Validation of DNA metabarcoding as a tool for diet analysis in reindeer (*Rangifer tarandus tarandus* L.)

Pernille Meyer

Master of Science Thesis

Centre of Ecological and Evolutionary Synthesis
Department of Biosciences
Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

June / 2019

2019

Title: “Validation of DNA metabarcoding as a tool for diet analysis in reindeer (*Rangifer tarandus tarandus* L.)”

Pernille Meyer

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Acknowledgements

The presented master thesis was written at the Centre for Ecological and Evolutionary Synthesis (CEES) at the Department of Biosciences, University of Oslo (UiO), under the supervision of main supervisor researcher Galina Gusarova (UiO), and co-supervisors Prof. Anne K. Brysting (UiO), postdoctoral fellow Stefaniya Kamenova (UiO), Prof. Monica A. Sundset (UiT) and Prof. Lars Folkow (UiT). The work in this master thesis was conducted at the Department of Biosciences, UiO and the Department of Arctic and Marine Biology at UiT - The Arctic University of Norway (UiT).

The sequencing service was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and supported by the "Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the Southeastern Regional Health Authorities.

The bioinformatics was performed on the Abel Cluster, owned by the University of Oslo and Uninett/Sigma2, and operated by the Department for Research Computing at USIT, the University of Oslo IT-department (http://www.hpc.uio.no).

There are many people who made this master degree possible. First of all, I would like to thank my main supervisor Galina Gusarova, who gave me the opportunity to join the REININ project and for supporting and guiding me throughout the thesis. Thank you to Anne K. Brysting, Stefaniya Kamenova, Monica A. Sundset and Lars Folkow for feedback, support, trust and advice, during the planning and execution of the feeding experiment, laboratory work and writing.

To Associate Prof. Tove Utsi, everyone involved in gathering the reindeer and to the slaughterer, I thank you and I truly appreciate all the work that was involved in the process of gathering these samples. Knowing how much hard work was involved and how difficult it is to handle the intestines, I am thankful and confident that this master thesis would not have been the same without these samples.

Thank you to everyone at the Department of Arctic and Marine Biology, UiT, for your hospitality, kindness, curiosity and advices. To Associate Prof. Gabi Wagner, thank you for all advice and support. To Leonardo Rescia, Nora Slåttebrekk and Jaymee van Dalum thank

you for all of your help in the execution of the feeding experiment. To Hans Lian, Hans Arne Solvang and Renate Thorvaldsen (technical staff at UiT), I am extremely grateful for all the help you provided.

To the three reindeer, 9/10, 10/10 and 12/10. Thank you for being patient and well behaved throughout the feeding experiment. It was a pleasure and an honour to work with you.

Thank you to Leidulf Nordang (Utviklingssjef drøvtyggarfôr, Felleskjøpet Fôrutvikling AS) for providing information on RF-80 and discussions regarding the results.

Thank you to UiO Helpdesk and to Sabry Razick for all of your help and solutions during the bioinformatics work. To Éric Coissac, whose knowledge and experience is unparalleled. I cannot thank you enough for the help with both bioinformatics and statistical analysis.

To Lilja, Marina, Elke and Linn; through all of my frustrations and complaints you have given me encouragement and tons of positive energy. I appreciate all of the conversations and discussions we have had.

Finally, I would like to thank my mother, my sister and my friends for all of you support and patience, for pushing me and believing in me. To my silver linings, Cani and Gaia, who always provided breaks, whenever I needed one.

Abstract

Free-ranging reindeer (*Rangifer tarandus* L.) eat a diverse diet of graminoids, herbs, shrubs, woody plants, mushroom, lichen and mosses. DNA metabarcoding is an established method in dietary studies and through e.g. non-invasive faecal collections it allows identification of the complete diet. However, there are methodological shortcomings regarding the application of DNA metabarcoding, especially for species with highly diverse diets and complex digestive systems such as the ruminant reindeer. In an effort to reveal the advantages and limitations of DNA metabarcoding, two experiments were conducted during this thesis.

In the first experiment, a controlled feeding trial was carried out at UiT - The Arctic University of Norway, three female semi-domesticated reindeer (*R. t. tarandus*) were fed (i) pelleted reindeer food (RF-80, containing a known mixture of eleven species), (ii) birch (*Betula pubescens* Ehrh.) and (iii) increasing amounts of lichens (mainly *Cladonia stellaris* (Opiz) Pouzar and Vezda). The methodological questions were: 1) Does DNA metabarcoding detect all species in a known diet and does it allow identification to species level for all species?; 2) For how long is DNA of a particular food item present in the digestive system?; and 3) How well is relative food biomass represented by output sequences? In the second experiment, samples were collected from different sections of the gastrointestinal (GI) tract of four male semi-domesticated reindeer on natural pasture in Finnmark, Norway to answer: 4) Does DNA detectability change through the GI tract?

The controlled feeding experiment detected eleven of the thirteen fed species, and the specific primers used in this experiment detected a higher diversity of species present in the diet than the general eukaryote primer. However, the short specific primer increased the number of wrongfully assigned taxa. DNA of *B. pubescens* was detected within 12 hours and it was present in the faeces for minimum 26 days after feeding (Question 2). It was possible to see a correlation between the fed biomass and the proportion of DNA sequences of *C. stellaris*, when using a general eukaryote primer (Question 3). In answer to Question 4, the data from this study indicate that for seed plant, eukaryote and fungal DNA, there is higher species richness in the distal colon than in the rumen, whilst the opposite is found for bryophyte DNA. Higher diversity is found in the cecum and distal colon for seed plant and eukaryote DNA, in the rumen for fungal DNA and possibly also in the rumen for bryophyte DNA. These results imply that non-invasive faecal samples are better to use, compared to invasive rumen samples, when studying species richness of seed plants in the diet or the overall diversity in the diet, regardless of targeted taxonomic group.

Table of contents

1 Introduction 1
2 Materials and methods 7
2.1 Feeding experiment 7
2.1.1 General remarks 7
2.1.2 Experimental setup 7
2.1.3 Experimentally controlled diet 8
2.1.4 Sample collection and subsampling 10
2.2 Gastrointestinal tract 13
2.2.1 Sample collection 13
2.3 DNA extraction 13
2.4 PCR amplification and high-throughput sequencing 14
2.5 Bioinformatics 16
2.6 Downstream statistical analysis 19
2.7 Ethical statement 22
3 Results 23
3.1 Does DNA metabarcoding identify all species to species level in a known diet? 24
3.2 For how long is DNA of a particular food item present in the digestive system? 27
3.3 How well is relative food biomass represented by output sequences? 30
3.4 Does DNA detectability change through the gastrointestinal tract? 34
4 Discussion 39
4.1 Does DNA metabarcoding identify all species to species level in a known diet? 39
4.2 For how long is DNA of a particular food item present in the digestive system? 41
4.3 How well is relative food biomass represented by output sequences? 44
4.4 Does DNA detectability change through the gastrointestinal tract? 45
5 Conclusion 47
6 Future studies 49
References 51
Appendix A – Feeding experiment 61
Appendix B - Gastrointestinal tract 89
Appendix C - Library preparation 95

1 Introduction

Dietary studies are important in understanding ecosystems and species interactions by connecting ecosystem levels (i.e. primary producers, primary consumers and secondary consumers). Thus, numerous studies have been conducted to increase knowledge about different aspects of diet, e.g. dietary niches, competition, dietary shifts etc (e.g. Soininen et al., 2014; Gebremedhin et al., 2016, Zhao et al., 2019). The potential impacts of ongoing climatic changes have emphasized the importance of understanding ecosystem functions and the need to make predictions regarding how an ecosystem may respond. An area vulnerable to rapid climate change is the Arctic (Arctic Climate Impact Assessment, 2004), and a keystone species in this region is the reindeer (*Rangifer tarandus* L.). To survive in periodically harsh conditions, with low forage availability, reindeer have adapted a unique and diverse diet consisting of graminoids, herbs, shrubs, woody plants, mushroom, lichen and mosses (Staaland and Sæbø, 1987; Mathiesen et al., 1999; Mathiesen et al., 2000; Inga, 2009), with a large seasonal variability (Kojola et al., 1995; Mathiesen et al., 2000; Heggberget et al., 2002). Additionally, indigenous people herd several of the subspecies in a migratory pattern (Turi, 2002; CAFF, 2006), one of which is the semi-domesticated reindeer (*Rangifer tarandus tarandus*) herded by the Saami people on the mainland of northern Scandinavia (Tyler et al., 2007). Variation between summer (coastal or mountains) and winter (lichen-dominated areas) pastures adds a seasonal aspect to the diversity in the diet of *R. t. tarandus* (Sundset et al., 2007; Tyler et al., 2007; Moen, 2008). The truly remarkable aspect of reindeer diet is their adaptation to utilize lichen (e.g. *Cladonia stellaris* (Opiz) Pouzar and Vezda) as food (Sundset et al., 2008, 2010; Węgrzyn et al., 2019), as they are potentially toxic to other animals, including ruminants (Cook et al., 2007; Dailey et al., 2008). However, it has been shown that climate change has a large negative effect on lichen recovery after reindeer foraging (Klein and Shulski, 2009) and that reindeer diet already has shifted due to plant response to increased CO_2 levels (Zhao et al., 2019). Because reindeer hold an important ecological role in the Arctic ecosystems, and in several areas are valued as livestock, it is important to study their diet in light of current climatic changes to develop solutions to potential negative impacts (Hansen et al., 2018).

Several approaches and methods can be used to study the diet of reindeer, and molecular technology development has created new opportunities for exploration of diets. Direct observations, macroscopic or microscopic identification of digested food remains in stomach

content or faeces are considered classical methods and still frequently used (Bjørkvoll et al., 2009; Nichols et al., 2016). Observational studies provide accurate information about vegetation types and habitats where the study animals forage at a certain time of the year. Certain observations also reveal specific forage events, for instance, kelp eaten by Svalbard reindeer (*R. t. platyrhynchus* Vrolik) during harsh winter conditions (Hansen and Aanes, 2012; Hansen et al., 2019). However, migration, rapid movement and the skittish nature of reindeer make it challenging to identify all the consumed species throughout a year. Microscopic identification can also be a challenging method to use when studying diet of e.g. ruminants, as highly digested plant remains in faeces are usually indistinguishable. For this reason, rumen samples from culled or fistulated animals are favoured for obtaining better taxonomic resolution, as they are more heterogeneous (Gaare and Skogland, 1975; Vavra et al., 1978). However, collecting samples from culled or fistulated animals, is highly invasive and newer methods such as stable isotope and DNA-based analysis could allow for non-invasive dietary identification based on the more homogenous faecal samples.

Stabile carbon isotopes allow differentiation of for example C3- and C4-plants in the diet (Codron et al., 2007), while measurements of nitrogen isotopes provide information about the trophic level of organisms (DeNiro and Epstein, 1981; Ambrose and DeNiro, 1986). In addition to faecal samples, stable isotope analysis of carbon and nitrogen ratios from hair, antler and whisker tissues can also be used to identify the diet of the study species (Zhao et al., 2019). Although the method does identify specific species present in the diet, it is especially useful for comparing functional plant groups in the diet, identifying pastures or seasons when animals were feeding as well as studying the effect of demography and climatic variables on species diet (Soininen et al., 2014; Hansen et al., 2019; Zhao et al., 2019). Stable isotope analysis is therefore a well-recognized method for use in ecological diet studies.

A popular DNA-based method allowing for the quantification of DNA from ingested prey is real-time PCR (RT-PCR). It allows investigation of presence/absence, or the relative quantity, of a given prey based on how much DNA has been amplified (Heid et al., 1996; Stephenson, 2016; Maddocks and Jenkins, 2017). With this method it is also possible to search for specific prey taxa (Mysterud et al., 2016), and to investigate microbial population densities (Denman and McSweeney, 2006; Salgado-Flores et al., 2016). RT-PCR is an acknowledged method in ecological studies, and has revealed vital information on the microbiota in digestive tracts of

ruminants (e.g. Sylvester et al., 2004; Denman and McSweeney, 2006; Salgado-Flores et al., 2016).

One of the latest DNA-based methods for diet analysis, and the main focus of this study, is DNA metabarcoding. DNA metabarcoding relies on short DNA fragments (i.e. "barcodes") to identify the species present in an environmental or dietary sample (Taberlet et al., 2012; Bohmann et al., 2014). Combined with next generation sequencing (NGS) it allows for the simultaneous detection of multiple species within complex mixtures of DNA. Metabarcoding primers are designed to amplify DNA from the target groups of organisms, and based on conserved priming sites, it can amplify a wide range of taxonomic groups (e.g. all eukaryotes) or be more specific (e.g. only seed plants). The DNA barcodes should also contain a highly variable sequence region that ideally is unique to each species, providing opportunity for species-level identifications (Taberlet et al., 2012). As a relatively recent method (Valentini et al., 2009; Taberlet et al., 2012), it is vital to study several aspects of its application and potential limitations, such as the capacity to obtain species level information. Due to the short length of the DNA barcodes and highly similar DNA sequences within taxa (i.e. no unique sequence regions adjacent to the conserved sequence regions, Sønstebø et al., 2010), identification for many taxa is restricted to genus or family level. The lack of reference sequences in the public reference databases is another reason for restricted taxonomic assignation (Taberlet et al., 2012).

Another disputed limitation with DNA metabarcoding is how well it estimates the relative biomass of ingested food, i.e. whether the number of sequenced reads (the total number of sequences for a given prey taxon, also referred to as number of output sequences) reflects biomass or if it merely is a representation of the DNA present in a sample (Deagle et al., 2013; Soininen et al., 2013; Thomas et al., 2014). Most ecological dietary studies focus on large, functional groups in the diet, and presence/absence data can be sufficient to answer questions related to diet diversity. However, to compare the relative amounts and importance of the functional groups within a diet, the relationship between sequenced reads and abundance or biomass must be investigated. This would provide a more detailed estimate of diet composition, and help to interpret the occurrence of dominant versus rare species in a diet based only on the DNA sequences.

Various studies have compared DNA metabarcoding results to those of other methods for diet assessment. Several studies show that DNA metabarcoding allows characterizing diets with a

similar or higher taxonomic precision compared to direct observations or macroscopic identification (e.g. Soininen et al., 2009; Pompanon et al., 2012; Kartzinel et al., 2015; Nichols et al., 2016). For instance, after comparing macroscopic identification and DNA metabarcoding of rumen content, Nichols et al. (2016) argued that the results from the two methods are comparable and positively correlated. But due to time- and cost-efficiency, one should favour DNA metabarcoding. Additionally, they showed that it is possible to make quantitative assessments about ingested prey biomass based on the number of output sequences. Another study, comparing microscopic identifications and DNA metabarcoding in herbivorous voles, did not manage to make quantitative assessments, but showed that DNA metabarcoding have a higher taxonomic resolution compared to microhistological analysis (Soininen et al., 2009). Kartzinel et al. (2015) compared the results from stable carbon isotopes and DNA metabarcoding, and showed a high degree of correlation between the two datasets in the relative contribution of shrub versus forb plant species in the diet of several species of African ungulates. In the same study they also demonstrated that DNA metabarcoding provides family, genus and species level identifications that can be used to show niche partitioning among several species with overlapping diets.

This thesis will focus solely on DNA metabarcoding using the Norwegian semi-domesticated reindeer (*R. t. tarandus*) as a model to address major methodological questions related to using DNA metabarcoding for diet analysis in ruminants. To my knowledge, no dietary experiment using DNA metabarcoding has been conducted on reindeer, and it is therefore of importance to clarify the following methodological questions: 1) Does DNA metabarcoding detect all species in a known diet and does it allow identification to species level for all species? 2) For how long is DNA of a particular food item present in the digestive system? 3) How well is relative food biomass represented by output sequences? 4) Does DNA detectability change through the gastrointestinal (GI) tract?

These questions were addressed by conducting a controlled feeding experiment (Questions 1-3) and by collecting samples from different sections of the reindeer GI tract (Question 4). To answer Question 1, different food items were fed to three captive reindeer allowing assessment of how well DNA metabarcoding performs in terms of species detection and taxonomic resolution. DNA from faecal samples and consumed food items were amplified using a multi-locus primer panel and sequenced alongside, to investigate if DNA degradation during digestion affects species detection and taxonomic resolution. For answering Question

2, birch (*Betula pubescens* Ehrh.) was introduced once on top of the standard pelleted feed (RF-80) that the reindeer received as part of their regular diet. By collecting faecal samples intensely the following days, the aim was to identify with a high precision when the DNA of *B. pubescens* first appeared in the faeces. Faecal samples were collected for the following ~2 months, to assess when *B. pubescens* DNA was no longer detectable in the faeces. For answering Question 3, lichens (mainly *Cladonia stellaris*) were added to the reindeer diet, in increased proportions through a period of 13 days. To answer Question 4, output sequences from samples collected across different sections of the reindeer GI tract were compared in order to identify which part of the GI tract had the highest species richness and diversity.

2 Materials and methods

2.1 Feeding experiment

2.1.1 General remarks

The feeding experiment was conducted at the Department of Arctic and Marine Biology, at the Arctic University of Norway (UiT), from the 22nd of January 2018 to the 18th of February 2018. Faecal samples were collected from the 22nd of January 2018 and continued until the 23rd of March 2018. The feeding experiment included three captive 7-year-old Norwegian semi-domesticated female reindeer (animal number 9/10, 10/10 and 12/10). The reindeer were half way through pregnancy at the time of the experiment. Deworming was done according to standard routine (Panacur given on the 5th of December 2017 and Ivomectin given on the 3rd of January 2018) prior to the experiment. Despite this, tapeworm was found in the faeces of animal 9/10 (31st of January 2018), but treatment was postponed until after the faecal sampling was completed (23rd of March 2018). The reindeer were weighed prior to and following the feeding experiment (19th of January 2018: 09/10: 102.4 kg, 10/10: 96.6 kg, 12/10: 104.4 kg. 18th of February 2018: 09/10: 104.0 kg, 10/10: 93.2 kg, 12/10: 103.2 kg).

Postdoctoral fellow Stefaniya Kamenova and I were responsible for feeding, cleaning and collected samples during the experiment. MSc Leonardo Rescia, MSc Nora Slåttebrekk and PhD Jaymee van Dalum aided with cleaning and preparations. Leonardo Rescia also helped collect samples during week 6 and conducted the sampling after the 22nd of February 2018. The technical staff at UiT brought the reindeer into the enclosures used in the experiment, and back out. They also taught us how to take care of the animals and helped with all of the practical aspects involved.

2.1.2 Experimental setup

The animals are normally kept in large outdoor enclosures (*E1*, encompassing approximately 2000 - 4000 m^2, Fig. A1 in appendix A), where the reindeer are fed *ad libitum* RF-80 and have access to several food taxa (e.g. *Betula pubescens*, lichens, bryophytes and graminoids). These food taxa can appear in the output sequences if ingested prior to the experiment. However, in an effort to minimize these food taxa, the experiment was conducted in winter

when a thick snow layer covered *E1* and thereby restricted foraging access. The snow is also the main water source for the reindeer during winter.

During the feeding experiment, smaller fan shaped enclosures (*E2*) were used (Fig. A1 in appendix A). *E2* consists of several corridors that progressively enlarge at the bottom (approximately 40 m^2), covered with tarmac, which was heated to prevent ice covering the ground. The animals were accustomed to *E2* prior to the experiment (4^{th}-22^{nd} of December 2017 and 19^{th} of January 2018), to limit the potential occurrence of undesired food items in their diet and in order to minimize their stress.

On the day before the animals were brought into *E2* for the feeding experiment (the 18^{th} of January), *E2* was cleaned with a high-pressure washer and propane burners to remove plants growing in cracks of the tarmac, other plant remains (e.g. pollen grains) or old faeces that might be potential sources of DNA contaminations.

After bringing the animals to *E2* (19^{th} of January 2018), the faecal material that was not sampled was removed with a broom and a dustpan several times a day (simultaneous as sample collection, Table 2). Most mornings the floor was cleaned thoroughly with water under high pressure in order to remove as many contamination sources as possible (e.g. urine and faecal remains). However, due to cold temperatures, this was not possible every day, as the water would freeze quickly and it was hazardous for both humans and animals involved.

At the start and throughout the experiment, all containers for food and water were cleaned with a 4% bleach solution and rinsed thoroughly with tap water. This procedure was repeated after introducing the new food items (see section 1.1.4) and thereafter a few times a week. The containers were assigned individually to each animal to limit contamination between the reindeer.

2.1.3 Experimentally controlled diet

RF-80

A commercially available pelleted food for reindeer (RF-80, FK Reinfôr, Felleskjøpet, Norway) was the main component in the diet of the reindeer. A detailed ingredient list was obtained from Felleskjøpet prior to the experiment (Table A1 in appendix A). Before the

feeding experiment started, the reindeer were kept together with a male reindeer and fed *ad libitum* RF-80 (7-9 kg per day, divided by four animals), and after the feeding experiment they were again fed *ad libitum* RF-80 (5-7 kg per day, divided by three animals).

RF-80 disintegrates if added to water, and therefore the amount fed to the reindeer was weighed in dry weight. Leftovers (intact RF-80) and residuals (crumbles and powder) of RF-80 were weighed approximately 24 hours after feeding, and subtracted from the total amount given to the reindeer, in order to calculate how much they actually ate (Table A2 in appendix A).

Betula pubescens

At the start of the feeding experiment (18th of January 2018), *Betula pubescens* was introduced to examine how long it takes for the DNA of a species to pass the gastrointestinal tract and for how long it is detectable in the faeces (Question 2). *B. pubescens* twigs were collected locally around the reindeer enclosures at the UiT campus (Fig. A1 in appendix A). The fact that the reindeer could have ingested *B. pubescens* in *E1* before they were moved to *E2* is not ideal. However, being aware of it and collecting reference faecal samples (i.e. samples collected before feeding *B. pubescens*) makes it possible to take this into account when analysing and interpreting the results. In order to accurately identify the time interval where the DNA of *B. pubescens* appears in the faeces, time zero of the experiment was set to after *B. pubescens* had been eaten (23rd of January 2018, 21.00) and from this point, all faeces were continuously collected in the following days (Table 2).

Cladonia stellaris

The reindeer were fed an increasing amount of lichen (a mixture containing mainly *Cladonia stellaris*) to investigate if the relative biomass of ingested taxa correlates with the number of output sequences for the taxa (Question 3). To enable the investigation of the relative biomass, RF-80 was fed continuously throughout the feeding experiment, and thereby, the number of output sequences from *C. stellaris* would be relative to the number of output sequences from RF-80

The lichens were collected from various mountain regions in southern Norway, and therefore contained several other undesired taxa (e.g. faeces from *Lepus timidus* L., *Betula pubescens*,

Salix herbacea L. and bryophytes). In an effort to reduce intake of these other taxa, preliminary sorting was conducted. *C. stellaris* is a favourite treat for the reindeer, and therefore, it was expected that the reindeer would consume it relatively rapidly. Another benefit of feeding the lichen mixture is that none of its component species are found in the pellets.

An overview of time and quantity of the introduction of lichens is given in Table 1. The lichen was wet when weighted and distributed, in order to ensure equal amounts fed to the three reindeer. The large differences between the three weights of input lichen biomasses offered to the reindeer over time (i.e. 20 g, 500 g and 2000 g) were chosen to test if it would be detectible in the output sequences. It was also of interest to investigate if 20 g of lichen (~2% of the average ingested diet) would be enough for detection with DNA metabarcoding.

Table 1 - Overview of the stepwise introduction of lichen mixture in the diet, during a period of thirteen days, including four days with only RF-80 in the diet.

Date	Wet weight (g)
24th of January 2018	20
25th of January 2018	20
26th of January 2018	20
27th of January 2018	-
28th of January 2018	-
29th of January 2018	500
30th of January 2018	500
31st of January 2018	500
1st of February 2018	-
2nd of February 2018	-
3rd of February 2018	2000
4th of February 2018	2000
5th of February 2018	2000

2.1.4 Sample collection and subsampling

Faecal samples

A first set of faecal samples was collected from each animal before any new food item was introduced. These reference samples were expected to only contain the species found in the pellets, but also to reveal if the reindeer had eaten any of the species growing in *E1*.

The first 10 days after *B. pubescens* was fed, faeces were collected every 2 hours during the daytime and every 3 hours during the night (Table 2) to be able to detect with a high precision

when the DNA of *B. pubescens* appeared. After the initial 10 days, faeces collection was carried out every ~3 hours during the day only. From the 23rd day to the 30th day of the experiment, one sample was collected each day. After the experiment was completed, samples were collected once or twice a week.

Table 2 – Overview of faecal sample collection intensity during the feeding experiment and in the following month.

Day	Hour intervals day (t)	Hour intervals night (t)
1-10	2	3
11-22	3	0
23-30	24	0
31-60	96-168	0

A total of 498 faecal samples were collected in clean zip lock bags, during the experiment and for the following month. Immediately after collection, all samples were stored at -20°C. From the 498 faecal samples, 234 samples were chosen for further analysis (Table A3-A5 in appendix A). Each day was split into four time intervals (00.00-06.00, 06.00-12.00, 12.00-18.00 and 18.00-24.00) and the last sample collected in each time intervals was chosen for further analysis.

Several of the samples that were collected were large faeces, and for practical reasons regarding transport of samples from UiT to the University of Oslo (UiO), where the remaining laboratory work was done, initial subsamples were taken. The samples were thawed and a smaller part of the faeces was placed in a new zip lock bag (no weights were taken at this stage). In an effort to reduce contamination when possible (i.e. when the faecal sample was a compact mass and not several small pellets), the outer layer was carefully removed and the inner part of the faeces was subsampled. Until transportation and immediately upon arrival at UiO, the samples were stored at -20°C.

For final subsampling prior to DNA extraction, the samples were thawed and thoroughly homogenized within the zip-lock bags. Then, ~250 mg from each sample was placed in a sterile Eppendorf tube and stored again at -20°C.

Food samples

To be able to address Question 1 (Does DNA metabarcoding identify all species to species level in a known diet?), subsamples were taken of all the food items fed to the reindeer (Table

3) to create a sequence reference list and to investigate the taxonomic resolution capacity of the DNA metabarcoding primers used in this study.

Two different production batches of RF-80 were fed to the reindeer and both were analysed to explore similarities or differences amongst them. RF-80 samples were also collected from the feeding tray of each animal to investigate potential contaminants. Additionally, samples from two lichen species occurring in *E1* and frequently consumed by reindeer, were collected in order to investigate whether these sequences were similar to the sequences of the fed lichen mixture.

The subsamples were placed in clean zip-lock bags and stored in a freezer at -20°C until transportation to UiO. At UiO, the food samples were thawed and homogenized using liquid nitrogen, and ceramic pestle and mortar. Around 20 mg of dry weight per sample were placed in sterile Eppendorf tubes for DNA extraction. Two sterile Eppendorf tubes were held open on the lab bench during the process in order to detect any potential cross-contamination during the procedure. These Open Air Negative Controls (Table 3) were treated similarly to the food and the faecal samples during the remaining laboratory steps.

Table 3 – Overview of the food subsampled during the feeding experiment (RF-80 = reindeer pelleted feed), including negative controls (DNA Negative control = DNA extraction control. Open Air Negative Control = eppendorf tubes left open on the lab bench during subsampling). Five subsamples of RF-80, two from different production batches, and three were taken from leftovers in the feeding trays of the reindeer (9/10, 10/10 and 12/10).

Food sample	**Details**
Betula pubescens	Fed to the reindeer (Question 2)
Cladonia stellaris	Lichen mixture fed to the reindeer (Question 3)
RF-80	Batch number 13543
RF-80	Batch number 13541
RF-80	9/10, from feeding tray
RF-80	10/10, from feeding tray
RF-80	12/10, from feeding tray
Lichens	Two lichen species from *E2*
NC_OA_1	Open Air Negative Control
NC_OA_2	Open Air Negative Control

2.2 Gastrointestinal tract

2.2.1 Sample collection

During traditional reindeer slaughtering (24^{th}-26^{th} of September 2017) in western Finnmark, Associate prof. Tove Utsi (UiT) collected samples from four 1.5-year-old male semi-domesticated reindeer. All reindeer were taken by hand into a large gathering pen where they stayed for several days, foraging on the natural pastures. After being culled and slaughtered by a skilled reindeer herder, the gastrointestinal (GI) tract was rapidly taken out of the four animals and one sample was taken from each stomach (rumen, reticulum, omasum and abomasum) and throughout the intestinal tract (Table 4). The samples were placed directly into clean plastic bags to avoid contamination from the environment. Samples from individuals #1, #2 and #4 were stored in a freezer within few hours, whilst samples from individual #3 was stored in a freezer the next day (within 24 hours).

Table 4 – Overview of samples taken from the gastrointestinal tract of four reindeer (1.5-year-old males) on natural pasture in western Finnmark, Norway.

Part of gastrointestinal tract	**Animal ID**			
	#1	#2	#3	#4
Rumen	✓	✓	✓	✓
Reticulum	✓	✓	✓	✓
Omasum	✓	✓	✓	✓
Abomasum	✓	✓	✓	✓
Small intestine 1 (Upper part)	✓	✓	✓	✓
Small intestine 2 (Lower part)	✓	✓		✓
Cecum	✓	✓	✓	✓
Spiral colon	✓	✓	✓	✓
Distal colon	✓	✓	✓	✓

Once transported to UiO for laboratory processing, the samples were stored at -20°C. Prior to subsampling, the samples were thawed and homogenized within the plastic bags. Subsamples were taken by scooping out 100-160 mg of the wet content and placing it in clean Eppendorf tubes. These tubes were stored at -20°C until DNA extraction.

2.3 DNA extraction

DNA was extracted from all samples, in random order (228 faecal samples from feeding experiment, 8 food samples, 35 samples from the GI tract and 2 negative “open air” controls)

using the Qiagen DNeasy® PowerSoil® Pro Kit. The manufactures' protocol was followed, except that solution C3 was heated at 56°C (maximum temperature of the heating cabinet), not at 65°C as recommended for dissolving the precipitates. Additionally, several of the GI tract samples had to be centrifuged repeatedly in order to move the lysate to the column due to the high viscosity of those samples. In every extraction round, 1-2 negative controls were included in order to monitor for possible contamination. One sample from the feeding experiment (X73) evaporated, and was re-eluted with 30 µl of solution C6 (less than the 100 µl recommended in the protocol) to secure high enough concentration of DNA.

2.4 PCR amplification and high-throughput sequencing

Ecological samples contain small amounts of DNA from many different species, and in faecal samples the DNA is usually degraded due to digestive processes. To increase the amount of DNA that is of interest, and thereby enhance the chances of identifying as many taxa as possible, PCR is used to target specific short DNA regions for amplification (Pompanon et al., 2012; Taberlet et al., 2012). Based on existing knowledge on reindeer diet (Nieminen and Heiskari, 1989; Storeheier et al., 2002b, 2003) and considering the food items offered during the feeding experiment, all samples were amplified using four different primer sets (Table 5): Sper01, Euka02, Fung01 and Bryo01, amplifying seed plants, eukaryotes, fungi and bryophytes, respectively. Each primer also contained an 8 or 9 bp sequence tags allowing identification of each sample (demultiplexing) after the sequencing process, where several samples were pooled together.

Table 5 – Overview of the four different primers used for DNA amplification, including the targeted taxonomic groups, primer names (Taberlet et al., 2018) and the gene regions they target, as well as forward and reverse primer sequences and original publication. (Abbreviations: cp = chloroplast, r = ribosomal, n = nuclear.)

Taxonomic group	**Primer name**	**Target**	**Forward primer**	**Reverse primer**	**Reference**
Eukaryotes	Euka02	18S, rDNA (V7 region)	TTTGTCTGSTTA ATTSCG	CACAGACCTGTT ATTGC	Guardiola et al., 2015
Fungi	Fung01	ITS1, nrDNA	GGAAGTAAAAG TCGTAACAAGG	CCAAGAGATCCG TTGYTGAAAGT	Epp et al., 2012
Spermatophyta (seed plants)	Sper01	p6 loop of the *trn*L intron, cpDNA	GGGCAATCCTG AGCCAA	CCATTGAGTCTCT GCACCTATC	Taberlet et al., 2007
Bryophytes	Bryo01	Part of the *trn*L intron, cpDNA	GATTCAGGGAA ACTTAGGTTG	CCATYGAGTCTC TGCACC	Epp et al., 2012; Taberlet et al., 2018

PCR plates with 96-well were used and because there was 228 faecal samples from the feeding experiment, they were distributed between three PCR plates (Table A6-A14 in appendix A). For the remaining samples (8 food samples, 35 samples from the GI tract), there were few enough to fit several replicates on one PCR plate (Table A15-A17 in Appendix A and Table B2 in appendix B).

Each PCR reaction contained 7.5 µl of AmpliTaq Gold 360 PCR Master Mix (Thermo Fisher Scientific, USA) at 1 mM, 0.6 µl of Bovine serum albumin (BSA, Sigma-Aldrich, USA) at 1 mM, 0.75 µl of tagged primers mix (Integrated DNA Technologies, USA), 4.5 µl of MilliQ H_20, and 2 µl of undiluted DNA. The primer mix was added by postdoctoral fellow Stefaniya Kamenova.

Positive controls were included in the PCR plates amplifying seed plant DNA from faecal and food samples in the feeding experiment (Table A12-A14 and A17 in appendix A), and in one of the PCR plates amplifying eukaryote DNA from faecal samples in the feeding experiment (Table A7 in appendix A), to confirm that the PCR reactions were successful. The DNA in the positive controls was artificially produced for the primers amplifying seed plants and eukaryotes. Positive controls were not included in the remaining PCR plates, because artificially produced DNA was not available.

To ensure reproducible results, each PCR reaction was replicated three times. However, PCR plate 3 of the feeding experiment (Table A8, A11 and A14 in appendix A) only had one replicate due to an error during the primer mix pipetting. Each PCR replicate contained 1-4 PCR negative controls, which were used to identify any potential contaminants. The following PCR conditions were used: Initial denaturation at 95°C for 10 min, followed by 35 cycles consisting of denaturation at 95°C for 30 s, annealing (temperature depending on the primers used; Euka02: 45°C; Fung01: 56°C; Sper01: 52°C; Bryo01: 54°C) for 30 s, elongation at 72°C for 1 min, and a final elongation at 72°C for 7 min before a final hold at 15°C. During PCR, several samples evaporated, possibly because the lids were not being properly sealed around the edges of the PCR plates. All of the samples that evaporated were continuously noted, to know which replicates were not included in pooling and sequencing (Table A18 in appendix A and Table B3 in appendix B).

To confirm that amplification was successful, between 8 and 16 samples from each PCR plate were visually inspected using a 1% agarose gel electrophoresis. A mix of 2 µl amplified DNA

and 1 µl loading dye was loaded into the wells of the gel. After the initial round of PCR, it was clear that bryophytes were not present in the samples from the feeding experiment. This was expected, as the reindeer were not fed any bryophytes. Therefore, for the remaining samples from the feeding experiment, only fungi, seed plants and eukaryotes were amplified. Bryophytes were still amplified for the GI tract samples because their diet was unknown.

Preparation of the samples for high-throughput sequencing was done by postdoctoral fellow Stefaniya Kamenova. The procedure involved pooling all the samples into eight different libraries (Table 6 and Table C4 in appendix C) based on DNA concentration measured with the Qubit® 2.0 Fluorometer (Qubit® dsDNA HS Assay Kit, Invitrogen™, Table C3 in Appendix C). The Norwegian Sequencing Centre prepared the libraries using a TruSeq PCR-free protocol (Illumina, USA) and carried out the 150bp pair-end sequencing using an Illumina HiSeq 4000 machine.

Table 6 – Overview of the content in the eight libraries. Sample type (faecal sample or food sample) is specified for the feeding experiment, as well as PCR plate number (P1, P2, P3) for the faecal samples in Library 1-3 and Library 9-11. Library 5 and Library 12 contain samples from both the feeding experiment and the gastrointestinal tract. (Abbreviations: GI tract = Gastrointestinal tract)

Library number	**Feeding experiment, sample type**	**Feeding experiment, primer**	**GI tract, primer**
1	Faeces, P1	Sper01	
2	Faeces, P2	Sper01	
3	Faeces, P3	Sper01	
5	Food	Sper01	Sper01 and Bryo01
9	Faeces, P1	Euka02 and Fung01	
10	Faeces, P2	Euka02 and Fung01	
11	Faeces, P3	Euka02 and Fung01	
12	Food	Euka02 and Fung01	Euka02 and Fung01

2.5 Bioinformatics

After receiving the sequences from the Norwegian Sequencing Centre, raw data analysis was conducted in UNIX, using the program OBITools (http://metabarcoding.org/obitools). The forward and reverse pair-end reads were aligned and merged using *illuminapairedend* (Boyer et al., 2016). Thereafter, reference files for each primer set used (i.e. Sper01, Euka02, ITS01 or Bryo01) were created, containing information about the primer set and all sample names as well as the combination of tags and primer sequences used. The reference files were used with the *ngsfilter* command (Boyer et al., 2016) to assign sequences to samples. For the

assignation, *ngsfilter* specifically uses the tags attached to the primers and compare them to the tags in the reference file.

Several types of errors can occur during PCR and sequencing, and different programs were implemented to deal with the errors. The OBITools program *obigrep* was used in UNIX to filter the data (i.e. discard chimeras or other artefact sequences (containing other bases than A, C, G and T)) and remove PCR errors (e.g. very short or very long sequences). To identify erroneous PCRs (i.e. sequences flagged as rare or chimeric sequences (singletons) or sequences flagged as PCR substitution or errors (internals); De Barba et al., 2014), the OBITools program *obiclean* was used in UNIX. However, the datasets containing fungal sequences were too complex for *obiclean*, and therefore erroneous PCRs were not identified and flagged in these datasets.

To construct a reference database for the taxonomic assignment of the sequences, the program *ecopcr* (OBITools program) was run to access the European Molecular Biology Laboratory (EMBL) Nucleotide Sequence Database (Baker et al., 2000). Specified minimum and maximum sequence lengths for each primer were set (Sper01=10-250 bp, Euka02=20-900 bp, Fung01=50-1000 bp, Bryo01=30-150 bp), and a maximum of three mismatch errors allowed on the priming sites. The minimum and maximum sequence lengths were based on the minimum and maximum DNA length of each taxonomic group targeted by each of the four primer sets (Taberlet et al., 2018). The program *ecotag* (OBITools program) was used to conduct the taxonomic assignment, using The EMBL Database. Once taxonomic assignation was completed, sequences from faecal samples from the feeding experiment amplified with the same primer (i.e. Library 1, 2 and 3 (Sper01), Library 9, 10 and 11 (Fung01), and Library 9, 10 and 11 (Euka02)) were merged. Then, *obiuniq* (OBITools program) was used to merge identical sequences, in order to remove duplicates and triplicates. However, as the datasets containing fungal sequences were too large for *obiuniq*, *sumaclust* (git.metabarcoding.org/obitools/sumaclust) was used instead for clustering this dataset. Additionally, *obigrep* was used to remove singletons in the datasets containing seed plant and eukaryote sequences.

After initial inspection of the faeces data from the feeding experiment in R Studio (R Studio Team, 2015), samples from PCR plate 3 (Table 6) were separated from the other two plates, as they lacked PCR replicates. The sequences from PCR plate 3 were investigated, and as the sequences in the controls (positive and negative) were only found in the control samples,

controls (positive and negative) were removed. Spurious sequences (i.e. generated by the PCR and sequencing errors) from the two plates with three PCR replicates (i.e. PCR plate 1 and 2) were filtered out using the packages ROBITools, ROBITaxonomy and Vegan (LECA - Laboratoire d'ecologie alpine, 2012a; LECA - Laboratoire d'ecologie alpine, 2012b; Oksanen et al., 2019). After filtering, the replicates in the samples from PCR plate 1 and 2 were merged together using the Dplyr package (Wickham et al., 2018). Finally, the two datasets (one containing samples from PCR plate 1 and 2, and one containing samples from PCR plate 3) were merged, also using the Dplyr package (Wickham et al., 2018).

For the datasets containing both the food samples fed during the feeding experiment and the samples from the GI tract, sequences flagged singletons or internals (De Barba et al., 2014) were removed using R Studio (except for the datasets containing fungal sequences). Thereafter, they were filtered in the same way as described above for PCR plate 2 and 3. Once filtering was completed, replicates were merged using the Dplyr package (Wickham et al., 2018).

Individual thresholds for identity match (i.e. how well the sequence matches with a sequence in the reference database) were set for each dataset (Table 7). For this, the length of the amplified sequences was taken into account (i.e. Euka02 and Fung01 amplify longer DNA fragments compared to Sper01 and Bryo01) as well as an evaluation of plots showing the maximum sequence read frequency in a sample against best identity score with the reference database. Additionally, rare sequences (less than 1% of the total number of reads in a sample) were filtered out. Once filtering was completed, additional information about the samples was added (e.g. time, date, names of the animals in the feeding experiment (9/10, 10/10 and 12/10), part of GI tract). For the GI tract datasets, samples were removed if there was only one sample left from that particular part of the GI tract after filtering.

Table 7 – Thresholds set for each dataset, for identity match with the database. The datasets are differentiated by content (i.e. experiment, sample type and targeted taxonomic group).

Experiment	Sample type	Taxonomic group	Threshold
Feeding experiment	Faeces	Eukaryotes	50 %
Feeding experiment	Faeces	Fungi	60 %
Feeding experiment	Faeces	Seed plants	95 %
Feeding experiment	Food	Eukaryotes	95 %
Feeding experiment	Food	Fungi	65 %
Feeding experiment	Food	Seed plants	90 %
Gastrointestinal tract		Eukaryotes	85 %
Gastrointestinal tract		Fungi	80 %
Gastrointestinal tract		Seed plants	90 %
Gastrointestinal tract		Bryophytes	90 %

For the files containing amplified seed plant sequences, and in some of the PCR replicates with amplified eukaryote sequences (Table A7 in appendix A), positive controls were used to verify the success of the PCR amplifications. The positive control sequences were identified in the positive control samples, confirming that PCR amplified successfully, and no cross-contamination was identified (i.e. no positive control sequences leaked into other wells on the PCR plate). The positive control sequences were then colour labelled in the plot showing the maximum frequency of number of reads in a sample against best identity score with the reference database. When filtering out rare sequences, the positive control sequences were removed, indicating that although the positive control sequences were detected, they were relatively rare compared to the food sequences identified by DNA metabarcoding.

2.6 Downstream statistical analysis

Éric Coissac created a general pipeline for statistical analyses in R Studio for the analysis of the GI tract samples and the faecal samples from the feeding experiment. I adapted and developed the pipeline for analysing each file corresponding to each of my research questions.

Investigation of Molecular Operational Taxonomic Units (MOTUs)

To answer Question 1 (Does DNA metabarcoding identify all species to species level in a known diet?), Question 2 (How long is the DNA of a particular food item present in the digestive system?), and Question 3 (How well is relative food biomass represented by output sequences?) the ROBITools package and the *is.subcladeof* command was used to identify the molecular operational taxonomic units (MOTUs) related to the species of interest; *Betulaceae*

(Question 1 and 2), Lecanoromycetes (Question 1 and 3), and all taxa from RF-80 (Question 1, Table A1 in appendix A). The *is.subcladeof* command uses the EMBL database to extract the taxonomic identity (taxid) for all the subclades of a given taxon. The EMBL database was uploaded in R Studio using the ROBITools package and the *read.taxonomy* command.

After identifying MOTUs related to the species of interest, I isolated all the sequences identified as subclades of *Betulaceae*, Lecanoromycetes, and as subclades of all taxa from RF-80 (Question 1). The MOTUs identified as subclades of *Betulaceae* were compared between the faecal and the food samples, for the primer sets Sper01 and Euka02. The MOTUs identified as subclades of Lecanoromycetes were compared between the faecal and the food samples, for the primer sets Fung01 and Euka02. In order to examine the accuracy of the species detection and identification, the MOTUs identified as subclades of taxa from the RF-80 samples were compared to the ingredient list of RF-80 (Table A1 in appendix). The results from these comparisons established which MOTUs were explored further in Question 2 and 3.

Also related to Question 1 (Does DNA metabarcoding identify all species to species level in a known diet?), the 20 most dominant MOTUs identified in the food samples were compared to the species in the ingredient list of RF-80, *B. pubescens* or *C. stellaris* to see how well they corresponded (Table A22-A31 in appendix A). The relative frequency of each sequence in the dataset was estimated using the *decostand* command from the Vegan package. A similar comparison was made with the datasets containing amplified sequences from the faecal samples (Table A19-A21 in appendix A).

Plots were made with the ggplot2 package (Wickham, 2016) for visualizing the proportion of output sequences over time for *Betulaceae* amplified with the Sper01 primer (Question 2). A similar plot was made using the dataset containing the proportion of output sequences over time for *Betula* DNA amplified with the Euka02 primer (Figure A2 in appendix A). To further investigate when *Betulaceae* DNA appeared, a plot was created using the ggplot2 package and the dataset containing faecal samples amplified with the seed plant primer (Sper01), selecting a time frame spanning between 40 hours prior to feeding and 100 hours after introducing *B. pubescens* in the reindeer diet. A comparison was made with the same time frames, using the Euka02 dataset for the sequences identified as *Betula*. To explore the decay of *B. pubescens*, I selected the samples collected 100-600 hours after feeding *B. pubescens*, and log-transformed the proportion of *Betulaceae* DNA (Sper01) and *Betula* DNA (Euka02) in the detected diet, using the ggplot2 package. By log-transforming the data, it is possible to

visualize the continuous decrease of DNA because the variation is magnified, in contrast to the non-transformed data. Different taxonomic assignations (*Betulaceae* and *Betula*) were used according to the dataset (Sperm01 versus Euka02) based on the results from Question 1.

To visualize the proportion of sequences, likely to represent *C. stellaris*, from the dataset with amplified fungi (Fung01) and eukaryotes (Euka02) over time, I used the ggplot2 package and targeted the family *Cladoniaceae* (Fung01 dataset) and the subclass of Lecanoromycetidae (Euka02 dataset). Different taxonomic assignations were used for the two datasets (*Cladoniaceae* and Lecanoromycetidae) based on the results from Question 1.

Decay time ($T_{1/2}$)

A linear model (1) was fit for the each animal results (9/10, 10/10 and 12/10) from the feeding experiment, to investigate the decay time ($T_{1/2}$, i.e. the time required for the proportion of DNA to be halved) of *B. pubescens* amplified with both the Sper01 and the Euka02 primers. The same model was applied for the *C. stellaris* amplified with the Fung01 and Euka02 primers.

$$lm(tX \sim log\ (X,2)) \quad (1)$$

In the linear model (1), X corresponds to the number of reads of *B. pubescens* (*Betulaceae* (Sper01) or *Betula* (Euka02)), and to the number of reads of *C. stellaris* (*Cladoniaceae* (Fung01) or Lecanoromycetes (Euka02)) within a given time interval for a specific animal; 2 – is the basis of logarithm. While tX defines the samples collected and analysed (i.e. observations) within a given time interval for a specific animal. For each of the datasets, the time intervals were set between the highest proportion of DNA and the following lowest proportion of DNA.

Multivariate analyses

The number of output sequences for each primer set for the GI tract analysis were Hellinger-transformed (Legendre & Gallagher, 2001) using the Vegan package. Then, Principal Component Analysis (PCA) was run, using the *dudi.pca* command from the Ade4 package (Dray and Dufour, 2007). The aim of PCA is to reduce dimensionality of a multivariate data set to a new set of variables called Principal Components (PCs). These variables summarize the maximum amount of variation in a dataset as linear combinations of the original variables

and can be visualized graphically, to compare samples with minimal loss of information. Colours were superimposed based on Animal ID or the parts of the GI tract, on plots of the two first PCs to compare the variation. The results from these analyses are shown in Fig. B1-B4 in appendix B.

Dietary diversity

Because it is impossible to directly compare diet across the GI tract (each part of the GI tract represents meals ingested at different time points), species (MOTU) richness and diversity for each part of the GI tract was estimated and compared among the different primer sets. For this, different diversity indices were calculated based on the Hill's number (Hill, 1973; Keylock, 2005). By plotting different q-values, visual representations of species (MOTU) richness (q=0), Shannon diversity index (q=1, H) and Gini coefficient (q=2) were made (Broms et al., 2015), using the ggplot2 package (Wickham, 2016). An additional diversity comparison was made by computing q=3. The results from the Gini coefficient (q=2) and q=3 are shown in boxplots in Fig. B5 and B6 in appendix B.

2.7 Ethical statement

DNA metabarcoding is a non-invasive method, for which ecological samples (e.g. soil, faeces) are used to detect all the species present.

Use of animals in research/experiments is regulated by the "Forskrift om bruk av dyr i forsøk" (FOR-2015-06-18-761) and usually requires a permit from the Norwegian Food Safety Authority. However, these regulations do not apply under conditions listed in §2, part 5, items 'a-f', which the present experiments were considered to conform with (§2, part 5, item 'f'). The feeding experiment therefore did not require any permits from The Norwegian Food Safety Authority.

Moreover, collection of samples from animals that are culled for other purposes is not considered to represent an animal experiment (§4, item 'a'), as long as an approved technique for euthanasia is employed (see §16 and Appendix C). The reindeer GI tract samples were collected during the traditional autumn slaughter in the reindeer herding family of Associate Prof. Tove Utsi and did not require extra animals being sacrificed. On this basis, an ethical permit was not required for the present study.

3 Results

The number of output sequences received from the Norwegian Sequencing Centre varied between ~13 million and ~17.8 million, for the different libraries (Table 8).

Table 8 – Overview of the number of output sequences in each library after Illumina Hiseq sequencing.

Library	Taxonomic group(s)	Number of output sequences
1	Seed plants	17 783 634
2	Seed plants	17 513 982
3	Seed plants	15 426 679
5	Seed plants and Bryophytes	17 779 834
9	Eukaryotes and Fungi	15 205 413
10	Eukaryotes and Fungi	15 657 638
11	Eukaryotes and Fungi	13 057 009
12	Eukaryotes and Fungi	16 868 076

After all the filtering steps were carried out in UNIX and R Studio, the number of MOTUs detected in the different datasets varied between 15 and 472 (Table 9). However, not all of the MOTUs identified in the samples from the feeding experiment represent food items fed to the reindeer (Table A19-A31 in appendix A). Although filtering removed the contaminant sequences found in the negative controls, there were still several sequences left in the samples that do not correspond to the fed species. For instance, some of the MOTUs identified in the dataset containing amplified fungal sequences from the faecal samples are most likely part of the microbiota in the GI tract, and were therefore not filtered out as contaminants. Because the controlled diet is known, the MOTUs corresponding to the thirteen fed species were extracted, in order to proceed with further analyses.

Table 9 – Overview of the number of MOTUs in each dataset after filtering. The datasets are differentiated by content (i.e. experiment, sample type and taxonomic group)

Experiment	Sample type	Taxonomic group	Number of MOTUs
Feeding experiment	Faeces	Eukaryotes	116
Feeding experiment	Faeces	Fungi	472
Feeding experiment	Faeces	Seed plants	67
Feeding experiment	Food	Eukaryotes	15
Feeding experiment	Food	Fungi	74
Feeding experiment	Food	Seed plants	77
Gastrointestinal tract		Eukaryotes	42
Gastrointestinal tract		Fungi	110
Gastrointestinal tract		Seed plants	56
Gastrointestinal tract		Bryophytes	68

3.1 Does DNA metabarcoding identify all species to species level in a known diet?

RF-80

The best detection of the species in RF-80 occurred within the reindeer faecal samples when amplified with the seed plant primer (Sper01, Table 10). Although no MOTUs were identified at species level, four MOTUs were assigned to genus level, three MOTUs at tribe level, two MOTUs at subfamily level and one MOTU at family level. Some of the identified MOTUs correspond to the same species in RF-80, e.g. three taxa in the ingredient list (Table A1 in appendix A) belong to Pooideae and are thus potential sources for the output sequences identified as Pooideae (Table 10). This means that potentially eight of the eleven species in RF-80 were identified in the faecal samples with the seed plant primer.

In comparison, only five species from the RF-80 ingredients list were identified in the food samples when amplified with Sper01 (Table 10). No MOTUs were identified at species level and only one MOTU identified at genus level, one MOTU at tribe level and one MOTU at subfamily level (Pooideae, as mentioned above). Additionally some MOTUs are identified to the PACMAD clade, which is a monophyletic group, not assigned a taxonomic rank in the National Center for Biotechnology Information (NCBI) taxonomy database (Sayers et al., 2009; www.ncbi.nlm.nih.gov/taxonomy). The PACMAD clade is one of the large clades within Poaceae, and includes the subfamily Panicoideae that *Saccharum officinarum* L. (on the RF-80 ingredient list) belongs to (Soreng et al., 2015).

Four of the species in RF-80 were identified in both faecal samples and food samples when amplified with the eukaryote primer (Euka02). Amongst the MOTUs detected from RF-80 (Table 10), only one was identified at species level (*Triticum aestivum* L.). The remaining MOTUs were identified at family and subfamily level in both faecal and food samples.

Table 10 – Overview of the taxonomic resolution (Rank) of the identified MOTUs that correspond to a species in the ingredient list of RF-80, including both faecal and food samples amplified with the Sper01 (seed plants) primer and the Euka02 (eukaryotes) primer.

Taxonomic group	Sample type	Identified MOTU	Rank	Scientific name (RF-80)
Seed plants	Faeces	*Picea*	Genus	*Picea abies*
		Pinus	Genus	*Pinus sylvestris*
		Chenopodioideae	Subfamily	*Beta vulgaris*
		Medicago	Genus	*Medicago sativa*
		Brassiceae	Tribe	*Brassica* sp.
		Brassicaceae	Family	*Brassica* sp.
		Avena	Genus	*Avena sativa*
		Poeae	Tribe	*Avena sativa*
		Triticeae	Tribe	*Triticum aestivum*
		Pooideae	Subfamily	*Avena sativa, Hordeum vulgare, Triticum aestivum*
	Food	*Pinus*	Genus	*Pinus sylvestris*
		Poeae	Tribe	*Avena sativa*
		Pooideae	Subfamily	*Avena sativa, Hordeum vulgare, Triticum aestivum*
		PACMAD clade	No rank	*Saccharum officinarum*
Eukaryotes	Faeces	Brassicaceae	Family	*Brassica* sp.
		Triticum aestivum	Species	*Triticum aestivum*
		Pooideae	Subfamily	*Avena sativa, Hordeum vulgare, Triticum aestivum*
	Food	Brassicaceae	Family	*Brassica* sp.
		Pooideae	Subfamily	*Avena sativa, Hordeum vulgare, Triticum aestivum*

Betula pubescens

The species detection accuracy of *B. pubescens* in faecal and food samples, for both the seed plant primer (Sper01) and the eukaryote primer (Euka02), is presented in Table 11. Neither of the primers identified *B. pubescens* to species level. However, higher taxonomic resolution was achieved with Euka02 (identification at genus level, *Betula*) than with Sper01 (identification at family level, *Betulaceae*). Additionally, the MOTU with the highest mean relative frequency in the food sample of *B. pubescens*, amplified with the seed plant primer was Dryadoideae (Table A26 in appendix A).

Table 11 – Overview of the taxonomic resolution (Rank) of the identified MOTUs that are related to *Betula pubescens*, including both faecal and food samples amplified with the Sper01 (seed plants) primer and the Euka02 (eukaryotes) primer.

Taxonomic group	Sample type	Identified MOTU	Rank
Seed plants	Faeces	*Betulaceae*	Family
	Food	*Betulaceae*	Family
Eukaryotes	Faeces	*Betula*	Genus
	Food	*Betula*	Genus

Cladonia stellaris

Species detection accuracy of *C. stellaris* in faecal and food samples, for both the fungal primer (Fung01) and eukaryote primer (Euka02), is shown in Table 12. In contrast to the results for *B. pubescens*, Euka02 was outperformed by Fung01 in terms of taxonomic resolution (*Cladonia* and Lecanoromycetidae, respectively). Still, none of the primers were able to identify *C. stellaris* to species level.

The lichen species from *E1* were identified as *Parmelia barrenoae* Divakar, Molina and Crespo, *Abrothallus* or Leotiomyceta, when using the fungal primer to amplify the food sample. Thus, DNA metabarcoding is able to differentiate the analysed lichen species when targeting fungal sequences.

Table 12 – Overview of the taxonomic resolution (Rank) of the identified MOTUs that are related to *Cladonia stellaris*, including both faecal and food samples amplified with the Fung01 (fungi) primer and the Euka02 (eukaryotes) primer.

Taxonomic group	Sample type	Identified MOTU	Rank
Fungi	Faeces	*Cladonia*	Genus
	Faeces	*Cladoniaceae*	Family
	Faeces	Lecanorineae	Suborder
	Faeces	Lecanoromycetidae	Subclass
	Food	*Cladoniaceae*	Family
Eukaryotes	Faeces	Lecanoromycetidae	Subclass
	Food	Lecanoromycetidae	Subclass

The food samples taken of *C. stellaris* were also amplified with the seed plant primer (Sper01), to investigate which plant species were present in the lichen mixture fed to the reindeer and that could be detected in the faecal samples. The MOTU with the highest mean relative frequency amongst the amplified seed plant (Sper01) sequences in the food samples taken of *C. stellaris* was Dryadoideae (Table A29 in appendix A).

3.2 For how long is DNA of a particular food item present in the digestive system?

By isolating the sequences identified as *Betulaceae* (Fig. 1) from the dataset containing faecal samples amplified with the seed plant primer (Sper01), it is apparent that 1) *Betulaceae* DNA was easily detectable, 2) it was possible to identify when *Betulaceae* DNA appeared (i.e. initial increase of *Betulaceae* DNA after time 0) and follow the decay of *Betulaceae* DNA and 3) *Betulaceae* DNA was present throughout the feeding experiment and during the rest subsequent sampling period. Fig. 1 shows nearly identical patterns for all three animals, with an initial increase after introduction of *B. pubescens*, followed by a decrease in the proportion of *Betulaceae* DNA. After 26 days the reindeer were reintroduced to *E1,* and from day 27 until the end of the sampling period *Betulaceae* DNA dominated the faecal samples. Similar patterns can be seen in the proportion of *Betula* DNA from the dataset containing faecal samples amplified with the eukaryote primer (Euka02, Fig. A2 in appendix A). Overall, the samples from PCR plate 3 (with no PCR replicates) followed the general pattern described above. Only one sample from PCR plate 3, collected at the 10th day of the feeding experiment, did not follow the general pattern. Approximately 75% of the seed plant DNA in the faecal sample from animal 9/10 was identified as *Betulaceae*, whilst the remaining samples collected around this time contained less than 10% of DNA from *Betulaceae* (Fig. 1).

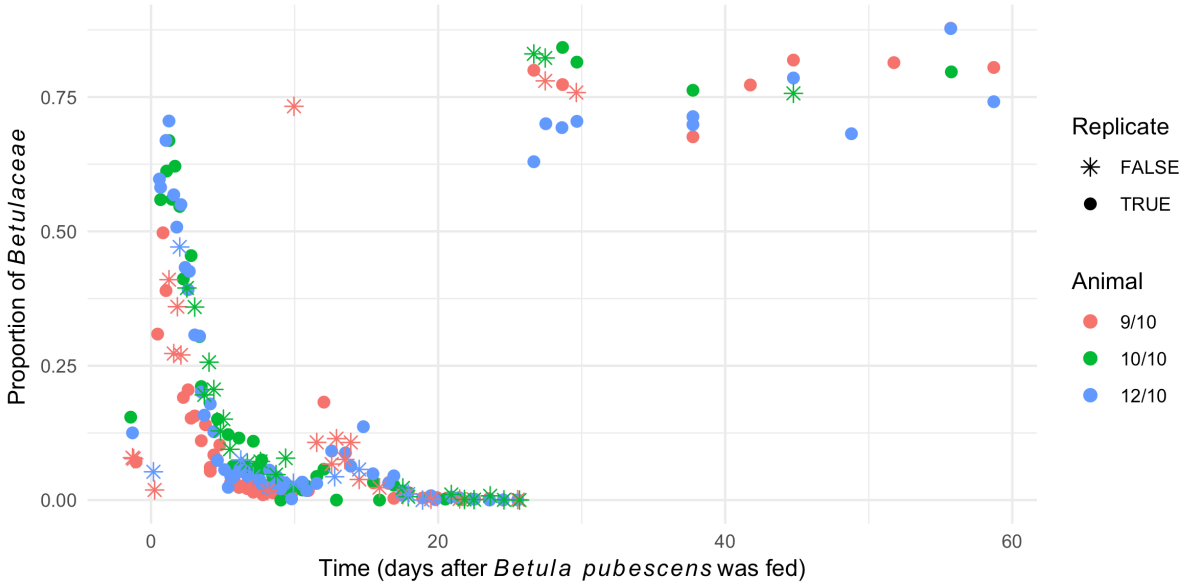

Figure 1 – Proportion of *Betulaceae* DNA in reindeer faecal samples, from the seed plant (Sper01) dataset. The three reindeer included in the experiment are represented by different colours. Samples without replicates (PCR plate 3) appear with a star sign, whilst the samples with three replicates are shown with a filled circle.

When looking at the 40 hours before and 100 hours after introducing *B. pubescens* (Fig. 1 and 2) it is clear that *Betulaceae* DNA was present in the faeces prior to the experiment. Approximately 12.5% of the seed plant DNA in the faecal samples collected before introducing *B. pubescens* (time 0) was identified as *Betulaceae* (Fig. 2A). For the samples amplified with the eukaryote primer (Euka02), less than 2.5% of the DNA was identified as *Betula* before time 0 (Fig. 2B). The faecal samples collected within the first 6 hours after introducing *B. pubescens* have an even lower proportion of *Betulaceae*/*Betula* in the faecal samples.

After introducing approximately 50 g of *B. pubescens*, it took ~12 hours for *Betulaceae*/*Betula* DNA to start appearing (i.e. initial increase of DNA after time 0) in the faecal samples, and ~25 hours to reach a maximum proportion of sequences for both primers (Fig. 2). The main difference between the two datasets (Sper01 and Euka02) is seen when focusing on individual reindeer. For the dataset containing faecal samples amplified with the seed plant primer (Sper01, Fig. 2A), animal 12/10 had the highest proportion of *Betulaceae* sequences (almost 75%), closely followed by animal 10/10, whilst animal 9/10 peaked at 50%. For the dataset containing faecal samples amplified with the eukaryote primer (Euka02, Fig. 2B), on the other hand, animal 9/10 had the highest proportion of *Betula* sequences (almost 12.5%), whilst animal 12/10 peaked at 7.5%. It is also interesting that the time between the highest proportions of *Betulaceae* sequences was relatively similar for the three animals when amplifying with the seed plant primer (Sper01, Fig. 2A), whereas with the eukaryote primer (Euka02, Fig. 2B) there was more variation with regard to when the highest proportion of *Betula* DNA occurred for the different animals.

Immediately after reaching the peak (i.e. highest proportion of *Betulaceae*/*Betula* DNA), the proportion of *Betulaceae*/*Betula* DNA dropped quickly (Fig. 2). Based on the results of the linear model of the initial decay of *Betulaceae* DNA in the dataset containing faecal samples amplified with the seed plant primer (Sper01), it appeared that the $T_{1/2}$ was 31 hours for animal 9/10, 37 hours for animal 10/10 and 27 hours for animal 12/10. For the dataset containing faecal samples amplified with the eukaryote primer (Euka02), the $T_{1/2}$ of *Betula* DNA was 22 hours for animal 9/10, 19 hours for animal 10/10 and 13 hours for animal 12/10.

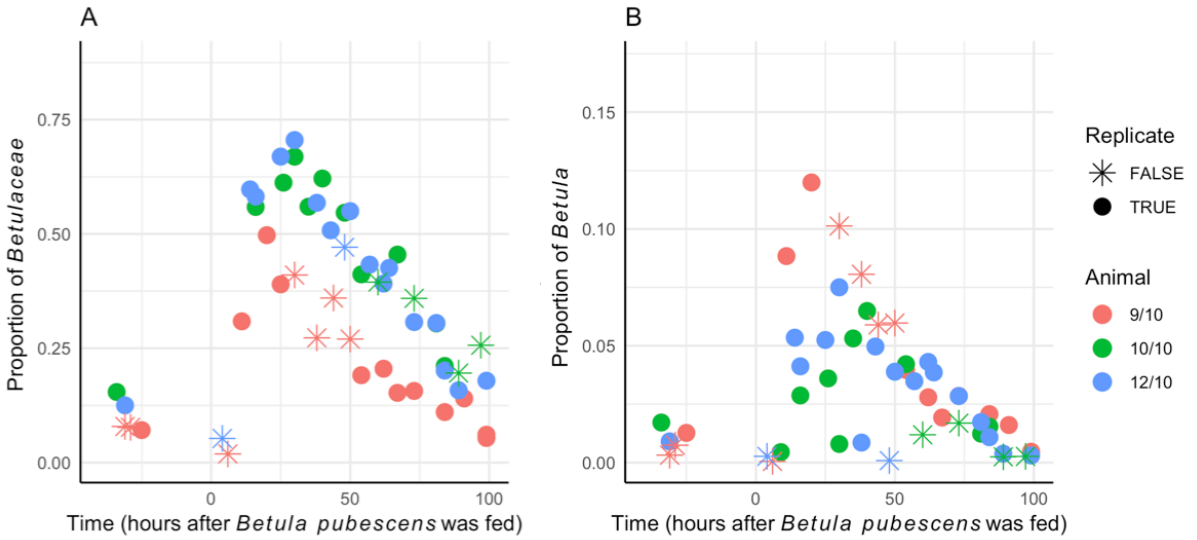

Figure 2 – Proportion of *Betulaceae* (A) and *Betula* (B) DNA during the 40 hours prior to feeding *B. pubescens* and 100 hours after introducing *B. pubescens*, in the datasets containing faecal samples amplified with seed plant (A, Sper01) and eukaryote (B, Euka02) primers. The three reindeer included in the experiment are represented by different colours. Samples without replicates (PCR plate 3) appear with a star sign, whilst the samples with three replicates are shown with a filled circle.

Betulaceae/*Betula* DNA never completely disappeared from the faeces (Fig. 1 and Fig. A2 in appendix A) but rather continuously appeared in low proportions. However, after log-transforming the values of the proportions of *Betulaceae*/*Betula* DNA, during the period with low proportions (100-600 hours after feeding *B. pubescens*), it was evident that the *Betulaceae*/*Betula* DNA was still decaying (Fig. 3). Although the log-transformed values of the proportion of *Betulaceae*/*Betula* DNA in the faecal samples were different between the two datasets, similar patterns of decay were observed for all the three animals, for both dataset containing faecal samples amplified with the seed plant primer (Sper01, Fig. 3A) and with the eukaryote primer (Euka02, Fig. 3B),

A small peak (i.e. increased proportion) was observed around 300 hours after *B. pubescens* was fed when looking at the dataset containing faecal samples amplified with the seed plant primer (Sper01, Fig. 1 and 3A). For the dataset containing faecal samples amplified with the eukaryote primer (Euka02), the small peak appeared between 350 and 400 hours after *B. pubescens* was fed (Fig. 3B and Fig. A2 in appendix A). From ~450 hours to ~620 hours after introduction of *B. pubescens*, less than 1% of the seed plant DNA corresponded to *Betulaceae* DNA (Fig. 3A). In the dataset containing faecal samples amplified with the eukaryote primer (Euka02), DNA of *Betula* constituted less than 1% of the total DNA from ~100 hours to ~620

hours after introduction of *B. pubescens*, and decreased to 0.1% of the total DNA after ~450 hours after introduction of *B. pubescens* (Fig. 3B).

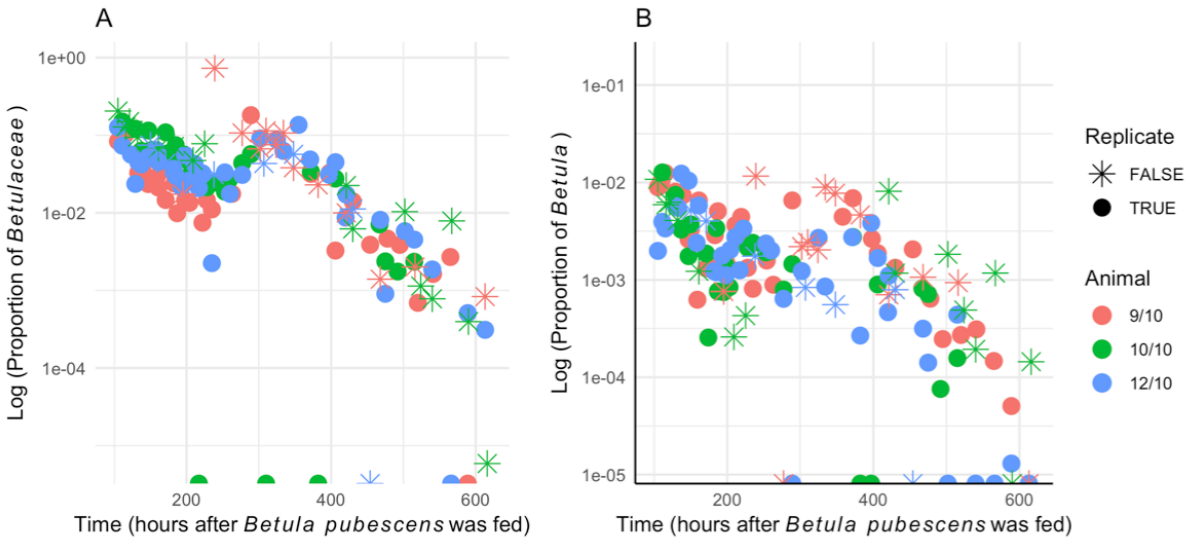

Figure 3 – Decay of *Betulaceae* (A) and *Betula* (B) DNA shown in log transformed proportion of all sequences in the datasets when amplified with seed plant (A, Sper01) and eukaryote (B, Euka02) primers. The three reindeer included in the experiment are represented by different colours. Samples without replicates (PCR plate 3) appear with a star sign, whilst the samples with three replicates are shown with a filled circle. The timeline (x-axis) shows the 40 hours prior to feeding *B. pubescens* and 100 hours after introducing *B. pubescens*.

3.3 How well is relative food biomass represented by output sequences?

By targeting sequences identified as subclades of *Cladoniaceae* in the dataset containing faecal samples amplified with the fungal primer (Fung01), and as subclades of Lecanoromycetes in the dataset containing faecal samples amplified with the eukaryote primer (Euka02), it was possible to visualise the changes in proportion of *Cladoniaceae*/Lecanoromycetes DNA during the days when different amounts of *C. stellaris* were fed (Fig. 4 and 5). Choice of primer (i.e. Fung01 or Euka02) had a large impact as the dataset containing faecal samples amplified with the eukaryote primer portrayed the gradual increase of fed *C. stellaris*, whilst the dataset containing faecal samples amplified with the fungal primer did not (Fig. 4 and 5).

The proportion of output sequences identified as subclades of *Cladoniaceae* in the dataset containing faecal samples amplified with the fungi primer (Fung01) was plotted to create a complete visualization of the taxa of interest, regardless of their taxonomic rank (Fig. 4). There was a very small amount of *Cladoniaceae* present in the faecal samples collected before *C. stellaris* was fed, but shortly after feeding 20 g of *C. stellaris* (within ~12 hours), the proportion of *Cladoniaceae* DNA in the faecal samples started increasing (Fig. 4). Throughout the time *C. stellaris* was fed (24th of January-5th of February 2018), there was a large variation in the proportion of *Cladoniaceae* DNA amongst the three reindeer (Fig. 4). However, it seems that the variation in the proportion of *Cladoniaceae* DNA amongst the three animals decreased after 2000g of *C. stellaris* was fed.

Both animal 9/10 and animal 12/10 had a small peak of *Cladoniaceae* DNA after ~60 hours, which then declined rapidly. Following the feeding of 500 g of *C. stellaris*, the proportion of *Cladoniaceae* DNA for animal 9/10 increased steeply (to ~50%). There was also a steep increase in the proportion of *Cladoniaceae* DNA for animal 12/10, however, it did not surpass 30% of the proportions. The third animal, 10/10, did not respond as quickly to the introduction of 20 g of lichens, but there was a steady increase towards the peak (~70%), approximately 24 hours after 500 g of *C. stellaris* was fed for the first time. Following this peak, the proportion of *Cladoniaceae* DNA for animal 10/10 decreased steadily. In contrast, the proportion dropped for both animal 9/10 and 12/10 between the feedings of 500 g and 2000 g, and peaks (highest proportions) appeared for these reindeer (9/10: ~50%, 12/10: ~40%) after feeding both 500 g and the 2000 g of *C. stellaris*.

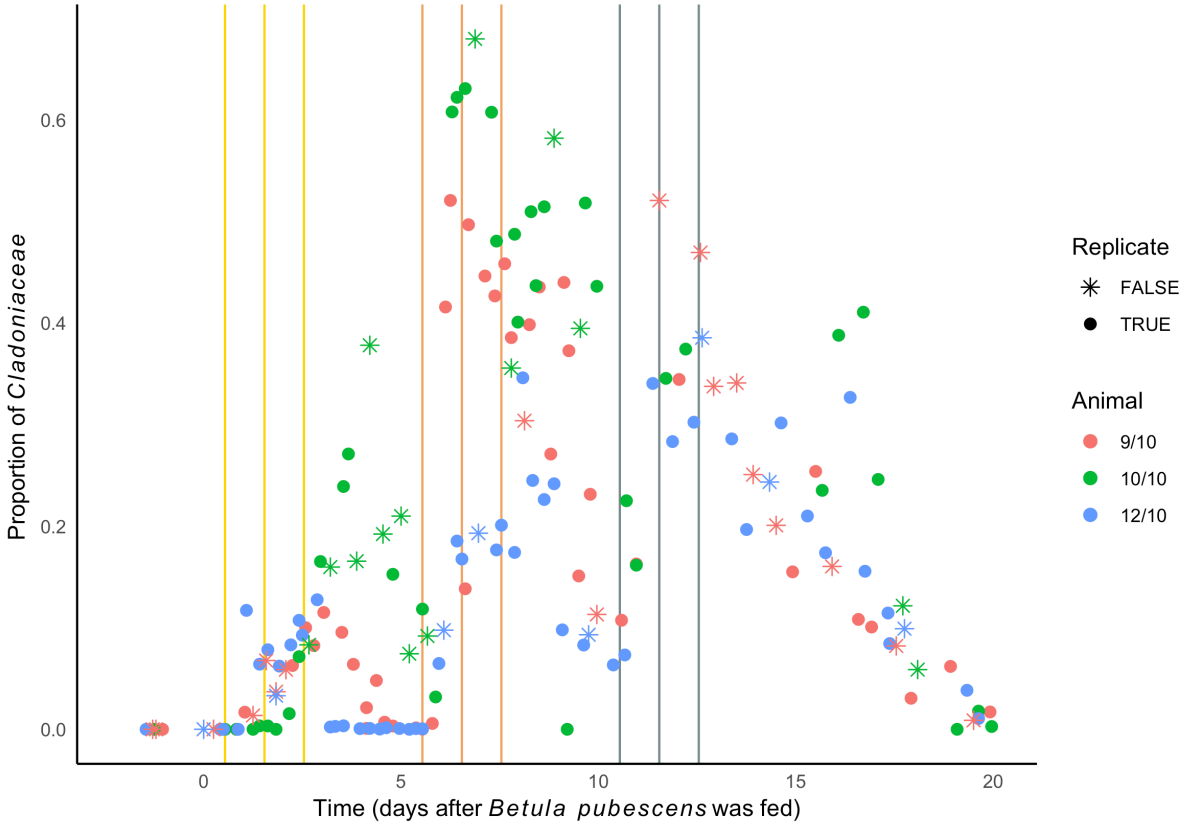

Figure 4 – Proportion of *Cladoniaceae* DNA in faecal samples from all sequences in the dataset containing amplified fungi (Fung01). The three reindeer included in the experiment are represented by colour, and the samples without replicates are shown with a star sign, whilst the samples with three replicates are shown with a filled circle. The three yellow lines to the left in the plot illustrate the 20g feeding, the three orange lines illustrate the 500g feeding and the three grey lines to the right illustrate the 2000g feeding of lichens.

For the dataset containing faecal samples amplified with the eukaryote primer (Euka02), the proportion of the sequences identified as subclades of Lecanoromycetidae (subclass containing to *C. stellaris*) showed that there was little effect of the first three feedings of 20 g of *C. stellaris*, with only a slight increase (Fig. 5). A few hours after the second feeding of 500 g of *C. stellaris*, there was an increase amongst all reindeer. The proportion of Lecanoromycetidae DNA for all reindeer dropped after the 500 g feeding, and following the 2000 g feeding, the proportion increased and reached the maximum proportion of Lecanoromycetidae DNA in the faecal samples for all three reindeer (Fig. 5). The estimated $T_{1/2}$ of *C. stellaris*, based on linear models, using the dataset containing amplified eukaryotes (Euka02), was 14 hours for animal 9/10, 16 hours for animal 10/10 and 9 hours for animal 12/10. In comparison, the $T_{1/2}$ using the dataset containing faecal samples amplified with the

fungal primer (Fung01) was 30 hours for animal 9/10, 20 hours for animal 10/10 and 22 hours for animal 12/10.

Although the three animals followed the same general pattern through the plot (Fig. 5), there was variation in the proportions amongst them. The proportion of Lecanoromycetidae was lowest for animal 12/10 from day 6 to day 15. Most of these days, animal 9/10 had the highest proportion; however, animal 10/10 had the highest proportion at day 9.

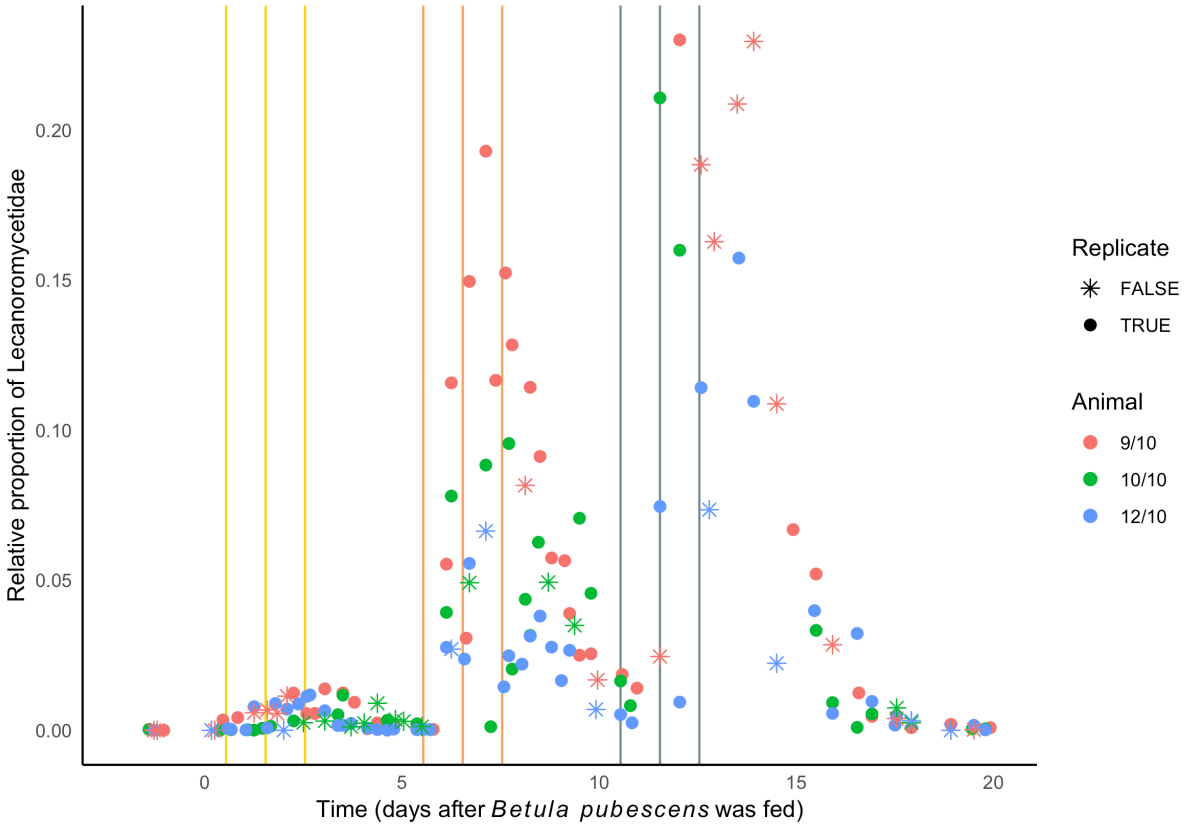

Figure 5 – Proportion of Lecanoromycetidae DNA in faecal samples from all sequences in the dataset containing amplified eukaryote sequences (Euka02). The three reindeer included in the experiment are represented by colour, and the samples without replicates are shown with a star sign, whilst the samples with three replicates are shown with a filled circle. The three yellow lines to the left in the plot illustrate the 20g feeding, the three orange lines illustrate the 500g feeding and the three grey lines to the right illustrate the 2000g feeding of lichens.

3.4 Does DNA detectability change through the gastrointestinal tract?

The principal component analyses (PCAs) generally showed no clear patterns (Fig. B1-B4 in appendix B). Based on the number of output sequences, the samples were neither grouped according to animal nor according to the parts of the GI tract from which they were collected, regardless of which primer was used. This can indicate that any potential differences in dietary composition between the collected samples are not explained by individual variation or by the different parts of the GI tract.

One exception might be that in the dataset containing samples amplified with the eukaryote primer (Euka02, Fig. B1 in appendix B), a weak structuring was seen based on the different parts of the GI tract (rumen and reticulum showing more similarity, and the omasum, abomasum and small intestine samples slightly separated from samples of the cecum, spiral loop and large intestine). However, the distances among samples were distorted by a horseshoe effect, indicating that a PCA was not optimal for the dataset. Another possible exception is that in the dataset containing samples amplified with the fungal primer (Fung01, Fig. B2 in appendix B), a weak structuring was seen based on the different animals (the samples from Animal 2 and 3 show more similarity, and the samples from animal 4 were slightly separated from the samples from the remaining animals).

Diversity

Boxplots of the species (MOTU) richness of the different samples from the GI tract are presented in Fig. 6. Here, the q-value in Hill's number is 0, meaning that all of the MOTUs weigh the same, regardless of abundance, and this is therefore a representation of number of MOTUs in each part of the GI tract.

Little variation was observed amongst the samples collected from the four stomachs (rumen, reticulum, omasum and abomasum) and the large intestines (cecum, spiral colon and distal colon) in the dataset containing samples amplified with the eukaryote primer (Euka02, Fig. 6A). The samples from the small intestine (small intestine 1 and small intestine 2) had lower median values (25 and <20 MOTUs, respectively) than the samples from the other parts of the GI tract. Additionally, there was much more variation amongst the samples collected from small intestine 2 (i.e. the lower part of the small intestine). The median values for all four

stomachs and the spiral colon were ~30 MOTUs. The highest median values were found in the cecum (35 MOTUs) and the distal colon (>35 MOTUs).

There was more variation amongst the samples in the dataset containing samples amplified with the fungal primer (Fung01, Fig. 6B) and the median values were higher than for the other taxonomic groups (Fig. 6A, C and D). The highest median value was found in the omasum (~70 MOTUs), followed by the spiral colon and the distal colon (~65 MOTUs). The median value was ~60 MOTUs in the rumen, reticulum and abomasum, whilst the lowest median value was found in the cecum (~55 MOTUs). The largest variance was found in the spiral colon, followed by the abomasum and cecum. The boxplot for the reticulum (Fig. 6B) also showed an outlier (black filled circle with yellow edge), indicating that this sample contained extremely high values compared the other samples from the reticulum.

In the dataset containing samples amplified with the seed plant primer (Sper01, Fig. 6C) there was large variation amongst the samples from the rumen, reticulum and distal colon. There were considerably less variation amongst samples from the other parts of the GI tract, with the exception of one outlier in the cecum boxplot (Fig. 6C), which had extremely low values compared to the other samples collected from the cecum. Apart from the spiral colon, it appears that the median value increased along the GI tract (i.e. more diversity the further down in the GI tract the samples were collected).

Nearly the opposite pattern is seen in the boxplots with the dataset containing samples amplified with the bryophyte primer (Bryo01, Fig. 6D). In these boxplots, the median value was highest in the rumen and decreased along the GI tract, before flattening out (except for the spiral colon). The variance was noticeably largest in the spiral colon, followed by reticulum and rumen.

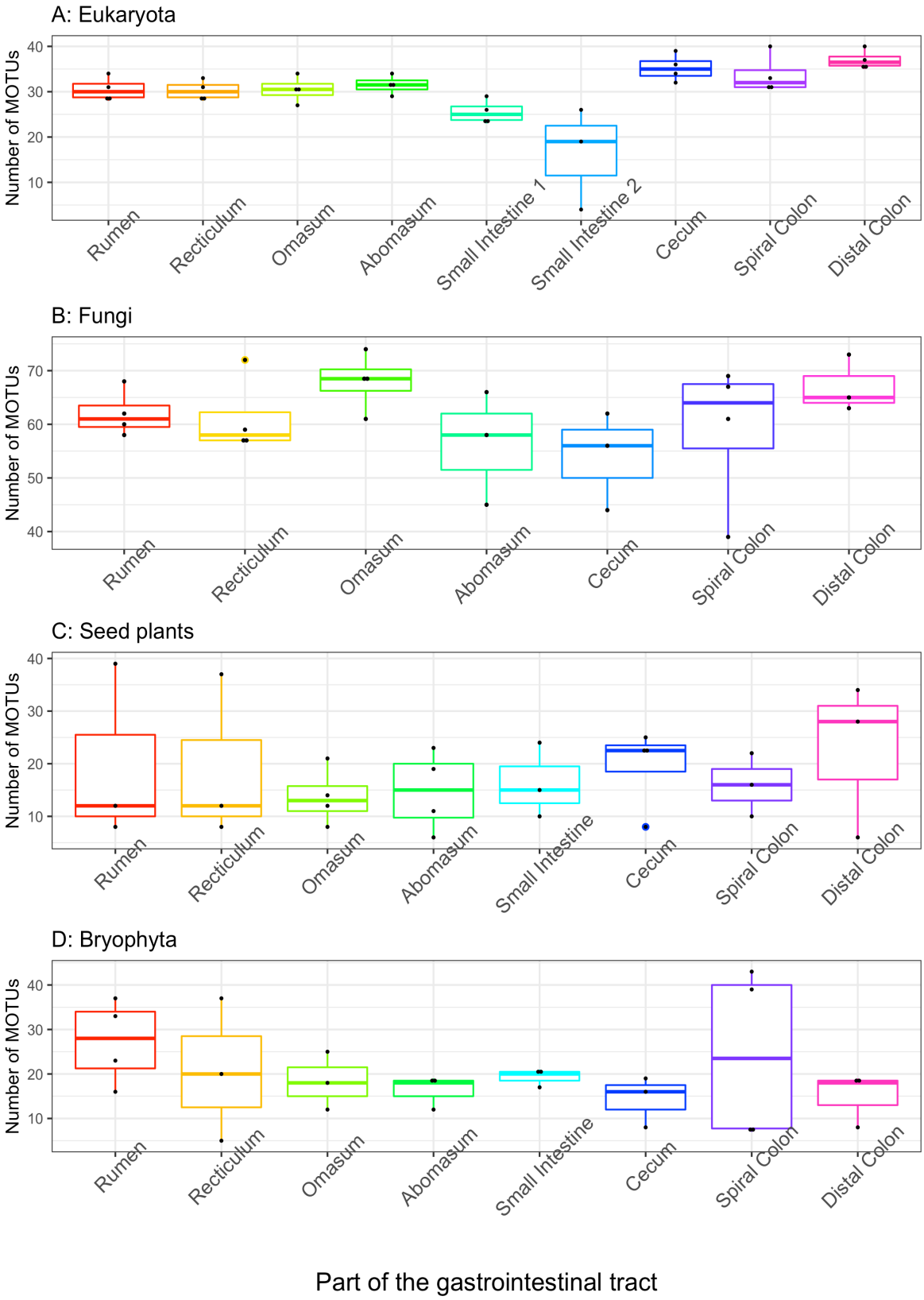

Figure 6 – Boxplots comparing the number of MOTUs (A: Eukaryota, B: Fungi, C: Seed plants (Spermatophyta), D: Bryophyta) along the gastrointestinal tract. Black, filled circles represent samples, and thick line within the quartile of the boxes shows the median. X-axis = Gastrointestinal parts. Y-axis =number of species (MOTUs).

By increasing the q-value in Hill's number to 1, the Shannon diversity index is applied. Here, the q-value gives more weight to the more abundant MOTUs (i.e. the more abundant MOTUs are considered more important) and thereby minimize the less abundant MOTUs. Thereby, Fig. 7 illustrates how diverse a part of the GI tract is, based on the most abundant and dominating MOTUs in the diet.

The samples from the large intestine (cecum, spiral colon and distal colon) were the most diverse, and the cecum had a higher diversity than any other part of the GI tract, when comparing the boxplots made from the dataset containing samples amplified with the eukaryote primer (Euka02, Fig. 7A). The lowest diversity was found in the lower part of the small intestine (Small Intestine 2), and generally, there was a lot of variation amongst the samples from the different parts of the GI tract, with the exception of the omasum.

In the dataset containing samples amplified with the fungal primer (Fung01, Fig. 7B,) there was a decrease in diversity along the GI tract (i.e. less diversity the further down in the GI tract the samples was collected). There was more variation in the rumen, reticulum and the distal colon than in the other parts of the GI tract.

In the boxplots made by using the dataset containing samples amplified with the seed plant primer (Sper01) there was more variation in the rumen, the cecum and the distal colon than in the other parts of the GI tract (Fig. 7C). The high variation between the samples collected from the rumen and in the distal colon, was largely affected by one sample, whilst the other samples in the respective parts of the GI tract were more similar. The highest median values for diversity were in the cecum and in the distal colon (~7 and ~10, respectively). In the remaining samples (rumen, reticulum, omasum and abomasum, the small intestine and the spiral colon), the median values were below 4.

Large variation was seen in the boxplots showing samples collected from the rumen and the spiral colon, when using the dataset containing samples amplified with the bryophyte primer (Bryo01, Fig. 7D). The median values decreased slightly along the tract of the stomachs and were a bit higher in the intestines than in the stomachs.

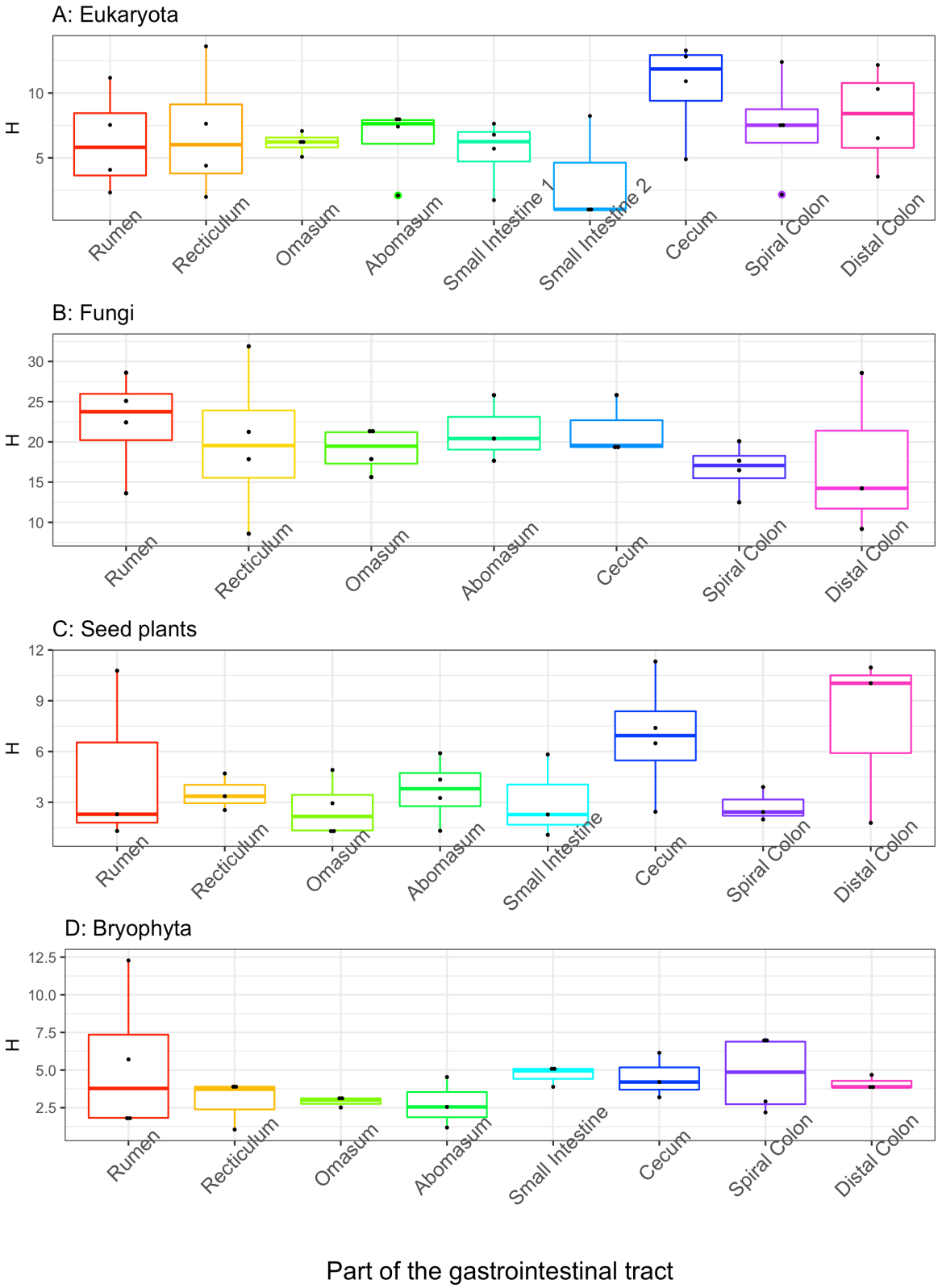

Figure 7 – Boxplots comparing the diversity (A: Eukaryota, B: Fungi, C: Seed plants (Spermatophyta), D: Bryophyta), based on the Shannon diversity index (H), along the gastrointestinal tract. Black, filled circles represent samples, and thick line within the quartile of the boxes shows the median. X-axis = Gastrointestinal parts. Y-axis = Shannon diversity (H).

4 Discussion

4.1 Does DNA metabarcoding identify all species to species level in a known diet?

The comparison between the identified output sequences from both faecal and food samples, and the species fed to the reindeer, showed that DNA metabarcoding did not identify all species to species level within a known diet. The main ingredients of RF-80 (*Hordeum vulgare* L., *Avena sativa* L. and *Triticum aestivum*) are species within the family Poaceae, which are difficult to separate by the P6 loop in the *trn*L intron (Sønstebø et al., 2010). *Triticum aestivum* was identified at species level in the dataset containing faecal samples amplified with the eukaryote primer (Euka02, Table 10). *Glycine max* Merr. and *Elaeis guineensis* Jacq. were not detected in any samples by any primer. The remaining species in RF-80, as well as *B. pubescens* and *C. stellaris*, were detected at higher taxonomical levels (e.g. genus or family level).

One possible explanation for the generally low taxonomic resolution (i.e. not to species level) is that the complete EMBL database was used as reference library. The EMBL database is the largest database and contains information on species from all over the world, and MOTUs will be identified at higher taxonomic levels because there are multiple possible matches for each sequence. Other studies have successfully identified large parts of the herbivore diet, with higher taxonomic resolution when modifying the EMBL reference database to the local area, or when combining the EMBL reference database with another reference database specific for a local area or a taxa (Quéméré et al., 2013; Soininen et al., 2013). A study by Soininen et al. (2015) address the issue of poor taxonomic resolution and emphasize that the overall impression of taxonomic resolution in the identified diet can be influenced by a dominant taxon with low taxonomic resolution (e.g. *Salicaceae*). By removing the taxa with low taxonomic resolution, it was possible to get a different impression of the taxonomic resolution.

Considerably more species from RF-80 were identified in the dataset containing faecal samples amplified with the seed plant primer (Sper01) than the dataset containing faecal samples amplified with the eukaryote primer (Euka02). A possible explanation for this is that RF-80 was boiled at 110°C during production (to prevent salmonellosis), causing DNA to

degrade. Degraded DNA makes species detection more challenging (Ghorbani et al., 2017), and using a shorter primer (e.g. Sper01) increase the chances of detection. However, the taxonomic resolution depends on the targeted gene region, and because the amplified eukaryote sequences are longer than the amplified seed plant sequences (Taberlet et al., 2018), it might explain why *Triticum aestivum* was only identified in the dataset containing faecal samples amplified with the eukaryote primer (Euka02). Although DNA extraction kits allow small quantities of samples to be used for DNA extraction, there might be too little DNA for amplification and sequencing in the RF-80 food samples (as a result of boiling at 110°C). The large difference between the amount of RF-80 eaten by the reindeer and the relatively small amount controlled (the food subsamples contained ~20 mg), might therefore affect the difference in detection between the two sample types.

Most ecological studies exploring diet are interested in detecting competition between or dietary niche overlap of target species in an ecosystem. For this, functional groups (e.g. shrubs, herbs, forbs, lichens etc.) are commonly used (e.g. Skogland, 1984; Kowalczyk et al., 2011; Soininen et al., 2013; Kartzinel et al., 2015; Soininen et al., 2015). To use functional plant groups in ecological studies usually requires detection at minimally order level (e.g. Poales). Other functional groups (e.g. mosses) can be categorized using identification to e.g. class level. Although DNA metabarcoding did not identify all of the fed species to species level in this study, it succeeded to provide taxonomic resolution at order level or higher, as most of the fed taxa were identified to family or genus level. Still, many of the MOTUs detected and identified (Table 9, Table A19-A31 in appendix A) were not part of the controlled diet. Some of these MOTUs might have been part of the reindeer diet (e.g. seed plants collected together with *C. stellaris*, contaminants from production of RF-80, or eaten by the reindeer in *E1*), others are part of the microbiota in the GI tract of the reindeer, and yet some might be contamination. For instance, *Dryadoideae* and *Saliceae* were identified with high mean relative frequency in the food samples taken of *B. pubescens* and in the fed lichen mixture (Table A26 and A29 in appendix A). One possibility is when *B. pubescens* was collected, branches of other tree species were accidentally collected simultaneously. However, *Saliceae* was not amongst the 20 MOTUs with the highest mean relative frequency in the faecal samples (Table A21 in appendix A). Another possibility is that cross-contamination happened during homogenizing and subsampling. *Saliceae* DNA was found in the open air negative controls, and although relative number of output sequences in the food samples equivalent to the relative number of output sequences from the open air negative controls

were removed, *Saliceae* DNA is not removed completely. *Dryadoideae*, on the other hand, was found in the food samples taken of *B. pubescens* and the fed lichen mixture but not in the open-air negative controls, the extraction negative controls or in any other food samples. This increases the confidence that the *Dryadoideae* DNA found in these samples is not present due to cross-contamination during subsampling or extraction. A possibility is that the p6 loop of the *trn*L gene is too short to differentiate and correctly identify the species. Thereby, the errors occurring during PCR have large negative impact on identification, i.e. a few wrongly incorporated base pairs in the sequence during PCR could result in the identification of a completely different species.

Fungal endophytes are another source of non-dietary identified MOTUs that would be unintentionally ingested by the reindeer (Soininen et al., 2013), and are identified in both faecal and food samples. Although the fungal endophytes affect the number and the proportions of MOTUs detected, they generally have little or no effect on herbivores (Faeth and Fagan, 2002), and should therefore ideally be removed prior to the statistical dietary analyses. The fungi from the microbiota in the GI tract are only detected in the faecal samples and should also be removed. However, because there is no reference database that allows for discrimination between intentionally eaten fungi (mushrooms and lichens), fungal endophytes and fungi from the microbiota in the GI tract, such elimination could not be completed for this study.

4.2 For how long is DNA of a particular food item present in the digestive system?

Few studies have identified how long it takes from food is ingested until it can be detected in the faeces, especially in large herbivores. A study by Sponheimer et al. (2003), using stable isotope analysis to detect dietary shifts of two horses and two alpacas from faecal samples. They showed that it takes 60 hours to detect dietary shifts of horses, whilst it takes almost 200 hours to detect dietary shifts of alpaca. Although the study showed that dietary shifts are more rapidly detected in horses, it still takes many hours more than it took to detect an increase of DNA of *B. pubescens* in the faecal samples of reindeer (within ~12 hours). This result suggests that DNA metabarcoding is a much more sensitive method for detecting when a dietary species appear in faecal samples, even in low quantities.

It was expected that the relative proportion of *Betulaceae* DNA would decrease after it was fed once at the start of the feeding experiment, and that remnants would be detectable over time. However, approximately 300 hours after *B. pubescens* was fed an unexpected re-increase of *Betulaceae* DNA emerged (Fig. 1 and 3). A likely explanation is that some parts of *B. pubescens* were stored in the cecum for further digestion. However, a similar re-increase of *Cladoniaceae* DNA was not observed (Fig. 4 and 5), and it is unclear what might cause the need for further digestion of some species, but not of others. Several studies have shown that different substances affect digestion. For instance, a study by Palo (1985) showed that water-soluble phenolic substances reduce rumen digestibility of *B. pendula* Roth, and another study showed how digestion of cellulose in *B. papyrifera* Marshall is inhibited by secondary metabolites in resin (Risenhoover et al., 1985). Several studies have also highlighted how tannins reduce digestibility (Hervás et al., 2000; Goel et al., 2005; Barboza et al., 2010). However, chemical analysis of the fed *B. pubescens* and *C. stellaris* has not been conducted in this study, and it is therefore not possible to determine if a substance, and if so, which substance might have affected the digestion *B. pubescens* and *C. stellaris* in this feeding experiment.

To my knowledge, no study has been done to examine how long DNA of a particular food item persists in the digestive system, and thus is detectable in the faeces, of a large herbivore. Therefore, a motivation to answer Question 2 was to reveal if *Betulaceae* DNA would be detected in the faeces over a long time period after ingestion of *B. pubescens*. By offering a small amount of *B. pubescens* once, it was possible to determine that the *Betulaceae* DNA was detected in the faeces for at least 26 days (Fig. 1 and A2). However, because the reindeer were reintroduced to *E1* and had access to *B. pubescens* again, it was not possible to conclude on the maximum duration that DNA from *B. pubescens* could be detected for. This result, combined with the result of rapid detection, can improve interpretation of faecal samples collected from free-ranging reindeer for ecological studies. In such studies, it will be important to emphasize that the faecal samples could represent the diet from at least the last 26 days when using DNA metabarcoding.

When comparing decay time ($T_{1/2}$) between sequences amplified by the different primers, it is expected that that the shorter sequences will be detected for a longer time period than the longer sequences (Valentini et al., 2009), and thereby have a higher $T_{1/2}$. This was found for *B. pubescens* DNA, when comparing $T_{1/2}$ in the output sequences from the faecal samples

amplified with the seed plant primer (Sper01 amplifies shorter sequences) and with the eukaryote primer (Euk02 amplifies longer sequences). However, the opposite was found for *C. stellaris* DNA when comparing $T_{1/2}$ in the output sequences from the faecal samples amplified with the fungal primer (Fung01) and amplified with the eukaryote primer (Euka02). The primer used to amplify fungal sequences provides higher mean and maximum sequence lengths than the primer used to amplify eukaryote sequences (Taberlet et al., 2018), and it is therefore an unexpected result to see a higher $T_{1/2}$ for the amplified fungal sequences. A possible explanation for this is that the fungal primer is more specific and was therefore able to detect the target taxon for a longer time.

By using a general primer amplifying eukaryote DNA (Euka02), it is possible to compare decay profiles of DNA from *B. pubescens* and *C. stellaris* in the dietary faecal samples (Fig. 5 and Fig. A2 in appendix A). Comparing these two plots reveal that Lecanoromycetidae DNA might decay slightly faster than *Betulaceae* DNA, as the data points (filled circles and star signs) are steeper and there was not a second peak (Fig. A2), as seen for *B. pubescens* (Fig 1 and A1 in appendix A). This is confirmed by comparing the $T_{1/2}$ of the two taxa ($T_{1/2\ Betula}$: 22 hours for animal 9/10, 19 hours for animal 10/10 and 13 hours for animal 12/10. $T_{1/2\ Lecanoromycetidae}$: 14 hours for animal 9/10, 16 hours for animal 10/10 and 9 hours for animal 12/10). Although the individual variability is large, the decay time of Lecanoromycetidae DNA decreases by several hours for each animal, compared to *Betula* DNA. A study by Storeheier et al. (2002a) emphasized that adaptation of microbiota in the GI tract has a large impact on digestibility, and that studies on digestibility should only be conducted when the microbiota is completely adapted to the food species tested. The reindeer had unrestricted access to *B. pubescens* in *E1* and it is assumed, based on the presence of *B. pubescens* DNA in the faecal samples collected prior to the feeding experiment, that the reindeer had eaten *B. pubescens* and that their microbiota was adapted for digesting *B. pubescens*. However, they were not fed the lichen mixture for several months prior to the feeding experiment. Although they had access to lichens in *E1* and *Cladoniaceae* was detected in the samples collected before time 0, their microbiota might not be fully adapted for digesting large amounts of *C. stellaris*. This then raises the question whether digestibility affects the speed at which DNA decays in the GI tract of reindeer.

4.3 How well is relative food biomass represented by output sequences?

The ability to estimate relative food abundance or biomass with DNA metabarcoding has been investigated by several studies (e.g. Deagle et al., 2013; Soininen et al., 2013; Thomas et al., 2014). Deagle et al. (2013) concluded that e.g. lowering the sequence quality and including short reads during filtering can be considered to improve detection of species proportions. However, Pompanon et al. (2012) described that DNA sequences are amplified exponentially during PCR, and that differences amongst prey species and digestion, DNA extraction, pooling and sequencing bias are listed as possible reasons for questioning results in terms of biomass or abundance.

In this study, there seems to be a threshold for the amount of *C. stellaris* that can be detected when using a specific primer amplifying fungal DNA (Fung01) in the faecal samples, as a higher proportion of *Cladoniaceae* DNA was detected during the three days when 500 g *C. stellaris* was fed than the days when 2000 g was fed (Fig. 4). By using the general primer amplifying eukaryote DNA (Euka02) in the faecal samples, a clear increase in the proportion of Lecanoromycetidae DNA can be seen between each feeding of *C. stellaris* (Fig. 5). Although the proportion of identified Lecanoromycetidae DNA is a representation of the proportion of DNA present in a sample, rather than a direct representation of the ingested biomass, it is interesting that proportion of Lecanoromycetidae DNA increase within ~24 hours after feeding an increased amount of *C. stellaris* (Fig. 5). The proportion of Lecanoromycetidae DNA increases much more between 20 g and 500 g, than between 500 g and 2000 g, which does not correspond to the increased amounts of *C. stellaris* (480 g (between 20 g and 500 g) and 1500 g (between 500 g and 2000 g)). However, by rather focusing on the fold change (i.e. there is a 25-fold increase between 20 g and 500 g, and fourfold increase between 500 g and 2000 g) it can possibly explain why the proportion of DNA increases much more between 20 g and 500 g, than between 500 g and 2000 g. Based on these results, it is fair to suggest that a general primer (e.g. Euka02), is better to use in studies where relative biomass in a diet is the focus. A possible reason for this is that Euka02 primer targeted and amplified a larger number of MOTUs (including seed plants, fungi etc.). Thereby the relative proportion of Lecanoromycetidae DNA was lower than the relative proportion of *Cladoniaceae* DNA amplified with the specific fungal primer (Fung01). A lower relative proportion would thereby eliminate the threshold (i.e. the maximum amount of

Lecanoromycetidae DNA that can be detected) in the dataset containing faecal samples amplified with the eukaryote primer (Euka02).

RF-80 was the main source of food throughout the feeding experiment and the intension was to compare the output sequences from RF-80 and *C. stellaris*, similarly to the study by Deagle et al. (2013) comparing biomass from different prey species in captive harbour seals. However, it was discovered after the feeding experiment was completed that RF-80 have been treated under very high temperatures, which clearly affects the DNA detectability (as described previously). Therefore, in future feeding experiments it is recommended using unprocessed food items and species that are easily identified by DNA metabarcoding (compared to complex mixtures of species such as RF-80). It should also be considered that some types of tissues and some taxonomic groups amplify better during PCR than others (Bradley et al., 2007; Valentini et al., 2009), and that direct biomass comparisons between e.g. seed plant taxa and lichen taxa might be biased.

4.4 Does DNA detectability change through the gastrointestinal tract?

As diet diversity is commonly the focus of ecological studies (e.g. Soininen et al., 2009; Valentini et al., 2009; Rayé et al., 2011; De Barba et al., 2014), it was of interest to investigate which part of the GI tract represents the diet diversity of reindeer best, and thus, recommend which part of the GI tract to use in ecological studies. There are several ways to measure diversity, and Hill´s number was used in this study to create a diversity spectrum based on the q-value, ranging from 0 to 3. The q-value has the power to minimise the importance of rare taxa, and was initially set at 0, thus allowing comparison of species (MOTU) richness along the GI tract of a reindeer (Spellerberg and Fedor, 2003). By increasing the q-value to 1 (i.e. Shannon diversity index), the heterogeneity in the samples is taken into account, meaning that the more abundant species are considered as more important when describing the diversity (Peet, 1974). When further increasing the q-value to 2 (i.e. Gini coefficient index, Fig. B5 in appendix B) and 3 (Fig. B6 in appendix B), the most abundant species are given even more importance and the rare species contribute even less in describing the diversity (Peet, 1974). The q-value can be increased infinitely ($q=\infty$) until only the taxa with maximal abundance in a sample are counted.

The most diverse part of the GI tract is different, according to the diversity measurement (i.e. q-value) and primer set (i.e. Euka02, Fungi01, Sper01 or Bryo01) used (Fig. 6 and 8, and Fig. B5 and B6 in appendix B). For instance, the highest median value in the boxplots showing the species (MOTU) richness for amplified bryophytes (Bryo01) are found in the rumen (Fig. 6D), whilst the highest median value in the boxplots showing Shannon diversity are found in the small intestine and spiral colon (Fig. 7D). However, there is a lot of individual variation in the rumen, and the highest Shannon diversity is found in one of the samples collected from this part of the GI tract. Both species (MOTU) richness and Shannon diversity of amplified seed plants (Sper01, Fig. 6C and 7C) possibly indicate that seed plant DNA accumulates in the cecum and distal colon. Thereby suggesting that it is beneficial to use faecal samples (assuming that the samples from the distal colon largely represent the same findings as if using faecal samples) when studying seed plant diet of ruminants with DNA metabarcoding. It is, however, possible that the reindeer consumed fewer seed plant species prior to being culled, and thus less seed plant diversity was observed in the stomachs.

The dataset containing samples amplified with the fungal primer (Fung01) had a much higher species (MOTU) richness and diversity (Fig. 6B and 7B) compared to the other datasets (containing amplified eukaryote, seed plant and bryophyte output sequences). This is most likely due to a co-amplification of microbiota in the GI tract and endophytes, in addition to the fungi eaten by the reindeer (lichens and mushrooms, see 4.1). Although microbiota in the GI tract is detected, further investigation of this dataset should be conducted to examine how accurately the samples collected from the distal colon (comparable to faecal samples) represent the microbiota in the GI tract of reindeer. Thereby, it could potentially motivate the use of faecal samples in future microbiota studies, instead of the commonly used rumen or cecum samples (Syrjälä et al., 1973; Mathiesen et al., 1987; Sundset et al., 2009; Salgado-Flores et al., 2016). This would be a huge benefit, as faecal samples are non-invasive and has the potential for repeated sample collections from the same animals without the need for e.g. rumen fistulation.

5 Conclusion

The overall aim of this thesis was to evaluate different aspects of using DNA metabarcoding to analyse reindeer diet. By conducting a controlled feeding experiment with three reindeer, it has been possible to show that the specific primer amplifying seed plant DNA detects a higher number of fed species than the general primer amplifying eukaryote DNA. Additionally, the specific primer amplifying fungal DNA gave a higher taxonomic resolution than the general primer amplifying eukaryote DNA. Although only one of the species in the controlled diet was detected at species level, most species were detected at a higher taxonomical level, within the taxonomic resolution often used in ecological studies.

The controlled feeding experiment also revealed an increase of *B. pubescens* DNA within 12 hours after feeding, demonstrating the ability for rapid detection when using DNA metabarcoding. Although it was not possible to assess when *B. pubescens* DNA was no longer detectable in the faeces, this study shows that *B. pubescens* DNA is present for at least 26 days after feeding. Additionally, results of $T_{1/2}$ suggest that *C. stellaris* DNA decay quicker than *B. pubescens* DNA. Generally, it is advised to be critical when interpreting biomass based on DNA metabarcoding results. Still, this study shows a correlation between fed biomass and proportion of output sequences when using a general primer amplifying eukaryote sequences. Nevertheless, it is of importance to clarify that the feeding experiment was conducted on captive reindeer, and that free-ranging reindeer have a diverse diet that is not reflected in the three food items fed during this experiment (RF-80, *B. pubescens* and *C. stellaris*).

The experiment using samples from the GI tract of four free-ranging reindeer indicates that there is a higher species (MOTU) richness and diversity in the distal colon, when using the datasets with samples amplified with the seed plant and eukaryote primers. The dataset with samples amplified with the bryophyte primer indicate the opposite; a higher species (MOTU) richness in the rumen, and possibly also a higher Shannon diversity in the rumen. The dataset with amplified fungi also imply a higher Shannon diversity in the rumen, but a slightly higher species (MOTU) richness in the distal colon. However, using DNA metabarcoding to identify the fungal part of the reindeer diet is complicated, as microbiota in the GI tract and endophytes are detected throughout the GI tract. There is a need for a reference database that can separate the different fungi into categories (e.g. potentially eaten mushroom or lichen,

unintentionally eaten endophytes and microbiota in the GI tract). In conclusion, the results presented from this experiment suggest that samples from the distal colon contain a large diversity of the reindeer diet. Thereby suggesting that faecal samples (provided that faecal samples largely represent the same as samples from the distal colon) should be used in ecological studies examining ruminant diet with DNA metabarcoding.

Overall, the results presented in this thesis show that DNA metabarcoding detects reindeer diet rapidly and allow for interpretation regarding relative abundance of diet. DNA metabarcoding is also able to detect a high diversity of ingested species, even species fed in low relative abundance. However, the feeding experiment only included three animals and samples from the GI tract were only collected from four animals. To increase the confidence in the findings, it would be necessary to include samples from more individuals.

6 Future studies

DNA metabarcoding is frequently used to address different ecological questions, but few experimental studies have been conducted to examine details of DNA metabarcoding, such as those examined in this study. However, the controlled feeding experiment included captive reindeer that were fed only three food items (RF-80, *B. pubescens* and *C. stellaris*), which do not reflect the diversity in the diet of free-ranging reindeer. It would therefore be interesting to collect and analyse faecal samples from free-ranging reindeer, within the first 24 hours after long-distance migration to a pasture with new species. This could potentially inform whether the diversity of the diet has any implication on how quickly DNA metabarcoding is able to detect a new species in the diet, or not. Further studies should be done to investigate the difference in longevity of DNA from several species or functional groups (graminoids, shrubs, herbs, mosses, lichen, kelp) in the GI tract of reindeer, as this can potentially reveal if one species or functional group is detectable for a longer time than others. Additionally, it might be of interest to examine the same questions, but focusing on other e.g. ruminant species in different geographic or climatic contexts, or non-ruminant herbivores e.g. ptarmigans that interact with reindeer.

In order to recommend DNA metabarcoding for ecological studies, it would be of beneficial to analyse the faecal samples collected during the feeding experiment with e.g. stable isotope analysis. By e.g. comparing the results regarding detection time of *B. pubescens* from DNA metabarcoding, with results from stable isotope analysis, it will be possible to make an assessment regarding which method is more precise (i.e. whether one method detects the introduced species faster or if they detect the introduced species at the same time).

Although the results from the GI tract provide interesting information, it would be advantageous to include more animals to increase the statistical confidence. Additional studies should therefore be conducted in a similar manor with more animals, to increase statistical confidence, and to possibly confirm the findings in this study. Further studies should also be conducted to examine if and how the microbiota in the GI tract changes and adapts to dietary changes.

References

Ambrose, S. H. and DeNiro, M. J. (1986) "The isotopic ecology of East African mammals" Oecologia, 69: 395-406

Arctic Climate Impact Assessment (2004). *Impacts of a warming Arctic: Arctic climate impact assessment*. Cambridge University Press, Cambridge

Baker, W., van den Broek, A., Camon, E., Hingamp, P., Sterk, P, Stoesser, G. and Tuli M. A., (2000) "The EMBL Nucleotide Sequence Database". Nucleic Acids Research, 28: 19-23

Barboza, P. S., Bennett, A., Lignot, J.-H., Mackie, R. I., McWhorter, T. J., Secor, S. M., Skovgaard, N., Sundset, M. A. and Wang T. (2010) "Digestive challenges for vertebrate animals: microbial diversity, cardiorespiratory coupling, and dietary specialization". Physiological and Biochemical Zoology 83:764-774

Bjørkvoll, E., Pedersen, B., Hytteborn, G., Jónsdóttir I. S. and Langvatn R. (2009) "Seasonal and interannual dietary variation during winter in female Svalbard reindeer (*Rangifer Tarandus Platyrhynchus*)". Arctic, Antarctic, and Alpine Research, 41:88-96

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., Yu, D. W. and de Bruyn, M. (2014) "Environmental DNA for wildlife biology and biodiversity monitoring". Trends in Ecology & Evolution, 29: 358-367

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P. and Coissac É. (2016), "Obitools: a unix-inspired software package for DNA metabarcoding". Molecular Ecology Resources, 16: 176-182

Bradley, B. J., Stiller, M., Doran-Sheehy, D. M., Harris, T., Chapman, C. A., Vigilant, L. and Poinar, H. (2007) "Plant DNA sequences from feces: potential means for assessing diets of wild primates". American Journal of Primatology, 69: 699-705

Broms, K. M., Hooten, M. B. and Fitzpatrick, R. M. (2015) "Accounting for imperfect detection in Hill numbers for biodiversity studies". Methods in Ecology and Evolution, 6: 99-108

CAFF (2006) World Reindeer Husbandry: CBMP EALÁT-Monitoring. Expert Network Monitoring Plan. Supporting publication to the CAFF Circumpolar Biodiversity Monitoring Program – Framework Document. By WRH and ICRH in cooperation with CAFF. CAFF CBMP Report No. 10, CAFF International Secretariat, Akureyri, Iceland.

Codron, D., Lee-Thorp, J. A., Sponheimer, M., Codron, J., De Ruiter D. and Brink, J. S. (2007) “Significance of diet type and diet quality for ecological diversity of African ungulates”. Journal of Animal Ecology, 76: 526-537

Cook, W. E., Raisbeck, M. F., Cornish, T. E., Williams, E. S., Brown, B., Hiatt, G. and Kreeger, T. J. (2007) “Paresis and death in elk (*Cervus elaphus*) due to lichen intoxication in Wyoming”. Journal of Wildlife Diseases, 43: 498-503.

Dailey, R. N., Montgomery, D. L., Ingram, J. T., Siemion, R., Vasquez, M. and Raisbeck, M. F. (2008) “Toxicity of the lichen decondary metabolite (+)-usnic acid in domestic sheep”. Veterinary Pathology, 45:19-25

Deagle, B. E., Thomas, A. C., Shaffer, A. K., Trites, A. W. and Jarman, S. N. (2013) “Quantifying sequence proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: which counts count?”. Molecular Ecology Resources, 13: 620-633

De Barba, M. , Miquel, C. , Boyer, F. , Mercier, C. , Rioux, D. , Coissac, É. and Taberlet, P. (2014) “DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet”. Molecular Ecology Resources, 14: 306-323

DeNiro, M. J., and Epstein, S. (1981) “Influence of diet on the distribution of nitrogen isotopes in animals”. Geochimica et Cosmochimica Acta, Volume 45: 341-351

Denman, S. E. and McSweeney, C. S.(2006) “Development of a real-time PCR assay for monitoring anaerobic fungal and cellulolytic bacterial populations within the rumen”. FEMS Microbiology Ecology, 58: 572-582,

Dray, S. and Dufour, A.-B. (2007) “The ade4 Package: Implementing the duality diagram for ecologists”. Journal of Statistical Software, 22: 1-20

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V. I., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kauserud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, É. and Brochmann, C. (2012) “New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems”. Molecular Ecology, 21: 1821-1833.

Faeth, S. H. and Fagan, W. F. (2002) “Fungal endophytes: common host plant symbionts but uncommon mutualists, integrative and comparative biology”. Integrative and Comparative Biology, 42: 360-368

Gaare E. and Skogland T. (1975) “Wild Reindeer Food Habits and Range Use at Hardangervidda”. In: Wielgolaski F.E. (eds) *Fennoscandian Tundra Ecosystems*. *Ecological Studies* (*Analysis and Synthesis*), vol 17. Springer, Berlin, Heidelberg

Gebremedhin B, Flagstad Ø, Bekele A, Chala D, Bakkestuen V, Boessenkool S, Popp, M., Gusarova, G., Schrøder-Nielsen, A., Nemomissa, S., Brochmann, C., Stenseth, N. C. and Epp, L. S. (2016) "DNA metabarcoding reveals diet overlap between the endangered walia ibex and domestic goats - implications for conservation". PLoS ONE 11: e0159133

Ghorbani, A., Gravendeel, B., Selliah, S., Zarré, S. and Boer, H. (2017) "DNA barcoding of tuberous Orchidoideae: a resource for identification of orchids used in Salep". Molecular Ecology Resources, 17: 342-352.

Goel, G., Puniya, A.K., Aguilar, C.N. and Singh, K. (2005) "Interaction of gut microflora with tannins in feed". Naturwissenschaften, 92: 497-503

Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, É., Wangensteen, O. S. and Turon, X. (2015) "Deep-sea, deep-sequencing: metabarcoding extracellular DNA from sediments of marine canyons". PLoS ONE 10: e0139633

Hansen, B. B. and Aanes, R. (2012) "Kelp and seaweed feeding by High-Arctic wild reindeer under extreme winter conditions". Polar Research, 31: 17258

Hansen, B. B., Lorentzen, J. R., Welker, J. M., Varpe, Ø., Aanes, R., Beumer, L. T., and Pedersen, Å. Ø. (2019) "Reindeer turning maritime: Ice-locked tundra triggers changes in dietary niche utilization". Ecosphere, 10: e02672

Hansen K. K., Sundset, M. A., Folkow, L. P., Nilsen, M. and Mathiesen, S. D. (2018) "Methane emissions are lower from reindeer fed lichens compared to a concentrate feed". Polar Research, 37: 1505396

Heggberget, T. M., Gaare, E. and Ball, J. P. (2002) "Reindeer (*Rangifer tarandus*) and climate change: Importance of winter forage". Rangifer, 22: 13-31.

Heid, C. A., Stevens, J., Livak K. J. and Williams P. M. (1996) "Real time quantitative PCR". Genome Research, 6: 986-994

Hervás, G., Frutos, P., Serrano, E., Mantecón A. and Giráldez, F. (2000) "Effect of tannic acid on rumen degradation and intestinal digestion of treated soya bean meals in sheep". The Journal of Agricultural Science, 135, 305-310

Hill, M. O. (1973) "Diversity and evenness: a unifying notation and its consequences". Ecology, 54: 427-432

Inga, B. (2009). "Reindeer (*Rangifer tarandus tarandus*) feeding on lichens and mushrooms: traditional ecological knowledge among reindeer-herding Sami in northern Sweden". Rangifer, 27: 93-106

Kartzinel, T. R., Chen, P. A., Coverdale, T. C., Erickson, D. L., Kress, W. J., Kuzmina, M. L., Rubenstein, D. I., Wang, W. and Pringle, R. M. (2015) "DNA metabarcoding illuminates dietary niche partitioning by African large herbivores". Proceedings of the National Academy of Sciences, 112: 8019-8024

Keylock, C. J. (2005) "Simpson diversity and the Shannon-Wiener index as special cases of a generalized entropy". Oikos, 109: 203-207

Klein, D. and Shulski, M. (2009). "Lichen recovery following heavy grazing by reindeer delayed by climate warming". Ambio, 38: 11-16

Kojola, I., Helle, T., Niskanen M. and Aikio P. (1995) "Effects of lichen biomass on winter diet, body mass and reproduction of semi-domesticated reindeer *Rangifer t. tarandus* in Finland". Wildlife Biology, 1: 33-38

Kowalczyk, R., Taberlet, P., Coissac, É., Valentini, A., Miquel, M., Kamiński T. and Wójcik J. M. (2011) "Influence of management practices on large herbivore diet - Case of European bison in Białowieża Primeval Forest (Poland)". Forest Ecology and Management, 261: 821-828

LECA - Laboratoire d'ecologie alpine (2012a). ROBITools: metabarcoding data biodiversity analysis. R package version 0.1.

LECA - Laboratoire d'ecologie alpine (2012b). ROBITaxonomy: metabarcoding data biodiversity analysis. R package version 0.1.

Maddocks, S. and Jenkins, R. (2017) "Quantitative PCR: things to consider". Maddocks and Jenkins (Ed.), *Understanding PCR* (pp 45-52), Academic Press

Mathiesen, S., Haga, Ø, Kaino, T. and Tyler, N. (2000) "Diet composition, rumen papillation and maintenance of carcass mass in female Norwegian reindeer (*Rangifer tarandus tarandus*) in winter". Journal of Zoology, 251: 129-138

Mathiesen, S. D., Orpin, C. G., Greenwood, Y. and Blix, A. S. (1987) "Seasonal changes in the cecal microflora of the high-arctic Svalbard reindeer (*Rangifer tarandus platyrhynchus*)". Applied and Environmental Microbiology, 53: 114-118.

Mathiesen, S. D., Utsi, T. H. Aa. and Sørmo, W. (1999) "Forage chemistry and the digestive system in reindeer (*Rangifer tarandus tarandus*) in northern Norway and on South Georgia". Rangifer, 19: 91-101

Moen, J. (2008) "Climate change: effects on the ecological basis for reindeer husbandry in Sweden". AMBIO: A Journal of the Human Environment, 37: 304-311

Mysterud, A., Easterday, W. R., Stigum, V. M., Aas, A. B., Meisingset, E. L. and Viljugrein, H. (2016) “Contrasting emergence of Lyme disease across ecosystems”. Nature Communications, 7: 11882

Nichols R. V., Åkesson M. and Kjellander P. (2016) “Diet assessment based on rumen contents: a comparison between DNA metabarcoding and macroscopy”. PLoS ONE 11: e0157977.

Nieminen, M. and Heiskari, U. (1989) “Diets of freely grazing and captive reindeer during summer and winter”. Rangifer, 9: 17-34

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E. and Wagner H. (2019). “Vegan: community ecology package”. R package version 2.5-4. CRAN.R-project.org/package=vegan

Palo, T. (1985) “Chemical defense in birch: inhibition of digestibility in ruminants by phenolic extracts”. Oecologia, 68: 10-14

Peet R. K. (1974) “The measurement of species diversity”. Annual Review of Ecology and Systematics, 5: 285-307

Pompanon, F., Deagle, B. E., Symondson, W. O., Brown, D. S., Jarman, S. N. and Taberlet, P. (2012) “Who is eating what: diet assessment using next generation sequencing”. Molecular Ecology, 21: 1931-1950

Quéméré E., Hibert F., Miquel C., Lhuillier E., Rasolondraibe E., Champeau J., Rabarivola, C., Nusbaumer, L., Chatelain, C., Gautier, L., Ranirison, P., Crouau-Roy, B., Taberlet, P. and Chikhi, L. (2013) “A DNA metabarcoding study of a primate dietary diversity and plasticity across its entire fragmented range”. PLoS ONE, 8: e58971

R Studio Team (2015) “RStudio: integrated development for R”. RStudio, Inc., Boston, MA

Rayé, G., Miquel, C., Coissac, É., Redjadj, C., Loison, A. and Taberlet, P. (2011) “New insights on diet variability revealed by DNA barcoding and high-throughput pyrosequencing: chamois diet in autumn as a case study”. Ecological Research, 26: 265-276

Risenhoover. K. L., Renecker, L. A. and Morgantini, L. E. (1985) “Effects of secondary metabolites from balsam poplar and paper birch on cellulose digestion”. Journal of Range Management, 38: 370-372

Salgado-Flores A., Hagen L. H., Ishaq S. L., Zamanzadeh M., Wright A.-D. G., Pope P. B. and Sundset, M. A. (2016) “Rumen and cecum microbiomes in reindeer (*Rangifer*

tarandus tarandus) are changed in response to a lichen diet and may affect enteric methane emissions. PLoS ONE 11: e0155213

Sayers E. W., Barrett T., Benson D. A., Bryant S. H., Canese K., Chetvernin V., Church D. M., DiCuccio M., Edgar R., Federhen S., Feolo M., Geer L. Y., Helmberg W., Kapustin Y., Landsman D., Lipman D. J., Madden T. L., Maglott D. R., Miller V., Mizrachi I., Ostell J., Pruitt K. D., Schuler G. D., Sequeira E., Sherry S. T., Shumway M., Sirotkin K., Souvorov A., Starchenko G., Tatusova T. A., Wagner L., Yaschenko E. and Ye J. (2009) "Database resources of the National Center for Biotechnology Information". Nucleic Acids Research, 37: D5-D15

Skogland, T. (1984) "Wild reindeer foraging-niche organization". Ecography, 7: 345-379

Soininen, E. M., Ehrich, D., Lecomte, N., Yoccoz, N. G., Tarroux, A., Berteaux, D., Gauthier, G., Gielly, L., Brochmann, C., Gusarova, G. and Ims, R. A. (2014) "Sources of variation in small rodent trophic niche: new insights from DNA metabarcoding and stable isotope analysis". Isotopes in Environmental and Health Studies, 50: 361-381

Soininen E. M., Gauthier G., Bilodeau F., Berteaux D., Gielly L., Taberlet P., Gusarova, G., Bellemain, E., Hassel, K., Stenøien, H. K, Epp, L. S., Schrøder-Nielsen, A., Brochmann, C. and Yoccoz, N. G. (2015) "Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding". PLoS ONE 10(1): e0115335.

Soininen E. M., Valentini, A., Coissac, É., Miquel, C., Gielly, L., Brochmann, C., Brysting, A. K. and Sønstebø, J. H. and Ims, R. A. and Yoccoz, N. G. and Taberlet, P. (2009) "Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures". Frontiers in Zoology, 6:16

Soininen, E.M., Zinger, L., Gielly, L., Bellemain, E., Bråthen, K. A., Brochmann, C., Epp, L. S., Gusarova, G., Hassel, K., Henden, J.-A., Killengre (2013) "Shedding new light on the diet of Norwegian lemmings: DNA metabarcoding of stomach content". Polar Biol, 36: 1069-1076

Soreng, R. J., Peterson, P. M., Romschenko, K., Davidse, G., Zuloaga, F. O., Judziewicz, E. J., Filgueiras, T. S., Davis, J.I. and Morrone, O. (2015) "A worldwide phylogenetic classification of the Poaceae (Gramineae)". Journal of Systematics and Evolution. 53: 117-137.

Spellerberg, I. F. and Fedor, P. J. (2003) "A tribute to Claude Shannon (1916-2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon-Wiener' Index". Global Ecology and Biogeography, 12: 177-179

Sponheimer, M., Robinson, T., Ayliffe, L., Passey, B., Roeder, B., Shipley, L., Lopez, E., Cerling, T., Dearing, D. and Ehleringer, J. (2003) "An experimental study of carbon-isotope fractionation between diet, hair, and feces of mammalian herbivores". Canadian Journal of Zoology, 81: 871-876

Staaland, H. and Sæbø, S. (1987) "Seasonal variations in mineral status of reindeer calves from Elgaa reindeer herding district, Norway". Rangifer, 7: 22-28

Stephenson F. H. (2016) "Real-Time PCR", Stephenson (Ed.), *Calculations for Molecular Biology and Biotechnology* (pp 215-320), Third Edition, Academic Press

Storeheier, P., Mathiesen, S., Tyler, N. and Olsen, M. (2002a) "Nutritive value of terricolous lichens for reindeer in winter". The Lichenologist, 34: 247-257

Storeheier, P., Mathiesen, S., Tyler, N., Schjelderup, I. and Olsen, M. (2002b) "Utilization of nitrogen- and mineral-rich vascular forage plants by reindeer in winter". The Journal of Agricultural Science, 139:151-160

Storeheier, P., Van Oort, B., Sundset, M., & Mathiesen, S. (2003) "Food intake of reindeer in winter". The Journal of Agricultural Science, 141: 93-101

Sundset, M.A., Barboza, P.S., Green, T.K. Folkow, L. P., Blix, A. S. and Mathiesen, S. D. (2010) "Microbial degradation of usnic acid in the reindeer rumen". Die Naturwissenschaften 97: 273-278.

Sundset, M.A., Edwards, J. E., Cheng, Y. F., Senosiain, R. S., Fraile, M. N., Northwood, K. S., Præsteng, K. E., Glad, T., Mathiesen, S. D. and Wright, A.-D. G. (2009) "Rumen microbial diversity in Svalbard reindeer, with particular emphasis on methanogenic archaea". FEMS Microbiology Ecology, 70: 553–562

Sundset, M. A., Kohn, A., Mathiesen, S. D. and Præsteng, K. E. (2008) "Eubacterium rangiferina, a novel usnic acid-resistant bacterium from the reindeer rumen". Naturwissenschaften, 95: 741-749

Sundset, M. A., Præsteng, K. E., Cann, I. K. O., Mathiesen, S. D., Mackie, R. I. (2007) "Novel Rumen Bacterial Diversity in Two Geographically Separated Sub-Species of Reindeer". Microbial Ecology, 54: 424-438

Sylvester, J. T., Karnati, S. K. R., Yu, Z., Morrison, M. and Firkins, J. L. (2004) "Development of an Assay to Quantify Rumen Ciliate Protozoal Biomass in Cows Using Real-Time PCR". The Journal of Nutrition, 134: 3378–3384

Syrjälä, L., Kossila, V. and Sipilä, H. (1973) "A study of nutritional status of Finnish reindeer (*Rangifer Tarandus L.*) in differents months: I. Composition and volume of the rumen microbiota". Agricultural and Food Science, 45: 534-541

Sønstebø, J. H., Gielly, L., Brysting, A. K., Elven, R., Edwards, M., Heile, J., Willerslev, E., Coissac, É., Rioux, D., Sannier, J., Taberlet, P. and Brochmann , C. (2010) "Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate". Molecular Ecology Resources, 10: 1009-1018

Taberlet, P., Bonin, A., Zinger, L., and Coissac, É. (2018) Environmental DNA: For Biodiversity Research and Monitoring. Oxford University Press

Taberlet, P., Coissac, É., Pompanon, F., Brochmann, C. and Willerslev, E. (2012) "Towards next-generation biodiversity assessment using DNA metabarcoding". Molecular Ecology, 21: 2045-2050.

Taberlet, P., Coissac, É., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C. and Willerslev, E. (2007) "Power and limitations of the chloroplast trn L (UAA) intron for plant DNA barcoding". Nucleic Acids Research, 35: e14

Thomas, A. C., Jarman, S. N., Haman, K. H., Trites, A. W. and Deagle, B. E. (2014) "Improving accuracy of DNA diet estimates using food tissue control materials and an evaluation of proxies for digestion bias". Molecular Ecology, 23: 3706-3718

Turi, J.M (2002) "The World Reindeer Livelihood – Current Situation, Threats and Possibilities". In Sakari Kankaanpää, Ludger Müller-Wille, Paulo Susiluoto, and Marja-Liisa Sutinen, (eds.), *Northern Timberline Forests: Environmental and Socio-economic Issues and Concerns* (pp. 70-75). Kolari, Finland: The Finnish Forest Research Institute

Tyler, N. J. C., Turi, J. M., Sundset, M. A., Bull, K. S., Sara, M. N., Reinert, E., Oskal, N., Nellemann, C., McCarthy, J. J., Mathiesen, S. D., Martello, M. L., Magga, O. H., Hovelsrud, G. K., Hanssen-Bauer, I., Eira, N. I., Eira, I. M. G. and Corell, R. W. (2007) "Saami reindeer pastoralism under climate change: applying a generalized framework for vulnerability studies to a sub-arctic social–ecological system". Global Environmental Change, 17: 191-206

Valentini, A., Miquel, C., Nawaz, M. A., Bellemain, E., Coissac, É., Pompanon, F., Gielly, L., Craud, C., Nascetti, G., Wincker, P., Swenson, J. E. and Taberlet, P. (2009) "New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach". Molecular Ecology Resources, 9: 51-60

Vavra, M., Rice, R. W. and Hansen, R. M. (1978) "A comparison of esophageal fistula and fecal material to determine steer diets". Journal of Range Management, 31: 11-13

Węgrzyn, M. H., Wietrzyk-Pełka, P., Galanty, A., Cykowska-Marzencka, B. and Sundset, M. A. (2019) "Incomplete degradation of lichen usnic acid and atranorin in Svalbard reindeer (*Rangifer tarandus platyrhynchus*)". Polar Research, 38

Wickham, H. (2016) ggplot2: elegant graphics for data analysis. Second edition, Springer

Wickham, H., François, R., Henry, L. and Müller, K. (2018) "dplyr: A Grammar of Data Manipulation". R package version 0.8.1

Zhao, L. Z., Colman, A. S., Irvine, R .J. Karlsen, S. R., Olack, G. and Hobbie, E. A. (2019) "Isotope ecology detects fine-scale variation in Svalbard reindeer diet: implications for monitoring herbivory in the changing Arctic". Polar Biology, 42: 793–805

Appendix A – Feeding experiment

Enclosures

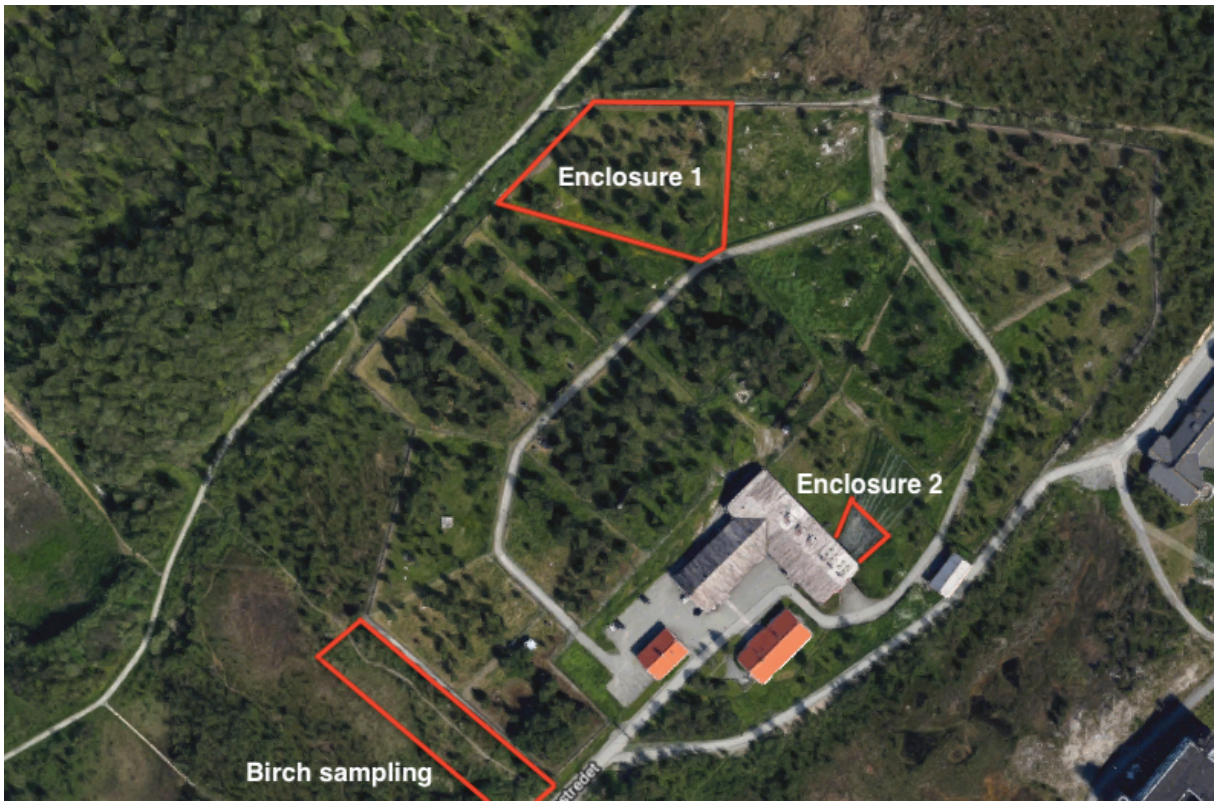

Figure A1– Photo of the Department of Arctic and Marine Biology, UiT - The Arctic University of Norway. Enclosure 1 (*E1*, large outdoor enclosure), enclosure 2 (*E2,* fan corridors used during experiment) and area where birch sampling took place are marked.

Food

Table A1 – List of ingredients in RF-80 (FK Reinfôr).

Content	Source	Latin name
Barley	Barley	*Hordeum vulgare*
Oat	Oat	*Avena sativa*
Wheat bran	Wheat	*Triticum aestivum*
Rape seed	Rape	*Brassica (napus* and/or *rapa)*
Rape fiber	Rape	*Brassica (napus* and/or *rapa)*
Beet pulp	Beet	*Beta vulgaris*
Soy flour	Soy	*Glycine max*
Molasses, sugarcane	Sugar cane	*Saccharum officinarum*
Molasses, beet	Sugar beet	*Beta vulgaris*
Alfalfa/Lucerne pellets	Alfalfa/Lucerne	*Medicago sativa*
Dried fat	African oil palm or *Brassica* sp.	*Elaeis guineensis* and *Brassica* sp.
"Lignobond"	Spruce or pine	*Picea abies* and *Pinus sylvestris*
Limestone flour		
Magnesium oxide		
Monocalcium phosphate		
Salt		
Micro-mineral mixture:		
Cu(Copper(II) sµLphate)		
Se (Sodium selenite)		
Zn (Zinc sµLfate)		
Mn (Manganese(II) sµLfate)		
I (Calcium iodate)		
Co (Cobalt(II) sµLphate)		
Vitamin mixture:		
Vitamin A		
Vitamin D3		
Vitamin E		

Table A2 – Overview of the amount of RF-80 eaten by the three reindeer (9/10, 10/10 and 12/10) during the feeding experiment. Prior to the feeding experiment and after the feeding experiment, the reindeer were fed RF-80 ad libitum.

Date	**Pellets eaten (g)**		
	9/10	**10/10**	**12/10**
19.01.18	1445	751	1081
20.01.18	1316	831	1264
21.01.18	1437	1128	1468
22.01.18	850	860	848
23.01.18	787	572	204
24.01.18	1158	980	1026
25.01.18	1427	1140	1217
26.01.18	1485	972	1144
27.01.18	1499	1213	1425
28.01.18	1500	1212	1478
29.01.18	1481	915	1236
30.01.18	1470	879	1144
31.01.18	1486	984	1217
01.02.18	1498	1308	1418
02.02.18	1498	1453	1446
03.02.18	1000	858	930
04.02.18	1000	770	808
05.02.18	1000	846	874
06.02.18	1500	979	1242
07.02.18	1500	1155	1282
08.02.18	1500	1225	1283
09.02.18	1500	1201	1083
10.02.18	1500	1490	1470
11.02.18	1500	869	1322
12.02.18	1500	1101	862
13.02.18	1700	1196	1638
14.02.18	2000	1558	1708
15.02.18	2000	2000	2000
16.02.18	2000	2000	2000
17.02.18	2000	2000	2000

Faecal samples

Table A3 – Overview of the faecal samples collected and analysed from animal 9/10, including information on date and time of faecal sampling. An additional column shows the number of hours from *Betula pubescens* was fed. The negative values indicate samples collected prior to feeding *B. pubescens.*

Sample ID	Date	Sample Time	Hours from *Betula pubescens* was fed
X_72	22.01.18	14:15	-31
X_73	22.01.18	16:00	-29
X_74	22.01.18	19:33	-25
X_1	24.01.18	03:12	6
X_2	24.01.18	08:16	11
X_3	24.01.18	17:00	20
X_4	24.01.18	22:27	25
X_5	25.01.18	03:15	30
X_6	25.01.18	11:00	38
X_7	25.01.18	17:10	44
X_8	25.01.18	23:17	50
X_9	26.01.18	03:25	54
X_10	26.01.18	11:05	62
X_11	26.01.18	16:34	67
X_12	26.01.18	22:00	73
X_13	27.01.18	06:00	81
X_14	27.01.18	09:00	84
X_15	27.01.18	16:06	91
X_16	27.01.18	23:59	99
X_17	28.01.18	06:00	105
X_18	28.01.18	11:32	110
X_19	28.01.18	16:37	115
X_20	28.01.18	23:59	123
X_21	29.01.18	06:00	129
X_22	29.01.18	10:38	133
X_23	29.01.18	16:25	139
X_24	29.01.18	23:57	147
X_25	30.01.18	03:00	150
X_26	30.01.18	11:58	159
X_27	30.01.18	14:21	161
X_28	30.01.18	23:55	171
X_29	31.01.18	05:56	177
X_30	31.01.18	11:55	183
X_31	31.01.18	16:07	187
X_32	31.01.18	23:58	195
X_33	01.02.18	03:01	198
X_34	01.02.18	09:21	204

Table A3 continued – Overview of the faecal samples collected and analysed from animal 9/10.

Sample ID	Date	Sample Time	Hours from *Betula pubescens* was fed
X_35	01.02.18	16:28	211
X_36	01.02.18	23:57	219
X_37	02.02.18	03:01	222
X_38	02.02.18	08:42	228
X_39	02.02.18	16:24	235
X_40	02.02.18	19:50	239
X_41	03.02.18	11:34	254
X_42	03.02.18	20:02	263
X_43	04.02.18	10:31	277
X_44	04.02.18	22:14	289
X_45	05.02.18	11:16	302
X_46	05.02.18	19:23	310
X_47	06.02.18	08:42	324
X_48	06.02.18	19:28	334
X_49	07.02.18	09:28	348
X_50	07.02.18	19:25	358
X_51	08.02.18	09:30	372
X_52	08.02.18	19:15	382
X_53	09.02.18	11:10	398
X_54	09.02.18	19:00	406
X_55	10.02.18	10:13	421
X_56	10.02.18	19:06	430
X_57	11.02.18	18:58	454
X_58	12.02.18	09:17	468
X_59	12.02.18	18:58	478
X_60	13.02.18	11:55	495
X_61	13.02.18	19:05	502
X_62	14.02.18	09:24	516
X_63	14.02.18	12:50	520
X_64	15.02.18	10:23	541
X_65	16.02.18	09:51	565
X_66	17.02.18	09:53	589
X_67	18.02.18	09:56	613
X_68	19.02.18	12:51	640
X_69	20.02.18	08:16	659
X_70	21.02.18	13:22	688
X_71	22.02.18	12:33	711
X_75	02.03.18	14:55	906
X_76	06.03.18	15:20	1002
X_77	09.03.18	15:08	1074
X_78	16.03.18	14:50	1242
X_79	20.03.18	14:38	1337
X_80	23.03.18	14:22	1409

Table A4 – Overview of the faecal samples collected and analysed from animal 10/10, including information on date and time of faecal sampling. An additional column shows the number of hours from *Betula pubescens* was fed. The negative values indicate samples collected prior to feeding *B. pubescens.*

Sample ID	Date	Sample Time	Hours from *Betula pubescens* was fed
Y_1	22.01.18	11:26	-31
Y_2	24.01.18	06:10	12
Y_3	24.01.18	13:30	19
Y_4	24.01.18	22:48	29
Y_5	25.01.18	03:19	33
Y_6	25.01.18	08:10	38
Y_7	25.01.18	13:01	43
Y_8	25.01.18	21:21	51
Y_9	26.01.18	03:21	57
Y_10	26.01.18	09:00	63
Y_11	26.01.18	16:38	70
Y_12	26.01.18	22:00	76
Y_13	27.01.18	06:00	84
Y_14	27.01.18	08:50	87
Y_15	27.01.18	14:35	92
Y_16	27.01.18	22:00	100
Y_17	28.01.18	06:00	108
Y_18	28.01.18	11:54	114
Y_19	28.01.18	16:43	119
Y_20	28.01.18	22:00	124
Y_21	29.01.18	06:00	132
Y_22	29.01.18	09:13	135
Y_23	29.01.18	14:14	140
Y_24	29.01.18	23:59	150
Y_25	30.01.18	03:22	153
Y_26	30.01.18	07:44	158
Y_27	30.01.18	14:30	164
Y_28	30.01.18	23:55	174
Y_29	31.01.18	02:45	177
Y_30	31.01.18	11:58	186
Y_31	31.01.18	16:12	190
Y_32	31.01.18	23:56	198
Y_33	01.02.18	03:00	201
Y_34	01.02.18	08:32	206
Y_35	01.02.18	13:43	212
Y_36	01.02.18	22:00	220
Y_37	02.02.18	06:01	228
Y_38	02.02.18	08:45	231
Y_39	02.02.18	16:29	238
Y_40	03.02.18	10:35	256
Y_41	03.02.18	16:35	262
Y_42	04.02.18	10:14	280
Y_43	04.02.18	22:12	292
Y_44	05.02.18	11:22	305
Y_45	05.02.18	19:25	313
Y_46	06.02.18	08:44	327

Table A4 continued – Overview of the faecal samples collected and analysed from animal 10/10.

Sample ID	Date	Sample Time	Hours from *Betula pubescens* was fed
Y_47	06.02.18	13:10	331
Y_48	07.02.18	09:30	351
Y_49	07.02.18	16:32	358
Y_50	08.02.18	09:13	375
Y_51	08.02.18	19:19	385
Y_52	09.02.18	10:15	400
Y_53	09.02.18	19:26	409
Y_54	10.02.18	09:52	424
Y_55	10.02.18	19:15	433
Y_56	11.02.18	18:51	457
Y_57	12.02.18	07:59	470
Y_58	12.02.18	16:00	478
Y_59	13.02.18	09:39	495
Y_60	13.02.18	19:01	505
Y_61	14.02.18	08:01	518
Y_62	14.02.18	16:51	527
Y_63	15.02.18	08:57	543
Y_64	16.02.18	11:48	570
Y_65	17.02.18	10:48	593
Y_66	18.02.18	12:50	619
Y_67	19.02.18	12:52	643
Y_68	20.02.18	08:33	662
Y_69	21.02.18	12:59	691
Y_70	22.02.18	13:26	715
Y_71	31.01.18	14:31	188
Y_72	02.03.18	15:25	909
Y_73	09.03.18	14:42	1077
Y_74	20.03.18	14:45	1341

Table A5 – Overview of the faecal samples collected and analysed from animal 12/10, including information on date and time of faecal sampling. An additional column shows the number of hours from *Betula pubescens* was fed. The negative values indicate samples collected prior to feeding *B. pubescens.*

Sample ID	Date	Sample Time	Hours from *Betula pubescens* was fed
Z_1	22.01.18	14:00	-29
Z_2	24.01.18	01:17	6
Z_3	24.01.18	11:10	16
Z_4	24.01.18	13:30	18
Z_5	24.01.18	22:28	27
Z_6	25.01.18	03:08	32
Z_7	25.01.18	11:00	40
Z_8	25.01.18	15:50	45
Z_9	25.01.18	21:13	50
Z_10	26.01.18	05:59	59
Z_11	26.01.18	11:07	64
Z_12	26.01.18	13:24	66
Z_13	26.01.18	22:00	75
Z_14	27.01.18	06:00	83
Z_15	27.01.18	08:50	86
Z_16	27.01.18	14:30	91
Z_17	27.01.18	23:59	101
Z_18	28.01.18	06:00	107
Z_19	28.01.18	11:55	113
Z_20	28.01.18	16:33	117
Z_21	28.01.18	23:59	125
Z_22	29.01.18	06:00	131
Z_23	29.01.18	10:42	135
Z_24	29.01.18	14:07	139
Z_25	29.01.18	23:50	149
Z_26	30.01.18	03:00	152
Z_27	30.01.18	11:41	160
Z_28	30.01.18	14:11	163
Z_29	30.01.18	23:55	173
Z_30	31.01.18	05:50	179
Z_31	31.01.18	11:07	184
Z_32	31.01.18	14:23	187
Z_33	31.01.18	21:57	195
Z_34	01.02.18	02:54	200
Z_35	01.02.18	09:24	206
Z_36	01.02.18	16:24	213
Z_37	01.02.18	22:21	219
Z_38	02.02.18	02:56	224
Z_39	02.02.18	10:10	231
Z_40	02.02.18	16:20	237
Z_41	02.02.18	19:46	240
Z_42	03.02.18	10:24	255
Z_43	03.02.18	17:40	262
Z_44	04.02.18	10:14	279
Z_45	04.02.18	22:09	291
Z_46	05.02.18	11:14	304

Table A5 continued – Overview of the faecal samples collected and analysed from animal 12/10.

Sample ID	Date	Sample Time	Hours from *Betula pubescens* was fed
Z_47	05.02.18	16:02	309
Z_48	06.02.18	10:40	327
Z_49	06.02.18	19:26	336
Z_50	07.02.18	09:43	350
Z_51	07.02.18	16:19	357
Z_52	08.02.18	08:23	373
Z_53	08.02.18	19:10	384
Z_54	09.02.18	10:30	399
Z_55	09.02.18	19:28	408
Z_56	10.02.18	10:17	423
Z_57	10.02.18	19:06	432
Z_58	11.02.18	18:51	456
Z_59	12.02.18	09:17	470
Z_60	12.02.18	16:00	477
Z_61	13.02.18	11:55	497
Z_62	13.02.18	19:07	504
Z_63	14.02.18	08:41	517
Z_64	14.02.18	16:38	525
Z_65	15.02.18	09:19	542
Z_66	16.02.18	11:21	568
Z_67	17.02.18	10:28	591
Z_68	18.02.18	09:56	615
Z_69	19.02.18	12:52	642
Z_70	20.02.18	09:14	662
Z_71	21.02.18	12:29	689
Z_72	22.02.18	13:03	714
Z_73	10.02.18	09:11	422
Z_74	25.01.18	23:13	52
Z_75	02.03.18	14:50	908
Z_76	02.03.18	14:57	908
Z_77	09.03.18	14:47	1076
Z_78	13.03.18	15:50	1173
Z_79	20.03.18	14:41	1339
Z_80	23.03.18	14:10	1411

PCR Plates – Faecal samples, Euka02

Table A6 – Plate 1 design from faecal samples from the feeding experiment, using the Euka02 primer. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control.

P1	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	X18	Y59	X24	X22	Y72	Z28	Z60	Z80	X57	X51	Z70	Z71
B	X2	X21	X20	X14	Z15	Z24	X80	Z77	Z3	Z48	Z74	Z55
C	Z17	X12	Y69	Y57	Z69	Z12	X75	Y24	X68	X23	Z27	Z35
D	Z18	Y56	X17	Z46	Z16	Z56	Z76	X79	Z52	Z11	X56	X59
E	X9	X15	Z34	X26	Z37	Z23	X76	Z75	Z78	X64	Z5	Z54
F	Z42	Z20	X28	Z7	Z53	Z68	Y74	Z44	Z65	Z49	X66	Z63
G	Z30	X19	Z61	Z45	Z66	Z32	Z79	X77	X11	X42	Z51	PCR_NC
H	Z40	DNA_NC	X16	DNA_NC	Z1	DNA_NC	X78	DNA_NC	Z72	DNA_NC	X4	PCR_NC

Table A7 – Plate 2 design from faecal samples from the feeding experiment, using the Euka02 primer. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control. PCR_POS = PCR positive control.

P2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Z67	X34	Y9	Z8	Z19	X50	Y70	Z4	Y8	Z36	X38	X36
B	Y33	X35	Y32	Y43	Z31	Y5	X74	Y21	Y40	Y50	Z13	Z38
C	X65	X25	Y53	Y14	Y49	Z33	X30	X37	Y23	Y39	Y28	Z59
D	X10	X33	Z25	Y13	X31	Y1	X44	X63	Y42	Y48	Y18	Y58
E	X3	X60	Z10	Y38	Y51	Y46	Z14	Z43	Y47	Y11	X29	PCR_NC
F	X70	X54	Y25	Y6	Y52	Y34	Y61	Y41	Y4	Y7	Y26	PCR_NC
G	X27	X53	Y3	Z73	Y36	Y44	Y71	Z6	Y31	Y45	Z21	PCR_POS
H	Z62	DNA_NC	Z22	DNA_NC	Y29	DNA_NC	X41	DNA_NC	Y2	DNA_NC	X39	PCR_POS

Table A8 – Plate 3 design from faecal samples from the feeding experiment, using the Euka02 primer. Only eight columns were used. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control.

P3	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
A	Z29	Y64	X67	Y35	Y15	X32	Y63	Y66
B	Z57	X6	X43	X8	Z9	X52	X71	Y17
C	Z50	Y54	X46	X45	Y19	Y65	Y68	Y73
D	Z26	Y12	Y22	X69	Y27	Y37	X49	X73
E	Y67	X72	Z47	Z41	X62	Y55	X1	PCR_NC
F	X7	Y30	X47	Y20	X48	X58	Y10	PCR_NC
G	Y62	Z58	Y16	X40	Y60	X55	X5	PCR_NC
H	Z2	DNA_NC	DNA_NC	DNA_NC	DNA_NC	DNA_NC	DNA_NC	PCR_NC

PCR Plates – Faecal samples, Fung01

Table A9 – Plate 1 design from faecal samples from the feeding experiment, using the Fung01 primer. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control.

P1	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	X18	Y59	X24	X22	Y72	Z28	Z60	Z80	X57	X51	Z70	Z71
B	X2	X21	X20	X14	Z15	Z24	X80	Z77	Z3	Z48	Z74	Z55
C	Z17	X12	Y69	Y57	Z69	Z12	X75	Y24	X68	X23	Z27	Z35
D	Z18	Y56	X17	Z46	Z16	Z56	Z76	X79	Z52	Z11	X56	X59
E	X9	X15	Z34	X26	Z37	Z23	X76	Z75	Z78	X64	Z5	Z54
F	Z42	Z20	X28	Z7	Z53	Z68	Y74	Z44	Z65	Z49	X66	Z63
G	Z30	X19	Z61	Z45	Z66	Z32	Z79	X77	X11	X42	Z51	PCR_NC
H	Z40	DNA_NC	X16	DNA_NC	Z1	DNA_NC	X78	DNA_NC	Z72	DNA_NC	X4	PCR_NC

Table A10 – Plate 2 design from faecal samples from the feeding experiment, using the Fung01 primer. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control.

P2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Z67	X34	Y9	Z8	Z19	X50	Y70	Z4	Y8	Z36	X38	X36
B	Y33	X35	Y32	Y43	Z31	Y5	X74	Y21	Y40	Y50	Z13	Z38
C	X65	X25	Y53	Y14	Y49	Z33	X30	X37	Y23	Y39	Y28	Z59
D	X10	X33	Z25	Y13	X31	Y1	X44	X63	Y42	Y48	Y18	Y58
E	X3	X60	Z10	Y38	Y51	Y46	Z14	Z43	Y47	Y11	X29	PCR_NC
F	X70	X54	Y25	Y6	Y52	Y34	Y61	Y41	Y4	Y7	Y26	PCR_NC
G	X27	X53	Y3	Z73	Y36	Y44	Y71	Z6	Y31	Y45	Z21	PCR_NC
H	Z62	DNA_NC	Z22	DNA_NC	Y29	DNA_NC	X41	DNA_NC	Y2	DNA_NC	X39	PCR_NC

Table A11 – Plate 3 design from faecal samples from the feeding experiment, using the Fung01 primer. Only eight columns were used. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control.

P3	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
A	Z29	Y64	X67	Y35	Y15	X32	Y63	Y66
B	Z57	X6	X43	X8	Z9	X52	X71	Y17
C	Z50	Y54	X46	X45	Y19	Y65	Y68	Y73
D	Z26	Y12	Y22	X69	Y27	Y37	X49	X73
E	Y67	X72	Z47	Z41	X62	Y55	X1	PCR_NC
F	X7	Y30	X47	Y20	X48	X58	Y10	PCR_NC
G	Y62	Z58	Y16	X40	Y60	X55	X5	PCR_NC
H	Z2	DNA_NC	DNA_NC	DNA_NC	DNA_NC	DNA_NC	DNA_NC	PCR_NC

PCR Plates – Faecal samples, Sper01

Table A12 – Plate 1 design from faecal samples from the feeding experiment, using the Sper01 primer. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control. PCR_POS = PCR positive control.

P1	1	2	3	4	5	6	7	8	9	10	11	12
A	X18	Y59	X24	X22	Y72	Z28	Z60	Z80	X57	X51	Z70	Z71
B	X2	X21	X20	X14	Z15	Z24	X80	Z77	Z3	Z48	Z74	Z55
C	Z17	X12	Y69	Y57	Z69	Z12	X75	Y24	X68	X23	Z27	Z35
D	Z18	Y56	X17	Z46	Z16	Z56	Z76	X79	Z52	Z11	X56	X59
E	X9	X15	Z34	X26	Z37	Z23	X76	Z75	Z78	X64	Z5	Z54
F	Z42	Z20	X28	Z7	Z53	Z68	Y74	Z44	Z65	Z49	X66	Z63
G	Z30	X19	Z61	Z45	Z66	Z32	Z79	X77	X11	X42	Z51	PCR_NC
H	Z40	DNA_NC	X16	DNA_NC	Z1	DNA_NC	X78	DNA_NC	Z72	DNA_NC	X4	PCR_POS

Table A13 – Plate 2 design from faecal samples from the feeding experiment, using the Sper01 primer. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control. PCR_POS = PCR positive control.

P2	1	2	3	4	5	6	7	8	9	10	11	12
A	Z67	X34	Y9	Z8	Z19	X50	Y70	Z4	Y8	Z36	X38	X36
B	Y33	X35	Y32	Y43	Z31	Y5	X74	Y21	Y40	Y50	Z13	Z38
C	X65	X25	Y53	Y14	Y49	Z33	X30	X37	Y23	Y39	Y28	Z59
D	X10	X33	Z25	Y13	X31	Y1	X44	X63	Y42	Y48	Y18	Y58
E	X3	X60	Z10	Y38	Y51	Y46	Z14	Z43	Y47	Y11	X29	PCR_NC
F	X70	X54	Y25	Y6	Y52	Y34	Y61	Y41	Y4	Y7	Y26	PCR_NC
G	X27	X53	Y3	Z73	Y36	Y44	Y71	Z6	Y31	Y45	Z21	PCR_POS
H	Z62	DNA_NC	Z22	DNA_NC	Y29	DNA_NC	X41	DNA_NC	Y2	DNA_NC	X39	PCR_POS

Table A14 – Plate 3 design from faecal samples from the feeding experiment, using the Sper01 primer. Only eight columns were used. Abbreviations: DNA_NC = DNA extraction negative control. PCR_NC = PCR negative control. PCR_POS = PCR positive control.

P3	1	2	3	4	5	6	7	8
A	Z29	Y64	X67	Y35	Y15	X32	Y63	Y66
B	Z57	X6	X43	X8	Z9	X52	X71	Y17
C	Z50	Y54	X46	X45	Y19	Y65	Y68	Y73
D	Z26	Y12	Y22	X69	Y27	Y37	X49	X73
E	Y67	X72	Z47	Z41	X62	Y55	X1	PCR_NC
F	X7	Y30	X47	Y20	X48	X58	Y10	PCR_NC
G	Y62	Z58	Y16	X40	Y60	X55	X5	PCR_POS
H	Z2	DNA_NC	DNA_NC	DNA_NC	DNA_NC	DNA_NC	DNA_NC	PCR_POS

PCR Plates – Food samples

Table A15 – Food sample plate design from the feeding experiment, using the Euka02 primer. All three replicates are on the same PCR plate, with replicate 1 in well columns 1-4, replicate 2 in well columns 5-8 and replicate 3 in well columns 9-12. Abbreviations: DNA_NC = DNA extraction negative control. NC_OA: Negative control open air. PCR_NC = PCR negative control.

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1.1	F3.3	F6.1	F8.3	F1.1	F3.3	F6.1	F8.3	F1.1	F3.3	F6.1	F8.3
B	F1.2	F4.1	F6.2	NC_OA_1	F1.2	F4.1	F6.2	NC_OA_1	F1.2	F4.1	F6.2	NC_OA_1
C	F1.3	F4.2	F6.3	NC_OA_2	F1.3	F4.2	F6.3	NC_OA_2	F1.3	F4.2	F6.3	NC_OA_2
D	F2.1	F4.3	F7.1	PCR_NC	F2.1	F4.3	F7.1	PCR_NC	F2.1	F4.3	F7.1	PCR_NC
E	F2.2	F5.1	F7.2	PCR_NC	F2.2	F5.1	F7.2	PCR_NC	F2.2	F5.1	F7.2	PCR_NC
F	F2.3	F5.2	F7.3		F2.3	F5.2	F7.3		F2.3	F5.2	F7.3	
G	F3.1	F5.3	F8.1		F3.1	F5.3	F8.1		F3.1	F5.3	F8.1	
H	F3.2	DNA_NC_1	F8.2		F3.2	DNA_NC_1	F8.2		F3.2	DNA_NC_1	F8.2	

Table A16– Food sample plate design from the feeding experiment, using the Fung01 primer. All three replicates are on the plate, with replicate 1 in well columns 1-4, replicate 2 in well columns 5-8 and replicate 3 in well columns 9-12. Abbreviations: DNA_NC = DNA extraction negative control. NC_OA: Negative control open air. PCR_NC = PCR negative control.

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1.1	F3.3	F6.1	F8.3	F1.1	F3.3	F6.1	F8.3	F1.1	F3.3	F6.1	F8.3
B	F1.2	F4.1	F6.2	NC_OA_1	F1.2	F4.1	F6.2	NC_OA_1	F1.2	F4.1	F6.2	NC_OA_1
C	F1.3	F4.2	F6.3	NC_OA_2	F1.3	F4.2	F6.3	NC_OA_2	F1.3	F4.2	F6.3	NC_OA_2
D	F2.1	F4.3	F7.1	PCR_NC	F2.1	F4.3	F7.1	PCR_NC	F2.1	F4.3	F7.1	PCR_NC
E	F2.2	F5.1	F7.2	PCR_NC	F2.2	F5.1	F7.2	PCR_NC	F2.2	F5.1	F7.2	PCR_NC
F	F2.3	F5.2	F7.3		F2.3	F5.2	F7.3		F2.3	F5.2	F7.3	
G	F3.1	F5.3	F8.1		F3.1	F5.3	F8.1		F3.1	F5.3	F8.1	
H	F3.2	DNA_NC_1	F8.2		F3.2	DNA_NC_1	F8.2		F3.2	DNA_NC_1	F8.2	

Table A17 – Food sample plate design from the feeding experiment, using the Sper01 primer. All three replicates are on the plate, with replicate 1 in well columns 1-4, replicate 2 in well columns 5-8 and replicate 3 in well columns 9-12. Abbreviations: DNA_NC = DNA extraction negative control. NC_OA: Negative control open air. PCR_NC = PCR negative control.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	F1.1	F3.3	F6.1	F8.3	F1.1	F3.3	F6.1	F8.3	F1.1	F3.3	F6.1	F8.3
B	F1.2	F4.1	F6.2	NC_OA_1	F1.2	F4.1	F6.2	NC_OA_1	F1.2	F4.1	F6.2	NC_OA_1
C	F1.3	F4.2	F6.3	NC_OA_2	F1.3	F4.2	F6.3	NC_OA_2	F1.3	F4.2	F6.3	NC_OA_2
D	F2.1	F4.3	F7.1	PCR_NC	F2.1	F4.3	F7.1	PCR_NC	F2.1	F4.3	F7.1	PCR_NC
E	F2.2	F5.1	F7.2	PCR_POS	F2.2	F5.1	F7.2	PCR_NC	F2.2	F5.1	F7.2	PCR_NC
F	F2.3	F5.2	F7.3		F2.3	F5.2	F7.3		F2.3	F5.2	F7.3	
G	F3.1	F5.3	F8.1		F3.1	F5.3	F8.1		F3.1	F5.3	F8.1	
H	F3.2	DNA_NC_1	F8.2		F3.2	DNA_NC_1	F8.2		F3.2	DNA_NC_1	F8.2	

Table A18 – Overview of samples that evaporated during PCR.

Sample type	**Primer**	**Plate**	**Replicate**	**Well evaporated**
Faecal sample	Sper01	P1	R1	H1; H12; H4
Faecal sample	Sper01	P1	R2	F2
Faecal sample	Sper01	P1	R3	A4; G3
Faecal sample	Euka02	P1	R1	H1, A5; A6; H4; G9
Faecal sample	Euka02	P1	R3	G3
Faecal sample	Bryo01	P1	R1	H3; H4; A12
Faecal sample	Euka02	P3	R2	G3
Faecal sample	Fung01	P3	R1	A7
Faecal sample	Fung01	P3	R2	F2
Faecal sample	Fung01	P3	R3	F2; G2; G3; G6
Faecal sample	Sper01	P3	R1	A1
Faecal sample	Sper01	P3	R2	F2
Faecal sample	Euka02	P2	R1	H12
Faecal sample	Euka02	P2	R2	A1; H12
Faecal sample	Fung01	P1	R1	A3; A11; A12; H2; H3; H6

Dominant MOTUs in faecal samples

Table A19 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified eukaryote sequences (Euka02) from the faecal samples. Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	Mean relative frequency
Brassicaceae	0.181350574
Neocallimastigaceae	0.077306871
Eukaryota	0.077204394
Iodamoeba sp. RL1	0.076681598
Eukaryota	0.057722441
Eukaryota	0.043444773
Chromadorea	0.042731348
Pooideae	0.039455065
Lecanoromycetidae	0.034795938
Desmonomata	0.031390539
Eukaryota	0.024795714
Juncus	0.023650741
Betula	0.021077768
Eukaryota	0.018434824
Entamoeba	0.014994416
Eukaryota	0.013491268
Gonostomatidae	0.013270827
Fallopia multiflora	0.011130499
Oribatulidae	0.010722280
Penicillium	0.009401209

Table A20 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified fungal sequences (Fung01) from the faecal samples. Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	**Mean relative frequency**
Cladoniaceae	0.06319851
Sporobolomyces ruberrimus	0.06020761
Cladoniaceae	0.05935724
Aspergillus	0.05078200
Cladosporium	0.03284640
Preussia	0.03106743
Eukaryota	0.02195416
Rhodotorula	0.02054341
Eukaryota	0.02014685
Preussia	0.01918650
Penicillium aff. *lividum*	0.01897601
Eukaryota	0.01738675
Abrothallus	0.01617349
Aspergillus fumigatus	0.01575310
Plectosphaerellaceae	0.01570691
Cladosporium herbarum	0.01324442
Vishniacozyma sp.	0.01280646
Eukaryota	0.01214058
Helotiales	0.01026486
Thelebolus	0.01017539

Table A21 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified seed plant sequences (Sper01) from the faecal samples. Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	**Mean relative frequency**
Betulaceae	0.20486684
Juncus	0.11317423
Polygoneae	0.10358756
Asterales	0.06698672
Medicago	0.05087890
Trifolium	0.04339782
Brassicaceae	0.04178214
Polygonum	0.03096070
Pisum	0.02665431
Polygonoideae	0.02518996
Rumiceae	0.02463336
Avena	0.02327361
Poeae	0.02269516
Pooideae	0.02125329
Pooideae	0.01843495
Loteae	0.01612375
Poeae Chloroplast Group 1 (Aveneae type)	0.01589193
Hordeinae	0.01453078
Persicaria	0.01358394
Chenopodioideae	0.01306447

Dominant MOTUs in food samples

Table A22 – Overview of the 15 MOTUs with the highest mean relative frequency in the dataset containing amplified eukaryote sequences (Euka02) from the samples taken of RF-80. Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	Mean relative frequency
Pooideae	7.131320e-01
Brassicaceae	2.408663e-01
Vigna angularis var. *angularis*	3.364172e-02
Brassicaceae	8.895405e-03
Atripliceae	1.927225e-03
Fungi	1.491811e-03
Carabodes labyrinthicus	4.547233e-05
Lecanoromycetidae	0.000000e+00
Betula	0.000000e+00
Cicadellidae	0.000000e+00
Dothidea sp. RWT7	0.000000e+00
Abrothallus	0.000000e+00
Oribatulidae	0.000000e+00
Carabodes labyrinthicus	0.000000e+00
Orbiliaceae	0.000000e+00

Table A23 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified fungi sequences (Fung01) from the samples taken of RF-80. Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	**Mean relative frequency**
Epicoccum	0.161877156
Cladosporium herbarum	0.144477578
Alternaria infectoria	0.133515832
Cladosporium	0.085347466
Fusarium sp. *UASWS1560*	0.071560042
Fusarium	0.059436326
Microdochium	0.053768215
Nectria	0.052951587
Alternaria	0.032698724
Microdochium	0.027982109
Fusarium	0.025170186
Alternaria infectoria	0.016025048
Fusarium	0.015965286
Parastagonospora nodorum	0.015124297
Neoascochyta exitialis	0.011702782
Aspergillaceae	0.011141546
Alternaria infectoria	0.009611790
Cystofilobasidium macerans	0.008268444
Ascochyta skagwayensis	0.007768463
Sordariomycetes	0.007738963

Table A24 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified seed plant sequences (Sper01) from the samples taken of RF-80. Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	Mean relative frequency
Betulaceae	0.304243417
Saliceae	0.112723926
Empetrum	0.076937100
Filipendula ulmaria	0.057461839
Athyrium	0.057227923
Poeae Chloroplast Group 2 (Poeae type)	0.049500079
Potentilleae	0.036228932
Rosaceae	0.035131567
Vaccinieae	0.034547072
Colurieae	0.027347755
Dryadoideae	0.021171543
Vaccinium ovalifolium	0.019029714
Cornus suecica	0.015586798
Pinus	0.014551947
Alchemilla	0.010982542
Geranium	0.009868240
Gymnocarpium dryopteris	0.008959834
Vaccinium	0.007486944
Cornus suecica	0.006729806
Gaylussacia	0.005827376

Table A25 – Overview of the 15 MOTUs with the highest mean relative frequency in the dataset containing amplified eukaryote sequences (Euka02) from the food samples taken of *Betula pubescens*. Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	**Mean relative frequency**
Betula	8.445860e-01
Cicadellidae	6.932063e-02
Dothidea sp. *RWT7*	3.844499e-02
Pezizomycotina	2.131606e-02
Orbiliaceae	1.755343e-02
Fungi	6.228601e-03
Lecanoromycetidae	2.485807e-03
Pooideae	3.287502e-05
Brassicaceae	2.609089e-05
Abrothallus	5.521066e-06
Vigna angularis var. *angularis*	0.000000e+00
Brassicaceae	0.000000e+00
Oribatulidae	0.000000e+00
Carabodes labyrinthicus	0.000000e+00
Atripliceae	0.000000e+00

Table A26 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified seed plant sequences (Sper01) from the food samples taken of *Betula pubescens*. Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	**Mean relative frequency**
Dryadoideae	0.279076350
Betulaceae	0.164114828
Saliceae	0.150479732
Maleae	0.079228721
Athyrium	0.074772445
Poeae Chloroplast Group 2 (Poeae type)	0.050115898
Potentilleae	0.038255984
Cornus suecica	0.035316172
Rosaceae	0.022028497
Rumiceae	0.016487005
Cornus suecica	0.015476351
Colurieae	0.010352720
Vaccinieae	0.006631912
Gymnocarpium dryopteris	0.004989637
Cornales	0.004852631
Vaccinium	0.004565967
Vaccinium ovalifolium	0.003983467
Saxifraga	0.003968236
Asteroideae	0.003836801
Loasaceae	0.003501848

Table A27 – Overview of the 15 MOTUs with the highest mean relative frequency in the dataset containing amplified eukaryote sequences (Euka02) from the food samples taken of lichen mixture (dominated by *Cladonia stellaris*). Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	**Mean relative frequency**
Lecanoromycetidae	9.638382e-01
Carabodes labyrinthicus	1.996394e-02
Fungi	1.558627e-02
Pezizomycotina	3.289728e-04
Brassicaceae	1.123143e-04
Abrothallus	7.804221e-05
Dothidea sp. *RWT7*	7.251424e-05
Pooideae	1.979722e-05
Betula	0.000000e+00
Vigna angularis var. *angularis*	0.000000e+00
Cicadellidae	0.000000e+00
Brassicaceae	0.000000e+00
Orbiliaceae	0.000000e+00
Oribatulidae	0.000000e+00
Atripliceae	0.000000e+00

Table A28 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified fungal sequences (Fung01) from the food samples taken of lichen mixture (dominated by Cladonia stellaris). Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that do not correspond to fed species are marked with black font.

Scientific name	**Mean relative frequency**
Cladoniaceae	4.682286e-01
Letharia vulpina	3.055494e-01
Cladoniaceae	2.262171e-01
Abrothallus	3.129871e-06
Eukaryota	1.793266e-06
Epicoccum	0.000000e+00
Cladosporium herbarum	0.000000e+00
Alternaria infectoria	0.000000e+00
Cladosporium	0.000000e+00
Fusarium sp. *UASWS1560*	0.000000e+00
Fusarium	0.000000e+00
Microdochium	0.000000e+00
Nectria	0.000000e+00
Alternaria	0.000000e+00
Microdochium	0.000000e+00
Eukaryota	0.000000e+00
Fusarium	0.000000e+00
Parmelia barrenoae	0.000000e+00
Leotiomyceta	0.000000e+00
Fusarium	0.000000e+00

Table A29 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified seed plant sequences (Sper01) from the food samples taken of lichen mixture (dominated by *Cladonia stellaris*). Repeating MOTUs appear when sequences are identified to different sequences in the reference database. MOTUs corresponding to the fed species are marked with grey font, and MOTUs that does not correspond to fed species are marked with black font.

Scientific name	**Mean relative frequency**
Dryadoideae	0.190529971
Betulaceae	0.154476490
Saliceae	0.145314522
Poeae Chloroplast Group 2 (Poeae type)	0.086075225
Empetrum	0.061768010
Rumiceae	0.051647965
Athyrium	0.046858447
Potentilleae	0.030373708
Rosaceae	0.027948516
Geranium	0.024973673
Cornus suecica	0.024105478
Saxifraga	0.019088180
Vaccinieae	0.018920408
Filipendula ulmaria	0.016519953
Vaccinium ovalifolium	0.013887692
Cornus suecica	0.009444064
Alchemilla	0.006466077
Gymnocarpium dryopteris	0.006304855
Gaylussacia	0.005612717
Asteroideae	0.005494742

Table A30 – Overview of the 15 MOTUs with the highest mean relative frequency in the dataset containing amplified eukaryote sequences from the food samples containing lichens collected in *E1*. Repeating MOTUs appear when sequences are identified to different sequences in the reference database.

Scientific name	**Mean relative frequency**
Lecanoromycetidae	9.285196e-01
Abrothallus	2.686601e-02
Fungi	2.295038e-02
Oribatulidae	1.729512e-02
Dothidea sp. *RWT7*	1.538802e-03
Pezizomycotina	1.180022e-03
Orbiliaceae	1.051116e-03
Carabodes labyrinthicus	5.915119e-04
Pooideae	7.456350e-06
Brassicaceae	0.000000e+00
Betula	0.000000e+00
Vigna angularis var. *angularis*	0.000000e+00
Cicadellidae	0.000000e+00
Brassicaceae	0.000000e+00
Atripliceae	0.000000e+00

Table A31 – Overview of the 20 MOTUs with the highest mean relative frequency in the dataset containing amplified fungal sequences from the food samples containing lichens collected in *E1*. Repeating MOTUs appear when sequences are identified to different sequences in the reference database.

Scientific name	**Mean relative frequency**
Parmelia barrenoae	0.136997010
Abrothallus	0.130205238
Leotiomyceta	0.124447804
Leotiomyceta	0.109751906
Parmelia barrenoae	0.073882523
Leotiomyceta	0.058505935
Leotiomyceta	0.052006918
Leotiomyceta	0.050155119
Abrothallus	0.039169507
Leotiomycetes	0.038269689
Glutinomyces	0.029004479
Abrothallus	0.028211106
Leotiomyceta	0.024515501
Leotiomyceta	0.022317303
Leotiomyceta	0.019639168
Leotiomyceta	0.016929579
Leotiomyceta	0.012434373
Leotiomyceta	0.010557720
Leotiomyceta	0.008045292
Leotiomyceta	0.007256102

Betula

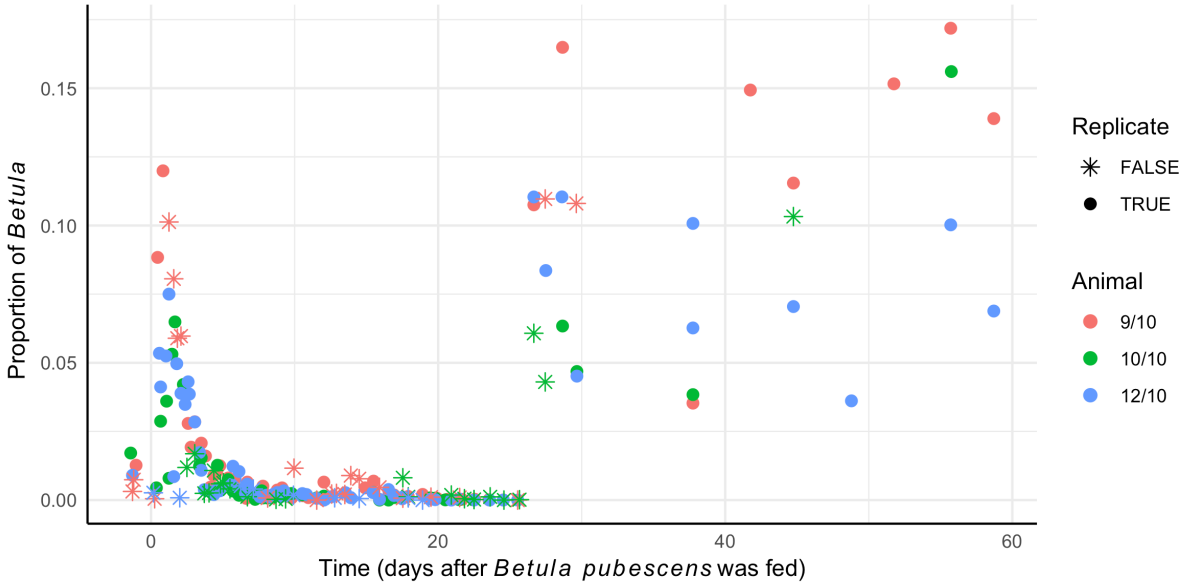

Figure A2 - Proportion of *Betula* DNA in the diet, based on all sequences in the dataset consisting of amplified eukaryote sequences. The three reindeer included in the experiment are represented by colour. The samples without replicates are shown with a star sign, whilst the samples with three replicates are shown with a filled circle.

Appendix B - Gastrointestinal tract

Samples

Table B1 – Overview of the samples collected from the GI tract. Liquid content refers to the consistency of the sample. Abbreviations: Small I. 1 = Small Intestine Upper part, Small I. 1 = Small Intestine Lower part.

Process ID	Sample Name	Animal ID	Part intestine	Subsampled material (mg)	Remarks 1
1.1	A1_S1	Animal_1	Rumen	127	
1.2	A1_S2	Animal_1	Reticulum	114	
1.3	A1_S3	Animal_1	Omasum	103	
1.4	A1_S4	Animal_1	Abomasum	104	Liquid content
1.5	A1_S5	Animal_1	Small I. 1	155	
1.6	A1_S6	Animal_1	Small I. 2	128	Liquid content
1.7	A1_S7	Animal_1	Cecum	118	
1.8	A1_S8	Animal_1	Spiral colon	139	
1.9	A1_S9	Animal_1	Distal colon	129	
1.10	DNANC1			NA	Extraction negative control
1.11	DNANC2			NA	Extraction negative control
2.1	A2_S1	Animal_2	Rumen	151	
2.2	A2_S2	Animal_2	Reticulum	151	
2.3	A2_S3	Animal_2	Omasum	130	
2.4	A2_S4	Animal_2	Abomasum	116	
2.5	A2_S5	Animal_2	Small I. 1	112	Liquid content
2.6	A2_S6	Animal_2	Small I. 2	154	
2.7	A2_S7	Animal_2	Cecum	148	
2.8	A2_S8	Animal_2	Spiral colon	141	
2.9	A2_S9	Animal_2	Distal colon	129	
2.10	DNANC3			NA	Extraction negative control
3.1	A3_S1	Animal_3	Rumen	150	
3.2	A3_S2	Animal_3	Reticulum	159	
3.3	A3_S3	Animal_3	Omasum	139	
3.4	A3_S4	Animal_3	Abomasum	157	
3.5	A3_S5	Animal_3	Small I. 1	150	
3.6	A3_S6	Animal_3	Cecum	114	
3.7	A3_S7	Animal_3	Spiral colon	112	
3.8	A3_S8	Animal_3	Distal colon	120	
3.9	DNANC4			NA	Extraction negative control
4.1	A4_S1	Animal_4	Rumen	159	
4.2	A4_S2	Animal_4	Reticulum	133	
4.3	A4_S3	Animal_4	Omasum	103	
4.4	A4_S4	Animal_4	Abomasum	104	
4.5	A4_S5	Animal_4	Small I. 1	129	
4.6	A4_S6	Animal_4	Small I. 2	137	
4.7	A4_S7	Animal_4	Cecum	111	
4.8	A4_S8	Animal_4	Spiral colon	107	
4.9	A4_S9	Animal_4	Distal colon	116	
4.10	DNANC5			NA	Extraction negative control

PCR Plates

Table B2 – PCR plate design for the samples collected from the GI tract. The numbers in the wells are the Process ID (Table B1) for each sample. The same PCR plate design was used for all primers (Euka02, Fungi01, Sper01 and Bryo01). Sample 1.10, 1.11, 2.10, 3.9 and 4.10 are negative DNA extraction controls. Abbreviations: PCR_NC = PCR negative control.

	1	**2**	**3**	**4**	**5**	**6**
A	1.1_	1.9_	2.6_	3.4_	4.3_	PCR_NC
B	1.2_	1.10_	2.7_	3.5_	4.4_	PCR_NC
C	1.3_	1.11_	2.8_	3.6_	4.5_	
D	1.4_	2.1_	2.9_	3.7_	4.6_	
E	1.5_	2.2_	2.10_	3.8_	4.7_	
F	1.6_	2.3_	3.1_	3.9_	4.8_	
G	1.7_	2.4_	3.2_	4.1_	4.9_	
H	1.8_	2.5_	3.3_	4.2_	4.10_	

Table B3 – Overview of samples that evaporated during PCR.

Primer	**Replicate**	**Well evaporated**
Fung01	R1/R2	F2; G2; G3
Euka02	R3	G3
Sper01	R1/R2	H1; H2
Sper01	R3	F2

Principal Component Analysis

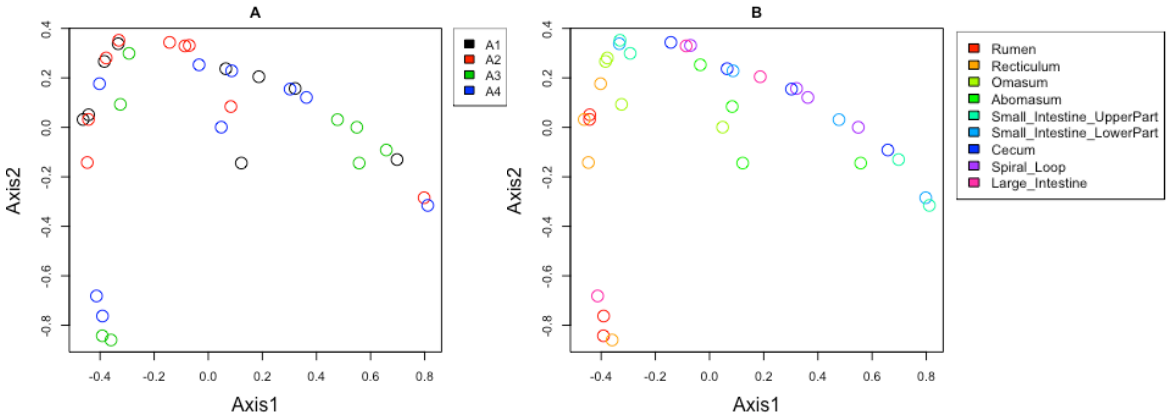

Figure B1 - Principal component analysis (PCA) of Hellinger-transformed number of sequences from the dataset with amplified eukaryotes (Euka02). A compares the sample distribution focusing on the four animals, whilst B compares the sample distribution focusing on the different parts of the gastrointestinal tract, indicated by different colours.

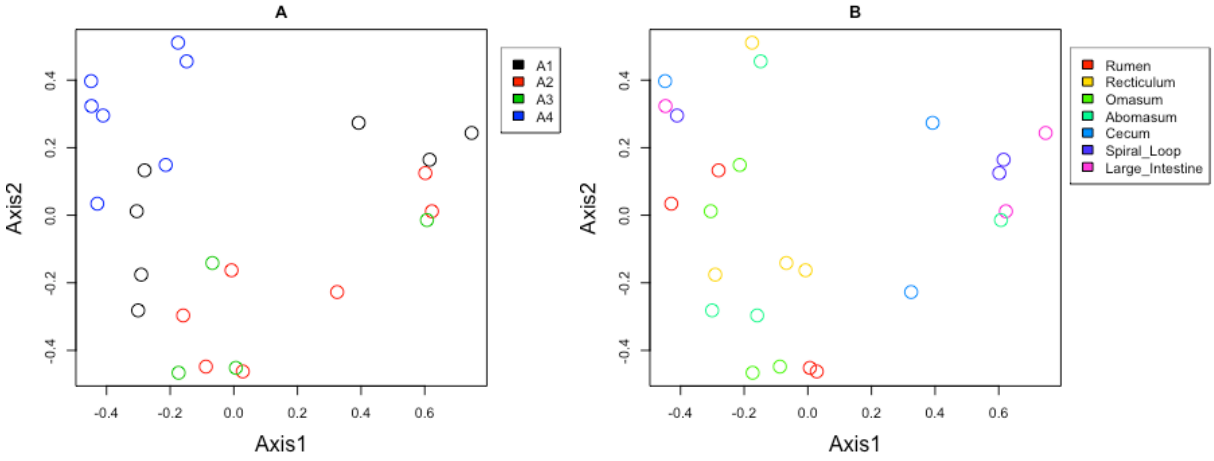

Figure B2 – Principal component analysis (PCA) of Hellinger-transformed number of sequences from the dataset with amplified fungi (Fung01). A compares the sample distribution focusing on the four animals, whilst B compares the sample distribution focusing on the different parts of the gastrointestinal tract, indicated by different colours.

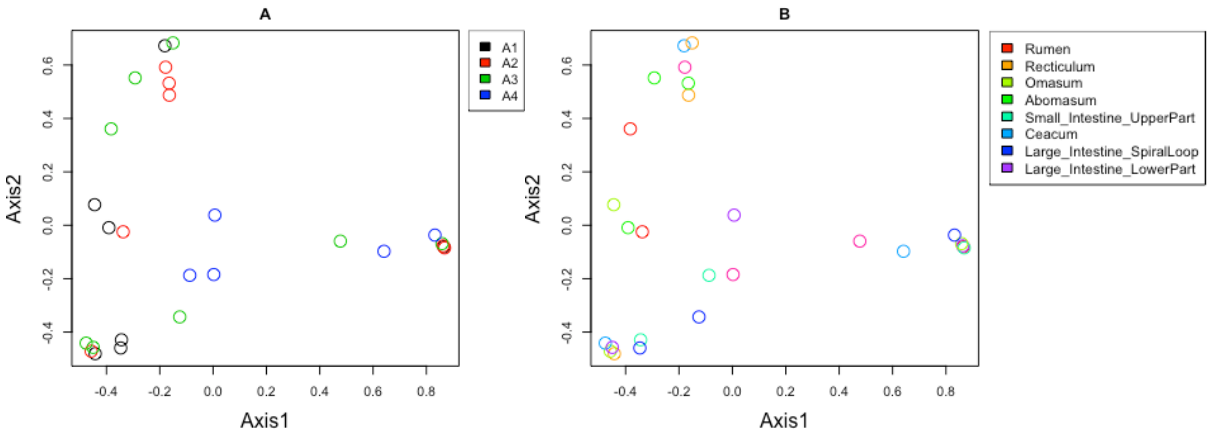

Figure B3 – Principal component analysis (PCA) of Hellinger-transformed number of sequences from the dataset with amplified seed plants (Sper01). A compares the sample distribution focusing on the four animals, whilst B compares the sample distribution focusing on the different parts of the gastrointestinal tract, indicated by different colours.

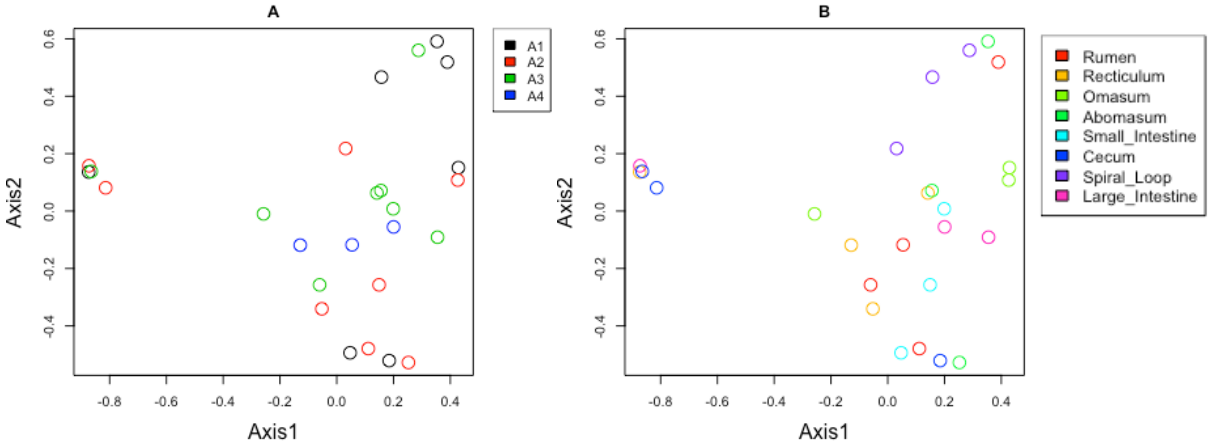

Figure B4 - Principal component analysis (PCA) of Hellinger-transformed number of sequences from the dataset with amplified bryophytes (Bryo01). A compares the sample distribution focusing on the four animals, whilst B compares the sample distribution focusing on the different parts of the gastrointestinal tract, indicated by different colours.

Diversity

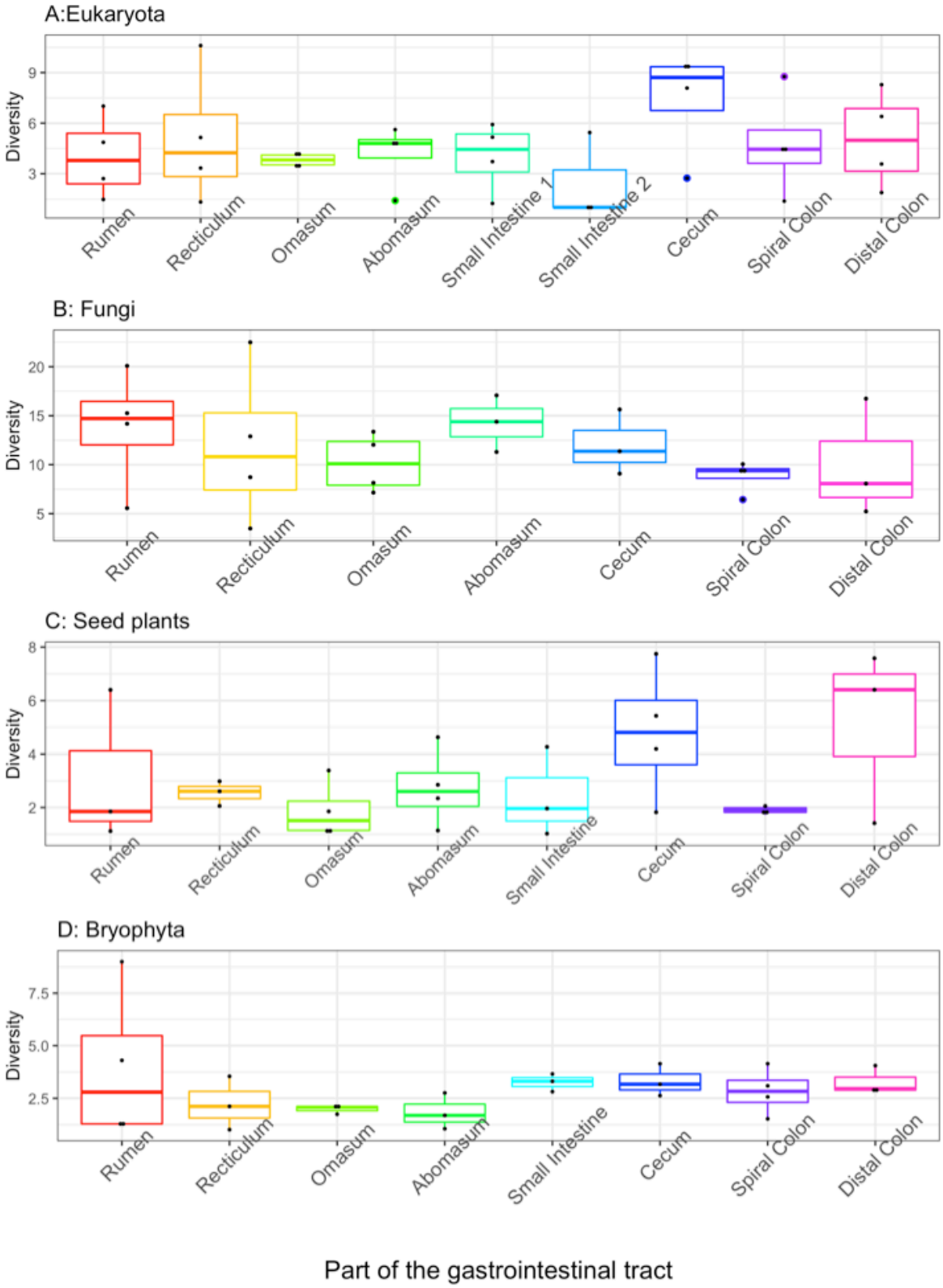

Figure B5 – Boxplots comparing the diversity (A: Eukaryota, B: Fungi, C: Seed plants (Spermatophyta), D: Bryophyta) along the gastrointestinal tract, when q=2 (i.e. Gini coefficient). Black, filled circles represent samples, and thick line within the quartile of the boxes shows the median. X-axis = Gastrointestinal parts. Y-axis = Diversity (Gini coefficient).

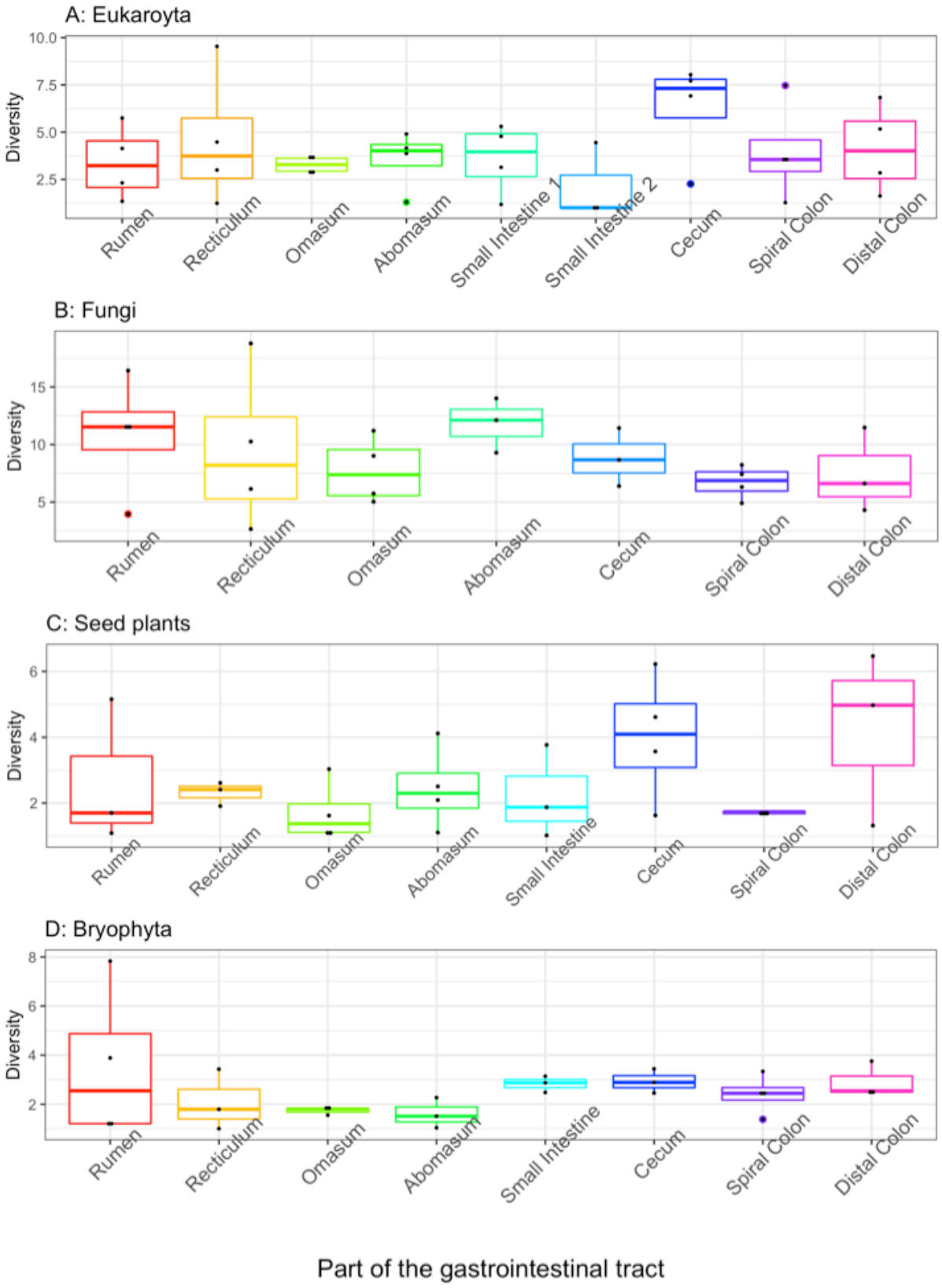

Figure B6 – Boxplots comparing the diversity (A: Eukaryota, B: Fungi, C: Seed plants (Spermatophyta), D: Bryophyta) along the gastrointestinal tract, when q=3. Black, filled circles represent samples, and thick line within the quartile of the boxes shows the median. X-axis = Gastrointestinal parts. Y-axis = Diversity.

Appendix C - Library preparation

Table C1 – Overview of pooling and purification of samples, with specific information on dates, volume and kit used. Abbreviations: GI tract = Gastrointestinal tract, FE = Feeding Experiment, P = PCR Plate.

Name pool	Date pooling	Total volume after pooling (µl)	Nb PCR replicates	Date PCR purification	Kit PCR purification
GI tract + Food (Sper01)	14-15.01.2019	1000	3	30.01.2019	QIAquick
GI tract (Bryo01)	14-15.01.2019	750	3	30.01.2019	QIAquick
GI tract + Food (Euka02)	14-15.01.2019	1000	3	30.01.2019	QIAquick
GI tract + Food (Fung01)	14-15.01.2019	1000	3	30.01.2019	QIAquick
FE_P1 (Sper01)	28.01.2019	1300	3	30.01.2019	QIAquick
FE_P2 (Sper01)	28.01.2019	1400	3	30.01.2019	QIAquick
FE_P3 (Sper01)	14-15.01.2019	1000	3	30.01.2019	QIAquick
FE_P1 (Euka02)	25.01.2019	1400	3	30.01.2019	QIAquick
FE_P2 (Euka02)	28.01.2019	1400	3	30.01.2019	QIAquick
FE_P3 (Euka02)	14-15.01.2019	900	3	30.01.2019	QIAquick
FE_P1 (Fung01)	25.01.2019	1300	3	30.01.2019	QIAquick
FE_P2 (Fung01)	28.01.2019	1300	3	30.01.2019	QIAquick
FE_P3 (Fung01)	25.01.2019	1000	3	30.01.2019	QIAquick

Table C2 – Overview of elution volume after purification and information on following gel electrophoresis. Abbreviations: GI tract = Gastrointestinal tract, FE = Feeding Experiment, P = PCR Plate.

Name pool	Elution volume after purification (µl)	Volume gel electrophoresis (1%) after PCR purification (µl)	Volume loading dye (µl)	Date gel electrophoresis
GI tract + Food (Sper01)	30	1	1	30.01.2019
GI tract (Bryo01)	30	1	1	30.01.2019
GI tract + Food (Euka02)	30	1	1	30.01.2019
GI tract + Food (Fung01)	30	1	1	30.01.2019
FE_P1 (Sper01)	30	1	1	30.01.2019
FE_P2 (Sper01)	30	1	1	30.01.2019
FE_P3 (Sper01)	30	1	1	30.01.2019
FE_P1 (Euka02)	30	1	1	30.01.2019
FE_P2 (Euka02)	30	1	1	30.01.2019
FE_P3 (Euka02)	30	1	1	30.01.2019
FE_P1 (Fung01)	30	1	1	30.01.2019
FE_P2 (Fung01)	30	1	1	30.01.2019
FE_P3 (Fung01)	30	1	1	30.01.2019

Table C3 – Results from Qubit measurement of the different pooled samples. Abbreviations: GI tract = Gastrointestinal tract, FE = Feeding Experiment, P = PCR Plate.

Name pool	Date	Type of measurement	Volume DNA (µl)	Measured DNA concentration (µg/ml)	Stock concentration (ng/µl)
GI tract + Food (Sper01)	31.01.2019	Broad Range	2	1.62	162
GI tract (Bryo01)	31.01.2019	Broad Range	2	0.421	42.1
GI tract + Food (Euka02)	31.01.2019	Broad Range	2	2.95	295
GI tract + Food (Fung01)	31.01.2019	Broad Range	2	3.61	361
FE_P1 (Sper01)	31.01.2019	Broad Range	2	1.16	116
FE_P2 (Sper01)	31.01.2019	Broad Range	2	1.16	116
FE_P3 (Sper01)	31.01.2019	Broad Range	2	1.10	110
FE_P1 (Euka02)	31.01.2019	Broad Range	2	2.82	282
FE_P2 (Euka02)	31.01.2019	Broad Range	2	3.47	347
FE_P3 (Euka02)	31.01.2019	Broad Range	2	2.30	230
FE_P1 (Fung01)	31.01.2019	Broad Range	2	4.01	401
FE_P2 (Fung01)	31.01.2019	Broad Range	2	3.93	393
FE_P3 (Fung01)	31.01.2019	Broad Range	2	3.60	360

Table C4 – Overview of final pooling, merging content in “Pool 1” with “Pool 2”. Final pool content is hereby referred to as “Library Process ID”. Abbreviations: GI tract = Gastrointestinal tract, FE = Feeding Experiment, P = PCR Plate.

Name Pool 1	Volume Pool 1 (µl)	Name Pool 2	Volume Pool 2 (µl)	Library Process ID
FE_P1 (Sper01)	No need for final pooling			Lib1
FE_P2 (Sper01)	No need for final pooling			Lib2
FE_P3 (Sper01)	No need for final pooling			Lib3
GI tract + Food (Sper01)	4	GI tract (Bryo01)	16	Lib5
FE_P1 (Euka02)	10	FE_P1 (Fung01)	10	Lib9
FE_P2 (Euka02)	10	FE_P2 (Fung01)	10	Lib10
FE_P3 (Euka02)	10	FE_P3 (Fung01)	10	Lib11
GI tract + Food (Euka02)	10	GI tract + Food (Fung01)	10	Lib12

Table C5 – Overview of pooled volume, Qubit measurement of content in libraries and volume sent for Illumina Hiseq sequencing.

Library Process ID	Volume final pool (µl)	Type of measurement	Volume DNA (µl)	Measured DNA concentration (µg/ml)	Stock concentration (ng/µl)	Volume final pool sent for sequencing (µl)
Lib1		Qubit (Broad range)	2	1.16	116	27
Lib2		Qubit (Broad range)	2	1.16	116	27
Lib3		Qubit (Broad range)	2	1.10	110	27
Lib5	20	Qubit (Broad range)	2	0.713	71.3	18
Lib9	20	Qubit (Broad range)	2	3.72	372	18
Lib10	20	Qubit (Broad range)	2	3.57	357	18
Lib11	20	Qubit (Broad range)	2	3.08	308	18
Lib12	20	Qubit (Broad range)	2	2.65	265	18

Table C6 – Final preparations before sequencing. General remark: Final qPCR during library prep run separately for GH/BRYO and EUKA/FUNGI pools. Desired sequencing depth per library corresponds to ~20.000 sequences per sample.

Library Process ID	Type of library preparation	Number of individually tagged samples	Desired sequencing depth per library	Volume final pool sent for sequencing (µl)
Lib1	TrueSeq PCR-free prep	288	5 760 000	27
Lib2	TrueSeq PCR-free prep	288	5 760 000	27
Lib3	TrueSeq PCR-free prep	192	3 840 000	27
Lib5	TrueSeq PCR-free prep	339	6 780 000	18
Lib9	TrueSeq PCR-free prep	576	11 520 000	18
Lib10	TrueSeq PCR-free prep	576	11 520 000	18
Lib11	TrueSeq PCR-free prep	384	7 680 000	18
Lib12	TrueSeq PCR-free prep	426	8 520 000	18