_text(size = 1),
  legend.position = "top",
  legend.title = element_blank(),
  legend.key.size = unit(8, "mm"),
  legend.text = element_text(size = 20))

COISamplesList
```

**Create non-metric multidimensional scaling (NMDS) graphs**

Graphs were created from relative frequency of occurrence data (i.e. the frequency a taxon occurred divided by the total number of MID-tag combinations, which was 361 for both barcoding regions) per MSCT.

First load in the vegan package and set the seed

```r
library(vegan)
set.seed(2)
```
**16S NMDS**

Read in the data, check format and set row names to be the same as column one (i.e. name of MSCT applied to the data represented by that row)

```r
CleanUp16S <- read.csv("16S methods NMDS FOO.csv", header = T)
summary(CleanUp16S)
colnames(CleanUp16S)
rownames(CleanUp16S) <- CleanUp16S[,1]
rownames(CleanUp16S)
```

Run NMDS with two-dimensions and a Bray-Curtis dissimilarity calculation, then check the stress and distribution of the data

```r
NMDS16S <- metaMDS(CleanUp16S[,2:106], k=2, trymax = 100, distance = "bray", autotransform = FALSE)
NMDS16S$stress
stressplot(NMDS16S)
plot(NMDS16S)
ordiplot(NMDS16S,type="n")
orditorp(NMDS16S,display="species",col="red",air=0.01)
orditorp(NMDS16S,display="sites",cex=1.25,air=0.01)
```

Plot ellipses coloured by filtering method and add labels. Ellipses include different thresholds tested for each proportional filtering method (e.g. percentage of total reads). Labels were required to show position of filters that were not proportional (e.g. singletons) and thus have only one data point (no ellipse). The object “treat” was used to set label names which were plotted using ordicenter.

```r
treat=c(rep("No Filter",1),
     rep("Singletons",1),
     rep("<10",1),
     rep("Maximum Contamination",1),
     rep("Total %",8),
     rep("Total % + Maximum Contamination",5),
     rep("Sample %",7),
     rep("Sample % + Maximum Contamination",5),
     rep("Taxa %",6),rep("Taxa % + Maximum Contamination",6),
     rep("Sample % + Taxa %",20))
ordiplot(NMDS16S,type="n")
ordiellipse(NMDS16S, groups=treat, draw="polygon", col=c("black","black","black", "cadetblue3","chocolate","goldenrod","black","darkorchid", "blue4","mediumseagreen","darkgreen"),
```
**COP**

Read in the data, check format and set rownames to be the same as column one (i.e. name of MSCT applied to the data represented by that row)

```r
CleanUpCOI <- read.csv("COI methods NMDS FOO.csv", header = T)
summary(CleanUpCOI)
colnames(CleanUpCOI)
rownames(CleanUpCOI) <- CleanUpCOI[, 1]
```

Run NMDS with two-dimensions and a Bray-Curtis dissimilarity calculation, then check the stress and distribution of the data

```r
NMDSCOI <- metaMDS(CleanUpCOI[,2:233], k=2, trymax = 100, distance = "bray", autotransform = F)
```

Plot ellipses coloured by filtering method and add labels. Ellipses include different thresholds tested for each proportional filtering method (e.g. percentage of total reads). Labels were required to show position of filters that were not proportional (e.g. singletons) and thus have only one data point (no ellipse). The object “treat” was used to set label names which were plotted using ordicenter.

```r
treat=c(rep("No Filter",1),
        rep("Singletons",1),
        rep("<10",1),
        rep("Maximum Contamination",1),
        rep("Total Threshold",6),
        rep("Total Threshold + Maximum Contamination",6),
        rep("Taxa Threshold",6),
        rep("Taxa Threshold + Maximum Contamination",6),
        rep("Sample Threshold",5),
        rep("Sample Threshold + Maximum Contamination",5),
        rep("Sample + Taxa Threshold",27))
```
ordiellipse(NMDSCOI, groups=treat, draw="polygon", col=c("black","black","black", "goldenrod","cadetblue3","chocolate","black","darkorchid", "blue4","mediumseagreen","darkgreen"), border=c("black","black","black","goldenrod","cadetblue3","chocolate","black", "darkorchid","blue4","mediumseagreen","darkgreen"), label=F, alpha = 0.25)

ordicenter(NMDSCOI,groups=treat)

4.7.5 Additional figures for visualising effectiveness of different minimum sequence copy thresholds
Figure S 4.1. Eurasian otter (*Lutra lutra*) diet 16S reads. Coloured lines show total reads in each sample (low reads = yellow, high reads = purple) following application of minimum sequence copy thresholds. ‘Low’, ‘Mid’ and ‘High’ depict the value utilised for proportional thresholds (‘Total %’, ‘Sample %’ and ‘Taxon %’), with ‘Opt + MC’ denoting the ‘optimal’ threshold combined with ‘Max Contamination’ methods.
Figure S 4.2. Eurasian otter (*Lutra lutra*) diet COI reads. Coloured lines show total reads in each sample (low reads = yellow, high reads = purple) following application of minimum sequence copy thresholds. 'Low', 'Mid' and 'High' depict the value utilised for proportional thresholds ('Total %', 'Sample %' and 'Taxon %'), with 'Opt + MC' denoting the 'optimal' threshold combined with 'Max Contamination' methods.
Figure S 4.3. Eurasian otters (*Lutra lutra*) with marine prey present in their diet following application of minimum sequence copy thresholds on 16S data (blue dots). Threshold applied is depicted above each graph, with ‘low’, ‘med’ and ‘high’ representing the value utilised for proportional thresholds (‘Total %’, ‘Sample %’ and ‘Taxon %’) whilst ‘Opt + MC’ represents the optimal proportional threshold value when combined with ‘Maximum Contamination’ filters.
Figure S 4.3. (continued)
Figure S 4.4. Eurasian otters (*Lutra lutra*) with marine prey present in their diet following application of minimum sequence copy thresholds on COI data (blue dots). Threshold applied is depicted above each graph, with ‘low’, ‘med’ and ‘high’ representing the value utilised for proportional thresholds (‘Total %’, ‘Sample %’ and ‘Taxon %’) whilst ‘Opt + MC’ represents the optimal proportional threshold value when combined with ‘Maximum Contamination’ filters.
Figure S 4.4. (continued)
Chapter Five – Otterly delicious: Investigating spatio-temporal variation in the diet of a recovering population of Eurasian otters (*Lutra lutra*) through DNA metabarcoding and morphological analysis of prey remains

5.1 Abstract

Generalist apex predators exhibit highly plastic diets, consuming a broad array of species and switching to alternative prey when the availability of preferred prey declines. Investigating the diet of these predators allows interactions within food webs and changes in species dynamics to be described over space and time, providing a valuable insight into both predator and prey populations. Here we investigated factors influencing dietary variation in a recovering population of Eurasian otters (*Lutra lutra*), an apex predator in freshwater ecosystems. We sampled 300 otters across England and Wales between 2007 and 2016, conducting both morphological analysis of prey remains, and DNA metabarcoding, on faeces collected from dead otters. Comparison of these methods showed that greater precision and detail could be achieved using DNA metabarcoding, however, combining data from both methodologies gave the most comprehensive description of the diet. Otters from all demographics exploited a broad range of taxa in a variety of habitats, and variation primarily reflected changes in prey distributions across the landscape, and seasonal changes in availability. This study provides novel insights into the adaptability of otters, which is likely to have aided their recent population recovery, and may increase resilience to future environmental changes.

*Keywords*: Dietary variation, DNA metabarcoding, Eurasian otter, faecal analysis, *Lutra lutra*, morphological analysis of prey remains

5.2 Introduction

Generalist apex predators have broad diets, and connect multiple energetic pathways by consuming prey from a variety of habitats and trophic levels (e.g. Rosenblatt *et al.* 2015; Berry *et al.* 2017; Vejřík *et al.* 2017). The foraging behaviour of apex predators has the potential to shape communities, by directly influencing prey populations and
indirectly impacting species at other trophic levels (Shurin et al. 2002; Knight et al.
2005; Wallach et al. 2015). Generalist predators tend to show high levels of plasticity in
their diet, switching to alternative prey when their preferred prey become less available
(Murdoch 1969; Erlinge 1983; Reif et al. 2001; Almeida et al. 2012). This plasticity
makes generalist apex predators more resilient to disturbance (Van Baalen et al. 2001;
Peers et al. 2014), although switching sometimes incurs fitness costs if alternative prey
are nutritionally sub-optimal (Ruiz-Olmo and Jiménez 2009; Cohen et al. 2014;
Moorhouse-Gann et al. 2020). Dietary shifts alter the rates at which different prey
species are taken, either increasing or decreasing predation pressure on alternative
prey species, and potentially negatively impacting their populations (Latham et al.
2013), which may be particularly important for threatened species. Dietary analysis of
apex predators allows interactions within food webs to be described, and changes in
dynamics to be analysed over space and/or time (Boyer et al. 2015; Bessey et al.
2019), providing an evidence base that is important for the conservation of both

Apex predators characteristically possess broad distributions and large home ranges
(Stier et al. 2016), resulting in dietary variation occurring across a range of spatio-
temporal scales (Lukasik and Alexander 2011; Almeida et al. 2012; Rosenblatt et al.
2015). Prey species differ in abundance and ease of capture between habitats and
times of the year (Čech et al. 2008; Rosenblatt et al. 2015), and variation in predator
diet typically reflects these changes in availability (Boyd and Murray 2001). Differences
in foraging behaviour between individuals can lead to differential exposure to diverse
threats, such as toxicological risk (e.g. consumption of prey species with high
contaminant load) or direct mortality (e.g. due to conflict with humans associated with
predation of farmed species (Stier et al. 2016). Shifts in diet can impact an individual’s
fitness over the short term (Ruiz-Olmo and Jiménez 2009; Lourenço et al. 2011) and
the persistence of the species in the long-term (Roos et al. 2001; Torres and Fonseca
2016), potentially affecting food web dynamics and ecosystem functioning (Wallach
et al. 2015; Hollings et al. 2016). Obtaining high resolution taxonomic dietary data,
alongside spatio-temporal and biotic data can give an insight into potential pressures
faced by generalist apex predators (Thomas et al. 2017) and their resilience to such
pressures, allowing both individual and population level inferences to be made
(Jeanniard-Du-Dot et al. 2017; Aizpurua et al. 2018).

Traditionally, identification of the dietary composition of predators has relied on the
morphological analysis of undigested remains in faeces and stomach contents (e.g.
Martins et al. 2011; McCully Phillips et al. 2019), but such methods may be biased.
Differences in digestion rates can lead to prey becoming over- or under-represented, as remains that are resistant to digestion are more likely to be successfully identified (Pompanon et al. 2012; Boyer et al. 2015). Soft-bodied prey (Arai et al. 2003), or prey that are only partially consumed (Granquist et al. 2018), are also likely to go undetected due to the lack of hard remains that can survive digestion. Where prey are morphologically similar to one another, identification can be difficult, potentially resulting in mis-identified remains or poor taxonomic resolution (e.g. identified to family rather than species; Spaulding et al. 2000). By detecting prey DNA in faeces and stomach contents, identifications can be made to a finer taxonomic resolution even where no visual traces are present (Bowser et al. 2013; Roslin and Majaneva 2016; Elbrecht, Vamos, et al. 2017). DNA metabarcoding achieves this by combining high-throughput sequencing (HTS) with DNA barcodes (short, variable regions of the genome) to simultaneously identify multiple taxa within many samples (Taberlet et al. 2018). Decreases in sequencing costs and the development of extensive reference databases have allowed DNA metabarcoding to be exploited by a greater range of studies (Hawlitschek et al. 2018), and it has become one of the primary methods for describing the diet of predators (e.g. Hardy et al. 2017; McInnes et al. 2017; Galan et al. 2018; Shi et al. 2018; Toju and Baba 2018).

The Eurasian otter (Lutra lutra, hereafter referred to as ‘otter’) is a generalist apex predator of European freshwater habitats, with a broad diet primarily consisting of fish (Kruuk 1995; Britton et al. 2006; Almeida et al. 2012; Krawczyk et al. 2016). Otter population declines across much of their European range in the 1950’s are generally attributed to habitat modification and acquisition of contaminants such as organochlorine pesticides and polychlorinated biphenyls (PCBs) (Strachan and Jefferies 1996; Roos et al. 2001; Clavero et al. 2010). In recent decades though, populations in Great Britain have increased and expanded their distribution, allowing otters to recolonise habitats from which they were once extirpated (Roos et al. 2001; Conroy and Chanin 2002; Sainsbury et al. 2019). Whilst otters have begun to return to habitats from which they have been absent in recent decades, it is likely that changes in the landscape and other factors have led to altered prey availability (Burns et al. 2016), freshwater contaminant loads (Harrad et al. 1994) and human disturbance, thereby potentially altering otter diet and foraging behaviour. Corresponding changes in the health and behaviour of individuals (Ruiz-Olmo and Jiménez 2009) are likely to alter selection pressures (Clavero et al. 2010) and thus impact the recent and continued recovery and distribution of the population (Stier et al. 2016).
Studies into the diet of otters have primarily focussed on morphological analysis of prey remains in spraint or stomach contents (e.g. Jędrzejewska et al. 2001; Ruiz-Olmo and Jiménez 2009; Almeida et al. 2012), potentially lacking information on a range of prey species. The current study utilised DNA metabarcoding alongside morphological analysis of undigested remains in order to directly compare the two methods, and to derive a more detailed and accurate analysis of the diet of otters. Analysis was conducted using faecal samples and data collected from dead otters found across England and Wales over a ten-year period, providing a unique opportunity to investigate dietary variation across spatial, temporal and biotic variables. Specifically we tested the hypotheses that: (i) DNA metabarcoding would detect a greater range of prey and identify prey to a finer taxonomic resolution than morphological analysis, (ii) DNA metabarcoding would detect more occurrences of prey, although the most frequent prey would remain the same between the two methods, (iii) changes in prey availability across the landscape would affect the composition of otter diet, (iv) temporal changes in prey availability would affect the composition of otter diet over seasonal and annual timescales, (v) dietary composition would vary between different demographic groups, and (vi) body condition would be associated with dietary variation, with better body condition in individuals consuming species with higher fat content.

5.3 Methods

5.3.1 Sample and data collection

Samples and associated metadata were acquired from 300 otters collected between 2007 and 2016, obtained from the Cardiff University Otter Project archive as described in Chapter Three. Calculation and assignment of body condition index, scored Scaled Mass Index (SMI; Peig and Green 2009; Peig and Green 2010; Guillemain et al. 2013), was carried out as described in Chapter Two. Spatial variables used for analysis included distance from the coast (km), latitude, longitude, urban land-use, altitude, slope and primary water habitat. These were calculated as described in Chapter Three, with the exception that a buffer with a 10 km diameter was used. A different buffer size was used due to the shorter timeframe represented by faecal samples; faecal samples reflect diet from the previous 24-72 hours (in mammals; Deagle et al. 2005; Casper et al. 2007; Thalinger et al. 2016), during which time otters can travel up to 10km (Chanin 2003), it was therefore deemed appropriate to use this distance to reflect the land used by otters within the sample timeframe. Buffers were used to calculate proportions of urban land-use (i.e. urban and suburban land use extracted from the UK land cover map from 2007; Morton et al. 2011), mean altitude and mean slope. Longitude, altitude and slope were found to be highly correlated (Fig. S 5.1) and so longitude was used as
a representative for the three variables. Otters within two and a half kilometres from a lake or transitional water were assigned to that habitat, whilst those further away were assumed to be feeding primarily in the river network and assigned to tributary or main river channel based upon weighted river length within the buffer designated to each otter (see Chapter Three for more details; as with overall buffer size, the distance used to assign otters to aquatic habitat was reduced for the current study).

5.3.2 Morphological analysis

A subsample of faecal material was taken for morphological analysis. This was thawed and soaked in a mixture of water and biological detergent for 24 hours. Samples were passed through sieves with a 0.5mm mesh and washed with water to make sure only hard parts remained. Remaining hard-parts were air-dried for 24 hours. A record was made of any samples that did not contain any hard-parts. Recognisable remains (bones, fish scales, feathers, fur) underwent microscopic identification using a range of keys (Libois and Hallet-Libois 1987; Miles and Gaglione 1992; Prenda and Granado-Lorencio 1992; Prenda et al. 1997; Watt et al. 1997; Miranda and Escala 2002; Conroy et al. 2005; Tercerie et al. 2019; University of Nottingham 2020). Prey remains were identified to the lowest possible taxonomic level and recorded as present or absent within a sample (Table S 5.1) by Segio Bedmar Castillo (see acknowledgements in 5.6.)

5.3.3 DNA metabarcoding analysis

Faecal samples were processed for HTS, and subsequent bioinformatic analysis was conducted, as described in Chapter Four (summarised in Fig. 5.1). Following bioinformatic processing, sequencing data underwent filtering steps to remove any remaining artefacts or contaminants in the data, following the protocol derived in Chapter Four. In brief, taxa that contributed to less than 0.5% of a sample’s total reads for 16S and 0.3% for COI were removed from each sample, and reads equal to or less than the maximum read count identified in unused MID-tag combinations or negative controls per taxon were removed. This method was conservative but was selected in order to avoid false positives which might lead to overrepresentation of some prey groups.
Figure 5.1. Summary of the metabarcoding workflow utilised in this study.

- **DNA extraction**
- **Amplification** of DNA using MID-tagged primers
- **Pool** samples
- **Create libraries** from pooled samples
- **Sequencing** on an Illumina MiSeq

**Lab work**

**Bioinformatics**

- **FastP** - quality check reads, discard poor quality and merge read pairs
- **Mothur** – assign reads to sample based upon MID-tags
- **Demultiplex** – separate reads into files based upon sample ID
- Edit reads headers so contain sample ID and merge back together
- **Denoise**, remove chimeras and make OTU table using Usearch11
- **Blast** OTUs

**Data filtering and statistics**

- **Filter** blast output keeping top hits only
- **Assign** taxonomic information to OTUs and aggregate
- **Remove** low percentage of reads from each sample and maximum read count in blank controls per taxon
- **Convert** to presence-absence matrix
- **Combine** 16S and COI matrices, and then combine with morphological analysis of hard parts matrix
- **Group** taxon into prey groups
  - Run **mvabund** to identify significant differences
  - Run **boral** and **bipartite** to visualise results
Reads were assigned to the lowest possible taxonomic level and recorded as present or absent within a sample, separately for 16S (Table S 5.2) and COI (Table S 5.3). Reads assigned to non-food items remaining in the analysis were removed, these included taxa not assigned to the animal kingdom (e.g. fungi and bacteria, which were deemed likely contaminants), those with low taxonomic resolution (e.g. Eutheria, which includes all extant British mammals and thus was not useful for further analyses), reads potentially from otters themselves (e.g. those assigned to Lutra lutra or mammalia) and taxa with a maximum size < 3 mm (e.g. diatoms, assumed to be due to secondary predation; Tables S 5.2 and S 5.3). Following removal of non-food items, data from the two datasets were combined to give a more complete representation of the diet of otters. If a taxon was present in either of the metabarcoding datasets, then that taxon was recorded as present in that sample. If a prey item was detected in a sample in both metabarcoding datasets, but at different levels of taxonomic resolution, only the presence with the greater taxonomic resolution was retained.

5.3.4 Comparison of methods

The frequency of occurrence for each prey item detected across the 300 otters screened was calculated for both morphological and metabarcoding datasets, allowing the two methods to be directly compared. Presences assigned to 'insect', 'beetle', 'mollusc' and 'snail' in morphological analysis were removed before comparing datasets; many identifications from these particular taxonomic groups were identified to a greater resolution through metabarcoding but removed as secondary predation (Table S 5.3), therefore these presences in the morphological analysis were also deemed likely to have occurred through secondary predation. Presences assigned to 'mammal' (identified from fur) in the morphological analysis were also removed before comparing datasets due to the uncertainty of fur coming from the otter grooming itself and metabarcoding identifying otter as the only mammal in these samples. Presence-absence matrices produced from each methodology were also combined in order to assess the overlap in data, revealing which data points were only detected by one method and which were detected by both (either at the same taxonomic level or at different resolutions; Fig. S 5.3).

5.3.5 Statistical analysis

In order to explore the associations between the composition of otter diet and biotic and abiotic drivers, we used general linear models for multivariate data on the combined data from morphological analysis and metabarcoding. Each taxon was assigned to a 'prey group' based upon similarities in taxonomy, morphology and ecological niche (Table 5.1). A small number of prey identified at low taxonomic
resolution could not be assigned to a group, and were removed from subsequent analyses (prey presences removed: 'Salmo genus', n = 5; 'Cyprinid', n = 2; 'Bird', n = 2). A prey group was recorded as present in an individual faecal sample if any one (or more) of the taxa assigned to that group were present. If a prey group occurred in less than three samples then the prey group was designated as rare and removed from subsequent analyses (Table 5.1).
<table>
<thead>
<tr>
<th>Latin name</th>
<th>Common name</th>
<th>Method of detection</th>
<th>Group</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lampetra fluviatilis</td>
<td>European river lamprey</td>
<td>Metabarcoding</td>
<td>Lamprey</td>
<td>Dissimilar in taxonomy to other species</td>
</tr>
<tr>
<td>Echithius vipera</td>
<td>Lesser weever</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taenialus bubalis</td>
<td>Longspined sea scorpion</td>
<td>Metabarcoding</td>
<td>Coastal Fish</td>
<td>Same taxonomic class and all found in coastal waters and near estuaries</td>
</tr>
<tr>
<td>Mugilida</td>
<td>Mullet</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pliureonectiformes</td>
<td>Flatfish</td>
<td>Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pliureonectidae</td>
<td>Right-eyed flounders</td>
<td>Metabarcoding</td>
<td>Flatfish</td>
<td>Same taxonomic order, similar morphology and similar life strategies</td>
</tr>
<tr>
<td>Platichthys flesus</td>
<td>European flounder</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scophthalmus rhombus</td>
<td>Brill</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pomatoschistus</td>
<td>Goby</td>
<td>Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pomatoschistus microps</td>
<td>Common goby</td>
<td>Metabarcoding</td>
<td>Goby</td>
<td>Same taxonomic genus, similar morphology, habitat use and life strategies</td>
</tr>
<tr>
<td>Pomatoschistus minutus</td>
<td>Sand goby</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barbatula barbatula</td>
<td>Stone loach</td>
<td>Metabarcoding/ Morphological</td>
<td>Loach</td>
<td>Same taxonomic order, similar morphology and similar life strategies</td>
</tr>
<tr>
<td>Cobitis taenia</td>
<td>spined loach</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gasterosteidae</td>
<td>Stickleback</td>
<td>Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gasterosteus aculeatus</td>
<td>Three-spined stickleback</td>
<td>Metabarcoding/ Morphological</td>
<td>Stickleback</td>
<td>Same taxonomic family, similar morphology, life strategies and habitat use</td>
</tr>
<tr>
<td>Pungitius puntitius</td>
<td>Nine-spined stickleback</td>
<td>Metabarcoding/ Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyprinidae</td>
<td>Cyprinidae</td>
<td>Morphological</td>
<td>Cyprinid*</td>
<td>Kept separate as not distinct enough taxonomically</td>
</tr>
<tr>
<td>Cyprinidae</td>
<td>Carp</td>
<td>Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carassius carassius</td>
<td>Crucian carp</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carassius auratus</td>
<td>Goldfish</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyprinus carpio</td>
<td>Common carp</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clenopahygyodon idella</td>
<td>Grass carp</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leuciscus</td>
<td>Ide or Dace</td>
<td>Metabarcoding</td>
<td>Leuciscus</td>
<td>Same taxonomic genus, similar morphology (often confused during identification), life strategies and habitat use</td>
</tr>
<tr>
<td>Leuciscus leuciscus</td>
<td>Dace</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rutilus rutilus</td>
<td>Roach</td>
<td>Metabarcoding/ Morphological</td>
<td>Rudd or Roach</td>
<td>Same taxonomic family, similar morphology (often confused during identification), life strategies and habitat use</td>
</tr>
<tr>
<td>Scardinus erythrophthalmus</td>
<td>Rudd</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scardinus erythrophalmus or Rutilus rutilus</td>
<td>Rudd or Roach</td>
<td>Metabarcoding</td>
<td>Rudd or Roach</td>
<td>Same taxonomic family, similar morphology (often confused during identification), life strategies and habitat use</td>
</tr>
<tr>
<td>Abramis brama</td>
<td>Common bream</td>
<td>Metabarcoding</td>
<td>Bream</td>
<td>Dissimilar morphology and habitat use to other cyprinids</td>
</tr>
<tr>
<td>Phoxinus phoxinus</td>
<td>Common Minnow</td>
<td>Metabarcoding/ Morphological</td>
<td>Minnow</td>
<td>Dissimilar morphology and habitat use to other cyprinids</td>
</tr>
<tr>
<td>Latin name</td>
<td>Common name</td>
<td>Method of detection</td>
<td>Group</td>
<td>Justification</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>---------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td>Gobio gobio</td>
<td>Gudgeon</td>
<td>Metabarcoding</td>
<td>Gudgeon</td>
<td>Dissimilar morphology and habitat use to other cyprinids</td>
</tr>
<tr>
<td>Tinca tinca</td>
<td>Tench</td>
<td>Metabarcoding/ Morphological</td>
<td>Tench</td>
<td>Dissimilar morphology and habitat use to other cyprinids</td>
</tr>
<tr>
<td>Barbus barbus</td>
<td>Common Barbel</td>
<td>Metabarcoding</td>
<td>Barbel**</td>
<td>Dissimilar morphology and habitat use to other cyprinids</td>
</tr>
<tr>
<td>Rhodeus amarus</td>
<td>European Bitterling</td>
<td>Metabarcoding</td>
<td>Bitterling**</td>
<td>Dissimilar morphology and habitat use to other cyprinids</td>
</tr>
<tr>
<td>Cottus gobio</td>
<td>European bullhead</td>
<td>Metabarcoding/ Morphological</td>
<td>Bullhead</td>
<td>Dissimilar morphology and habitat use to other species</td>
</tr>
<tr>
<td>Percidae</td>
<td>Percidae</td>
<td>Morphological</td>
<td>Percidae</td>
<td>Same taxonomic family, similar morphology and habitat use</td>
</tr>
<tr>
<td>Perca fluviatilis</td>
<td>European perch</td>
<td>Metabarcoding/ Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gymnocephalus cemua</td>
<td>Eurasian ruffe</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esox lucius</td>
<td>Northern pike</td>
<td>Metabarcoding/ Morphological</td>
<td>Pike</td>
<td>Dissimilar morphology and habitat use to other species</td>
</tr>
<tr>
<td>Anguilla anguilla</td>
<td>European eel</td>
<td>Metabarcoding/ Morphological</td>
<td>Eel</td>
<td>Dissimilar taxonomy to other species and catadromous life strategy</td>
</tr>
<tr>
<td>Salmo</td>
<td>Salmonid</td>
<td>Morphological</td>
<td>Brown trout or Atlantic Salmon*</td>
<td>Kept separate as not distinct enough taxonomically</td>
</tr>
<tr>
<td>Salmo salar</td>
<td>Atlantic salmon</td>
<td>Metabarcoding</td>
<td>Salmon</td>
<td>Anadromous, therefore has a very different life strategy to other salmonis species</td>
</tr>
<tr>
<td>Salmo trutta</td>
<td>Brown trout</td>
<td>Metabarcoding/ Morphological</td>
<td>Brown trout</td>
<td>Although some are anadromous, not all are, therefore kept separate from atlantic salmon as this different in individual life strategy may affect availability to otters</td>
</tr>
<tr>
<td>Thymallus thymallus</td>
<td>Grayling</td>
<td>Metabarcoding</td>
<td>Grayling</td>
<td>Different life strategies to other salmonids and stocked fish</td>
</tr>
<tr>
<td>Oncopterus mykiss</td>
<td>Rainbow trout</td>
<td>Metabarcoding</td>
<td>Rainbow trout</td>
<td>Different life strategies to other salmonids and stocked fish</td>
</tr>
<tr>
<td>Anura</td>
<td>Frog or toad</td>
<td>Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bufo bufo</td>
<td>Common toad</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ranidae</td>
<td>Frog</td>
<td>Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rana temporaria</td>
<td>Common frog</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salamandridae</td>
<td>Newt</td>
<td>Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lissotriton helveticus</td>
<td>Palmate newt</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lissotriton vulgaris</td>
<td>Smooth newt</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triturus cristatus</td>
<td>Great crested newt</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anas platyrynchos</td>
<td>Mallard</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anser anser</td>
<td>Greylag goose</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aythya fuligula</td>
<td>Tufted duck</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cygnus olor</td>
<td>Mute swan</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tachybaptus ruficollis</td>
<td>Little grebe</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fulica atra</td>
<td>Eurasian coot</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gallinula chloropus</td>
<td>Common moorhen</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

197
Table 5.1. (continued)

<table>
<thead>
<tr>
<th>Latin name</th>
<th>Common name</th>
<th>Method of detection</th>
<th>Group</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columba palumbus</td>
<td>Wood pigeon</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corvus monedula</td>
<td>Jackdaw</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larus argentatus</td>
<td>Chicken</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gallus gallus</td>
<td>Seagull</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sylvia curruca</td>
<td>White throat</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aves</td>
<td>Bird</td>
<td>Morphological</td>
<td>Bird*</td>
<td>Kept separate as not distinct enough taxonomically</td>
</tr>
<tr>
<td>Rattus norvegicus</td>
<td>Brown rat</td>
<td>Metabarcoding</td>
<td>Mammal**</td>
<td>Different taxonomic class, morphology and habitat use to other taxa in the study</td>
</tr>
<tr>
<td>Astacidae</td>
<td>Crayfish</td>
<td>Morphological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austropotamobius pallipes</td>
<td>White-clawed crayfish</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacifastacus leniusculus</td>
<td>Signal crayfish</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stigmatogaster subterranea</td>
<td>Centipede</td>
<td>Metabarcoding</td>
<td>Centipede**</td>
<td>Dissimilar taxonomy to other species</td>
</tr>
<tr>
<td>Eiseniella tetraedra</td>
<td>Square-tailed worm</td>
<td>Metabarcoding</td>
<td>Earthworm**</td>
<td>Same taxonomic family, similar morphology and similar life strategies</td>
</tr>
<tr>
<td>Lumbricidae</td>
<td>Earthworm</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumbricus variegatus</td>
<td>Califonia blackworm</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Styolidius heringianus</td>
<td>Earthworm</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deroceras</td>
<td>Smooth land slug</td>
<td>Metabarcoding</td>
<td>Gastropod</td>
<td>Same taxonomic class and similar morphology</td>
</tr>
<tr>
<td>Lymnaea stagnalis</td>
<td>Great pond snail</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anodonta anatina</td>
<td>Duck mussel</td>
<td>Metabarcoding</td>
<td>Mussel**</td>
<td>Same taxonomic family, similar morphology and similar life strategies</td>
</tr>
<tr>
<td>Pseudanodonta complanata</td>
<td>Depressed river mussel</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anax imperator</td>
<td>Emperor dragonfly</td>
<td>Metabarcoding</td>
<td>Odonata</td>
<td>Same taxonomic order, similar morphology, habitat use and life strategies</td>
</tr>
<tr>
<td>Ischnura elegans</td>
<td>Blue-tailed damselfly</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hediste diversicolor</td>
<td>Ragworm</td>
<td>Metabarcoding</td>
<td>Marine invertebrates</td>
<td>Dissimilar habitat use to other invertebrates</td>
</tr>
<tr>
<td>Palaemon varians</td>
<td>Common ditch shrimp</td>
<td>Metabarcoding</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
All statistical analyses were performed in R [version 3.6.0] and R Studio [version 1.2.1335] (R Core Team 2019) using ‘mvabund’, ‘bipartite’ and ‘boral’ packages (scripts available in Supplementary information 5.7.2). The ‘mvabund’ package allows model-based analysis of multivariate data to test hypotheses about the effects of environmental variables on the composition of dietary data (Wang et al. 2012). A multivariate generalised linear model (MGLM) approach provides increased statistical power for detecting differences in communities with less abundant taxa and is less prone to mis-interpretations of data due to mean-variance effects, compared to distance-based methods (Warton et al. 2012), making it more reliable in this context. The ‘many.glm’ function was used to create a MGLM using a binomial family and a ‘cloglog’ link function. The presence-absence matrix of prey groups from the combined morphological and metabarcoding dataset was used to create an ‘mvabund’ object, which was used as the response variable in the glm. The global models included the following fixed variables: sex, size of otter, scored SMI, year, season, distance from the coast (km), primary water habitat, percentage of urban land-use, latitude and longitude. Interactions between sex and size of otter, distance from the coast and sex, distance from the coast and size, primary water habitat and sex, primary water habitat and size, and between latitude and longitude were also include in the global model. Model assumptions were checked on the global model before conducting model selection via Akaike’s Information Criterion (AIC) using the stepwise algorithm in the step function (Hastie and Pregibon 1992; Venables and Ripley 2002). The final model included the fixed variables season, longitude and distance from the coast. The function ‘anova.manyglm’ was used to infer significance of fixed variables in the final model on dietary composition at the community level and prey group specific level. This was conducted using a Monte Carlo resampling method, likelihood ratio test and corrected univariate p-values for multiple testing, to identify significant associations for changes across the dietary composition and for each prey group individually.

To complement the mvabund analysis, the boral package was used to plot significant variables. The boral package conducts Bayesian ordination and regression analysis on multivariate data (Hui 2016). Binomial models for boral analysis included the same fixed and response variables as in the final mvabund model. The number of latent variables was set as two. Model assumptions were checked and latent variable values extracted. Latent variables were plotted against significant fixed variables to visualise the individual samples and the indicator species that best described their position in a low-dimension ordination plot. Bipartite graphs were also created to visualise the
composition of otter diet at the community level and the effects of significant fixed variables, using the plotweb function in the bipartite package (Dormann et al. 2008).

5.4 Results

Otters consumed a range of vertebrate and invertebrate taxa (Table 1; Tables S 5.1 to S 5.3). Vertebrate taxa were primarily identified as freshwater fish, but amphibians, birds (primarily waterfowl), coastal fish and mammals were also identified. Invertebrate taxa were primarily identified as crayfish, with some molluscs, insects, earthworms and marine invertebrates also identified at low frequencies. Taxonomic classifications within each prey group varied between morphological and metabarcoding analyses.

5.4.1 Morphological analysis

Of the 300 otters screened, morphological analysis recovered 279 occurrences of prey from 23 taxa (Table 5.1) in 172 otters, with an average of 1.62 taxa per otter. Dietary data were not recovered from 128 otters due to the absence of hard parts suitable for morphological analysis, prey remains being assigned to secondary prey items or due to identifications having a poor taxonomic resolution (Table S 5.1). Of the taxa detected, 22 were identified as vertebrates (11 to species level, eight to family, two to order and one to class) and one was identified as an invertebrate (family level describing crayfish, Astacidae).

5.4.2 DNA metabarcoding analysis

Sequencing yielded 17.6 million paired-end reads for the 16S library and 13.7 million for the COI library, which was reduced to 5 million for 16S and 1.1 million for COI following data processing (Fig. 5.2). Of the 300 otters screened, dietary data was recovered for 241 otters using 16S, with an average of 20,618 reads and 2.87 taxa per otter, and 149 using COI, with an average of 7,509 reads and 1.6 taxa per otter. Dietary data was not recovered in 42 otters due to poor amplification of DNA, DNA being assigned to non-food items or due to identifications having a poor taxonomic resolution (Tables S 5.2 and S 5.3). Reads were assigned to 54 vertebrate taxa (48 to species level, one to genus and four to family) in the 16S data (Table 5.1; Table S 5.2), whilst COI data was assigned 21 vertebrate taxa (18 to species level, one to genus and one to family; Table 5.1; Table S 5.3) and 15 invertebrate taxa (14 to species level and one to genus). Combined results from metabarcoding datasets produced 799 occurrences of prey from 70 taxa in 258 otters, with an average of 3.08 taxa per otter. There were 567 occurrences and 34 taxa only detected using 16S primers, 109
occurrences and 17 taxa only detected using COI primers and 123 occurrences and 18 taxa detected using both primer sets (Fig. S 5.2).

**Figure 5.2.** Total number of reads remaining in both 16S and COI datasets at each stage of the metabarcoding and data filtering process.

### 5.4.3 Comparison of methods

Dietary data were recovered for 268/300 otters in total; prey items were identified only by morphological analysis from 10 otters, only by metabarcoding from 96 otters and by both methods from 162 otters. Following removal of suspected secondary prey items (Tables S 5.1 to S 5.3), metabarcoding identified 20 taxa that were not detected using morphological analysis, 39 taxa were identified to a greater resolution by metabarcoding and 11 taxa were identified to the same taxonomic level using both methods (Fig. 5.3). Of the nine taxa ‘only’ identified by morphological analysis, all were identified by metabarcoding at a greater taxonomic resolution, e.g. where crayfish were identified only by morphological analysis, two species of crayfish were identified via metabarcoding (Fig. 5.3). Metabarcoding identified 528 prey item presences that were not detected using morphological analysis, 144 presences were detected at a greater resolution by metabarcoding and 122 were identified to the same taxonomic resolution using both methods (Fig. S 5.3). Morphological analysis detected 45 prey item presences that were not detected by metabarcoding, but only detected one presence to a greater taxonomic resolution (one metabarcoding identification of ‘rudd/roach’ was distinguished to ‘rudd’ using morphological analysis; Fig. S 5.3). Taxa that were identified by both methods were detected at a greater frequency of occurrence using
metabarcoding. The frequency of occurrence of each prey group differed with method of dietary analysis; based on morphological analysis, bullhead was the most frequently detected taxon (14%), followed by amphibian (12%) and stickleback (11%); based on metabarcoding, brown trout and stickleback were most frequently detected taxa (both at 37%), followed by eel (27%) and bullhead (23%) (Fig. 5.3).
Figure 5.3. Taxon presences in the diet of Eurasian otters (*Lutra lutra*) identified using morphological analysis of prey remains (orange) and DNA metabarcoding (blue) on faecal samples. Grey boxes show similar taxonomic groups from (a) fish and (b) other prey groups. Faecal samples were obtained from dead otters collected from across England and Wales from 2007 – 2016.
5.4.4 Dietary variation

Combining data from morphological analysis and metabarcoding increased the amount of dietary data recovered, therefore subsequent analyses to assess dietary variation (and investigate hypotheses iii - vi) were carried out on a combined dataset. Following aggregation of taxa into prey groups and removal of groups with less than three presences, data input consisted of 765 occurrences of prey from 26 groups (Table 5.1; Fig. 5.4) across 268 otters, with an average of 2.85 prey groups per otter. The most frequent prey groups in the diet of otters were stickleback (39%), brown trout (37%), eel (26%) and bullhead (24%). Model-based ordination showed that most prey groups cluster close together, suggesting most otters have a similar dietary composition (Fig. 5); although, marine and coastal prey (‘coastal fish’, ‘marine inverts’, ‘flatfish’ and ‘goby’) appeared to cluster closer together in both ordinations and Cyprinidae (‘roach/rudd’, ‘ide/dace’, ‘carp’ and bream ‘Abramis brama’) clustered together.
Figure 5.4. Presences of prey groups (grey) and the taxa that contributed to each prey group (orange) in the diet of Eurasian otters (*Lutra lutra*). Data were obtained by combining identifications made through morphological analysis of prey remains and DNA metabarcoding of faeces obtained from dead otters collected from across England and Wales from 2007 – 2016.
Figure 5.5. Model-based unconstrained residual ordination biplot for Eurasian otter (Lutra lutra) diet. Latent variable models using an unconstrained ordination with (a) no predictors and (b) a residual ordination after controlling for the effect of season, coastal proximity and longitude. Each number represents one otter and taxon labels represent prey items; numbers closer together represent otters with more similar diets, and taxon labels closer together represent prey items more likely to co-occur in the diet of otters. In both ordinations most prey items and otters cluster close together, showing no clear pattern in dietary variation. However, in the unconstrained ordination, marine and coastal fish cluster closer to each other and further away from other taxa and Cyprinidae cluster close to one another. Data were obtained by combining identifications from morphological analysis of prey remains and DNA metabarcoding of faeces obtained from dead otters collected across England and Wales between 2007 – 2016.
At the community level (i.e. changes in overall composition of otter diet rather than prey specific associations), distinct otter diets were significantly associated with season (MGLM: Dev = 65.46, p = 0.001), longitude (MGLM: Dev = 70.55, p = 0.001) and distance from the coast (MGLM: Dev = 80.59, p = 0.001). Most prey species were observed in every season, at all longitudes and all distances from the coast, however, subtle changes in occurrences of certain species drove changes in the composition of otter diets across these variables. Across the seasons, no prey groups had a higher frequency of occurrence in spring, but birds and brown trout occurred more frequently in summer, whilst Atlantic salmon, cyprinids, percids, loach and marine/estuarine prey occurred more frequently in autumn and winter, with crayfish occurring less in winter (Fig. 5.6). Longitudinal variation appeared to be primarily driven by greater frequency of occurrences for salmonids, amphibians and marine/estuarine prey in the west, with more cyprinids and percids occurring in the east (Fig. 5.7), whilst variation with distance from the coast was primarily driven by greater occurrences of marine/estuarine prey and eels near the coast and bullhead occurring more inland (Fig. 5.8).

Prey specific associations were only found for season and distance from the coast; presences of rudd/roach in the diet of otters were significantly associated with season (MGLM: Dev = 10.45, p = 0.035; Fig. S 5.4), with more occurring in the diet in winter, and the presence of eels significantly declined with increasing distance from the coast (MGLM: Dev = 15.54, p = 0.004; Fig. S 5.5), whilst bullhead presences significantly increased with distance from the coast (MGLM: Dev = 12.22, p = 0.02; Fig. S 5.5). No specific prey were associated with longitude and no significant associations were found between dietary variation of otters and sex, length, body condition, proportion of urban land-use, water habitat type or year.
Figure 5.6. Frequency of occurrence of prey items in the diet of Eurasian otters (Lutra lutra) in each season. Data were obtained by combining identifications made through morphological analysis of prey remains and DNA metabarcoding of faeces collected from dead otters across England and Wales from 2007 – 2016. The width of the upper boxes is proportional to the number of otters sampled from each season, the width of the lower boxes is proportional to the frequency of occurrence of each taxon in the diet of otters, and the width of each line connecting the upper and lower boxes is proportional to the number of otters from a season that consumed that prey item. Highlighted prey groups represent those with greater frequencies in summer (yellow), autumn (red), winter (blue) or autumn/winter (red/blue), and * shows specific prey groups that were significantly associated with season.
Figure 5.7. Frequency of occurrence of prey items in the diet of Eurasian otters (*Lutra lutra*) at different longitudes. Data were obtained by combining identifications made through morphological analysis of prey remains and DNA metabarcoding of faeces collected from dead otters across England and Wales between 2007 – 2016. The width of the lower boxes is proportional to the frequency of occurrence of each taxon in the diet of otters and the width of each line connecting the upper and lower boxes is proportional to the number of otters from a particular longitude that consumed that prey item. Highlighted prey groups represent those with greater frequencies in western regions (purple) or in eastern regions (yellow).
Figure 5.8. Frequency of occurrence of prey items in the diet of Eurasian otters (*Lutra lutra*) at different coastal proximities. Data were obtained by combining identifications made through morphological analysis of prey remains and DNA metabarcoding of faeces collected from dead otters across England and Wales between 2007 – 2016. The width of the lower boxes is proportional to the frequency of occurrence of each taxon in the diet of otters and the width of each line connecting the upper and lower boxes is proportional to the number of otters from a particular distance from the coast that consumed that prey item. Highlighted prey groups represent those with greater frequencies near the coast (purple) or inland (yellow), and * shows specific prey groups that were significantly associated with proximity to the coast.
5.5 Discussion

A broad range of prey were identified in the diet of otters in England and Wales, reflecting their generalist foraging behaviour and ability to take prey from a range of habitats. By utilising road killed otters, we were able to analyse dietary composition alongside biotic and abiotic data, revealing spatial and temporal dietary variation. These findings display the opportunistic foraging behaviours of otters and are likely to provide an indication of the variations in prey abundances across different habitats (Boyer et al. 2015; Deiner et al. 2017; Hawlitschek et al. 2018).

5.5.1 Comparison of methods

DNA metabarcoding detected a greater range and frequency of prey, and to a greater taxonomic resolution, than morphological analysis of prey remains. This allowed detection of easily digested prey (e.g. European river lamprey) and more presences of typically larger fish that may have only been partially consumed (e.g. brown trout, *Salmo trutta*). Some prey presences in some individuals were only detected through morphological analysis, which may occur through differential gut retention times (Carss and Parkinson 1996) resulting in prey remains surviving longer than DNA (Casper et al. 2007; Tollit et al. 2009). These findings align with previous comparison studies (Casper et al. 2007; Hope et al. 2014; Thalinger et al. 2016; Jeanniard-Du-Dot et al. 2017); however, where other studies found the identities of the most frequently occurring prey were the same between the two methods, our study found they differed (Fig. 3). Morphological analysis under-estimated frequently consumed prey (e.g. brown trout) and attributed a large proportion of the diet to lower frequency prey (e.g. loach), reflecting a finding by Lanszki et al. (2015) that less important food types are more frequently morphologically identified in faeces due to differential gut retention times of prey remains (Carss and Parkinson 1996; Carss and Nelson 1998). Choice of method is thus likely to affect ecological conclusions made from the data, and whilst prey were more likely to be detected using metabarcoding, a combined approach gave a more comprehensive description of otter diet.

5.5.2 Dietary Composition

Otters primarily predated on freshwater fish, with the most frequently consumed prey identified as stickleback, brown trout, eel and European bullhead (*Cottus gobio*). When freshwater fish are less available, otters will switch to alternative prey (e.g. Britton et al. 2006; Remonti et al. 2010; Almeida et al. 2012; Krawczyk et al. 2016), a similar behaviour to that which is exhibited by other generalist predators (e.g. Xu et al. 2012; Yeager et al. 2014; Rosenblatt et al. 2015; Tobajas et al. 2016; Spencer et al. 2017). In the current study, amphibians (predominantly common frog, *Rana temporaria*) were the
most frequent non-fish prey consumed, followed by waterfowl (predominantly common moorhen, *Gallinula chloropus*) and crayfish (predominantly the invasive signal crayfish, *Pacifastacus leniusculus*). Consumption of signal crayfish (and grass carp, *Ctenopharyngodon idella*) highlights the ability of otters to consume, and potentially assist biocontrol, of abundant invasive freshwater species. These results largely agree with previous studies, suggesting that the composition of otter diet may be reflective of prey abundances in Britain (Copp and Roche 2003; Miranda *et al.* 2008; Almeida *et al.* 2012).

Protected species (e.g. great crested newt, *Triturus cristatus*, and white-clawed crayfish, *Austropotamobius pallipes*; Stroud 2017) often have low abundances and are less likely to be encountered by otters, with otters more frequently taking the more common species within these groups. Predation on protected species only made up a small proportion of otter diet in this study, suggesting these are rare predation opportunities and are unlikely to threaten protected species persistence. An exception to this is the European eel, a critically endangered species with a declining population (Bark *et al.* 2007; Aprahamian and Walker 2008; ICES 2019). Eels have long been reported as the favoured prey of otters (Copp and Roche 2003; Britton *et al.* 2006; Miranda *et al.* 2008), but studies have found as eel populations decline so does predation by otters (Copp and Roche 2003; Almeida *et al.* 2012; Kruuk 2014; Moorhouse-Gann *et al.* 2020). Here we found otters are still frequently consuming eels regardless of their decline, this disparity between studies suggests further research is required into otter-eel dynamics and the threats otters may present to future eel recruitment. Otters were also observed to consume species stocked by fish farms (e.g. carp and rainbow trout, *Oncorhynchus mykiss*), which is a concern for anglers and aquaculture management, as well as a sources of risk for otters given their conflict with these parties (Vaclavikova *et al.* 2011; Poledňíková *et al.* 2013; Grant and Harrington 2015). Stocked fish did not constitute a large proportion of the diet though, therefore whilst otters may consume and impact stocked fish populations, they are more likely to consume wild counterparts, particularly smaller bodied fish such as bullhead (Britton *et al.* 2006; Grant and Harrington 2015; Lanszki *et al.* 2015; Lyach and Cech 2017).

### 5.5.3 Spatial variation

Greater frequencies of marine prey were observed in the diet of otters closer to the coast, reflecting the ability of otters to opportunistically consume prey from different habitats (Beja 1991; Jędrzejewska *et al.* 2001; Clavero *et al.* 2004; Reid *et al.* 2013; Krawczyk *et al.* 2016). Otters utilise marine prey to different extents, with individuals in the Scottish Isles specialising on marine prey (e.g. Kruuk and Moorhouse 1990; Watt...
1995) whilst coastal otters in mainland Britain and Europe consume these prey less frequently (Beja 1991; Heggberget and Moseid 1994; Clavero et al. 2004; Parry et al. 2011; Moorhouse-Gann et al. 2020). In this study, consumption of marine prey only constituted a small proportion of the diet, thus implying that most otters in England and Wales exploit marine species infrequently or not at all. As otter populations recover and expand their distribution, it is possible that exploitation of marine habitats will increase, either due to increased competition for freshwater prey, or as coastal individuals gain experience hunting marine prey. Proximity to the coast was also associated with prevalence in the diet of two of the most dominant prey; eel consumption declined and bullhead increased inland. Whilst bullhead are abundant in a variety of habitats (both upland and lowland; Tomlinson and Perrow 2003) eel abundances tend to decline with increasing distances from the tidal limit (Ibbotson et al. 2002), leading to otters switching prey as bullhead become more available than eels. Previous studies suggest otters switch from eel to common species, such as bullhead and trout, as eel populations decline (e.g. Almeida et al. 2012; Moorhouse-Gann et al. 2020), however, our observations suggest that despite declines, eel were still taken (slightly) more frequently than bullhead between 2007 and 2016.

Variation in otter diet with longitude reflected changing prey distributions, with Salmonidae consumed more frequently in the west and Cyprinidae and Percidae in the east, consistent with population densities of these families (e.g. Common carp, *Cyprinus carpio*: NBN atlas 2020b; European perch, *Perca fluviatilis*: NBN atlas 2020d, Atlantic salmon, *Salmo salar*: NBN atlas 2020a). These findings reflect the opportunistic foraging behaviour of otters, with individuals more likely to encounter and consume abundant prey. We also observed more amphibian and marine species being consumed by western otters, potentially suggesting greater reliance on these species as alternative prey, or greater availability in these regions (e.g. increased opportunity to feed on marine prey due to more coastline in western regions). Opportunistic foraging was further implied by the lack of dietary differences between otters from different aquatic habitat types, suggesting that otters are utilizing prey from a variety of habitats within their range, rather than focusing on the nearest habitat. There was also no association with the degree of urban / rural habitat in the otter vicinity, suggesting that neither availability of prey, or otter feeding behaviour, vary considerably where waterways pass through urban areas.

### 5.5.4 Temporal variation

We observed seasonal changes in the diet of otters, with greater frequencies of Percidae and Cyprinidae (particularly rudd, *Scardinius erythrophthalmus*, and roach,
*Rutilus rutilus*) consumed during autumn and winter. This is a similar finding to Grant and Harrington (2015) who suggested decreases in motility made Percidae and Cyprinidae prey easier to catch in winter. Otters also consumed more loach in autumn and primarily consumed Atlantic salmon in autumn and winter; this may reflect greater availability, with autumnal abundance peaks in loach (Hofmann and Fischer 2001) and Atlantic salmon migrating upstream to breed (Hendry and Cragg-Hine 2003; Everard 2013; Atlantic Salmon Trust 2018). Marine prey were also primarily consumed in autumn and winter, suggesting otters are more likely to use marine prey as an alternative resource during these seasons.

More birds occur in otter diets during summer and autumn, potentially reflecting opportunistic foraging on young birds that are vulnerable to predation as they leave the nest. Brown trout were also consumed in greater numbers in summer compared to the other seasons, potentially reflecting greater availability of trout during summer. Surprisingly, amphibians showed only a weak seasonal trend in otter diets, whereas many otter studies using morphological analysis have found distinct seasonal peaks in spring and winter (e.g. Clavero *et al.* 2005; Parry *et al.* 2015; Moorhouse-Gann *et al.* 2020). We found broadly similar frequencies across the seasons, with slightly higher frequencies in winter when amphibians are more vulnerable and spring when they aggregate for breeding (Beebee 2013). The weaker association may be due the improved detection of fish species found using metabarcoding thus altering the relative importance of amphibians during these months. Similarly, the invasive signal crayfish, whilst consumed, was not preferentially taken during a particular season nor comprised a large proportion of the diet. In Mediterranean regions, invasive crayfish (primarily red swamp crayfish, *Procambarus clarkii*) are an important dietary element for otters (Adrian and Delibes 1987; Beja 1996; Correia 2001; Barrientos *et al.* 2014) particularly during droughts when fish are less available. The lack of relationship between British otters and signal crayfish may be a consequence of greater environmental stability in temperate regions, providing otters with the opportunity to frequently consume fish species throughout the year.

We expected the diet of otters to reflect annual changes in prey populations and distributions as species undergo population increases or declines (Hayhow *et al.* 2019); however, no significant time trends were observed over the ten year study period. Earlier studies have found fewer eels in the diet of otters in line with eel population declines (Copp and Roche 2003; Almeida *et al.* 2012; Kruuk 2014; Moorhouse-Gann *et al.* 2020; respectively reporting years 1991 - 2000, 1970 - 2010, 2003 - 2013, and 1994 - 2010). The apparent consistency in eel predation shown by
the current study may reflect a stabilisation in eel populations in later years (2007 – 2016), albeit at lower abundances. We also expected to observe greater consumption of invasive species by otters over time as invasive species become more abundant with population increases (e.g. signal crayfish; Sibley et al. 2002; Holdich et al. 2014), yet invasive species contributed only a small proportion of otter diet. This may indicate a preference by otters for native species, as observed in Mediterranean otters (Blanco-Garrido et al. 2008), or lower abundance of invasive compared to native prey. However, as invasive species continue to undergo population expansions and become more available to otters, greater consumption may be observed (Balestrieri et al. 2013).

5.5.5 Biotic variation

Our data suggest that there were no demographic (i.e. sex, size or body condition) differences in the diet of otters. This contrasts a recent study by Moorhouse-Gann et al. (2020) who found an association between high value prey and body condition of otters. Whilst the discrepancy between studies may be due to the shorter time frame or smaller sample size investigated in this study, it may also reflect the difference in methods applied. It is possible that the increased frequency of higher quality prey species revealed by metabarcoding reflects detection of smaller (e.g. juvenile) prey individuals not distinguished morphologically. Although identified as high-quality species, such prey may represent relatively little nutritional gain. Whilst metabarcoding provides a greater insight into the species consumed by a predator, it cannot reveal the size or number of prey consumed (Deagle et al. 2013; Elbrecht and Leese 2015; Pawluczyk et al. 2015; Piñol et al. 2015; Hawlitschek et al. 2018; Mata et al. 2019), potentially overlooking an important aspect of demographic variation. For example, adult otters might consume primarily large trout whereas young otters might focus on small fry. Metabarcoding cannot differentiate between size or number of prey consumed, and although morphological analyses can, (Britton et al. 2006; Grant and Harrington 2015; Lyach and Cech 2017), it is extremely laborious and relies on particular hard parts being present within a sample (e.g. fish vertebrae used to estimate size), which may be misleading where, for example, otters have only consumed the flesh of prey and not hard parts (Ruiz-Olmo, Jiménez and Margalida 1998; Adámek et al. 2003; Kortan et al. 2007). These findings emphasise the need to choose and combine appropriate methods in order to interpret differences in prey consumed by otters of different demographic groups.
5.5.6 Limitations
Given the difficulty associated with accurate morphological identification of prey remains from faecal matter, and similar issues with DNA barcodes of closely-related prey, some identifications were not resolved to species level. Equally, the common reliance of metabarcoding on public reference databases can introduce errors resulting from misidentification of barcoded specimens, the presence of only partial sequences, or the omission of some species altogether (Gerwing et al. 2016; Zinger et al. 2019). Such issues with metabarcoding will likely be alleviated by ongoing initiatives to comprehensively barcode British fauna and flora (The Darwin Tree of Life 2020), after which the accuracy of these methods will further improve and fewer mis-identifications will be made (Hibert et al. 2013; Gerwing et al. 2016). It is also likely that some occurrences reflect secondary predation (Sheppard et al. 2005; Pompanon et al. 2012; Bowser et al. 2013; Galan et al. 2018), although DNA degradation prior to consumption (Kamenova et al. 2018; Nielsen et al. 2018) and the use of minimum sequence copy thresholds likely minimise this potential source of error in metabarcoding data. Methods employed here allow us to quantify frequency of occurrence, but because prey abundance data were not available at sufficient spatial or temporal resolution it was not possible to conduct comparative analyses of prey preference.

5.5.7 Conclusions
Metabarcoding provides a methodological advance for the study of generalist apex predator diets, providing greater precision for the identities and frequencies of species consumed compared to traditional morphological methods. We have shown DNA for metabarcoding can be successfully amplified from faeces collected from dead otters, providing a unique approach for studying the diet of a protected, elusive species across a range of spatio-temporal and biotic variables. Otters were shown to exploit a broad range of prey from different habitats, with dietary variation representing the adaptability of otters to both seasonal changes in prey availability and differences in prey distributions across the landscape. The dietary plasticity of otters observed here has likely aided the recovery of British populations (Van Baalen et al. 2001; Peers et al. 2014) and may increase resilience of these populations to future environmental stressors. Greater dietary resolution also provided an insight into prey population dynamics within the environment, supporting the use of metabarcoding studies of generalist predators to help guide biodiversity management, especially where surveying may be difficult (Deiner and Altermatt 2014; Boyer et al. 2015; Deiner et al. 2017; Hawlitschek et al. 2018).
5.6 Acknowledgements

This work was funded by KESS II (Knowledge Economy Skills Scholarship) and the Wildlife Trust of South and West Wales who partnered this project. Sample collection was conducted by Cardiff University Otter Project employees and placement students during post mortems of otter carcasses. Identification of morphological prey remains from faecal samples were performed by Sergio Bedmar Castillo.
5.7 Supplementary Information

5.7.1 Correlation between landscape variables

![Correlation Figure](image)

**Figure S 5.1.** Associations between spatial variables. Plot shows the correlation using a Pearson’s correlation coefficient. Spatial variables represent the landscape within a 10 km buffer of each individual otter: ‘Longitude’ is the position an otter was found at (center of the buffer) and ‘Altitude’ and ‘Slope’ are the mean values within each buffer. Blue circles represent positive correlations whilst red circles represent negative correlations. Darker colours and larger circles represent stronger correlations.

5.7.2 Scripts for analysing data in R

R Code used for analysing metabarcoding data acquired from Eurasian otter faecal samples. Code was run using R [version 3.6.0] and R studio [version 1.2.1335] (R Core Team 2019) and converted into document format using R markdown (Xie et al. 2018; Allaire et al. 2020). Executable code is presented in grey boxes.

### Load packages

```r
library("corrplot")
library("ggplot2")
library("plyr")
library("reshape2")
library("viridis")
library("scales")
```
library("RColorBrewer")
library("mvabund")
library("boral")
library("birpartite")

**Correlation plot**
Load data and extract necessary information

```r
buffer <- read.csv("Otter10kmVariables.csv", header = T)
summary(buffer)
colnames(buffer)
corrcheck <- buffer[, c(12, 19, 13, 9, 10)]
str(corrcheck)
```

Check correlation between landscape variables acquired from 10km buffers around each otter

```r
cor1 <- cor(corrcheck)
corrplot(cor1, type = "upper", order = "hclust", tl.srt = 45, tl.col = "black")
```

**Create heat charts to showing taxonomic identifications by each method**

*Compare metabarcoding 16S and COI identifications*

Load in data and format for plotting

```r
HeatChart <- read.csv("Combined Heat Chart.csv", header = T)
summary(HeatChart)
meltedHeatChart <- melt(HeatChart)
```

Plot heat chart

```r
ggchart <- ggplot(meltedHeatChart, aes(variable, y = reorder(Taxon, desc(Taxon)))) +
  geom_tile(aes(fill = value), colour = "white") +
  scale_fill_gradientn(na.value = "white", colours = viridis(3),
    values = rescale(c(1, 2, 3)),
    breaks = c(1, 2, 3), labels = c("16S Only", "COI Only", "16S and COI"),
    guide = "legend") +
  coord_fixed(ratio = 3, xlim = NULL, ylim = NULL, expand = TRUE) +
  theme(axis.title.y = element_blank(),
    axis.title.x = element_blank(),
    axis.text.x = element_blank(),
    axis.text.y = element_text(size = 9, colour = "black"),
    legend.position = "none")
ggchart
```
Compare identifications from morphological analysis of prey remains and metabarcoding

Load in data and format for plotting

```r
HeatMethods <- read.csv("HTSvsHP_HeatChart.csv", header = T)
HeatMethods$Group <- as.character(HeatMethods$Group)
HeatMethods$Group <- factor(HeatMethods$Group, levels=unique(HeatMethods$Group))
HeatMethods$Group <- factor(HeatMethods$Group, levels = rev(levels(HeatMethods$Group)))
summary(HeatMethods)
meltedHeatMethods <- melt(HeatMethods)
```

Plot heat chart with all data

```r
ggplot(meltedHeatMethods, aes(variable, Group)) %+% geom_tile(stat = "identity", aes(fill = value), colour = "white") %+% scale_fill_gradientn(na.value = "white", colours=viridis(4, direction = -1), values=rescale(c(1,2,3,4)), breaks = c(1,2,3,4), labels = c("Only molecular","Only morphological", "Molecular and morphological at same taxonomic level", "Molecular and morphological at different taxonomic level"), guide = "legend") %+% theme(legend.position="none", axis.title.x = element_blank(), axis.title.y = element_blank(), axis.text.x = element_blank(), axis.text.y = element_text(size=10, colour = "black"), plot.margin=unit(c(0.5,0.5,1,1),"cm"))
```

Load in data for just fish taxa and reformat

```r
HeatMethods <- read.csv("HTSvsHP_HeatChart_FishPrey.csv", header = T)
HeatMethods$Group <- as.character(HeatMethods$Group)
HeatMethods$Group <- factor(HeatMethods$Group, levels=unique(HeatMethods$Group))
HeatMethods$Group <- factor(HeatMethods$Group, levels = rev(levels(HeatMethods$Group)))
summary(HeatMethods)
meltedHeatMethods <- melt(HeatMethods)
```
Plot heat chart for just fish taxa

```r
ggplot(meltedHeatMethods, aes(variable, Group)) +
geom_tile(stat = "identity", aes(fill = value), colour = "white") +
scale_fill_gradientn(na.value = "white", colours=viridis(4, direction = -1),
values=rescale(c(1,2,3,4)),
breaks = c(1,2,3,4),
labels = c("Only molecular","Only morphological",
"Molecular and morphological at same taxonomic level",
"Molecular and morphological at different taxonomic level"), guide = "legend") +
theme(legend.position="none",
    axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    axis.text.x = element_blank(),
    axis.text.y = element_text(size=12, colour = "black"),
    plot.margin=unit(c(0.5,0.5,1,1),"cm"))
```

Load in data for non-fish taxa and reformat

```r
HeatMethods2 <- read.csv("HTSvsHP_HeatChart_AltPrey.csv", header = T)
HeatMethods2$Group <- as.character(HeatMethods2$Group)
HeatMethods2$Group <- factor(HeatMethods2$Group,
    levels=unique(HeatMethods2$Group))
HeatMethods2$Group <- factor(HeatMethods2$Group,
    levels = rev(levels(HeatMethods2$Group)))
summary(HeatMethods2)
meltedHHeatMethods2< - melt(HeatMethods2)
```

Plot heat chart for just fish taxa

```r
ggplot(meltedHHeatMethods2, aes(variable, Group)) +
geom_tile(stat = "identity", aes(fill = value), colour = "white") +
scale_fill_gradientn(na.value = "white", colours=viridis(4, direction = -1),
values=rescale(c(1,2,3,4)),
breaks = c(1,2,3,4),
labels = c("Only molecular","Only morphological",
"Molecular and morphological at same taxonomic level",
"Molecular and morphological at different taxonomic level"), guide = "legend") +
theme(legend.position="none",
    axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    axis.text.x = element_blank(),
    axis.text.y = element_text(size=12, colour = "black"),
    plot.margin=unit(c(0.5,0.5,1,1),"cm"))
```
Model based analysis of dietary data

Load in data and check format

```r
HTSHP2 <- read.csv("HTS+HPThreePlusNoInsectORMollusc.csv", header = T)
summary(HTSHP2)
colnames(HTSHP2)
rownames(HTSHP2) <- HTSHP2[,1]
rownames(HTSHP2)
str(HTSHP2)
```

Create object of only species consumed

```r
dietHTSHP2 <- mvabund(HTSHP2[,21:47])
```

Create model and check asumptions

```r
MVdietHTSHP2 <- manyglm(dietHTSHP2 ~ Sex + Size + Size:Sex + Scored.SMI + Year2 + Season + Sex:Season + Size:Season + long + lat + lat:long + KmRiverDist + WaterClass + Urban + Sex:WaterClass + Size:WaterClass + KmRiverDist:Sex + KmRiverDist:Size, family = binomial(link="cloglog"), data = HTSHP2)
plot(MVdietHTSHP2)
```

Conduct model simplification by stepwise deletion by AIC and extract significance of variables on the general prey composition and specific prey groups

```r
step(MVdietHTSHP2, test = "Chisq")
MVdietHTSHP2.2 <- manyglm(dietHTSHP2 ~ Season + long + KmRiverDist, family = binomial(link="cloglog"), data = HTSHP2)
plot(MVdietHTSHP2.2)
anovaMVdietHTSHP2.2 <- anova(MVdietHTSHP2.2, resamp = "montecarlo", test = "LR", p.uni="adjusted")
anovaMVdietHTSHP2.2
```

Bayesian Ordination And Regression Analysis (BORAL)

Create an object with only significant variables from the mvabund analysis and then run BORAL with and without this object

```
```
X <- as.data.frame(HTSHP2[, c(9,16,20)])
dietHTSHP2Boral <- boral(dietHTSHP2, family = "binomial", lv.control=list(num.lv=2))
dietHTSHP2Boral2 <- boral(dietHTSHP2, X=X, family = "binomial",
    lv.control=list(num.lv=2))

Check BORAL objects and plot ordinations (first need to run the lvspolt2 function code; O’Hara et al. 2016)

summary(dietHTSHP2Boral)
plot(dietHTSHP2Boral)
lvsplot(dietHTSHP2Boral, return.vals = T)
lvsplot2(dietHTSHP2Boral, alpha=0.5, main="", cols.lvs = "grey78", cols.coefs = "springgreen4", a=1.3, jitter = TRUE)

summary(dietHTSHP2Boral2)
plot(dietHTSHP2Boral2)
lvsplot(dietHTSHP2Boral2, return.vals = T)
lvsplot2(dietHTSHP2Boral2, alpha=0.5, main="", cols.lvs = "grey78", cols.coefs = "springgreen4", a=1.3, jitter = TRUE)

Bipartite plots
Load in data and check format

BipartiteHTSHP <- read.csv("Combined_HTS_HP_Bipartite.csv", header = T)
summary(BipartiteHTSHP)
rownames(BipartiteHTSHP) <- BipartiteHTSHP[, 1]
rownames(BipartiteHTSHP)
colnames(BipartiteHTSHP)

Plot data for season

plotweb(BipartiteHTSHP[,3:6], text.rot=90,
    col.high = c("darkolivegreen3", "gold2", "indianred2", "steelblue3"),
    bor.col.high = c("darkolivegreen3", "gold2", "indianred2", "steelblue3"),
    col.interaction = c("darkolivegreen3", "gold2", "indianred2", "steelblue3"),
    bor.col.interaction = c("darkolivegreen3", "gold2", "indianred2", "steelblue3"),
    bor.col.low = "ivory4", col.low = "ivory4", y.width.low = 0.1, high.xoff = F,
    low.y = 0.7, high.y = 1.7, labsize = 2.3, low.spacing = 0.0175,
    high.spacing = 0.15, method = "normal", text.low.col = "black")

Plot data for longitude

plotweb(BipartiteHTSHP[,7:272], text.rot = 90, col.high = viridis(263),
    bor.col.high = viridis(263), method = "normal", col.interaction = viridis(263),
Plot data for distance from the coast

```r
plotweb(BipartiteHTSHP[,273:531], text.rot = 90, col.high = viridis(256),
        bor.col.high = viridis(256), method = "normal", col.interaction = viridis(256),
        bor.col.interaction = viridis(256), high.spacing = 0.00145, low.spacing = 0.015,
        low.y = 0.7, high.y = 1, text.low.col = "black", labsize = 2.3)
```
### 5.7.3 Taxonomic identifications and filtering of dietary data

**Table S 5.1.** Taxa obtained from morphological analysis of prey remains in Eurasian otter (*Lutra lutra*) faecal samples. Latin names and common names are given for each taxon identified, along with whether the taxon was removed before data interpretation and the reason for removal of the taxon.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinopterygii</td>
<td>Fish</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>30</td>
</tr>
<tr>
<td>Anguilla anguilla</td>
<td>European eel</td>
<td>N</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Anura</td>
<td>Anura</td>
<td>N</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>Astacidae</td>
<td>Crayfish</td>
<td>N</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>Aves</td>
<td>Bird</td>
<td>N</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>Barbatula barbatula</td>
<td>Stone loach</td>
<td>N</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Coleoptera</td>
<td>Beetle</td>
<td>Y</td>
<td>Suspected secondary predation (following comparison with metabarcoding data)</td>
<td>6</td>
</tr>
<tr>
<td>Cottus gobio</td>
<td>European Bullhead</td>
<td>N</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>Cyprinidae</td>
<td>Carp</td>
<td>N</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Cyprinidae</td>
<td>No ID Cyprinid</td>
<td>N</td>
<td></td>
<td>26</td>
</tr>
<tr>
<td>Diptera</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (following comparison with metabarcoding data)</td>
<td>2</td>
</tr>
<tr>
<td>Esox lucius</td>
<td>Northern pike</td>
<td>N</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Gasterosteidae</td>
<td>Stickleback</td>
<td>N</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Gasterosteus aculeatus</td>
<td>Three-spined stickleback</td>
<td>N</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>Gastropoda</td>
<td>Snail</td>
<td>Y</td>
<td>Suspected secondary predation (following comparison with metabarcoding data)</td>
<td>9</td>
</tr>
<tr>
<td>Insecta</td>
<td>Insect</td>
<td>Y</td>
<td>Suspected secondary predation (following comparison with metabarcoding data)</td>
<td>41</td>
</tr>
<tr>
<td>Invertebrate</td>
<td>Invertebrate</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>2</td>
</tr>
<tr>
<td>Mammalia</td>
<td>Mammal</td>
<td>Y</td>
<td>Not distinguishable from otter</td>
<td>2</td>
</tr>
<tr>
<td>Mollusca</td>
<td>Mollusc</td>
<td>Y</td>
<td>Suspected secondary predation (following comparison with metabarcoding data)</td>
<td>3</td>
</tr>
<tr>
<td>Not assigned</td>
<td>Unidentifiable remains</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>4</td>
</tr>
<tr>
<td>Perca fluviatilis</td>
<td>European perch</td>
<td>N</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Percidae</td>
<td>Percidae</td>
<td>N</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Phoxinus phoxinus</td>
<td>Minnow</td>
<td>N</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>Pleuronectiformes</td>
<td>Flatfish</td>
<td>N</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Pomatoschistus</td>
<td>Goby</td>
<td>N</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Pangitius pungitius</td>
<td>Nine-spined stickleback</td>
<td>N</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Ranidae</td>
<td>Frog</td>
<td>N</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>Rutilus rutilus</td>
<td>Roach</td>
<td>N</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Salamandridae</td>
<td>Newt</td>
<td>N</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Salmo</td>
<td>Salmo</td>
<td>N</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>Salmo trutta</td>
<td>Brown trout</td>
<td>N</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Tinca tinca</td>
<td>Trench</td>
<td>N</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Trichoptera</td>
<td>Caddisfly</td>
<td>Y</td>
<td>Suspected secondary predation</td>
<td>1</td>
</tr>
<tr>
<td>Vertebrata</td>
<td>Vertebrate</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>10</td>
</tr>
</tbody>
</table>
Table S 5.2. Taxa obtained from sequencing Eurasian otter (*Lutra lutra*) faecal samples using 16S primers FN2199 (5'- yayaagcagaagacct -3') and R8B7 (5'- ttatcctrggtarctgg -3') (modified from Deagle *et al.* 2009). Latin names and common names are given for each taxon identified following bioinformatic analysis, along with whether the taxon was removed before data interpretation and reason for removal of the taxon. Total read counts and presences per taxa shown were calculated following bioinformatic analysis and artefact removal.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Cyprinus carpio 'mirror' x Cyprinus carpio 'singuonensis') x Carassius auratus red var</td>
<td>Carp hybrid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>147061</td>
<td>88</td>
</tr>
<tr>
<td>Abramis brama</td>
<td>Common bream</td>
<td>N</td>
<td></td>
<td>47656</td>
<td>8</td>
</tr>
<tr>
<td>Anas platyrhynchos</td>
<td>Mallard</td>
<td>N</td>
<td></td>
<td>104851</td>
<td>10</td>
</tr>
<tr>
<td>Anguilla anguilla</td>
<td>European eel</td>
<td>N</td>
<td></td>
<td>541334</td>
<td>83</td>
</tr>
<tr>
<td>Anser anser</td>
<td>Greylag goose</td>
<td>N</td>
<td></td>
<td>200</td>
<td>1</td>
</tr>
<tr>
<td>Armoglossus laterna</td>
<td>Mediterranean scaldfish</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aythya ferina</td>
<td>Common Pochard</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aythya fuligula</td>
<td>Tufted duck</td>
<td>N</td>
<td></td>
<td>527</td>
<td>1</td>
</tr>
<tr>
<td>Barbatula barbatula</td>
<td>Stone loach</td>
<td>N</td>
<td></td>
<td>124226</td>
<td>31</td>
</tr>
<tr>
<td>Barbus barbus</td>
<td>Common barbel</td>
<td>N</td>
<td></td>
<td>467</td>
<td>1</td>
</tr>
<tr>
<td>Bilateria</td>
<td>Bilateria</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bufo bufo</td>
<td>Common toad</td>
<td>N</td>
<td></td>
<td>4541</td>
<td>3</td>
</tr>
<tr>
<td>Buglossidium luteum</td>
<td>Solenette</td>
<td>Y</td>
<td>Mock Community (potential contamination)</td>
<td>1535</td>
<td>3</td>
</tr>
<tr>
<td>Callionymus lyra</td>
<td>Common dragonet</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Canis lupus familiaris</td>
<td>Domestic dog</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Carassius auratus</td>
<td>Goldfish</td>
<td>N</td>
<td></td>
<td>37377</td>
<td>4</td>
</tr>
<tr>
<td>Carassius carassius</td>
<td>Crucian carp</td>
<td>N</td>
<td></td>
<td>25101</td>
<td>5</td>
</tr>
<tr>
<td>Chelidonichthys lucernus</td>
<td>Tub gurnard</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Chelon labrosus or Mugil cephalus or Liza aurata or Oedalechilus labeo</td>
<td>Mullet</td>
<td>N</td>
<td></td>
<td>1242</td>
<td>1</td>
</tr>
<tr>
<td>Cobitis taenia</td>
<td>Spined loach</td>
<td>N</td>
<td></td>
<td>1886</td>
<td>1</td>
</tr>
<tr>
<td>Coccothraustes coccothraustes</td>
<td>Hawfinch</td>
<td>Y</td>
<td>Contamination</td>
<td>2092</td>
<td>3</td>
</tr>
<tr>
<td>Columba palumbus</td>
<td>Common wood pigeon</td>
<td>N</td>
<td></td>
<td>520</td>
<td>2</td>
</tr>
<tr>
<td>Conger conger</td>
<td>Conger eel</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Corvus monedula monedula</td>
<td>Jackdaw</td>
<td>N</td>
<td></td>
<td>13880</td>
<td>1</td>
</tr>
<tr>
<td>Taxon</td>
<td>Common Name</td>
<td>Remove</td>
<td>Reason for removing</td>
<td>Total Read Count</td>
<td>Total Presences</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>------------------------</td>
<td>--------</td>
<td>---------------------------------------</td>
<td>------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Cottus gobio</td>
<td>European bullhead</td>
<td>N</td>
<td></td>
<td>1270682</td>
<td>72</td>
</tr>
<tr>
<td>Ctenopharyngodon idella</td>
<td>Grass carp</td>
<td>N</td>
<td></td>
<td>272</td>
<td>1</td>
</tr>
<tr>
<td>Cygnus olor</td>
<td>Mute swan</td>
<td>N</td>
<td></td>
<td>10441</td>
<td>1</td>
</tr>
<tr>
<td>Cyprinus carpio</td>
<td>Common carp</td>
<td>N</td>
<td></td>
<td>193457</td>
<td>28</td>
</tr>
<tr>
<td>Dicentrarchus labrax</td>
<td>European bass</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Echiichthys vipera</td>
<td>Lesser weever</td>
<td>N</td>
<td></td>
<td>2223</td>
<td>1</td>
</tr>
<tr>
<td>Esox lucius</td>
<td>Northern pike</td>
<td>N</td>
<td></td>
<td>100393</td>
<td>11</td>
</tr>
<tr>
<td>Eupercaria</td>
<td>Eupercaria</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Eutheria</td>
<td>Eutheria</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fulica atra</td>
<td>Eurasian coot</td>
<td>N</td>
<td></td>
<td>8982</td>
<td>1</td>
</tr>
<tr>
<td>Gadus morhua or Pollachius pollachius</td>
<td>Atlantic cod or European pollock</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gallinula chloropus</td>
<td>Common moorhen</td>
<td>N</td>
<td></td>
<td>192881</td>
<td>18</td>
</tr>
<tr>
<td>Gallus gallus</td>
<td>Chicken</td>
<td>N</td>
<td></td>
<td>1301</td>
<td>2</td>
</tr>
<tr>
<td>Gasterosteus aculeatus</td>
<td>Three-spined stickleback</td>
<td>N</td>
<td></td>
<td>434184</td>
<td>98</td>
</tr>
<tr>
<td>Gobio gobio</td>
<td>Gudgeon</td>
<td>N</td>
<td></td>
<td>77466</td>
<td>12</td>
</tr>
<tr>
<td>Gymnocephalus cernua</td>
<td>Ruffe</td>
<td>N</td>
<td></td>
<td>4574</td>
<td>3</td>
</tr>
<tr>
<td>Hymenocephalus striatissimun</td>
<td>Hymenocephalus striatissimun</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Larus argentatus</td>
<td>European herring gull</td>
<td>N</td>
<td></td>
<td>4380</td>
<td>1</td>
</tr>
<tr>
<td>Lepidopsetta bilineata</td>
<td>Rock sole</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lepidorhombus whiffiagonis</td>
<td>Megrim</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Leptostichaeus pumilus</td>
<td>Neck banded blenny</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Leuciscus idus or Leuciscus leuciscus</td>
<td>Ide or Common dace</td>
<td>N</td>
<td></td>
<td>5384</td>
<td>2</td>
</tr>
<tr>
<td>Leuciscus leuciscus</td>
<td>Common dace</td>
<td>N</td>
<td></td>
<td>1805</td>
<td>2</td>
</tr>
<tr>
<td>Limanda limanda</td>
<td>Common dab</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lissotriton helveticus</td>
<td>Palmate newt</td>
<td>N</td>
<td></td>
<td>93381</td>
<td>8</td>
</tr>
<tr>
<td>Lissotriton vulgaris</td>
<td>Smooth newt</td>
<td>N</td>
<td></td>
<td>17612</td>
<td>5</td>
</tr>
<tr>
<td>Lutra lutra</td>
<td>Eurasian otter</td>
<td>Y</td>
<td>Predator</td>
<td>4941349</td>
<td>211</td>
</tr>
<tr>
<td>Lutrinae</td>
<td>Otter</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table S 5.2. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammalia</td>
<td>Mammal</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Melanogrammus aeglefinus</td>
<td>Haddock</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Meleagris gallopavo</td>
<td>Wild Turkey</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Merlangius merlangus</td>
<td>Whiting</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Merlangius merlangus or Gadus morhua</td>
<td>Whiting or Atlantic cod</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Microchirus vaniegatus</td>
<td>Thickback sole</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mullus surmuletus</td>
<td>Striped red mullet</td>
<td>Y</td>
<td>Mock Community (potential contamination)</td>
<td>334</td>
<td>1</td>
</tr>
<tr>
<td>Mustela putorius furo</td>
<td>Ferret</td>
<td>Y</td>
<td>Contamination</td>
<td>422</td>
<td>1</td>
</tr>
<tr>
<td>Mustelus asterias</td>
<td>Starry smooth-hound</td>
<td>Y</td>
<td>Mock Community</td>
<td>53451</td>
<td>7</td>
</tr>
<tr>
<td>Not assigned</td>
<td>Not assigned</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oncorhyncus mykiss</td>
<td>Rainbow trout</td>
<td>N</td>
<td></td>
<td>82972</td>
<td>16</td>
</tr>
<tr>
<td>Parophrys vetulus</td>
<td>English sole</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Perca fluviatilis</td>
<td>European perch</td>
<td>N</td>
<td></td>
<td>49608</td>
<td>15</td>
</tr>
<tr>
<td>Phalacrocoryx carbo</td>
<td>Great cormorant</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Phoxinus phoxinus</td>
<td>Common minnow</td>
<td>N</td>
<td></td>
<td>82389</td>
<td>30</td>
</tr>
<tr>
<td>Platichthys stellatus</td>
<td>Starry flounder</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pleuronectes platessa</td>
<td>European plaice</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pleuronectidae</td>
<td>Righteye flounders</td>
<td>N</td>
<td></td>
<td>129979</td>
<td>14</td>
</tr>
<tr>
<td>Pollachius pollachi</td>
<td>European pollock</td>
<td>Y</td>
<td>Mock Community</td>
<td>151386</td>
<td>11</td>
</tr>
<tr>
<td>Pomatoschistus microps</td>
<td>Common goby</td>
<td>N</td>
<td></td>
<td>45665</td>
<td>10</td>
</tr>
<tr>
<td>Pomatoschistus minutus</td>
<td>Sand goby</td>
<td>N</td>
<td></td>
<td>38610</td>
<td>6</td>
</tr>
<tr>
<td>Carassius gibelio</td>
<td>Prussian carp</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pungitius laevis</td>
<td>Smoothtail nine-spined stickleback</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pungitius pungitus</td>
<td>Nine-spined stickleback</td>
<td>N</td>
<td></td>
<td>62857</td>
<td>16</td>
</tr>
<tr>
<td>Raja</td>
<td>Skate</td>
<td>Y</td>
<td>Mock Community</td>
<td>115659</td>
<td>11</td>
</tr>
<tr>
<td>Rajidae</td>
<td>Skate</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rana temporaria</td>
<td>Common frog</td>
<td>N</td>
<td></td>
<td>28626</td>
<td>12</td>
</tr>
<tr>
<td>Rattus norvegicus</td>
<td>Brown rat</td>
<td>N</td>
<td></td>
<td>486</td>
<td>1</td>
</tr>
</tbody>
</table>
Table S 5.2. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhodeus amarus</td>
<td>European bitterling</td>
<td>N</td>
<td></td>
<td>2732</td>
<td>2</td>
</tr>
<tr>
<td>Rutilus rutilus</td>
<td>Common roach</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Salmo salar</td>
<td>Atlantic salmon</td>
<td>N</td>
<td></td>
<td>32829</td>
<td>11</td>
</tr>
<tr>
<td>Salmo trutta</td>
<td>Brown trout</td>
<td>N</td>
<td></td>
<td>835809</td>
<td>116</td>
</tr>
<tr>
<td>Salmonidae</td>
<td>Salmonid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sardina pilchardus</td>
<td>European pilchard</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Scardinius erythropthalmus or Rutilus rutilus</td>
<td>Common rudd or Common roach</td>
<td>N</td>
<td></td>
<td>196348</td>
<td>25</td>
</tr>
<tr>
<td>Scardinius erythropthalmus</td>
<td>Common rudd</td>
<td>N</td>
<td></td>
<td>100114</td>
<td>12</td>
</tr>
<tr>
<td>Scomber scombrus</td>
<td>Atlantic mackerel</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Scophthalmus</td>
<td>Turbot</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Scophthalmus rhombus</td>
<td>Brill</td>
<td>N</td>
<td></td>
<td>15827</td>
<td>2</td>
</tr>
<tr>
<td>Scyliorhinus canicula</td>
<td>Small-spotted catshark</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Solea solea</td>
<td>Common sole</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sparus aurata</td>
<td>Gilt-head Bream</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sprattus sprattus</td>
<td>European sprat</td>
<td>Y</td>
<td>Mock Community</td>
<td>5646</td>
<td>7</td>
</tr>
<tr>
<td>Squalius cephalus or Leuciscus leuciscus or Rutilus rutilus</td>
<td>Chub or Common dace or Common roach</td>
<td>N</td>
<td></td>
<td>11395</td>
<td>3</td>
</tr>
<tr>
<td>Sylvia currucosa</td>
<td>Lesser whitethroat</td>
<td>N</td>
<td></td>
<td>208</td>
<td>1</td>
</tr>
<tr>
<td>Tachybaptus ruficollis</td>
<td>Little grebe</td>
<td>N</td>
<td></td>
<td>24095</td>
<td>3</td>
</tr>
<tr>
<td>Taurulus bubalis</td>
<td>Long-spined sea scorpion</td>
<td>N</td>
<td></td>
<td>10898</td>
<td>3</td>
</tr>
<tr>
<td>Thymallus thymallus</td>
<td>Grayling</td>
<td>N</td>
<td></td>
<td>12409</td>
<td>3</td>
</tr>
<tr>
<td>Tinca tinca</td>
<td>Tench</td>
<td>N</td>
<td></td>
<td>155681</td>
<td>15</td>
</tr>
<tr>
<td>Trachurus trachurus</td>
<td>Atlantic horse mackerel</td>
<td>Y</td>
<td>Mock Community (potential contamination)</td>
<td>1109</td>
<td>3</td>
</tr>
<tr>
<td>Triturus cristatus</td>
<td>Great crested newt</td>
<td>N</td>
<td></td>
<td>10833</td>
<td>3</td>
</tr>
<tr>
<td>Zeus faber</td>
<td>John dory</td>
<td>Y</td>
<td>Mock Community</td>
<td>202250</td>
<td>12</td>
</tr>
</tbody>
</table>
Table S 5.3. Taxa obtained from sequencing Eurasian otter (*Lutra lutra*) faecal samples using COI primers Mod_mCOIintF (5' - ggwacwggwtgaacwgtwaycc -3') (modified from Leray *et al.* 2013) and HCO-2198 (5' - taaacctcaggtgacaaaaatatca -3') (Folmer *et al.* 1994). Latin names and common names are given for each taxon identified following bioinformatic analysis, along with whether the taxon was removed before data interpretation and the reason for removal of the taxon. Total read counts and presences per taxa shown were calculated following bioinformatic analysis and artefact removal.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>NA</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>3649618</td>
<td>109</td>
</tr>
<tr>
<td>Abramis brama orientalis</td>
<td>Common bream</td>
<td>N</td>
<td></td>
<td>15290</td>
<td>4</td>
</tr>
<tr>
<td>Abrostola tripertita</td>
<td>Moth</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>1138</td>
<td>1</td>
</tr>
<tr>
<td>Acilius sulcatus</td>
<td>Water beetle</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Acrolepia autumnilletia</td>
<td>Moth</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Agapetus fuscipes</td>
<td>Caddisfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>216</td>
<td>1</td>
</tr>
<tr>
<td>Allacma fusca</td>
<td>Springtail</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>533</td>
<td>1</td>
</tr>
<tr>
<td>Alloteuthis sp. BOLD:AAB2767</td>
<td>Squid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Anax imperator</td>
<td>Emperor dragonfly</td>
<td>N</td>
<td></td>
<td>144</td>
<td>2</td>
</tr>
<tr>
<td>Anguillicola crassus</td>
<td>Eel parasitic nematoad</td>
<td>Y</td>
<td>Parasite</td>
<td>134</td>
<td>1</td>
</tr>
<tr>
<td>Anodonta anatina</td>
<td>Duck mussel</td>
<td>N</td>
<td></td>
<td>432</td>
<td>1</td>
</tr>
<tr>
<td>Anthaxia istiana</td>
<td>Beetle</td>
<td>Y</td>
<td>Foreign</td>
<td>337</td>
<td>1</td>
</tr>
<tr>
<td>Aphidiae sp. BOLD-2016</td>
<td>Aphid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aphrodes marakovi</td>
<td>Spittlebug</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aphrophora</td>
<td>Spittlebug</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apis mellifera mellifera</td>
<td>Honey bee</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aporrectodea</td>
<td>Earthworm</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Arcitalius domeni</td>
<td>Landhopper</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>1943</td>
<td>1</td>
</tr>
<tr>
<td>Arion flagellus</td>
<td>Durham slug</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Arion owenii</td>
<td>Tawny soil slug</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Arthropoda</td>
<td>Arthropod</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Asellus aquaticus</td>
<td>Waterlouse</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>68111</td>
<td>17</td>
</tr>
<tr>
<td>Aspergillus versicolor</td>
<td>Fungi</td>
<td>Y</td>
<td>Not Animal Kingdom</td>
<td>228</td>
<td>1</td>
</tr>
<tr>
<td>Asterias rubens</td>
<td>Common starfish</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aurelia</td>
<td>Jellyfish</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Austropotamobius pallipes</td>
<td>White-clawed crayfish</td>
<td>N</td>
<td></td>
<td>12951</td>
<td>3</td>
</tr>
</tbody>
</table>
Table S 5.3. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bætis</td>
<td>Mayfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>6676</td>
<td>7</td>
</tr>
<tr>
<td>Bætis rhodani</td>
<td>Dark olive mayfly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bætis vemos</td>
<td>Medium olive mayfly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bos</td>
<td>Cow</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Brachionus calyciflorus</td>
<td>Rotifer</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bulweria bulwerii</td>
<td>Bulwers petrel</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Byturnus tomentosus</td>
<td>Raspberry beetle</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>971</td>
<td>1</td>
</tr>
<tr>
<td>Caillonymus lyra</td>
<td>Common dragonet</td>
<td>Y</td>
<td>Mock Community</td>
<td>17688</td>
<td>13</td>
</tr>
<tr>
<td>Calliphora vicina</td>
<td>Blow fly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Calliphora vomitoria</td>
<td>Blue bottle fly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>232</td>
<td>1</td>
</tr>
<tr>
<td>Calyptratae</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>737</td>
<td>3</td>
</tr>
<tr>
<td>Canthocamptidae sp. BOLD:ACJ8158</td>
<td>Copepod</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>177</td>
<td>1</td>
</tr>
<tr>
<td>Carassius carassius</td>
<td>Crucian carp</td>
<td>N</td>
<td></td>
<td>1364</td>
<td>2</td>
</tr>
<tr>
<td>Chaetogaster diastrophus</td>
<td>Annelid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Chaetopteryx</td>
<td>Caddisfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>140</td>
<td>1</td>
</tr>
<tr>
<td>Chironomidae sp. PA2_1</td>
<td>Chironomid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Chironomus &lt;subgenus&gt;</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>Chironomus luridus</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>1217</td>
<td>2</td>
</tr>
<tr>
<td>Chironomus nudibasis</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>87</td>
<td>1</td>
</tr>
<tr>
<td>Chironomus palidivittatus</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>97</td>
<td>1</td>
</tr>
<tr>
<td>Chironomus riparius</td>
<td>Chironomid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Chrysaora hyoscella</td>
<td>Compass jellyfish</td>
<td>Y</td>
<td>Mock Community</td>
<td>4765</td>
<td>13</td>
</tr>
<tr>
<td>Cladosporium herbarum</td>
<td>Fungus</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cobitis taenia</td>
<td>Spined loach</td>
<td>N</td>
<td></td>
<td>603</td>
<td>1</td>
</tr>
<tr>
<td>Colletotrichum coccodes</td>
<td>Fungi</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Conger conger</td>
<td>Conger eel</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Corophium volutator</td>
<td>Amphipod</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>2386</td>
<td>1</td>
</tr>
<tr>
<td>Corystes cassivelaunus</td>
<td>Sand crab</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cottus</td>
<td>European bullhead (relabel)</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table S 5.3. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cottus microstomus</td>
<td>European bullhead (relabel)</td>
<td>N</td>
<td></td>
<td>68</td>
<td>1</td>
</tr>
<tr>
<td>Cottus perifretum</td>
<td>European bullhead</td>
<td>N</td>
<td></td>
<td>133858</td>
<td>32</td>
</tr>
<tr>
<td>Crangonyx floridanus</td>
<td>Amphipod (invasive)</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>5671</td>
<td>1</td>
</tr>
<tr>
<td>Crennocia irrorata</td>
<td>Caddisfly</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cycloidae sp. BOLD:AAG9780</td>
<td>Copepod</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>195</td>
<td>1</td>
</tr>
<tr>
<td>Cyclops strenuus</td>
<td>Copepod</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>2852</td>
<td>1</td>
</tr>
<tr>
<td>Cyclostephanos sp. WTC16</td>
<td>Diatom</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cygnus olor</td>
<td>Mute swan</td>
<td>N</td>
<td></td>
<td>18856</td>
<td>1</td>
</tr>
<tr>
<td>Cyprinidae</td>
<td>Cyprinid</td>
<td>N</td>
<td></td>
<td>5722</td>
<td>2</td>
</tr>
<tr>
<td>Cypriniformes</td>
<td>Cyprinid</td>
<td>N</td>
<td></td>
<td>204</td>
<td>1</td>
</tr>
<tr>
<td>Daphnia curvirostris</td>
<td>Daphnia</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>1881</td>
<td>1</td>
</tr>
<tr>
<td>Demodex folliculorum</td>
<td>Mite on human skin</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Deroceras</td>
<td>Smooth land slug</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Deroceras sp. BOLD:AAI9663</td>
<td>Smooth land slug</td>
<td>N</td>
<td></td>
<td>182</td>
<td>1</td>
</tr>
<tr>
<td>Didymella pinodes</td>
<td>Fungi</td>
<td>Y</td>
<td>Not Animal Kingdom</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>Dinocras cephalotes</td>
<td>Stoneflies</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Diphyllobothrium dendriticum</td>
<td>Tapeworm</td>
<td>Y</td>
<td></td>
<td>1951</td>
<td>1</td>
</tr>
<tr>
<td>Dissotrocha macrostyla</td>
<td>Rotifer</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dromius quadrimaculatus</td>
<td>Ground beetle</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>902</td>
<td>1</td>
</tr>
<tr>
<td>Drusus annulatus</td>
<td>Caddisfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>110</td>
<td>1</td>
</tr>
<tr>
<td>Ecdyonurus sp. EC-37-FR(MV)</td>
<td>Mayfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>690</td>
<td>1</td>
</tr>
<tr>
<td>Echiichthys vipera</td>
<td>Lesser weaver</td>
<td>N</td>
<td></td>
<td>1055</td>
<td>1</td>
</tr>
<tr>
<td>Eiseniella tatraedra</td>
<td>Earthworm</td>
<td>N</td>
<td></td>
<td>351</td>
<td>1</td>
</tr>
<tr>
<td>Enallagma cyathigerum</td>
<td>Common blue damselfly</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Endochironomus albipennis</td>
<td>Chironomid</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Enterococcus faecalis</td>
<td>Bacteria</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Erpobdella octoculata</td>
<td>Leech</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>158</td>
<td>1</td>
</tr>
<tr>
<td>Esox lucius</td>
<td>Northern Pike</td>
<td>Y</td>
<td>No Reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Euchlanis dilatata</td>
<td>Rotifer</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>343</td>
<td>2</td>
</tr>
</tbody>
</table>
Table S 5.3. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eudonia mercurella</td>
<td>Moth</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>237</td>
<td>1</td>
</tr>
<tr>
<td>Eurycercus lamellatus</td>
<td>Water flea</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>28728</td>
<td>3</td>
</tr>
<tr>
<td>Fulica atra</td>
<td>Eurasian coot</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gammarus duebeni</td>
<td>Gammarus</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gammarus pulex</td>
<td>Amphipod</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>14422</td>
<td>4</td>
</tr>
<tr>
<td>Gasterosteus</td>
<td>Three-spined stickleback (relabel)</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gasterosteus aculeatus</td>
<td>Three-spined stickleback</td>
<td>N</td>
<td></td>
<td>276541</td>
<td>40</td>
</tr>
<tr>
<td>Glomastix murorum var. felina</td>
<td>Fungi</td>
<td>Y</td>
<td>Not Animal Kingdom</td>
<td>287</td>
<td>1</td>
</tr>
<tr>
<td>Glomeris marginata</td>
<td>Pill millipede</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>178</td>
<td>1</td>
</tr>
<tr>
<td>Gymnocephalus cernua</td>
<td>Ruffe</td>
<td>N</td>
<td></td>
<td>3845</td>
<td>2</td>
</tr>
<tr>
<td>Habroplebia fusca</td>
<td>Mayfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>91</td>
<td>1</td>
</tr>
<tr>
<td>Habrotricha elisa elusa</td>
<td>Rotifer</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Halesus digitatus</td>
<td>Caddisfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>169</td>
<td>1</td>
</tr>
<tr>
<td>Harmothoe glabra</td>
<td>Annelid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hediste diversicolor</td>
<td>Ragworm</td>
<td>N</td>
<td></td>
<td>217</td>
<td>1</td>
</tr>
<tr>
<td>Heptageniidae</td>
<td>Mayfly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Holometabola</td>
<td>Endopterygota</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>Human</td>
<td>Y</td>
<td>Contamination</td>
<td>12835</td>
<td>28</td>
</tr>
<tr>
<td>Hydactinia echinata</td>
<td>Hydrozoa</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hydropsyche instabilis</td>
<td>Caddisfly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hydropsyche siltalai</td>
<td>Caddisfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>1449</td>
<td>2</td>
</tr>
<tr>
<td>Hypogastrudidae sp. BOLD:ACE1775</td>
<td>Springtail</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>334</td>
<td>1</td>
</tr>
<tr>
<td>Ichneumonidae</td>
<td>Parasitoid wasp</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>180</td>
<td>1</td>
</tr>
<tr>
<td>invertebrate environmental sample</td>
<td>Invert</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>133</td>
<td>1</td>
</tr>
<tr>
<td>Ischnura elegans</td>
<td>Blue-tailed damselfly</td>
<td>N</td>
<td></td>
<td>1014</td>
<td>1</td>
</tr>
<tr>
<td>Isoperla grammatica</td>
<td>Yellow sally stonefly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>3045</td>
<td>1</td>
</tr>
<tr>
<td>Isopoda sp. BOLD:AAH4103</td>
<td>Isopod</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>802</td>
<td>1</td>
</tr>
<tr>
<td>Lampetra fluviatilis</td>
<td>European river lamprey</td>
<td>N</td>
<td></td>
<td>1234</td>
<td>3</td>
</tr>
<tr>
<td>Leuckartiara octona</td>
<td>Hydrozoa</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Leucothoe incisa</td>
<td>Amphipod</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table S 5.3. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limnephilus lunatus</td>
<td>Caddisfly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Limnodynastes claperedianus</td>
<td>Worm</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lissotriton helveticus helveticus</td>
<td>Palmate newt</td>
<td>N</td>
<td></td>
<td>332</td>
<td>2</td>
</tr>
<tr>
<td>Lochmaea capreae</td>
<td>Leaf beetle</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>206</td>
<td>1</td>
</tr>
<tr>
<td>Lucilia</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>553</td>
<td>1</td>
</tr>
<tr>
<td>Lucilia caesar</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>153</td>
<td>1</td>
</tr>
<tr>
<td>Lumbriicidae</td>
<td>Earthworm</td>
<td>N</td>
<td></td>
<td>1049</td>
<td>1</td>
</tr>
<tr>
<td>Lumbriculus variegatus</td>
<td>California blackworm</td>
<td>N</td>
<td></td>
<td>8404</td>
<td>1</td>
</tr>
<tr>
<td>Lutra lutra</td>
<td>Otter</td>
<td>Y</td>
<td>Predator</td>
<td>112</td>
<td>1</td>
</tr>
<tr>
<td>Lymnaea stagnalis</td>
<td>Great pond snail</td>
<td>N</td>
<td></td>
<td>136</td>
<td>1</td>
</tr>
<tr>
<td>Macropelopia sp. G_BA30</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>6154</td>
<td>3</td>
</tr>
<tr>
<td>Macrosiphini</td>
<td>Aphid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Melanopsichium pennysylvanicum 4</td>
<td>Fungi</td>
<td>Y</td>
<td>Not Animal Kingdom</td>
<td>77</td>
<td>1</td>
</tr>
<tr>
<td>Melita palmata</td>
<td>Amphipod</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Microplanus sp. 3 MAP-2016</td>
<td>Flatworm</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Microplanus terestris</td>
<td>Flatworm</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>81</td>
<td>1</td>
</tr>
<tr>
<td>Microspectra atrofasciata</td>
<td>Fly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Microspectra contracta</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>1116</td>
<td>2</td>
</tr>
<tr>
<td>Microspectra notescens</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>199</td>
<td>1</td>
</tr>
<tr>
<td>Microspectra pallidula</td>
<td>Chironomid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mitrocomella brownei</td>
<td>Hydrozoa</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mollusca</td>
<td>Mollusc</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>148</td>
<td>1</td>
</tr>
<tr>
<td>Mustelus</td>
<td>Smooth hound</td>
<td>Y</td>
<td>Mock Community</td>
<td>1726</td>
<td>13</td>
</tr>
<tr>
<td>Mustelus asterias</td>
<td>Stary smooth hound</td>
<td>Y</td>
<td>Mock Community</td>
<td>194850</td>
<td>13</td>
</tr>
<tr>
<td>Nabis limbatusi</td>
<td>Damsel bug</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nais alpina</td>
<td>Worm</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>176</td>
<td>1</td>
</tr>
<tr>
<td>Nais barbata</td>
<td>Annelid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nannochloropsis oculata</td>
<td>Algae</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nemoura avicularis</td>
<td>Stoneflies</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>853</td>
<td>1</td>
</tr>
<tr>
<td>Neoconodeillum brachytarsum</td>
<td>Proturan</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table S 5.3. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neopipo cinnamomea</td>
<td>Cinnamon neopipo</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nevemannia incertae sedis</td>
<td>Unknown</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>229</td>
<td>1</td>
</tr>
<tr>
<td>Not assigned</td>
<td>Not assigned</td>
<td>Y</td>
<td>Not distinct enough</td>
<td>34579</td>
<td>49</td>
</tr>
<tr>
<td>Notonecta glauca</td>
<td>Common backswimmer</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Obelia sp. 1 SL-2013</td>
<td>Hydrozoa</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Opalimosa mirabilis</td>
<td>Fly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ophiura ophiura</td>
<td>Serpent star (brittle star)</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Orchestia gammarellus</td>
<td>Shrimp</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>599</td>
<td>1</td>
</tr>
<tr>
<td>Ornyctolagus cuniculus</td>
<td>European rabbit</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pacifastacus leniusculus</td>
<td>Signal crayfish</td>
<td>N</td>
<td></td>
<td>10670</td>
<td>15</td>
</tr>
<tr>
<td>Pacifastacus leniusculus klamathensis</td>
<td>Signal crayfish</td>
<td>N</td>
<td></td>
<td>4048</td>
<td>2</td>
</tr>
<tr>
<td>Pacifastacus leniusculus leniusculus</td>
<td>Signal crayfish</td>
<td>N</td>
<td></td>
<td>45823</td>
<td>14</td>
</tr>
<tr>
<td>Pacifastacus leniusculus towbridgii</td>
<td>Signal crayfish</td>
<td>N</td>
<td></td>
<td>559</td>
<td>1</td>
</tr>
<tr>
<td>Palaemon varians</td>
<td>Common ditch shrimp</td>
<td>N</td>
<td></td>
<td>574</td>
<td>2</td>
</tr>
<tr>
<td>Pancrustacea</td>
<td>Crustacean</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Penicillium rubens Wisconsin 54-1255</td>
<td>Fungi</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Penicillium sclerotornum</td>
<td>Fungi</td>
<td>Y</td>
<td>Not Animal Kingdom</td>
<td>1604</td>
<td>2</td>
</tr>
<tr>
<td>Pentatoma ruipes</td>
<td>Shield bug</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>344</td>
<td>1</td>
</tr>
<tr>
<td>Phaenopspectra flavipes</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>358</td>
<td>1</td>
</tr>
<tr>
<td>Phoxinus phoxinus</td>
<td>Common minnow</td>
<td>N</td>
<td></td>
<td>73127</td>
<td>20</td>
</tr>
<tr>
<td>Phytosciara flavipes</td>
<td>Fungus gnat</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Plathichthys flesus</td>
<td>European flounder</td>
<td>N</td>
<td></td>
<td>250</td>
<td>1</td>
</tr>
<tr>
<td>Plecostroemia conspersa</td>
<td>Caddisfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>548</td>
<td>1</td>
</tr>
<tr>
<td>Pleuronecetes platea</td>
<td>European plaice</td>
<td>N</td>
<td></td>
<td>277</td>
<td>2</td>
</tr>
<tr>
<td>Pomatoschistus</td>
<td>Goby</td>
<td>N</td>
<td></td>
<td>583</td>
<td>1</td>
</tr>
<tr>
<td>Prodiamesa olivacea</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>206</td>
<td>1</td>
</tr>
<tr>
<td>Pseudanodonta complanata</td>
<td>Depressed river mussel</td>
<td>N</td>
<td></td>
<td>232</td>
<td>1</td>
</tr>
<tr>
<td>Pscooptera sp. BOLD:ACC1555</td>
<td>Booklice</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pterostichus melanarius</td>
<td>Ground beetle</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table S 5.3. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pterostichus nigrita</td>
<td>Black ground beetle</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>5361</td>
<td>1</td>
</tr>
<tr>
<td>Pungitius</td>
<td>Nine-spined stickleback</td>
<td>N</td>
<td></td>
<td>101176</td>
<td>21</td>
</tr>
<tr>
<td>Pungitius pungitius</td>
<td>Nine-spined stickleback</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pythium adhaerenus</td>
<td>Pathogen</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Radix</td>
<td>Snail</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>228</td>
<td>2</td>
</tr>
<tr>
<td>Radix balthica</td>
<td>Wandering snail</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>929</td>
<td>1</td>
</tr>
<tr>
<td>Radix sp. OUM1_3</td>
<td>Snail</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rana temporaria</td>
<td>Common frog</td>
<td>N</td>
<td></td>
<td>233456</td>
<td>40</td>
</tr>
<tr>
<td>Reesa vesputae</td>
<td>Beetle</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rheoecpectus atipes</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>1699</td>
<td>1</td>
</tr>
<tr>
<td>Rheotanytarsus pentapoda</td>
<td>Chironomid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rhodeus amarus</td>
<td>European bitterling</td>
<td>N</td>
<td></td>
<td>590</td>
<td>2</td>
</tr>
<tr>
<td>Rhyacodrilus cocineus</td>
<td>Worm</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>430</td>
<td>1</td>
</tr>
<tr>
<td>Rhychocemus sylvaticus</td>
<td>Flatworm</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>256</td>
<td>1</td>
</tr>
<tr>
<td>Rotaria sordida</td>
<td>Rotifer</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rutilus rutilus</td>
<td>Common roach</td>
<td>N</td>
<td></td>
<td>4278</td>
<td>5</td>
</tr>
<tr>
<td>Salmo trutta</td>
<td>Brown trout</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Saprolegnia ferax</td>
<td>Water mould</td>
<td>Y</td>
<td>Not Animal Kingdom</td>
<td>53</td>
<td>1</td>
</tr>
<tr>
<td>Scardinius erythrophthalmus</td>
<td>Rudd</td>
<td>N</td>
<td></td>
<td>131475</td>
<td>7</td>
</tr>
<tr>
<td>Scopelochirinus hopei</td>
<td>Amphipod</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Scophthalmus rhombus</td>
<td>Brill</td>
<td>N</td>
<td></td>
<td>73761</td>
<td>20</td>
</tr>
<tr>
<td>Scyllorhinus canicula</td>
<td>Small spotted catshark</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Serratella ignita</td>
<td>Blue winged olive</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>208</td>
<td>1</td>
</tr>
<tr>
<td>Sialis litoria</td>
<td>Alderfly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>63</td>
<td>1</td>
</tr>
<tr>
<td>Sida crystallina</td>
<td>Water flea</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>6644</td>
<td>2</td>
</tr>
<tr>
<td>Simcephalus</td>
<td>Daphnia</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>325</td>
<td>1</td>
</tr>
<tr>
<td>Simcephalus himalayensis microdus</td>
<td>Daphnia</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>2127</td>
<td>1</td>
</tr>
<tr>
<td>Simulium &lt;subgenus&gt;</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>984</td>
<td>1</td>
</tr>
<tr>
<td>Simulium angustians</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>166</td>
<td>1</td>
</tr>
<tr>
<td>Simulium aureum</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (&lt;=3cm)</td>
<td>725</td>
<td>1</td>
</tr>
</tbody>
</table>
Table S 5.3. (continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Common Name</th>
<th>Remove</th>
<th>Reason for removing</th>
<th>Total Read Count</th>
<th>Total Presences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulium cryophilum group sp. Scrgplng-01</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>787</td>
<td>1</td>
</tr>
<tr>
<td>Simulium intermedium</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>540</td>
<td>1</td>
</tr>
<tr>
<td>Simulium noelleri</td>
<td>Fly</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Simulium trifasciatum</td>
<td>Fly</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>136</td>
<td>1</td>
</tr>
<tr>
<td>Spinachia spinachia</td>
<td>Sea stickleback</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stenus impressus</td>
<td>Rove beetle</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>602</td>
<td>1</td>
</tr>
<tr>
<td>Stigmatogaster subteranea</td>
<td>Centipede</td>
<td>N</td>
<td></td>
<td>258</td>
<td>1</td>
</tr>
<tr>
<td>Stramenopiles</td>
<td>Algae</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Styliaria lacustris</td>
<td>Marine worm</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>105</td>
<td>1</td>
</tr>
<tr>
<td>Stylopidius heringianus</td>
<td>Wurm</td>
<td>N</td>
<td></td>
<td>678</td>
<td>1</td>
</tr>
<tr>
<td>Suberites</td>
<td>Sea sponge</td>
<td>Y</td>
<td>Mock Community</td>
<td>1085</td>
<td>10</td>
</tr>
<tr>
<td>Suberites domuncula</td>
<td>Sea sponge</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Suberites pagurorum</td>
<td>Sea sponge</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tanytarsus brundini</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>77</td>
<td>1</td>
</tr>
<tr>
<td>Tanytarsus ejuncidus</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>148</td>
<td>1</td>
</tr>
<tr>
<td>Tanytarsus heudensis</td>
<td>Chironomid</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>62</td>
<td>1</td>
</tr>
<tr>
<td>Tanytarsus sp. BOLD:AAV3526</td>
<td>Chironomid</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Terebrantes</td>
<td>Parasitoid wasp</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>445</td>
<td>1</td>
</tr>
<tr>
<td>Thallassiosira pseudonana</td>
<td>Diatom</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Thanatophilus micans</td>
<td>Carion beetle</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Thymallus thymallus</td>
<td>Grayling</td>
<td>N</td>
<td></td>
<td>911</td>
<td>2</td>
</tr>
<tr>
<td>Tinca tinca</td>
<td>Tench</td>
<td>N</td>
<td></td>
<td>4797</td>
<td>5</td>
</tr>
<tr>
<td>Tylolepbus moshonensis</td>
<td>Parosite</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>uncultured fungus</td>
<td>Fungi</td>
<td>Y</td>
<td>Not Animal Kingdom</td>
<td>304</td>
<td>1</td>
</tr>
<tr>
<td>Upogobia deltaura</td>
<td>Mud lobster</td>
<td>Y</td>
<td>No reads after artefact removal</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Velia caprai</td>
<td>Water cricket</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>382</td>
<td>1</td>
</tr>
<tr>
<td>Zeus faber</td>
<td>John Dory</td>
<td>Y</td>
<td>Mock Community</td>
<td>10576</td>
<td>12</td>
</tr>
<tr>
<td>zooplankton environmental sample</td>
<td>Zooplankton</td>
<td>Y</td>
<td>Suspected secondary predation (=3cm)</td>
<td>2839</td>
<td>3</td>
</tr>
</tbody>
</table>
Figure S 5.2. Taxon presences in the diet of Eurasian otters (*Lutra lutra*) using DNA metabarcoding on faecal samples. Purple lines represent presences only identified using the 16S primer set, green presences only identified using the COI primer set, and yellow identified using both primer sets. Each row is a taxon and each column is an individual faecal sample.
Figure S 5.3. Taxon presences in the diet of Eurasian otters (*Lutra lutra*) using morphological analysis of prey remains and DNA metabarcoding on faecal samples for fish taxa (a) and non-fish taxa (b). Colour of squares depicts which method led to the identification: metabarcoding only (yellow), morphological analysis only (green), both at same taxonomic level (blue) and both but at different taxonomic levels (purple). Rows represent taxon and columns represent individual faecal samples. Faecal samples were obtained from dead otters collected from across England and Wales from 2007 – 2016.
Figure S 5.3. (continued)
Figure S 5.4. Frequency of occurrence of common rudd (*Scardinius erythrophthalmus*) and common roach (*Rutilus rutilus*) in the diet of Eurasian otters (*Lutra lutra*) in each season. Data were obtained by combining identifications made through morphological analysis of prey remains and DNA metabarcoding of faeces collected from dead otters across England and Wales from 2007–2016. The width of the upper boxes and lines connecting upper and lower boxes is proportional to the number of otters from each season that had consumed Rudd/Roach, and the width of the lower boxes is proportional to the total frequency of occurrence of Rudd/Roach.
Figure S 5.5. Frequency of occurrence of European eel (*Anguilla anguilla*) and European bullhead (*Cottus gobio*) in the diet of Eurasian otters (*Lutra lutra*) at different coastal proximities. Data were obtained by combining identifications made through morphological analysis of prey remains and DNA metabarcoding of faeces collected from dead otters across England and Wales from 2007 – 2016. The width of the upper boxes and lines connecting upper and lower boxes is proportional to the number of otters at each proximity to the coast that had eel or bullhead, and the width of the lower boxes is proportional to the total frequency of occurrence of eel or bullhead.
Chapter Six – General Discussion

6.1 Aims and Overview

The overarching aim of this PhD was to analyse variation in the trophic ecology of a freshwater apex predator across large spatial and temporal scales. The project focussed on distinguishing variation in the foraging ecology of Eurasian otters in England and Wales, utilising molecular techniques to advance our understanding.

Specific aims were to investigate dietary variation of otters (i) across spatial scales, as species assemblages change regionally and between habitats; (ii) across temporal scales, as the availability of species changes seasonally and annually; and (iii) between different demographic groups (i.e. sex, size and body condition). To explore these ecological questions, the project aimed to (v) develop a metabarcoding approach suitable for studying otters; and (v) compare differences in dietary descriptions using traditional and molecular techniques; with different techniques allowing (vi) interpretation of otter diet over both long and short-term time-scales.

To address these aims, several different methodologies were used to reflect different aspects of dietary variation. Chapter Two began by exploring the influence of marine derived nutrients (MDN) on the isotopic composition of otters. MDNs were found to contribute to the diet of otters, however, such factors only produced weak associations, suggesting that other, unexplored variables were contributing to isotopic signatures. Chapter Three expanded the study to a greater spatial distribution and tested a wider range of potential drivers of variation, including habitat and land use. Findings from Chapter Three supported conclusions of the previous chapter, suggesting MDNs only contribute to low proportions of the diet, and also suggested nutrient acquisition of otters was influenced by fertiliser inputs into the environment and the availability of high trophic level prey. Both isotopic studies allowed changes in foraging ecology to be interpreted across a range of spatio-temporal and demographic variables, however, questions still remained about whether isotopic changes were due to differences in prey consumed or underlying isotopic changes in the environment.

In order to gain a greater insight into the specific species consumed by otters, DNA metabarcoding was conducted on faecal samples taken from the same individuals as in Chapter Three. First, laboratory (e.g. primer choice) and bioinformatic (e.g. quality thresholds) protocols were optimised, and the influence of post-bioinformatic filters were tested in order to achieve the most suitable representation of the data (Chapter
Once appropriate filters were chosen and applied to the data, trophic interactions were identified to a high resolution and investigated against a suite of spatio-temporal and demographic variables (Chapter Five). By carrying out morphological analysis of prey remains on the same samples, Chapter Five also examined differences in data interpretation between traditional and metabarcoding methodologies. Below we discuss findings of this study in regards to our aims, and suggest how stable isotope and metabarcoding data relate to one another.

6.2 Summary of findings and comparison of SIA with NGS

Freshwater fish were identified as the primary prey of otters, with individuals opportunistically consuming a variety of alternative prey, including species from marine and terrestrial habitats. Results from this project support findings from previous literature (Almeida et al. 2012; Krawczyk et al. 2016), demonstrating the adaptability of otters. However, previous research has primarily focussed on morphological analysis of prey remains in spraint (e.g. Jędrzejewska et al. 2001; Ruiz-Olmo and Jiménez 2009; Almeida et al. 2012), often lacking detailed prey identifications, individual biotic information and large spatio-temporal variation. This study thus expanded our knowledge of otter diet not only through the use of molecular methodology, but also through utilisation of samples and data collected from roadkill otters during a long-term national monitoring scheme, allowing dietary differences to be identified over a broader range of spatio-temporal and biotic variables.

6.2.1 Associations between otter diet and spatial changes in nutrient influxes and prey availability

Variation in carbon isotopes indicated that nutrient acquisition by otters was influenced by MDNs, with greater proportions acquired by otters in close proximity to the coast (Chapters Two and Three), or those inhabiting coastal environments (Chapter Three). Isotopic findings suggested that otters may obtain MDNs directly through consumption of marine species, or indirectly through consumption of freshwater prey in food webs enriched by MDNs. Metabarcoding identified increased frequencies of marine species in the diet of otters closer to the coast, however, no significant differences were identified between coastal and freshwater habitats. Results from metabarcoding thus suggested that direct acquisition of MDNs through predation on marine species was influenced primarily by proximity to the coast, whereas differences in isotopic signatures between habitats were driven by indirect pathways (e.g. MDN contributions to the underlying isotopic baseline of freshwater food chains). Such findings are reflective of the large ranges occupied by otters, with individuals from all habitats able to opportunistically consume marine prey providing their range includes coastal...
habitats. Otters in coastal regions also displayed the greatest individual variation (Chapter Three) in carbon signatures, indicating these individuals exploit both marine and freshwater prey; however, such variation shows individuals differ in their use of each prey source to a variable degree. Stable isotopes did not reflect a strong marine signature though, supported by the low frequencies of marine prey identified in the diet of otters. Such findings suggest MDNs acquired from coastal habitats form a low proportion of the diet of otters in England and Wales and are likely acquired through opportunistic foraging events, which is a contrast to conspecifics inhabiting the Scottish isles which primarily forage on marine prey (Kruuk and Moorhouse 1990; Watt 1995).

Spatial variation in the consumption of MDNs by otters was also influenced by predation on diadromous fish migrating from marine habitats into freshwaters. Acquisition of MDNs via diadromous fish was suggested to primarily occur through predation on Atlantic salmon and brown trout, identified by the greater enrichment in $^{13}$C in otters from regions with greater anadromous fish abundance (Chapters Two and Three). Additionally, metabarcoding found Atlantic salmon and brown trout occurred more frequently in the diet of otters inhabiting areas with larger anadromous fish runs (i.e. West Britain compared to East), thus suggesting that otters obtain MDNs directly through predation of anadromous fish. Metabarcoding also identified lower frequencies of eels and higher frequencies of bullhead further upstream; this represents a dietary switch from catadromous to freshwater fish, decreasing the chances of otters obtaining MDNs and potentially contributing to the observed decline in heavier isotopes further upstream. Interpretation of MDN acquisition via diadromous fish using metabarcoding should be considered with caution though, as freshwater life stages appear genetically identical to marine individuals and therefore cannot be distinguished from one another. However, as isotopic signatures did not reflect a strong marine signal, acquisition of MDNs by otters through consumption of diadromous fish was suggested to occur infrequently, due to transient opportunities presented during fish migration into freshwaters.

Landscape scale variation in nitrogen isotopes indicated that nutrient acquisition by otters was influenced by changes in land use and the availability of high trophic level prey. Otters inhabiting regions with greater arable and horticultural land use not only possessed nitrogen signatures more enriched in $^{15}$N, but also displayed greater individual variation in nitrogen signatures (Chapter Three). Such differences were suggested to reflect utilisation of food webs altered by fertiliser run-off from arable and horticultural land, since fertilisers produce more enriched (Bedard-Haughn et al. 2003; Anderson and Cabana 2005; Urton and Hobson 2005; Diebel and Zanden 2009;
Hoffman et al. 2012) and more variable (Vander Zanden and Rasmussen 1999; Anderson and Cabana 2005; Loomer et al. 2014) nitrogen signatures in the surrounding environment. Nitrogen signatures enriched in $^{15}$N were also observed in otters closer to the coast (Chapter Two) and those in lowland regions (Chapter Three), suggesting greater consumption of higher trophic level prey in waterways with greater discharge, due to the larger ecosystem size supporting a greater range of species and thus increasing availability of high trophic level prey (Holt et al. 1999; Srivastava et al. 2008). Although metabarcoding revealed landscape changes in dietary composition across a longitudinal gradient, dietary shifts did not reflect changes in trophic levels indicated by stable isotope analysis; many species were taken in similar frequencies at all longitudes, with salmonids, amphibians and marine species more frequently predated on in the west, and cyprinids and percids in the east. Differences in trophic level of prey consumed therefore seem an unlikely driver of the observed variance in nitrogen signatures of otters, suggesting such variance primarily reflects anthropogenic inputs of nitrogenous runoff. Interpretation of trophic level consumption through metabarcoding should be considered with caution though, as results may be confounded by shifts in trophic level within species (e.g. where larger fish of the same species may feed at higher trophic levels).

### 6.2.2 Associations between otter diet and temporal changes in prey availability

Changes in isotopic signatures and prey consumed supported the hypothesis that otter diet varies temporally, however, this variance was only observed with seasonal, and not annual, changes. Temporal findings from this study support previous literature, suggesting seasonal differences in the diet of otters reflects changes in the availability of prey in the environment (e.g. Jędrzejewska et al. 2001; Clavero et al. 2003; Miranda et al. 2008; Moorhouse-Gann et al. 2020), with predation rates on particular species increasing when they become more abundant (e.g. during breeding periods; Chapter Two; Chapter Three; Chapter Five; Carss et al. 1990; Clavero et al. 2005; Parry et al. 2015) or easier to catch (e.g. decreased motility of Cyprinidae in colder months; Chapter Five; Grant and Harrington 2015). In particular, greater enrichment in heavier isotopes in autumn and winter corresponded with metabarcoding data showing increased frequencies of Atlantic salmon, suggesting otters acquire MDNs through consumption of anadromous fish during breeding migrations. Metabarcoding also identified autumn and winter peaks in marine species and stocked fish (which may have been fed marine fish bait; Jackson et al. 2013; Bašić et al. 2015; Gutmann Roberts et al. 2017), potentially indicating that seasonal variation in MDN acquisition is facilitated through multiple pathways (e.g. Nolan et al. 2019).
The lack of dietary variation in otters over the years investigated (1993 – 2007 in Chapter Two and 2007 – 2016 in Chapters Three and Five) was surprising, as changes in abundances and distributions of prey through population expansions (e.g. invasive signal crayfish; Sibley et al. 2002; Holdich et al. 2014) and declines (e.g. eel; Bark et al. 2007; Aprahamian and Walker 2008; ICES 2019) are likely to affect availability to otters. Similarities in isotopic signatures, both within individuals and across the population, suggested otters did not vary their use of prey from different trophic levels or basal resources within the study period. Whilst consumption of isotopically similar prey can obscure dietary changes, metabarcoding indicated that there were no significant changes in species consumed between 2007 and 2016. Such consistencies in diet potentially reflect greater stability of trophic interactions within temperate food webs when compared against the far more variable diet of otters in Mediterranean regions (Clavero et al. 2003; Remonti et al. 2007; Ruiz-Olmo and Jiménez 2009; Krawczyk et al. 2016), allowing otters in England and Wales to exploit similar ranges of prey in each year. Alternatively, the lack of long-term dietary variation may indicate dietary preferences, with otters in England and Wales choosing to consume particular native prey even when those species are less abundant.

6.2.3 Identifying differences in diet between otters of different demographic groups

Differences in diet between demographic groups were only observed through the analysis of stable isotopes, with no such variation detected using metabarcoding. Carbon signatures in Chapters Two and Three suggest males are more likely to consume MDNs, with Chapter Two indicating that this is largely driven by males consuming more anadromous fish. Such differences in diet between the sexes may have been missed by metabarcoding analysis due to freshwater resident and anadromous brown trout being genetically identical and therefore indistinguishable. Nitrogen variation in Chapter Three suggested that females consume higher trophic level prey in all seasons and in all aquatic habitats except tributaries. The lack of such dietary differences between the sexes in metabarcoding data may thus suggest that nitrogen variation reflects differences in the size of prey consumed (larger fish of same species feeding at higher trophic levels), rather than differences in the species consumed. Chapter Three also suggested that over short time periods dietary differences between the sexes were associated with size class, with small males possessing a more similar isotopic signature to females than larger males. Size class differences were not observed over longer time periods represented in Chapters Two or Three though, nor were they observed in metabarcoding data, suggesting size
differences are masked over longer periods and are not driven by changes in prey species identities. Size class was also identified as the only demographic factor to influence individual variation in isotopic composition (Chapter Three), indicating large otters consume a smaller range of trophic levels than smaller otters, but with no change in range of basal resources utilised.

The lack of association between isotopic or prey species composition and body condition suggests that otters in different conditions did not differ in their foraging ecology. This was surprising as we expected better condition otters to consume more nutritionally valuable prey, as suggested in other studies; Moorhouse-Gann et al. (2020) found a positive association between consumption of fat-rich eels, and improved body condition. The lack of such an association here may be due to the lower sample size and more recent time period assessed in this study. Alternatively, the lack of association may be due to body condition indices being a lifetime representation, thus dietary analyses over short time spans (muscle in Chapter Two, whiskers in Chapter Three and faecal samples in Chapter Five) may not reflect the condition of an individual (Lanszki et al. 2015). The only sample type utilised in the PhD that represented foraging over a lifetime were bone samples (Chapter Two), although, isotopic analysis of bone also suggested lifetime isotopic composition was not associated with body condition.

6.2.4 Optimising metabarcoding protocols to decrease the impact of artefacts on dietary data

Metabarcoding provides a means for analysing specific trophic interactions between predators and their prey (Deiner et al. 2017; Taberlet et al. 2018), however, a variety of factors can influence metabarcoding data (Alberdi et al. 2018; Jusino et al. 2019). Such factors include primer choice (Deagle et al. 2014; Piñol et al. 2014), laboratory protocols (Murray et al. 2015; Schnell et al. 2015) and bioinformatic thresholds (Edgar 2016), all of which were optimised for this study (Chapter Four) to increase the detection rates of prey taxa whilst limiting generation of artefacts and amplification of otter DNA. Artefacts remaining in data following precautionary steps can be removed using post-bioinformatic filters; the use of these filters is prevalent in the literature (e.g. Gebremedhin et al. 2016; Guardiola et al. 2016; Hänfling et al. 2016; McInnes et al. 2017), but there is no consensus on which is the most appropriate. To assess the effectiveness of post-bioinformatic filters, Chapter Four compared a selection of different filtering methods and revealed how filter choice can skew the interpretation of metabarcoding data. The most appropriate method involved removing a low percentage of each sample’s reads as well as the maximum reads identified in a blank
per taxon, with percentage threshold chosen through analysis of false positives in mock communities. This combination of filters successfully removed artefacts produced by tag-jumping, contamination and mis-assignment, whilst reducing the occurrence of false negatives. Combining these two filters was thus concluded to give the most accurate representation of the data and was recommended for use in future metabarcoding studies.

6.2.5 Comparison of methods for analysing diet of otters

By collecting faecal samples from otters post mortem and carrying out both metabarcoding and morphological analysis of prey remains on the same samples, the two methods were able to be directly compared (Chapter Five). Although broad descriptions of the diet generally aligned, showing otters primarily consumed fish followed by amphibians and crayfish, the diversity and relative importance of prey differed. Metabarcoding identified more prey presences from a broader range of species and to a greater taxonomic resolution than morphological methods, supporting the findings of previous studies comparing the two approaches (Casper et al. 2007; Hope et al. 2014; Thalinger et al. 2016; Jeanniard-Du-Dot et al. 2017). Our findings differed to previous studies though, as the relative importance of prey varied between methodologies; greater detection rates by metabarcoding revealed morphological analysis underestimated frequently consumed prey and attributed a greater proportion of the diet to less important food types. Such findings therefore suggest metabarcoding gives a more accurate description of the diet of otters, providing greater clarity into trophic interactions (e.g. by distinguishing between Cyprinidae) and the species responsible for driving spatio-temporal variation (discussed in 6.2.1 and 6.2.2). However, as some presences were only identified through morphological analysis, it was concluded that combining data from each technique provides the most comprehensive description of predator diet.

Metabarcoding and morphological analyses could not be directly compared to stable isotope analyses due to the differences in samples types utilised. General conclusions remained consistent between the different methodologies, displaying the broad diet of otters and reliance on freshwater prey. However, stable isotope analysis provided a unique insight into the acquisition of marine derived nutrients and inputs from arable and horticultural land use, which was not possible through the other methods utilised. This PhD thus highlights the benefits of using different approaches to study generalist predator diets, with each method investigating a distinct aspect of dietary variation.
6.2.6 Interpreting changes in otter diet over different timescales

Using both methodologies allowed differences in the diet of otters to be analysed over a range of timescales, from snapshot descriptions by metabarcoding (reflecting consumption within 24-72 hours in mammalian carnivores; Deagle et al. 2005; Casper et al. 2007; Thalinger et al. 2016) through to nutrient assimilation over weeks, or even lifetimes, using stable isotope analysis (Tieszen et al. 1983; Hobson and Clark 1992; Dalerum and Angerbjörn 2005; Nielsen et al. 2018). By comparing isotopic signatures of different sample types (Chapter Two), or full whisker with basal segment signatures (Chapter Three), differences in nutrient assimilation over long and short timescales could be determined. Such comparisons suggested MDNs contribute more to long-term nutrient assimilation than short-term, whilst factors such as anthropogenic inputs from fertilisers affected nutrient assimilation more consistently. Together, isotopic and metabarcoding findings highlight the benefit of analysing samples that represent a range of time periods, allowing short-term changes in diet that might be obscured over long timescales to be identified alongside long-term variation which may be missed in short-term analyses, giving a more complete representation of the foraging ecology of predators.

The use of roadkill otters allowed a range of samples to be collected, however, this meant opportunities to investigate individual changes in diet through consecutive sampling were limited. In this study, individual variation could only be assessed using serial sampling along the length of each whisker, revealing changes in isotopic signatures within an individual over time (Darimont and Reimchen 2002; Newsome et al. 2009). Consecutive sampling was not possible using faecal samples for metabarcoding, or bone and muscle samples for stable isotope analysis, as each tissue type was only sampled once. Individual dietary shifts could therefore only be assessed in regards to nutrient assimilation and not specific changes in prey species consumed. Isotopic signatures between whisker segments showed little variation in most individuals (Chapter Three), suggesting foraging habits largely remained stable over time, with some individual variation occurring depending upon life stage and habitat use (discussed in 6.2.1 and 6.2.2).
6.3 Implications for species conservation and ecosystem monitoring

6.3.1 Monitoring habitats

By understanding trophic interactions over broad spatio-temporal scales, threats to both predators and prey can be assessed, as well as vulnerabilities to the trophic network (Polis and Winemiller 2013; Berg et al. 2015), whilst reducing the effect of local extremes that may skew interpretation of the data (Aizpurua et al. 2018). Stable isotope analysis revealed differences in nutrient loads detected in otters reflected changes throughout the trophic network due to allochthonous nutrient influxes from natural (e.g. MDN) and anthropogenic sources (e.g. fertilisers). These findings displayed pathways by which marine and terrestrial ecosystems contribute to nutrient loads in freshwater ecosystems and influence species at multiple trophic levels. Allochthonous subsidies can increase productivity in freshwater systems, but they can also have detrimental impacts if nutrient inputs are too high, e.g. via eutrophication (Khan and Ansari 2005). Findings in this study thus show analysing isotopic signatures of generalist apex predators can provide a means to monitor changes in nutrient loads within freshwater trophic networks and help guide management of freshwater habitats. Metabarcoding indicated foraging by otters was influenced by differences in availability of prey in the surrounding environment, providing an insight into prey population dynamics and biodiversity within freshwater habitats. As generalist predators are more likely to consume abundant prey, metabarcoding of generalist predator diet can provide an insight into changes in prey populations within the environment, helping guide biodiversity management especially where surveying may be difficult (Deiner and Altermatt 2014; Boyer et al. 2015; Deiner et al. 2017; Haltitschek et al. 2018). Whilst this thesis has focussed on describing dietary variation in otters, results suggest that the combined approach of isotopic and metabarcoding data may be more widely applicable, allowing other systems to be monitored through studying apex predator diets.

6.3.2 Impacts of variable diets for otter populations

Findings from this project showed how dietary plasticity allows otters to adapt to changes in prey abundances over time and space. Many habitats have undergone modification since otters last inhabited them, yet similarities in the composition of diet between altered habitats (e.g. urban, or arable and horticultural land use) and more natural habitats indicate otters are able to successfully inhabit and forage in modified landscapes. Such dietary variation likely aided the recent population recovery of British otter populations by allowing them to adapt to various environments with different
species compositions as they expanded their distribution and recolonised habitats. Additionally, dietary plasticity may lead to resilience to future environmental and species abundance changes, such as due to climate change, aiding persistence of otter populations. Dietary shifts may also lead to changes in contaminant acquisition; switching to prey with high contaminant loads is likely to lead to increased contaminant burdens for otters, whereas consumption of prey with lower contaminant loads, or predation on a range of prey, is likely to reduce an otter's exposure to contaminants. The dietary switch observed with coastal proximity in Chapter Five may result in decreases in contaminant acquisition, as otters further inland consume fewer long-lived fatty eels that are likely to have high contaminants loads, and more short-lived bullhead with lower contaminant loads (Moorhouse-Gann et al. 2020). Such dietary shifts may thus improve the health of individuals and aid persistence of otter populations. Whilst the perceived benefits of dietary variation indicate high resilience of otters to changes in the environment and trophic network, previous declines in population were rapid, it is therefore critical to continue to monitor otter populations to prevent such a decline occurring again.

### 6.3.3 Predation impacts on prey species

This project proposed that otters are unlikely to pose a major threat to the persistence of protected prey populations, due to the observation that otters primarily consume species with high abundances in England and Wales. This was supported by the low occurrence, or even absence, of protected species identified using metabarcoding. Findings from this study suggest that whilst prey populations remain low, otters are less likely to encounter and consume such prey, thus reducing the impacts of otter predation on protected species. However, there was one exception; the European eel is a critically endangered species with a declining population, yet it was found to be taken by a large proportion of the otters in the study. High occurrence of eels in otter diet suggests that otters have a preference for eel, supporting findings from the literature (Copp and Roche 2003; Britton et al. 2006; Miranda et al. 2008), and are likely to hunt them regardless of their abundance in the environment. Otter predation may thus pose a threat to current eel populations in Britain. Metabarcoding also identified invasive species in the diet of otters (e.g. signal crayfish); although these predation events did not make up a large proportion of the diet, such foraging suggests otters are able to adapt to increasing abundances of invasive species and potentially aid biocontrol of freshwater invasive species.
6.3.4 Predation impacts on stocked fish and conflict with aquaculture

As otters recover and expand into modified landscapes, conflict with humans is likely to increase. Such conflict primarily occurs between otters and anglers or aquaculture management due to predation of stocked fish. Previous studies investigating the diet of otters often lacked clarity (Jędrzejewska et al. 2001; Ruiz-Olmo and Jiménez 2009; Almeida et al. 2012; Moorhouse-Gann et al. 2020), limiting interpretation of predation on stocked fish. Through the use of metabarcoding, this study was able to identify many fish to species level, revealing that frequencies of stocked fish in the diet of otters were low (Chapter Five). By combining findings from this study with those of other studies that demonstrate otters utilise different sized fish to anglers (Grant and Harrington 2015; Lyach and Čech 2017), predation threats by otters on stocked fish are suggested to be low. Consumption of stocked fish did occur more frequently in autumn and winter though, therefore we recommend that anglers and aquaculture managers focus mitigation during these seasons (e.g. using electric fences or stocking diversion ponds with uneconomic species; Bodner 1995; Kucerová 1999). This study shows how investigating dietary variation in generalist predators can reveal potential conflicts with humans and help guide mitigation.

6.4 Future directions

6.4.1 Expanding analyses to other populations

Otter populations from West Europe through to East Asia experienced rapid population declines in the 20th century (Strachan and Jefferies 1996; Roos et al. 2001; Clavero et al. 2010). Whilst some populations, particularly in western Europe, have since begun to recover and recolonise habitats (Strachan and Jefferies 1996; Mason 1998; Roos et al. 2001; Conroy and Chanin 2002; Sainsbury et al. 2019), other populations are still in decline (Roos et al. 2015). Understanding the dietary variation of a recovering population, as studied here in England and Wales, can thus provide valuable information into the factors aiding a successful population increase and expansion. Utilising molecular methods to describe the diet in other recovering otter populations, as well as declining or stable populations, will provide a greater insight into differences in trophic dynamics between populations. Such analyses may thus indicate why some otter populations are recovering whilst others continue to experience declines, helping guide conservation measures across the whole distribution of the species.

6.4.2 Ecosystem level assessments

Expanding analyses to include data obtained from both predators and prey would allow comprehensive analysis of nutrient flow and trophic interactions within the ecological network. In this study, comparisons between isotopic and metabarcoding data
suggested which prey may be contributing to changes in nutrient assimilation by otters, but differences in time periods represented by each sample type prevented definitive conclusions from being made. By collecting isotopic data from prey species alongside predators, mixing models can be employed to assess the contribution of prey to a predator's overall isotopic signature (Phillips 2012). Isotopic analyses of diet remain confounded by similarities in prey isotopic signatures, reducing taxonomic detail by clustering species together (Crawford et al. 2008), but allows direct conclusions to be made regarding nutrients assimilated through predation on different prey categories. Collecting data on prey abundances in the surrounding environment also allows prey choice to be assessed through comparisons with metabarcoding data (Vaughan et al. 2018). Collecting information on prey abundance from the same locations and time periods as otter roadkill can indicate whether individual otters focus foraging on the most abundant prey in the surrounding environment or possess preferences for particular prey regardless of abundance.

Collection of prey samples and abundance data was not employed in this study due to the complexity of data collection over the broad spatio-temporal scales which otters were sampled from. Focussing data collection on more localised scales may aid data recovery from predators and prey, providing more clarity into nutrient assimilation and prey choice in otter populations. Through such analyses, implications of dietary changes in otters could be expanded to other trophic levels, providing a greater insight into pathways for nutrient flow within freshwater ecosystems and impacts of predators on prey species. Comparing dietary and abundance data would thus provide an indication of the resilience of the trophic network to future environmental changes, helping guide conservation measures to protect a range of species with ongoing global change.

### 6.4.3 Assessing otter health in relation to dietary variation

Dietary shifts between prey with different nutritional value and contaminant load have the potential to threaten future persistence of predator populations through impacts to individual fitness (Österblom et al. 2008; Ruiz-Olmo and Jiménez 2009; Lourenço et al. 2011). Investigations into nutritional value and contaminant load of prey species are thus important considerations for conserving otter populations. Findings from this project can guide such investigations by indicating the primary prey consumed by otters, allowing studies to focus analyses on prey that contribute to large proportions of otter diet and are thus likely to have the biggest impacts.
6.4.4 Individual dietary variation assessments

Ambiguity related to the rate at which otter whiskers grow (Chapter Three) limited the interpretation of dietary shifts in individual otters over time. Whisker growth rates can be obtained through the use of rhodamine B in captive feeding trials; as species consume food infused with rhodamine B, it is incorporated into their tissues and leaves a detectable sign in metabolically inert tissues (Spurr 2002; Purdey et al. 2003; Robertson et al. 2013; Weerakoon et al. 2013). The rate at which a whisker grows can be calculated by measuring the length of the whisker between the base and the rhodamine B signature. Whisker growth rates would provide greater precision regarding the time period each whisker segment represented, thus allowing isotopic changes to be depicted across specific timescales. In order to investigate individual dietary variation using metabarcoding, spraint left by living otters would need to be collected consecutively over time, rather than using roadkill. Greater precautions must be made when collecting spraint samples though, as exposure to UV light and contamination from environmental sources can decrease the amount of dietary DNA acquired from each sample (Oehm et al. 2011; Alberdi et al. 2019). Spraint would also require genotyping alongside metabarcoding to identify if samples came from the same individual. Whilst spraint analysis would require more intensive sample processing, it would provide an indication into changes in prey species consumed by individuals over time and show whether individuals consume a range of species or focussing foraging on particular prey.

6.4.5 Investigating eel-otter dynamics

Chapter Five showed that eel remain a commonly consumed prey species, regardless of the ongoing population decline experienced by British eel populations. Whilst such a finding suggests that otters potentially pose a threat to eel populations, this does not expose the full extent of predation threats. Further investigations into eel-otter dynamics are therefore required to assess if recovering otter populations have a large impact on eel recruitment and future persistence of the species. Such analyses would be best investigated alongside the diet of other eel predators (e.g. great cormorant, Phalacrocorax Carbo) in order to investigate the full impact different predators have on eels.

6.4.6 Improving metabarcoding analyses

Artefacts in metabarcoding data have largely been associated with low read counts (Brown et al. 2015; Leray and Knowlton 2017), however, Chapter Four found some artefacts can occur in greater abundances than expected. Whilst Chapter Four
suggested a suitable method for removing the majority of artefacts that remain in metabarcoding datasets, further research into the relationship between read count and assignment to taxon is required in order to fully assess artefact production and further recommend how to reduce impacts of erroneous reads. Such assessments may also provide an indication into the impacts of secondary predation, which have also been suggested to occur at low read counts, allowing data to be filtered to only represent primary consumption. Future studies into read count abundances may therefore lead to greater accuracy for the interpretation of dietary metabarcoding studies, whilst also providing an insight into the potential nutrition provided by secondary predation.

6.5 Conclusions
This thesis showed investigating the diet of an apex predator provides great detail not only into predator ecology, but also the ecology of their prey and changes impacting the entire ecosystem. The use of molecular methods furthered current understanding of otter diet, with metabarcoding providing greater detail into trophic interactions and stable isotope analysis depicting changes in nutrient assimilation. The highly generalist foraging behaviour of otters described in this study implies otters are resilient to changes in the environment, not only facilitating ongoing population recovery but potentially aiding persistence of the species with future environmental changes. Both metabarcoding and stable isotope analysis indicated dietary changes in otters occurred in response to landscape, seasonal and demographic scale changes. Spatio-temporal changes revealed ecosystem level changes due to allochthonous nutrient influxes, and gave an indication of the diversity of prey present within freshwater environments. The findings of this study thus highlight the use of dietary variation in apex predators to indicate broad scale changes within the environment.
Bibliography


Berry, T.E., Osterriether, S.K., Murray, D.C., Coghlan, M.L., Richardson, A.J., Grealy,


Codron, D., Codron, J., Lee-Thorp, J.A., Sponheimer, M., De Ruiter, D. and Brink, J.S.


Ficetola, G.F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M.,


Gutmann Roberts, C., Bašić, T., Amat Trigo, F. and Britton, J.R. (2017). Trophic consequences for riverine cyprinid fishes of angler subsidies based on marine-


Harper, L.R., Watson, H.V., Donnelly, R., Hampshire, R., Sayer, C.D., Breithaupt, T.


NBN atlas (2020d). Perch (*Perca fluviatilis* Linnaeus, 1758). Available at:

Newsome, S.D., Tinker, M.T., Monson, D.H., Oftedal, O.T., Ralls, K., Staedler, M.M.,
diet specialization in California sea otters (*Enhydra lutris nereis*). *Ecology*, 90(4),
pp. 961–974. Available at:

and Dickman, C.R. (2015). The ecological effects of providing resource subsidies
to predators. *Global Ecology and Biogeography*, 24(1), pp. 1–11. doi:
10.1111/geb.12236.

Green, R.E. and Shapiro, B. (2018). Minimizing polymerase biases in
metabarcoding. *Molecular Ecology Resources*, 18(5), pp. 927–939. doi:
10.1111/1755-0998.12895.

Nichols, S.J., Kefford, B.J., Campbell, C.D., Bylemans, J., Chandler, E., Bray, J.P.,
routine DNA metabarcoding of macroinvertebrates using bulk samples for
freshwater bioassessment: Effects of debris and storage conditions on the
recovery of target taxa. *Freshwater Biology*, (September), pp. 1–14. doi:
10.1111/fwb.13443.


Nolan, E.T., Gutmann Roberts, C. and Britton, J.R. (2019). Predicting the contributions
of novel marine prey resources from angling and anadromy to the diet of a
freshwater apex predator. *Freshwater Biology*, 64(8), pp.1542-1554. Available at:

O’Brien, P.J. and Wehr, J.D. (2010). Periphyton biomass and ecological stoichiometry


O’Hara, B. (2016). Fish Parasites. Available at: http://rstudio-pubs-
static.s3.amazonaws.com/170545_88f06a76ecc94b91ba27d505085546a0.html


Rytkönen, S., Vesterinen, E.J., Westerdunin, C., Leviäkangas, T., Vatka, E., Mutanen,


The Darwin Tree of Life (2020). The Darwin Tree of Life Project. Available at: https://www.darwintreeoflife.org [Accessed: 12 October 2020].


