Australian National University

A thesis submitted for partial fulfilment of the degree

Masters of Biotechnology (Advanced)

Research School of Biology

The Australian National University

Uncovering the microbiome of sympatric European brown hares and European rabbits

Somasundhari Shanmuganandam

Supervised by

Dr Benjamin Schwessinger Dr Robyn Hall

May 2019

Word count: 9,926

Declaration of original work

I declare that the work presented in this thesis is my original work unless stated otherwise.

Acknowledgement

I would like to thank my supervisors Dr Benjamin Schwessinger and Dr Robyn Hall for their immense support and motivation throughout this project. Without their guidance and patience, completion of this thesis would not have been possible. Ben's insightful knowledge on various fields helped me learn new things each day and I am grateful to have learnt basic python from him. Robyn's cheery attitude and her vast knowledge about lagomorphs, had me inspired all through this year. I have gained immerse knowledge in multiple fields and received hands-on experience in sequencing strategies this year all thanks to them. I am also grateful that they chose me to work on this project they designed even though my background in bioinformatics was limited.

I am thankful to Dr Louis Ranjard for his support and suggestions. Though I knew him only for a short duration of time, his remarks on my work always helped me improve. I want to thank my examiners Dr David Gordon, Dr Jana Sperschneider and Dr Sasha Mikheyev who gave me valuable comments during the examiners meetings, which helped shape this project. I also want to thank Maria, Sam and Ash for reading a draft of my thesis and providing me with some valuable comments.

I would also like to thank Nina and Elena for making me feel welcomed at CSIRO. I am also thankful to Tegan for guiding me in different laboratory techniques during my first few days at CSIRO. I am greatly indebted to Yiheng who guided me, gave me valuable suggestions and helped me in designing a script in python all through his tight schedule. I am grateful to all the lab members both in CSIRO and ANU who provided such a fun and friendly environment to work in.

Lastly, I want to thank my amazing friends who were always available for me and pushed me to work harder. I want to specifically thank Akshai and Supria for taking time to read my thesis and for constantly making me laugh at tense situations. I also want to thank Prabhakar who helped me fix errors and understand different tools which I used in python even though biology was not his background. I am more grateful to have wonderful parents who are responsible for who I am today. Without their constant support and motivation, nothing would have been possible. I want to thank God for giving me a wonderful life which I am thankful for each day.

Table of Contents

List of Figures

List of Tables

Abbreviations

rRNA	Ribosomal ribonucleic acid
DNA	Deoxyribonucleic acid
ONT	Oxford Nanopore Technologies
PacBio	Pacific Biosystems
ACT	Australian Capital Territory
CSIRO	Commonwealth Scientific and Industrial Research Organisation
PCR	Polymerase Chain Reaction
NTC	No-template control
BLAST	Basic local alignment search tool
OTU	Operational taxonomic units
ASV	Amplicon sequence variant
P or p-value	Probability value
t-stat	Test statistic
Gbps	Giga base pairs
PUFA	Polyunsaturated fatty acids
PERMANOVA	Permutational multivariate analysis of variance
BMR	Basal metabolic rate
BMI	Body mass index

Abstract

Microbiome refers to a vast community of micro-organisms confined to a certain region either inside a host or in an environment. In Australia, both European brown hares and European rabbits are invasive species with latter being a major pest destroying native flora and even fauna. Despite their increasing population, no previous studies had worked on evaluating their microbial diversity, physiology and epidemiology in a natural sympatric environment. In this research, I aimed to study and compare the bacterial diversity and composition in faecal pellet of nine hares and twelve rabbits using two sequencing strategies, Illumina MiSeq and nanopore MinION. Diversity analysis revealed significant difference between the microbiome of hares and rabbits (p-value = 0.001; PERMANOVA pseudo-F test). Furthermore, all hare samples had similar bacterial diversity and composition whereas rabbit samples had varying level of similarity/dissimilarity between each other. This level of variation did not appear to be correlated with age, sex, lactation and month of sample collection (season) in this study. Even with these differences, the dominant phyla were *Firmicutes* (mean relative frequency of 69 ± 5% in hares and 83 ± 9% in rabbits) and *Bacteroidetes* (23 ± 5% in hares and 6.65 ± 4.8% in rabbits) in both populations. However differences were observed in phylum *Verrcomicrobia* which was present only in rabbit populations while phyla *Lentisphaerae* and *Synergistetes* were present only in hare populations. This difference in bacterial diversity and composition could be associated with diet, behaviour and health. This study is the first to compare microbiome of sympatric hares and rabbits in Australia and hence will provide a foundation for future studies that aim to understand diet and lifestyle associated microbiome patterns in these lagomorphs.

Chapter 1: Introduction

Humans, animals and most of the eukaryotes depend on a vast community of micro-organisms to protect them against pathogens and to fulfill important metabolic pathways that are required for proper body functioning (Ana Popovic and Parkinson, 2018). This vast community of micro-organisms confined to a certain niche are referred to as the microbiome. The microbiome of the host is believed to share a symbiotic relationship with the body where low diversity and imbalance could potentially lead to several diseases (Haque and Haque, 2017; Valdes et al., 2018). Since each part of the body perform a different function, the microbiome associated also varies throughout the body (Conlon and Bird, 2014). Gut microbiomes of most animals and some insects are composed of trillions of beneficial bacteria, predominantly of phyla *Bacteroidetes* and *Firmicutes,* which aid in breaking down complex molecules and also help maintain immunity (Bull and Plummer, 2014). However, the relative frequency or abundance of each phylum present within the gut varies among different hosts as it is largely influenced by the host's diet (Singh et al., 2017). For example, diets rich in animal proteins result in increase of phyla *Bacteroidetes* and *Proteobacteria* while *Firmicutes* decrease in abundance (David et al., 2013). Skin microbiomes, on the other hand, are studied extensively in recent times to develop treatment for skin cancer (Nakatsuji et al., 2018). Faecal microbiomes not only give an estimate of the gut microbiome but also helps in understanding the diet and lifestyle pattern of the host (Wu et al., 2016; Valdes et al., 2018). Though faecal microbiomes are widely studied in several animal species, some species are comparatively less focused on. Hence in this study, I focused on estimating the microbiome of hares and rabbits.

1.1 Microbiome of European brown hares and European rabbits

Over the years, few studies have focused on identifying and analysing the gut or faecal microbiome of different breeds of European rabbits such as domestic New Zealand white rabbits, caldes and Rex rabbits (Eshar and Weese, 2014; Arrazuria et al., 2018; O' Donnell et al., 2017; Velasco-Galilea et al., 2018; Zeng et al., 2015). However, only one study focused on estimating the microbiome of wild European brown hares (Stalder et al., 2019). The predominant bacteria in European rabbits observed were of phyla *Firmicutes, Tenericutes* and *Bacteroidetes* while the predominant bacteria in European brown hares were of phyla *Firmicutes, Bacteroidetes* and *Spirochaetes*. Phyla *Firmicutes* and *Bacteroidetes*

are the dominant bacteria present in both hares and rabbits, making them important to study. Bacteria in the phylum *Firmicutes* are involved in metabolizing dietary plant polysaccharides (David et al., 2013); since rabbits and hares are herbivorous they have increased abundance of *Firmicutes* in their faecal microbiome. Bacteria in the phylum *Bacteroidetes* are involved in fermentation of indigestible polysaccharides to generate energy (Johnson et al., 2017). Our understanding of wild hare and rabbit microbiomes in a sympatric environment is limited since no previous study had focused on performing a comparative analysis. In a sympatric environment, both hares and rabbits are likely to be exposed to the same food source and abiotic factors which may influence their microbiome in a similar way and hence more accurate comparison can be performed (Bahrndorff et al., 2016; Conlon and Bird, 2014). Australia is home to only one species of rabbit, the European rabbit (*Oryctolagus cuniculus*) and one species of hare, the European brown hare (*Lepus europaeus*), both of which were introduced during the year 1788 and 1862 respectively (Myers, 1970; Stott, 2003). Rabbits quickly grew in numbers, reaching population of 200 million by the year 2017, colonising every territory around Australia except for northern regions (All that's interesting, 2016; Invasive species fact sheet, 2011). They consume seedlings of several exotic flora resulting in destruction of plants; The family *Casuarinaceae* is a group of Australian native trees that are at risk due to rabbit grazing (Bird et al., 2012). Populations of native animals, such as the pig-footed bandicoot and the greater bilby are observed to be greatly reduced since they compete for similar food source as rabbits (All that's interesting, 2016). Hares are comparatively lower in number as it took them a relatively longer time to colonise and are therefore considered only a minor pest (Stott, 2003). The difference in the colonising pattern of hares and rabbits could be associated with difference in their behaviour and dietary habits. Understanding their microbiome may facilitate in developing better control strategies than the ones currently available.

1.2 Common differences between hares and rabbits in terms of diet and lifestyle

Though hares and rabbits belong to the same family *Leporidae*, they differ from each other on the basis of lifestyle, behaviour and diet. Rabbits prefer to remain in groups of five to sometimes even 20 and live in an extensive burrow system while hares usually prefer to live alone above ground and only interact with other hares during mating and grazing (Cowan, 1987, Monaghan and Metcalfe, 1985). Rabbits also produce more offspring than hares each year (Gibb, 1990). Another difference between them is that hares use their long hind legs to

run away from predators while rabbits hide in burrows (Geggel, 2016). These behavioural differences could impact their dietary pattern as hares may require higher energy for thermoregulation than rabbits to survive above ground (Schai-Braun et al., 2015). In terms of diet, rabbits prefer grasses, soft stems and vegetable plant rich in carbohydrates (Zhu et al., 2015; Green et al., 2017) while hares prefer arable crops and consume plants rich in crude fat especially during winter (Schai-Braun et al., 2015; Reichlin et al., 2006). This diet regime varies based on season (Reichlin et al., 2006). Pregnant rabbits/hares may require comparatively more energy in the form of crude fat which is either obtained through diet or from metabolising the fat stored in body fat reserves (Gidenne, 2010; Hackländer et al., 2002). A major difference is in the size of caecum, a fermenter organ rich in bacteria associated with breaking down of complex food substances; European brown hares have a smaller caecum than European rabbits and gut passage rate is comparatively faster in hares (Stott, 2008). Both hares and rabbits are hindgut fermenters; they have a simple single chambered stomach and thus find it hard to digest crude fibre (Hernández-Martínez et al., 2017). Hence, indigestible substances are excreted while only fine food particles are retained in the caecum for fermentation (Kuijper et al., 2004). This strategy results in the production of two types of faeces; a hard faeces high in fibres and soft faeces high in proteins (Pairet et al., 1986). The soft faeces released from the caecum are consumed immediately from their anus while the hard faeces are excreted out (Kuijper et al., 2004). This process of consuming their own droppings is termed as coprophagy (Bugle and Rubin, 1993). Soft faeces of rabbits have tough outer membrane which allows faeces to be retained in the stomach for long microbial digestion while hares produce soft faeces without any intact membrane, thus dissipating almost immediately inside the stomach (Kuijper et al., 2004). Though there is only a minor difference between microbiomes of hard faeces and soft faeces (Zeng et al., 2015), coprophagy may still have an influence on faecal microbiomes of hares and rabbits. Without coprophagy, sufficient nutrient requirement for proper functioning of the body will not be met, hence will result in malnutrition which in turn has a negative effect on the microbiome (Hirakawa, 2001; Subramanian et al., 2014). These behavioural and dietary differences between hares and rabbits may result in considerable difference in their microbiomes as well, thus making it important to study and compare.

1.3 Second and third generation sequencing to study the microbiome

Since not all species of bacteria can be cultured easily due to their varying growth preferences, culture-independent techniques like next generation sequencing and long read sequencing are regularly used to unravel the complex microbiome of several species across animal kingdom (Simon Bahrndorff, 2016).

1.3.1 Next-generation sequencing: Illumina MiSeq sequencing system

Next generation (NGS) or second generation sequencing technologies such as Illumina, Roche 454, Ion Torrent and SOLiD are widely used to generate enormous amount of data cheaply (Metzker, 2009). These systems produce high throughput results in a shorter time as opposed to first-generation sequencing technique like Sanger sequencing (Alekseyev et al., 2018). Although NGS came with several advantages, read length limited its application; maximum being 600 bp generated by Ion Torrent (Ion S5/Ion S5 Plus/Ion S5 Prime) and Illumina MiSeq paired end 2x300 bp V3 chemistry (Speranskaya et al., 2018). The shorter read lengths are a consequence of limitation in technology, thus compelling the users to target only selected region of a gene for sequencing (Pollard et al., 2018). Currently two NGS approaches are widely being used, Illumina sequencing platform and Ion torrent platform, each having advantages and disadvantages. Though the cost of both instruments are similar, Ion Torrent has shorter run time (~ 2 hours) and offers automated library preparation approach reducing technician errors (Bahassi and Stambrook, 2014; Alekseyev et al., 2018). Though the run time for Illumina sequencing platform is longer (~56 hours), sequence yield per run is more in Illumina than Ion Torrent (Speranskaya et al., 2018). Hence in this study, I chose Illumina MiSeq reagent kit V3 sequencer to obtain more yield at long read length (600 bp) over Ion torrent and Illumina HiSeq sequencer. Although Illumina HiSeq platform produces more yield at comparatively low error rate, it has few limitations such as high cost, long run time (~ a week) and produces short reads (500 bp) (Besser et al., 2018). Working of Illumina sequencer is given in Supplementary Figure 2.

1.3.2 Long-read sequencing: Oxford's Nanopore Technology's MinION

Third generation sequencing or long read sequencing like Pacific Biosystems (PacBio) and Oxford Nanopore Techniques (ONT) have similar or lower cost (in case of multiplexing) per sample compared to second generation sequencing technologies (Wick et al., 2017). ONT's MinION sequencer produce reads with an average length over 6-8 kb and PacBio Sequel produce reads over 10–14 kb in length but both with an error rate of 10-20% (Bleidorn, 2016, Zhang et al., 2019, Ardui et al., 2018). The error rate is low in PacBio compared to ONT but ONT produces more yield due to its ability to sequence multiple molecules (Weirather et al., 2017). Due to lower cost and higher yield, I used ONT's MinION sequencing platform in my study along with Illumina MiSeq platform. The working of ONT's MinION is described in Supplementary Figure 3.

1.3.3 16S rRNA gene for microbiome analysis

Either conserved loci or whole genome metagenomics of host can be sequenced to study their microbiome. Generally, microbial studies are focused on 16S rRNA gene since it contains both hypervariable and highly conserved regions that are found in most bacterial and archaeal species (Janda and Abbott, 2007). This ~1500 bp long gene is composed of nine hypervariable regions that provide species-specific sequences targeted for scientific investigations (Clarridge, 2004). Each hypervariable region has varying efficiency in distinguishing between different bacterial species (Chakravorty et al., 2007). Due to the presence of conserved regions flanking either end of hypervariable regions in the 16S rRNA gene, it is possible to use a pair of universal primer to target most of the bacterial species present. The availability of a large database (such as SILVA 16S database) made 16S amplicon sequencing analysis easier (Ranjan et al., 2016). Although 16S rRNA is restricted to identifying only bacterial species and not fungi or viral species, it has an advantage over whole genome sequencing which fails to identify rare taxa because of low sequence depth and inadequately curated databases (Stefanini and Cavalieri, 2018). Hence in this study, I sequenced V3-V4 region of 16S rRNA and the entire 16S rRNA gene using Illumina MiSeq and Nanopore MinION sequencer, respectively.

1.4 Significance

This thesis will make number of significant contributions to the field of microbiome analysis in animals using popular sequencing strategies currently available. It will increase our understanding of the microbiome in European brown hares and European rabbits, while also pointing out the key features both positive and negative of each sequencing pipelines used. Since no previous studies had focused on comparing microbiome of sympatric European rabbits and European brown hares in Australia, it is possible to hypothesise that hares and rabbits could be reservoir to various harmful pathogens that can cause diseases like tularaemia and brucellosis which affects livestock and even humans (Ohara et al., 1974). Though these diseases are uncommon in Australia, hares and rabbits are known to be the reservoirs of these pathogens in Europe and other countries (Ditchfield et al., 1960). Additionally, it is also possible to gain better insight on how diet and/or lifestyle regulates the microbiome and the type of diet best preferred by these animals in a natural sympatric environment.

1.5 Aim

The aim of this study can be divided into two sections. Firstly, I aimed to study and compare the faecal microbiome of hares and rabbits using both Illumina MiSeq and Nanopore MinION 16S sequencing pipelines. Secondly, I focused on comparing the data obtained between Illumina and Nanopore platform while also evaluating the pros and cons of these sequencing platforms.

Chapter 2: Materials and methods

2.1 General methods

2.1.1 Sample Collection

The hares and rabbits used in this study were shot as part of routine vertebrate pest control operations in Mulligan's Flat region, ACT, between the months of January and September, 2016. During the same year, hares and rabbits were eradicated from this region. Animals were shot from a vehicle using a 0.22-caliber rifle targeting the head or chest. The faecal pellets from the descending colon were then collected during post-mortem and stored at -20°C. All the sampling was conducted according to the Australian Code for the Care and Use of Animals for Scientific Purposes as approved by the CSIRO Wildlife and Large Animal ethics committee (approvals #12-15 and #16-02). In this study, I used faecal pellet samples of nine hares and twelve rabbits to study their microbiomes. Details of the hare and rabbit faecal tissue samples used in this study are given in Table 1.

Table 1: Hare and Rabbit sample information[1]

Sample	Sex	Month of sample collection	Weight	Lactating
Hare-1	NA	January	NA	No
Hare-2	F	May	3100	No
Hare-3	M	May	3700	No
Hare-4	M	June	3040	No
Hare-5	M	June	2920	No
Hare-6	F	June	3460	No
Hare-7	F	July	3250	No
Hare-8	F	July	3350	No
Hare-9	F	September	4900	Pregnant and lactating
Rabbit-1	F	February	NA	Pregnant and lactating
Rabbit-2	M	March	1080	No
Rabbit-3	F	March	960	No
Rabbit-4	F	March	1380	Yes
Rabbit-5	F	March	1440	Yes
Rabbit-6	NA kitten	March	400	No
Rabbit-7	M	March	1400	No
Rabbit-8	F	March	1500	Pregnant and lactating
Rabbit-9	F	April	NA	Yes
Rabbit-10	F	April	NA	No
Rabbit-11	M	June	1500	No
Rabbit-12	M	June	1400	No

2.1.2 DNA extraction and Quality Control

I had performed all steps for DNA extraction at room temperature (RT~20°C) unless stated otherwise.

I had performed DNA extraction of nine hare and twelve rabbit faecal pellet samples following the protocol from DNeasy Blood and Tissue Kit (Qiagen). I optimized the original protocol to

[1] NA refers to sample with unavailable information

improve DNA yield by doubling the initial tissue input (~50 mg) while keeping the volume of buffer ATL (Qiagen) and Proteinase K (Qiagen) constant at 324 µl and 36 µl, respectively. This increased the concentration of the extracted DNA considerably (from < 0.1 ng/µl to ~2.3 ng/µl). I also included RNase treatment step because the previously extracted DNA had high RNA/protein contamination (A260/A280 > 2). For this, I added 4 µl of RNase (Qiagen) and incubated for 5 min at RT prior to addition of buffer AL (Qiagen). This stage brought down the A260/A280 ratio to ~1.8. During DNA extraction from the faecal samples, I included reagent-only controls, one for hare and one for rabbit samples, to check for contamination present within the reagents used. I had split the eluted DNA into two aliquots to minimize repeated freeze-thaw cycle. I stored one set of aliquot (half of the total volume of eluted DNA) at -20°C freezer and the other at +4°C fridge for each sample. I used 5 µl of +4°C aliquot of each sample to run 0.8% agarose gel electrophoresis for 40 min at 100V. I also used the same aliquots for DNA quantification using both Qubit 2.0 fluorometer broad range assay kit (Invitrogen) and Nanodrop (Thermofisher). The DNA concentration, A260/A280 and A260/A230 of each sample were recorded and compared to estimate DNA quality and to ensure the sample is free from RNA/protein contamination.

2.2 Illumina MiSeq 16S sequencing

I had performed all stages of library preparation at room temperature (RT~20°C) unless stated otherwise using -20°C aliquots of DNA.

2.2.1 Library Preparation

I targeted the V3-V4 (~460 bp) region of the 16S rRNA gene for library preparation since it is one of the best studied regions with the most extensive reference database such as SILVA.

Prior to PCR, I diluted all the samples to 2.3 ng (the sample with least amount as measured by Qubit) using buffer EB (Qiagen; 10 mM Tris-Cl, pH 8.5). The primers used to target the V3-V4 region are given in Table 2.

Table 2: Illumina 16S V3-V4 region specific primer with highlighted overhang sequence. The bold sequence denotes the overhang sequences which is where Illumina index barcodes attach. Used for multiplexing purposes.

Primers (Klindworth et al., 2013)	**Sequence 5'-3'**
S-D-Bact-0341-b-S-17 Forward Primer	5'-**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**CCTACGGGNGGCWGCAG -3'
S-D-Bact-0785-a-A-21 Reverse Primer	5'-**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**GACTACHVGGGTATCTAATCC -3'

I performed PCR amplification of the V3-V4 region following the protocol given in the sample preparation guide from Illumina (Illumina, 2013). I replaced 2x KAPA HiFi HotStart ReadyMix with Platinum SuperFi PCR Master Mix (Invitrogen™) due to latter's low error rate (Hallmaier-Wacker et al., 2018). For 25 µl PCR reaction, I added 2X Platinum SuperFi PCR Master Mix, 0.5 µM of each forward and reverse primer, 4.6 ng of template DNA and made up to final volume using nuclease free water. I included a negative PCR no-template control (NTC) to verify that the product obtained was free from contamination. Unlike reagent-only controls which contains reagents used for DNA extraction, no-template PCR control comprises only of nuclease free water as template with PCR reagents in it. I took equal volume of control and sample DNA for amplification in PCR. The PCR conditions used were as follows: initial denaturation at 98°C for 30 sec, followed by 25 cycles of 98°C (10 sec), 55°C (15 sec), and 72°C (30 sec) with a final extension of 10 min at 72°C. I ran 5 µl of PCR product on 1% agarose gel for 30 min at 100V to observe the band size. I performed cleanup of PCR product (stage 1) using AMPure XP beads at 1x ratio and eluted in buffer EB to obtain final volume of 50 µl following the protocol from Illumina (Illumina, 2013). I ran the samples on 1% agarose gel for 30 min at 100V to check whether the cleaned up DNA was free from primers. Next, I used the Nextera XT index kit (N7XX and S5XX, Illumina) for PCR based barcoding of all samples following the protocol from Illumina (Illumina, 2013) . I took 5 µl of template DNA with 0.2 µM each of forward and reverse primer and made it up to 50 µl using nuclease free water to get a final PCR volume of 50 µl. I used the same PCR conditions as above but limited the number of cycles to eight instead. I ran 5 µl of the obtained PCR product on a 1.2% agarose gel for 20 min at 100V to visualize the size of product. I then cleaned up the PCR product (stage 2) using

AMPure XP beads at 1x ratio and eluted in buffer EB following the same protocol (Illumina, 2013). I ran both stage 1 and stage 2 cleaned up DNA product on 1.2% agarose gel for 20 min at 100V alongside each other to check whether the product size was lengthened due to addition of barcodes. I performed quantification of barcoded DNA using Qubit high sensitivity assay and I diluted a random selection of eight samples to a concentration of 0.9 ng/µl to run on a tapestation (D1000 high sensitivity; Agilent technologies).

Based on the observed results, I pooled all the samples together in equimolar concentration to get a final concentration of 4 nM (0.004 pmol/uL) which was the suggested library input concentration for MiSeq system. Additionally, I quantified the pooled DNA on qubit HS assay to ensure 4 nM. The next few steps were followed from the same protocol (Illumina, 2013). I brought down the concentration of both denatured DNA and PhiX (control) to 9.5 pM based on the recommended concentration. I replaced 10% of the total volume of DNA library with PhiX control and then loaded the sample into MiSeq 2 x 300 bp paired end V3 reagent cartridge.

2.2.2 QIIME 2 for data analysis

MiSeq run produced raw demultiplexed and adapter trimmed fastq reads which I analyzed using open-source QIIME 2-2019.1 software (Bolyen et al., 2018). I first viewed the raw demultiplexed sequences using q2-demux plugin. I then filtered the raw data using the DADA2 pipeline (Callahan et al., 2016) (via q2-dada2) in QIIME 2. DADA2 performs multiple functions such as filtering out noisy reads, correcting errors in marginal sequences, removing chimeric sequences, removing singletons, joining denoised paired-end reads and then dereplicating these sequences (Callahan et al., 2016). For DADA2, I trimmed the first 7 bases in both forward and reverse reads and truncated the sequence to a total length of 298 bp and 256 bp for forward and reverse reads, respectively. This length was chosen based on the demux graph to allow sufficient overlapping (~50 bp) of both reads. The data received from DADA2 were in form of amplicon sequence variants (ASV) which was shown to outperform operational taxonomical unit (OTU) as it can go down to single-nucleotide differences over the sequenced gene region at better resolution (Callahan et al., 2017). Using rarefaction, I chose a subsampling depth of 35,766 per sample, which retained more than 50% of the species while eliminating only two controls, no-template control and one reagent-only control (rabbit), both of which had low number of sequences. The rarified dataset was then used in analyzing alpha and beta

diversity. Alpha diversity measures the diversity of bacterial communities within a sample while beta diversity measures the bacterial diversity between different samples.

I used three alpha diversity metrics, Faith's phylogenetic diversity (Faith, 1992), Shannon index and observed ASVs (calculated using Kruskal-Wallis test (Kruskal, 1952)), each of which computed for phylogenetic diversity, species richness/evenness and number of ASVs observed, respectively per sample. For beta diversity analysis, I performed weighted UniFrac (Lozupone et al., 2007) and bray-curtis dissimilarity test. All of which were constructed using the q2-diversity plugin available in QIIME 2. I used beta diversity group significance pipeline to perform permutation-based statistical analysis.

For taxonomic analysis, I assigned the taxonomy to ASVs using the q2-feature-classifier VSEARCH, a global alignment tool, to align the query sequence against the SILVA 16S database (April, 2018) at 99% sequence identity (Bokulich et al., 2018; Rognes et al., 2016; Quast et al., 2013). I then used the resulting output to construct taxa bar plots and heat maps to clearly show the bacterial species identified using QIIME 2 plugins (q2-taxa). I initially tried using Greengenes database (DeSantis et al., 2006), but switched to SILVA since the former was last updated in the year 2013.

For statistical analysis of taxonomic results, I used student's t-test which estimates statistical difference between means of two groups. The csv table generated from the taxonomic bar plot in QIIME 2 was first downloaded. This contains the information about the abundance of various taxa in each sample. With this data, I performed F-test in Microsoft excel. F-test estimates whether the variance between two sample groups are even or uneven. I then performed either t-test assuming unequal variance or t-test assuming equal variance based on the observation. I set the null hypothesis value to be zero, assuming no significant difference between two groups. If the observed p-value (probability value) was less than 0.05, I rejected the null hypothesis and stated significant difference was observed between the groups.

2.3 Nanopore MinION 16S sequencing

I had performed all steps for library preparation at room temperature (RT~20°C) unless stated otherwise using aliquots of DNA that were stored at -20°C freezer.

2.3.1 Library preparation

For library preparation, I had targeted the entire 1.5 kbp region of the 16S rRNA gene. The primers used at this stage are given in Table 3.

Table 3: Nanopore 16S Primers for entire 1.5 kbp 16S rRNA region

Primers	**Sequence (5'-3')**
27F	AGAGTTTGATCMTGGCTCAG
1492R	CGGTTACCTTGTTACGACTT

I observed that by choosing an input PCR volume of 100 µl per sample, the concentration of the product obtained after PCR was greatly increased compared to choosing either 25 µl or 50 µl input PCR volume. For a final volume of 100 µl, I added 2X Platinum SuperFi PCR Master Mix (Invitrogen), 0.4 µM each of forward and reverse primer, 5 µl of template DNA (2.3 ng/µl) and made up to final volume using nuclease free water. PCR conditions I used were as follows, initial denaturation at 98°C for 30 sec, followed by 28 cycles of 98°C (10 sec), 55°C (15 sec) and 72°C (40 sec) with final extension at 72°C for 5 min.

I ran 5 µl PCR product on 1% agarose gel for 30 min at 100V to confirm the band size of ~1.5 kbp. I purified the PCR products using AMPure XP beads at 1x ratio and eluted in 15 µl nuclease free water following the protocol (Hu and Schwessinger, 2018). I ran 3 µl purified eluted DNA on 1% agarose gel for 30 min at 100V to ensure no primers were present. I quantified the samples using Qubit high sensitive assay (1 µl) (Invitrogen) and Nanodrop (Thermofisher). For library preparation, I used the Ligation sequencing kit 1D (SQK-LSK108) in combination with native barcoding kit 1D (EXP-NBD103) (ONT). I prepared the samples following a previous protocol (Hu and Schwessinger, 2018) with slight modifications. For twelve samples (five of hare and seven of rabbit), input amount of DNA taken for end

preparation step was 500 ng per sample and I recovered more than 75% of the input as measured by Qubit high sensitivity assay. For the addition of barcodes, I took 80 ng of end prepped DNA per sample as input in 9 µl of nuclease free water and I diluted the barcoded DNA obtained in 12 µl of nuclease free water. Approximately 75% of the input was recovered at this stage as measured by Qubit. I then took equimolar amounts of each barcoded sample (36.4 ng) and pooled them together to generate a total amount of 437 ng. Since the total volume exceeded 50 µl, I reduced the volume by performing a 2.5x AMPure bead cleanup. I performed adapter ligation on pooled barcoded DNA following the protocol (Hu and Schwessinger, 2018). Prior to loading the sample, I performed QC of the MinION flow cell and number of active pores observed was around 1300. I then prepared the priming mix and DNA library and added it into the MinION R.9.4.1 flow cell following the same protocol (Hu and Schwessinger, 2018). I processed the remaining samples (four of hare and five of rabbit) in the same way to generate a final library amount of ~200 ng to obtain maximum pore occupancy. I did not include any reagent controls or no template control (NTC) for MinION sequencing since the initial concentration of controls were too low to process.

2.3.2 16S data analysis

The raw reads obtained from sequencing run were in fast5 format. I first demultiplexed the reads using deepbinner (Wick et al., 2018). I ran the demultiplexed fast5 reads in Guppy 2.3.7 (ONT) to perform basecalling, adapter trimming and conversion into fastq format. I checked the quality of fastq reads obtained using pycoQC (Leger and Leonardi, 2019). I then analyzed the reads in EPI2ME desktop agent using the 16S analysis workflow at a threshold value of 7. This workflow aligned the reads against the NCBI 16S database using BLAST resulting in the generation of two csv files per sample. One file contained information about reads passing the threshold quality and the other contained taxid information. I extracted taxonomic ranks from these taxids using ete3 module (Huerta-Cepas et al., 2016) and added to create a table. I built another table by merging the reads from all samples together into a single dataframe based on taxid/runid. Then I filtered these reads based on quality score (greater than 7) and sequence length (greater than 1400bp) using pandas v0.24.2, a python data analysis library (Perez and Granger, 2007). These two tables (dataframe) constructed were in Biological Observation Matrix (BIOM) format that was compatible with QIIME 2. I wrote all python scripts in Jupyter Notebook interface (McKinney, 2010) and then stored it in github repository

(https://github.com/SomaAnand/Master_thesis_ANU). I then imported these two tables into QIIME 2 for constructing bar plots and performing diversity analysis using the same pipelines I used for Illumina data.

I performed the statistical analysis of taxonomic results using the same test I used in analyzing Illumina MiSeq data.

Chapter 3: Results

3.1 Illumina MiSeq 16S sequencing results

3.1.1 Raw data produced from Illumina MiSeq sequencing pipeline

After sequencing on MiSeq platform (Illumina), I obtained an output of approximately 16.7 million reads across 24 samples and out of these, ~14.56 million reads (84.5% of the total) had a quality score greater than 30 (probability of incorrect base call is 1 in 1000 bp). Out of the quality filtered reads, only ~11 million reads were assigned to the barcodes. Cluster density was found to be 619 ± 23 K/mm^2 with a total yield (number of bases sequenced) of ~8.94 Gbp. After filtering and merging reads using DADA2, I acquired the following read statistics as shown in Table 4.

Table 4: Summary of Illumina MiSeq reads for 24 samples after filtering, denoising, merging and removing chimeras using DADA2 pipeline

Samples	Input reads	Reads after filtering and denoising	Reads after merging	Reads after removing chimeras
Hare-1	519,092	348731 (67%)	345725 (67%)	73094 (14%)
Hare-2	559,301	379100 (68%)	344293 (62%)	58271 (10%)
Hare-3	569,441	393267 (69%)	351765 (62%)	60413 (11%)
Hare-4	543,060	342154 (63%)	316256 (58%)	64186 (12%)
Hare-5	622,569	427761 (69%)	385270 (62%)	64187 (10%)
Hare-6	419,506	280210 (67%)	249004 (59%)	46487 (11%)
Hare-7	511,631	339603 (66%)	307831 (60%)	53240 (10%)
Hare-8	514,414	332739 (65%)	317316 (62%)	65425 (13%)
Hare-9	571,837	388060 (68%)	365750 (64%)	67697 (12%)

Samples	Input reads	Reads after filtering and denoising	Reads after merging	Reads after removing chimeras
Hare-control	265,118	167260 (63%)	166216 (63%)	44455 (17%)
Rabbit-1	462,357	316038 (68%)	285188 (62%)	66318 (14%)
Rabbit-2	291,996	197374 (68%)	167564 (57%)	35766 (12%)
Rabbit-3	453,056	300140 (66%)	293078 (65%)	72228 (16%)
Rabbit-4	421,016	285040 (68%)	271624 (65%)	59718 (14%)
Rabbit-5	419,297	293973 (70%)	272399 (65%)	61557 (15%)
Rabbit-6	530,705	352939 (67%)	338367 (64%)	52497 (10%)
Rabbit-7	508,406	359827 (71%)	338520 (67%)	71963 (14%)
Rabbit-8	554,885	390559 (70%)	371926 (67%)	70698 (13%)
Rabbit-9	505,058	350984 (69%)	328058 (65%)	63806 (13%)
Rabbit-10	479,822	325289 (68%)	295563 (62%)	61549 (13%)
Rabbit-11	509,936	363206 (71%)	349930 (69%)	113310 (22%)
Rabbit-12	572,686	390660 (68%)	339046 (59%)	77282 (13%)
Rabbit-control	6,279	3426 (55%)	3303 (53%)	2327 (37%)
Stage1-NTC	2,158	466 (22%)	437 (20%)	228 (11%)

A total of 10,457 amplicon sequence variants (ASV) were generated with total sequence count (reads) of 1,406,702 after DADA2. From Table 4, the stage 1-NTC control had the lowest read count of 228 and rabbit-11 had the highest read count of 113,310. The overall mean read count was 58,612.58. The mean number of reads per ASV was 134.5, with the minimum being 1 and maximum frequency being 22,157. The merged average read length produced were 430.98 bp, with minimum length of 291 bp and maximum of 471 bp at a standard deviation of 9.96 bp.

3.1.2 Alpha and Beta diversity analysis of hare and rabbit microbiome

To test for phylogenetic diversity within each hare and rabbit sample, I used an alpha diversity metric called Faith's phylogenetic diversity (Faith's pd). Faith's pd measures biodiversity that incorporates phylogenetic difference within a sample. Using this metric, I observed that bacterial diversity in hares was significantly different from rabbits (krushkal-wallis test; sample size: 21; p-value = 0.01; H value = 5.83) as shown in Figure 1.

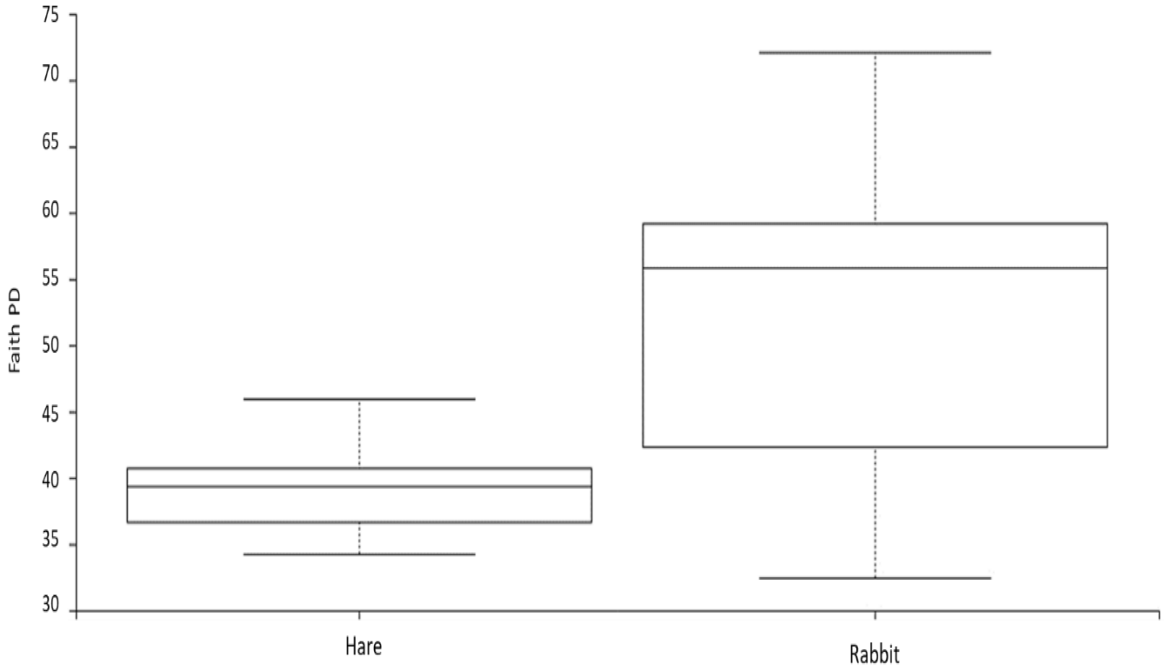


Figure 1: Variation in distribution of values (y-axis) indicate significant difference in phylogenetic diversity between hare and rabbit populations (Kruskal-Wallis test; p-value = 0.01; H = 5.83). Evaluated using faith's phylogenetic diversity metric, an alpha diversity analysis, which measures bacterial diversity within each sample. Small box indicates similar diversity between all hare samples and large box indicates difference in diversity between each rabbit sample.

In Figure 1, the median (horizontal line inside the box) of rabbit was aligned outside the hare box which indicate difference in diversity between hares and rabbits. Small box size of hare indicate that all hare samples had similar diversity. The rabbit box was observed to be taller indicating that each rabbit sample had varying degree of similarity and/or dissimilarity in terms of diversity.

To measure bacterial species richness and evenness (based on ASV) within a sample, I used Shannon index shown in Figure 2.

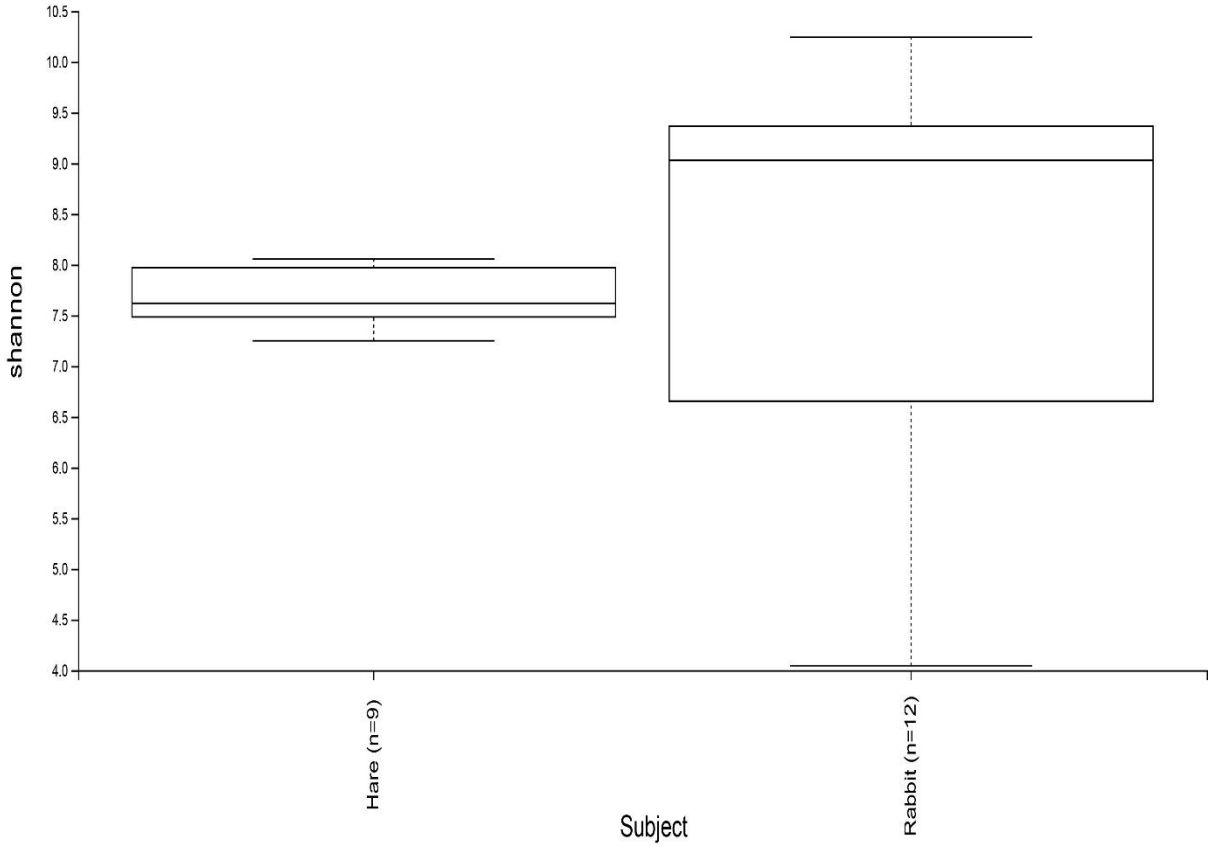


Figure 2: No variation in bacterial species richness and evenness between hare and rabbit populations (Kruskal-Wallis test; sample size: 21; H = 1.13; p-value = 0.28). Constructed using alpha diversity metric, Shannon index, which measures species richness and abundance. Hare box is observed to be aligned next to the rabbit box indicating no difference in richness/evenness.

Figure 2 shows hare box to be aligned next to the rabbit box having similar shannon index values (y-axis), this indicated no significant difference in terms of bacterial richness/evenness between the two populations (Kruskal-Wallis test; sample size: 21; H = 1.13; p-value = 0.28). However, taller rabbit box indicate variation in richness/evenness between rabbit samples.

I used another alpha diversity metric called observed ASVs to estimate total ASVs associated with each sample as shown in Table 5.

Table 5: Number of ASVs associated with each sample as measured by an alpha diversity metric, observed ASVs

Samples	Observed ASVs (p-value = 0.06)
Hare-1	682
Hare-2	745
Hare-3	733
Hare-4	712
Hare-5	820
Hare-6	651
Hare-7	679
Hare-8	635
Hare-9	689
Rabbit-1	1345
Rabbit-2	1026
Rabbit-3	939
Rabbit-4	1216
Rabbit-5	1264
Rabbit-6	465
Rabbit-7	631
Rabbit-8	806
Rabbit-9	1136
Rabbit-10	1226
Rabbit-11	444
Rabbit-12	1845

Table 5 indicates no significant difference between hare and rabbit samples in terms of observed ASVs (Kruskal-Wallis test; sample size: 21; H = 3.41; p-value = 0.06). However, variation in observed ASVs was observed between different rabbit samples.

Beta-diversity analysis measures difference in bacterial communities between different samples. I used a beta-diversity metric called weighted Unifrac distance metric which computes samples based on their bacterial abundance and phylogeny as shown in Figure 3.

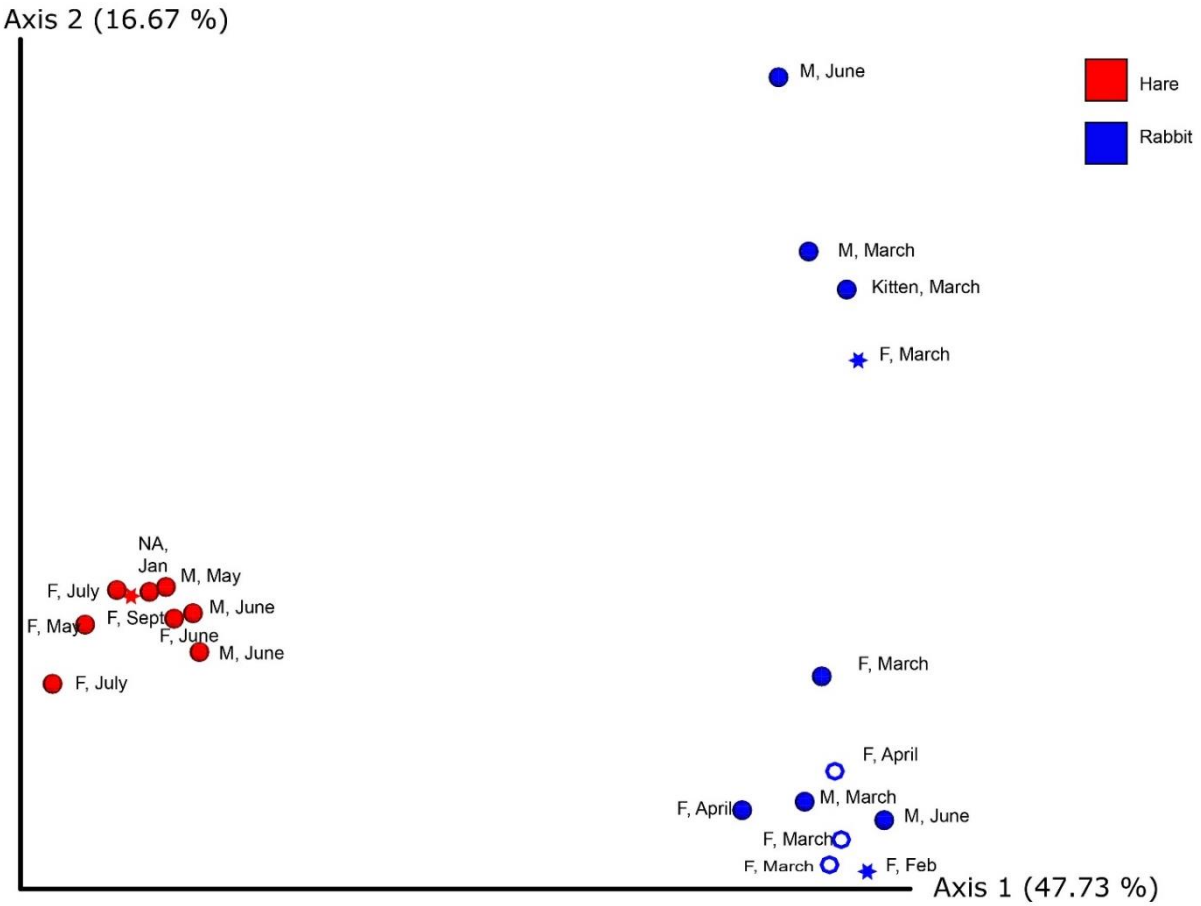


Figure 3: Segregation of hare samples from rabbit samples indicate significant difference in terms of bacterial phylogenetic diversity and abundance between two populations (PERMANOVA pseudo F test; sample size: 21; t-stat = 15.95; p-value = 0.001). This figure was constructed using weighted UniFrac distance matrix which measures variation in bacterial phylogenetic diversity and abundance between each sample. Axes indicate % variance; F and M stands for female and male respectively; * (star) -pregnant and lactating female, o (ring) - lactating female, color red - hare and color blue - rabbit.

I observed the bacterial diversity between hare and rabbit samples to be significantly different (PERMANOVA pseudo-F test; sample size: 21; t-stat = 15.9; p-value = 0.001). Axes in Figure 3 indicate variance which is the degree of difference observed between each sample. Axis 1 at variance of 47.73%, clearly distinguished hares from rabbits. Axis 2 at variance of 16.67%, clustered hare samples together indicating similar diversity between all nine hare samples whereas the rabbit samples were spread out indicating difference in diversity between twelve

rabbit samples. Comparing both axes, the degree of variance was more in axis 1, hence the difference in diversity between hare and rabbit populations was more compared to difference within rabbit population. Also, no correlation was observed in both rabbit and hare samples in terms of sex, age, lactation/pregnancy and month of sample collection (season).

I performed another Beta-diversity analysis, Bray-Curtis dissimilarity metric, which computes for bacterial compositional dissimilarity between different samples (Figure 4). This test groups samples with similar bacterial composition together while samples with dissimilar composition are spread out.

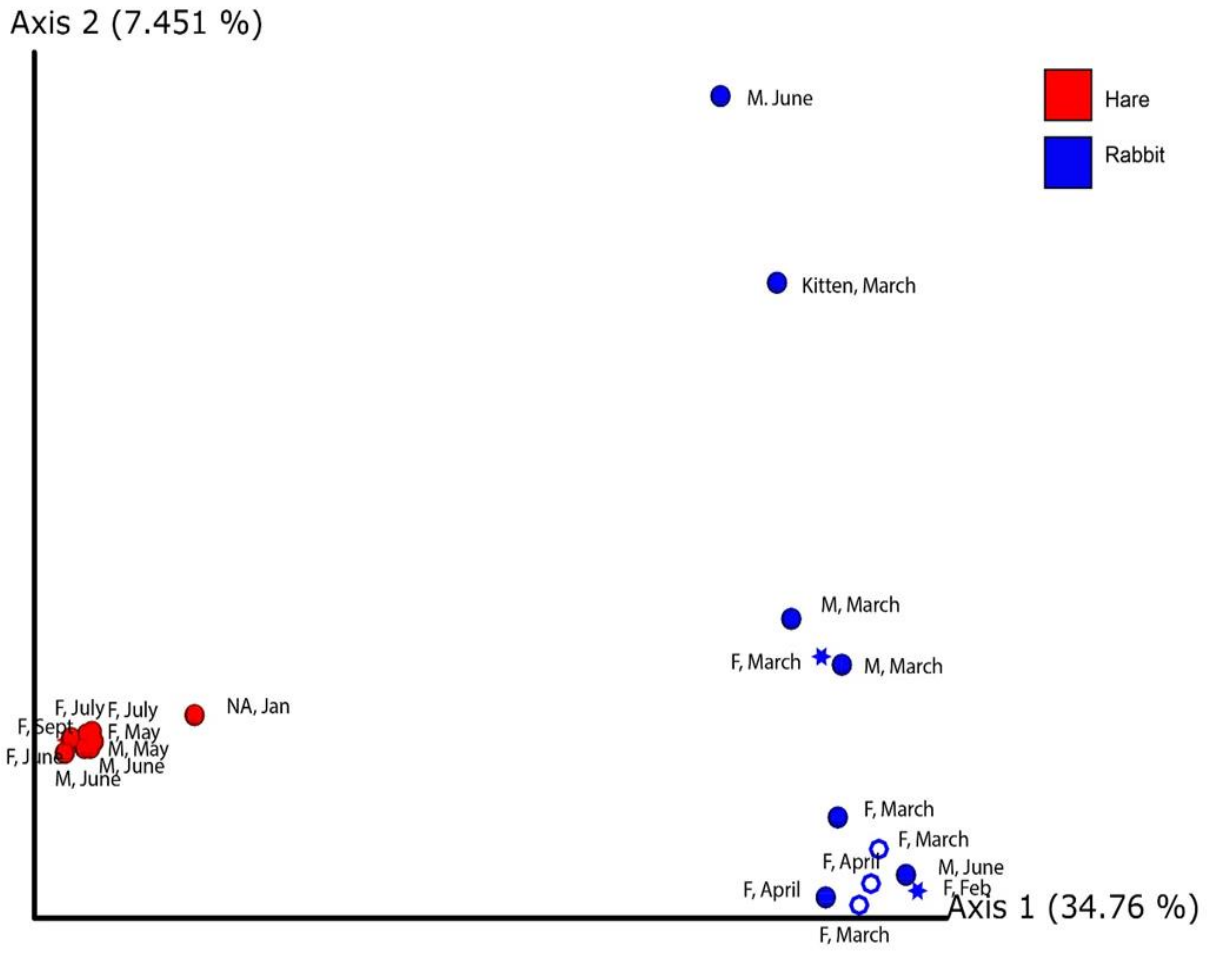


Figure 4: Segregation of hare samples from rabbit samples indicate significant difference in bacterial composition between two populations (PERMANOVA pseudo F test; sample size: 21; t-stat = 9.73; p-value = 0.001). This figure was constructed using Bray-Curtis dissimilarity metric which groups samples together or away from each other on the basis of similarity/dissimilarity in bacterial composition. Axes indicates % variance; F and M stands for female and male respectively; * (star) - pregnant and lactating female, o (ring) - lactating female, color red - hare and color blue - rabbit.

I observed in Figure 4 that hare samples were significantly different from the rabbit samples (p-value 0.001; PERMANOVA pseudo-F test). Axis 1 (variance 34.76%) clearly distinguished rabbits from hares. This indicate difference in compositional diversity between hares and rabbits. Axis 2 (variance 7.451%) separated rabbit samples from each other while the hare samples were clustered together. This indicate that rabbit samples vary in composition among each other while hare samples had similar bacterial composition and no correlation was observed between any factors.

3.1.3 Taxonomical classification of hare and rabbit faecal microbiomes

Assigning the ASVs produced to the SILVA 16S database using VSEARCH at 99% sequence identity, produced taxonomic classification only till genus rank with several bacteria denoted as unclassified or uncultured.

Taxonomical classification at phylum rank clearly distinguished between faecal microbiome of hares and rabbits as shown in Figure 5-A. *Firmicutes* and *Bacteroidetes* were the dominant phyla in both hares and rabbits as seen by over-representation of dark and light shade of purple in Figure 5-A. Phylum *Firmicutes* was present significantly more in rabbits (mean relative frequency/abundance of 83 ± 9%) than hares (mean relative frequency of 69 ± 5%) (Student's t-test; sample size: 21; t-stat = 2.27; $p = 0.04$). On the other hand, phylum *Bacteroidetes* was present significantly more in hares (mean relative frequency of 23 ± 5%) than rabbits (mean frequency of 6.65 ± 4.8%) (Student's t-test; sample size: 21; t-stat = 6.17; $p < 0.001$). As seen in the figure, phyla *Tenericutes*, *Actinobacteria* and *Cyanobacteria* were more abundant in rabbits with mean relative frequency of 3.29 ± 1.72%, 2.27 ± 0.9% and 1.6 ± 1.84%, respectively than hares with mean relative frequency of 0.348 ± 0.1%, 1.44 ± 0.5% and 0.31 ± 0.256%, respectively (Student's t-test; sample size: 21; t-stat = 5.65, 2.29, 2.25, respectively; $p < 0.05$). On the contrary, phylum *Proteobacteria* was more abundant in hares with mean relative frequency of 1.24 ± 0.35% compared to rabbits with mean relative frequency of 0.59 ± 0.45% (Student's t-test; sample size: 21; t-stat = 3.38; $p = 0.003$). Major difference was observed in phylum *Verrucomicrobia* (represented by dark red color in Figure 5-A) which was present only in rabbit samples at mean relative frequency of 2.12 ± 2.2%. Phyla *Lentisphaerae* and *Synergistetes* (represented by blue shades in Figure 4-A) were present only in hare samples at a mean relative frequency of 1.54 ± 0.69% and 0.75 ± 0.3%, respectively. Additionally, a superphylum *Patescibacteria* was also observed in both hare and rabbit samples as shown by presence of light green color in Figure 5-A.

Reagent-only controls and no-template control (NTC) were included to test for bacterial abundance present within the reagents and/or for contamination during sequencing run. However, considerable abundance of bacterial phyla were detected as depicted in Figure 5-B. *Proteobacteria* was the only phylum of bacteria present in NTC. *Firmicutes* and *Bacteroidetes* were dominant phyla present in both hare and reagent controls. Phyla *Dependentiae, Chloroflexi, Chlamydiae* and *Fibrobacteres* were present only in reagent controls and absent in samples as indicated by presence of * (star) in Figure 5-B.

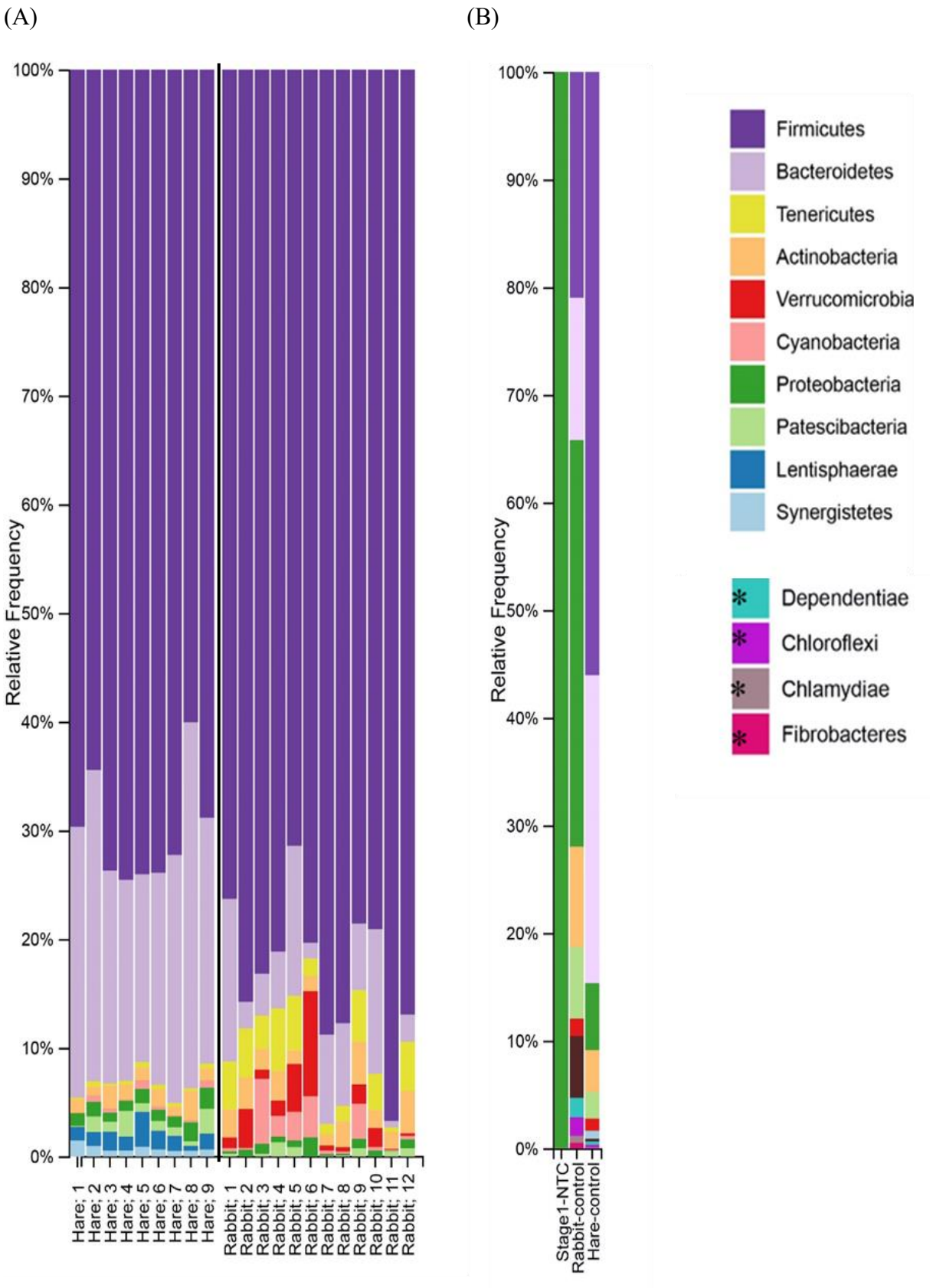


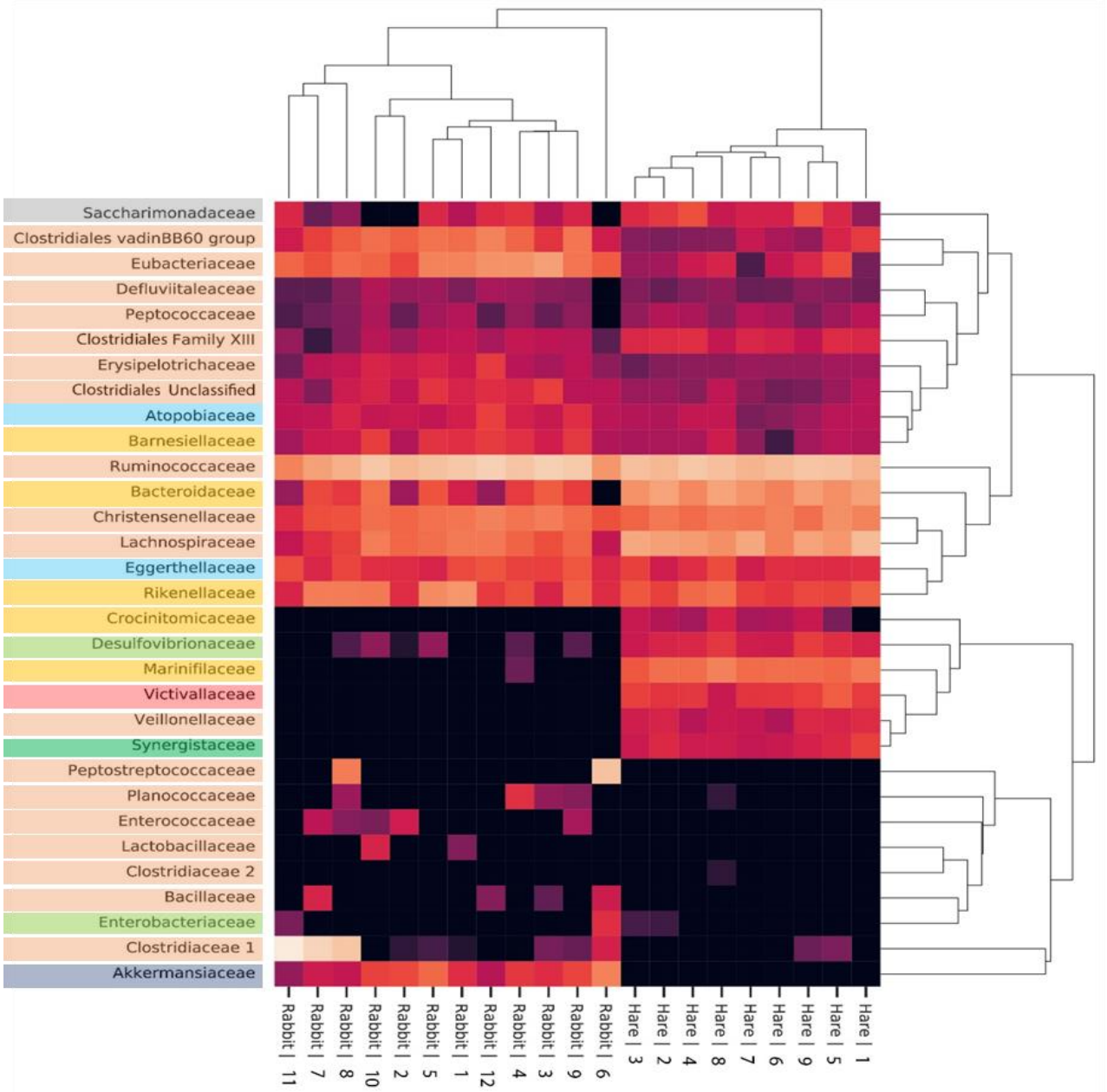
Figure 5: Difference in taxonomic classification between hare, rabbit and control samples at phylum rank. Taxonomic bar plot generated by analysing V3-V4 region of 16S rRNA gene depicting relative frequency of (A) all rabbit and hare samples (B) reagent-only controls of hare and rabbit along with a PCR no-template control (NTC) at a cut-off > 0.5% relative frequency. Bacterial phylum marked with * (star) represents phyla present only in reagent-only controls.

Under family and genus rank, clear cluster differentiating hare and rabbit samples was observed indicating difference in microbiomes as shown in Figure 6-A and 7-A. Figure 6 and 7 represents abundance of each bacterial family/genus in a form of heat map. Samples with more abundant bacteria (log10 frequency) were represented by orange shade, low abundant bacteria were in purple shade and black shade denoted absence of bacteria. Figure 5 and 6 were clustered on the basis of relative abundance/frequency of samples and ASVs which can be seen with the dendrogram. In figure 6-A and 7-A, Genus *Ruminococcus 1* (*Ruminococcaceae*) was the most dominant followed by genus *Bacteroides* (*Bacteroidaceae*), both of which were present significantly more in hares than rabbits (student's t-test; sample size: 21; t-stat = 9.5 and 8.1; p-value < 0.001) as shown by increase in log10 frequency. *Ruminococcus 1* was present at mean relative frequency of 27.3 ± 7.57% in hares and 2.35 ± 2.16% in rabbits whereas *Bacteroides* was present at mean relative frequency of 14.31 ± 3.71% in hares and 1.44 ± 1.47% in rabbits. Genera *Victivallis* (*Victivallaceae*), *Fluviicola* (*Crocinitomicaceae*), *Oribacterium* (*Lachnospiraceae*), *Oscillibacter* (*Ruminococcaceae*), *Pyramidobacter* (*Synergistaceae*) and *Cloacibacillus* (*Synergistaceae*) were only present in hare samples while *Akkermansia* (*Akkermansiaceae*) was only present in rabbit samples. Some bacterial genera were only present in a few samples; *Paraclostridium* (*Peptostreptococcaceae*) was present only in rabbit-8, *Lactobacillus* (*Lactobacillaceae*) was present only in rabbit-10, *Bacillus* (*Bacillaceae*) was present in only four rabbit samples and *Paeniclostridium* (*Peptostreptococcaceae*) was present in only rabbit-6.

In controls depicted in figure 6-B and 7-B, unclassified genera from the family *Burkholderiaceae* were primarily present in NTC and rabbit reagent-control. However, the hare reagent-control was composed primarily of family *Ruminococcaceae*.

Several bacterial families and genera remained unclassified, or belonged to a genus with uncultured species. Hence, they were not included in Figure 6 and 7. Genera *Ruminococcus 1* and *Clostridium 1* were not depicted in Figure 7.

(A)


Saccharimonadaceae
Clostridiales vadinBB60 group
Eubacteriaceae
Defluviitaleaceae
Peptococcaceae
Clostridiales Family XIII
Erysipelotrichaceae
Clostridiales Unclassified
Atopobiaceae
Barnesiellaceae
Ruminococcaceae
Bacteroidaceae
Christensenellaceae
Lachnospiraceae
Eggerthellaceae
Rikenellaceae
Crocinitomicaceae
Desulfovibrionaceae
Marinifilaceae
Victivallaceae
Veillonellaceae
Synergistaceae
Peptostreptococcaceae
Planococcaceae
Enterococcaceae
Lactobacillaceae
Clostridiaceae 2
Bacillaceae
Enterobacteriaceae
Clostridiaceae 1
Akkermansiaceae
Rabbit | 11
Rabbit | 7
Rabbit | 8
Rabbit | 10
Rabbit | 2
Rabbit | 5
Rabbit | 1
Rabbit | 12
Rabbit | 4
Rabbit | 3
Rabbit | 9
Rabbit | 6
Hare | 3
Hare | 2
Hare | 4
Hare | 8
Hare | 7
Hare | 6
Hare | 9
Hare | 5
Hare | 1

(B)

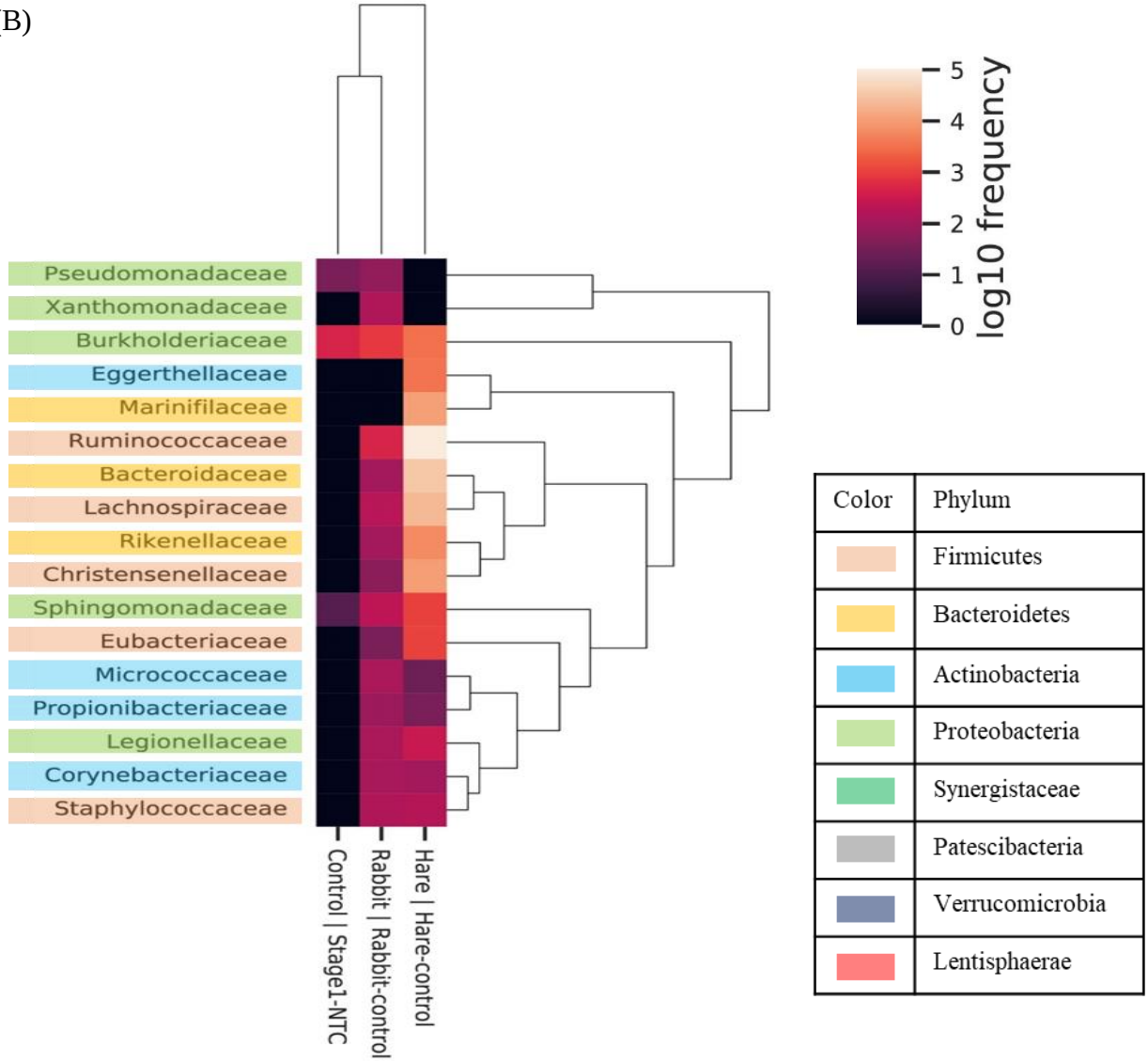
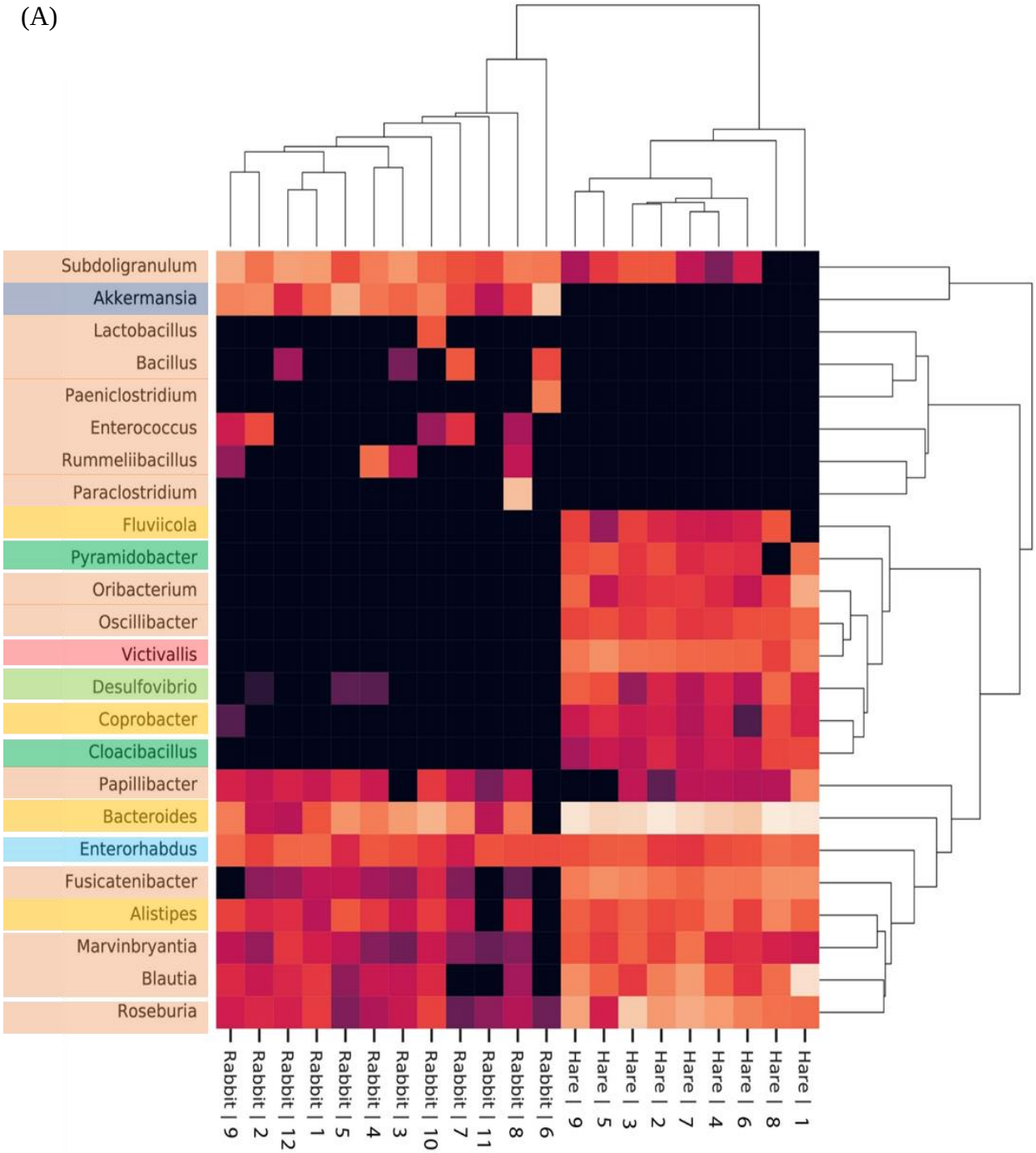


Figure 6: Similarity and dissimilarity in identified bacterial families between hare, rabbit and control samples. Normalized heat map clustered with respect to samples and ASVs indicated by the presence of dendrogram in (A) hare and rabbit samples (B) reagent-only controls and NTC at cut-off > 0.5% relative frequency. Each color highlighted in bacterial family corresponds to the phylum represented in the adjacent table.

(A)


Subdoligranulum
Akkermansia
Lactobacillus
Bacillus
Paeniclostridium
Enterococcus
Rummeliibacillus
Paraclostridium
Fluviicola
Pyramidobacter
Oribacterium
Oscillibacter
Victivallis
Desulfovibrio
Coprobacter
Cloacibacillus
Papillibacter
Bacteroides
Enterorhabdus
Fusicatenibacter
Alistipes
Marvinbryantia
Blautia
Roseburia
Rabbit | 9
Rabbit | 2
Rabbit | 12
Rabbit | 1
Rabbit | 5
Rabbit | 4
Rabbit | 3
Rabbit | 10
Rabbit | 7
Rabbit | 11
Rabbit | 8
Rabbit | 6
Hare | 9
Hare | 5
Hare | 3
Hare | 2
Hare | 7
Hare | 4
Hare | 6
Hare | 8
Hare | 1

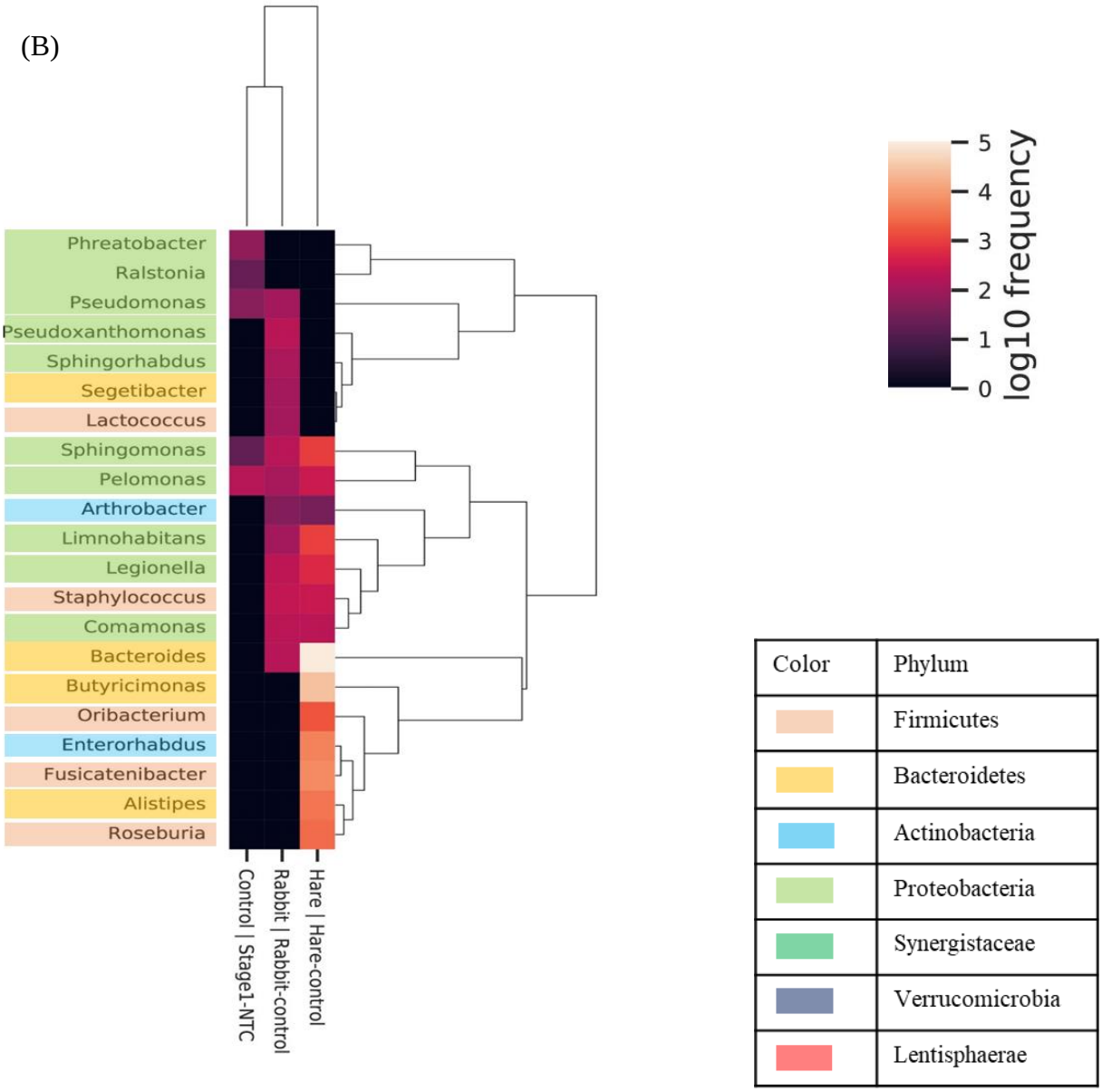


Figure 7: Similarity and dissimilarity in identified bacterial genera between hare, rabbit and control samples. Normalized heat map clustered with respect to samples and ASVs indicated by the presence of dendrogram in (A) hare and rabbit samples (B) reagent-only controls and NTC at cut-off > 0.5% relative frequency. Each color highlighted in bacterial genera corresponds to the phylum represented in the adjacent table.

Overall taxonomical classification using Illumina MiSeq sequencing platform showed significant difference between hare and rabbit samples at phyla, family and genera level (Kruskal-Wallis test; sample size: 21; H = 3.68; p-value < 0.05) with most of the sequences assigned to phylum *Firmicutes*.

3.2 Nanopore MinION 16S sequencing results

3.2.1 Raw data produced from nanopore MinION 16S sequencing pipeline

More than 18 Gb of sequenced data was generated from 21 samples in two MinION sequencing runs. The summary of data in terms of total reads produced, reads passing the standard threshold score for MinION of seven, total bases, median read length and average quality score for all samples are given in Table 6.

Table 6: Summary of reads after filtering for 21 samples after two MinION sequencing run[2]

Sample	Total reads	Reads passing quality score of 7	Total bases	Mean read length (of pass reads)	Average quality score (of pass reads)
Hare-1	590,923	444,779 (75%)	759,853,201	1,553	9.41
Hare-2	763,006	593,083 (78%)	1,032,389,762	1572	9.53
Hare-3	155,925	141,604 (91%)	231,322,214	1,585	11.03
Hare-4	198,989	180,449 (91%)	294,785,999	1,588	11.06
Hare-5	127,431	115,724 (91%)	187,198,038	1,579	11.04
Hare-6	233,751	211,670 (91%)	349,814,705	1,588	11.06
Hare-7	195,121	177,213 (91%)	289,448,816	1,585	10.98
Hare-8	568,551	430,423 (76%)	744,343,036	1,557	9.71
Hare-9	810,361	616,848 (76%)	1,077,263,865	1,569	9.47
Hare-control	NA	NA	NA	NA	NA
Rabbit-1	250,270	226,033 (90%)	381,751,639	1,588	11.05
Rabbit-2	179,169	163,172 (91%)	263,250,543	1,580	11
Rabbit-3	181,509	164,699 (91%)	278,965,227	1,581	11.12
Rabbit-4	305,081	275,920 (90%)	467,309,173	1,584	11.09

[2] NA refers to samples (reagent-only controls and NTC) which had low initial concentration and hence avoided in MinION run.

Sample	Total reads	Reads passing quality score of 7	Total bases	Mean read length (of pass reads)	Average quality score (of pass reads)
Rabbit-5	237,499	214,948 (91)	360,282,854	1,583	11.04
Rabbit-6	275,414	248,677 (90%)	418,306,053	1,575	11.15
Rabbit-7	268,807	243,668 (91%)	408,879,313	1,582	11.13
Rabbit-8	611,100	460,587 (75%)	831,527,583	1,557	9.78
Rabbit-9	580,937	431,531 (74%)	803,348,396	1,572	9.63
Rabbit-10	495,476	381,084 (77%)	652,263,795	1,550	9.51
Rabbit-11	520,445	393,958 (76%)	722,443,100	1,559	9.69
Rabbit-12	575,643	428,700 (74%)	766,118,450	1,557	9.48
Rabbit-control	NA	NA	NA	NA	NA
Stage 1-NTC	NA	NA	NA	NA	NA

I observed a total of 3,175 OTU (features) in 21 samples after importing into QIIME 2 and a total of 4,612,436 reads were generated. The minimum and maximum number of reads were 91,172 (hare-5) and 387,967 (hare-9) with a mean of 219,639.8 over all samples. Minimum and maximum number of reads per OTU was 1 and 892,095, respectively with an observed mean of 1,461.

3.2.2 Alpha and Beta diversity analysis of hare and rabbit microbiome

I only performed non-phylogenetic diversity analysis for MinION data due to unavailability of phylogenetic information. The phylogenetic information are accessed through the representive sequence (consensus sequence from the reference OTU cluster) in QIIME 2 which is unavailable for MinION data since it was processed in EPI2ME platform.

To test for difference in bacterial diversity within samples, I used an alpha diversity metric called Shannon index (Figure 8) since Faith's pd requires phylogenetic information as it

estimates diversity by taking sum of length of branches. Shannon index estimates bacterial richness and evenness on the basis of abundance of OTUs.

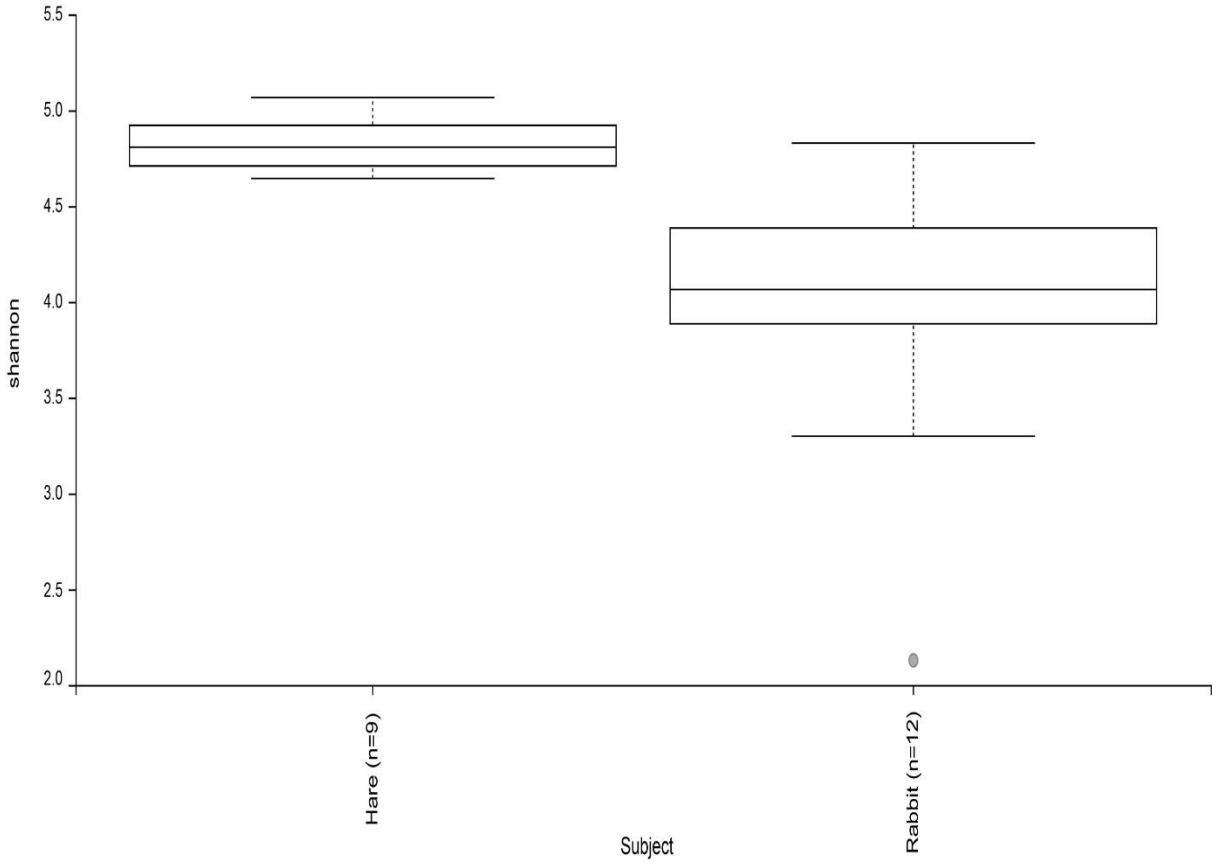


Figure 8: Considerable variation in distribution of values indicate significant difference in species richness and evenness between hare and rabbit populations (Kruskal-Wallis test; sample size =21; H = 12.12; p-value = 0.0004). Evaluated using Shannon index metric, an alpha diversity analysis, which measures bacterial species richness and evenness with respect to abundance of OTUs. Small box indicates similar richness/evenness between all hare samples and the large box indicates difference in richness/evenness between each rabbit sample.

Figure 8 shows that all the hare samples lie between the range of 4.5 to 5.1 and had smaller box symbolising similar level of evenness and richness between hare samples. Rabbit samples were between the range of 3.3 to 4.8 with one outlier and had larger box symbolising varying level of sample evenness and richness between twelve rabbit samples. A p–value less than 0.05 indicates that there were significant difference between hare and rabbit samples with respect to Shannon index.

To further estimate number of OTUs that were associated with each sample, I used another alpha diversity metric called observed OTUs which is shown in Table 7.

Table 7: Number of OTUs associated with each sample as measured by an alpha diversity metric, observed OTUs.

Sample	Observed OTU (p-value = 0.13)
Hare-1	585
Hare-2	552
Hare-3	457
Hare-4	532
Hare-5	642
Hare-6	533
Hare-7	535
Hare-8	580
Hare-9	504
Rabbit-1	771
Rabbit-2	759
Rabbit-3	832
Rabbit-4	880
Rabbit-5	834
Rabbit-6	494
Rabbit-7	500
Rabbit-8	468
Rabbit-9	841
Rabbit-10	851
Rabbit-11	251
Rabbit-12	892

I observed from Table 7 that there were no significant difference between hare and rabbit samples with respect to OTUs present (Kruskal-Wallis test; sample size: 21; $H = 2.22$; $p = 0.13$). Number of observed OTUs was considerably more in rabbit samples than hare samples, however, large variation in OTU was observed between different rabbit samples.

Bray-Curtis dissimilarity test was one beta-diversity analysis I performed to estimate the level of similarity/dissimilarity between each sample on the basis of bacterial composition (Figure 9)

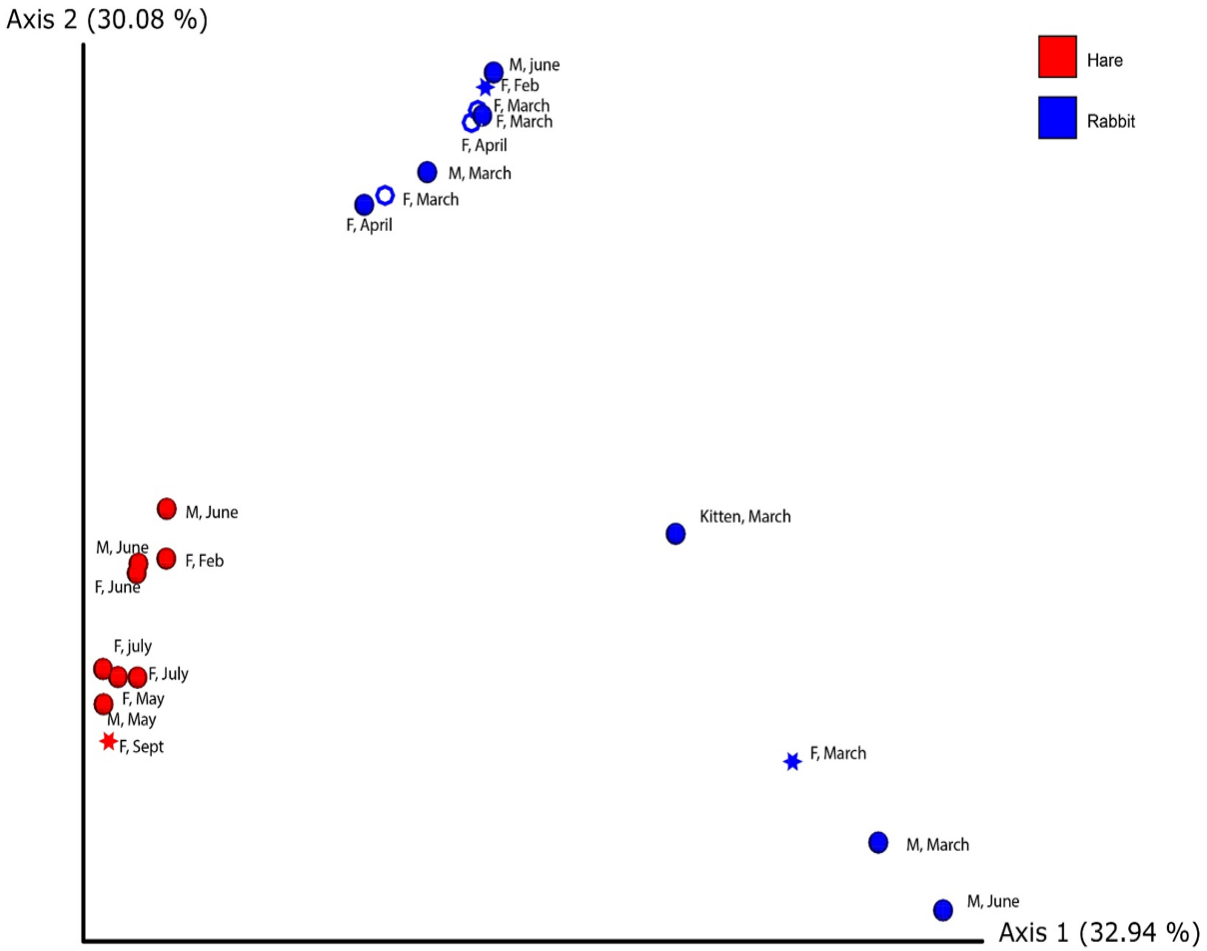


Figure 9: Segregation of hare samples from rabbit samples indicate significant difference in bacterial composition between two populations (PERMANOVA pseudo F test; sample size: 21; t-stat = 8.84; p-value = 0.001). This figure was constructed using Bray-Curtis dissimilarity metric which groups samples together or away from each other on the basis of similarity/dissimilarity in bacterial composition. Axes indicates % variance; F and M stands for female and male respectively; * (star) - pregnant and lactating female, o (ring) - lactating female, red color - hare and blue color - rabbit.

In Figure 9, axis 1 (32.94% variance) showed hare samples to be clustered together while rabbit samples were spread out. Axis 2 (30.08% variance) showed certain degree of variation between rabbit samples but all hare samples were present relatively close to each other. Hence, this plot shows significant difference between rabbit and hare samples on the basis of bacterial composition (PERMANOVA pseudo F test; sample size: 21; t-stat = 8.84; p-value = 0.001).

3.2.3 Taxonomical classification of rabbit and hare faecal microbiomes

EPI2ME classified reads against NCBI 16S database using BLAST. This produced species-level classification as described below.

In both hare and rabbit samples, *Firmicutes* and *Bacteroidetes* were the most dominant phyla as indicated by over-representation of two shades of purple color in Figure 10. *Firmicutes* was present at a mean relative frequency of 61.1 ± 8.02% in hares and 83.8 ± 10.26% in rabbits yet there were no significant difference between the means of two groups (Student's t-test; sample size = 21; t-stat = 2.24; p = 0.26). I observed three rabbit sample (rabbit-7,8 and 11) to be composed primarily of only *Firmicutes* (>96%). Phylum *Bacteroidetes* was present at a mean relative frequency of 33.44 ± 9.09% in hares and 5.77 ± 5.01% in rabbits showing significant difference (Student's t-test; sample size = 21; t-stat = 3.63; p = 0.006). However, only 0.1% of *Bacteroidetes* was associated with sample rabbit-11. Phylum *Verrucomicrobia* (4.39 ± 7.34%) was present only in rabbit samples while phyla *Lentisphaerae* (0.78 ± 0.57%) and *Synergistetes* (0.3 ± 0.15) were present only in hare samples. Phylum *Proteobacteria* was present significantly more in hares with mean relative frequency of 0.78 ± 0.27% than rabbits with mean relative frequency of 0.2 ± 0.26% (Student's t-test; sample size = 21; t-stat = 2.69; p = 0.02). Phyla *Cynaobacteria*, *Actinobacteria* and *Tenericutes* were present significantly more in rabbits with mean frequency of 0.95 ± 1.37%, 0.49 ± 0.52% and 1.9 ± 1.4% respectively than hares (observed only in three samples at less than 0.05%) (Student's t-test; sample size = 21; t-stat = 2.71, 2.89 and 3.92, respectively; $p < 0.02$) as depicted in Figure 10.

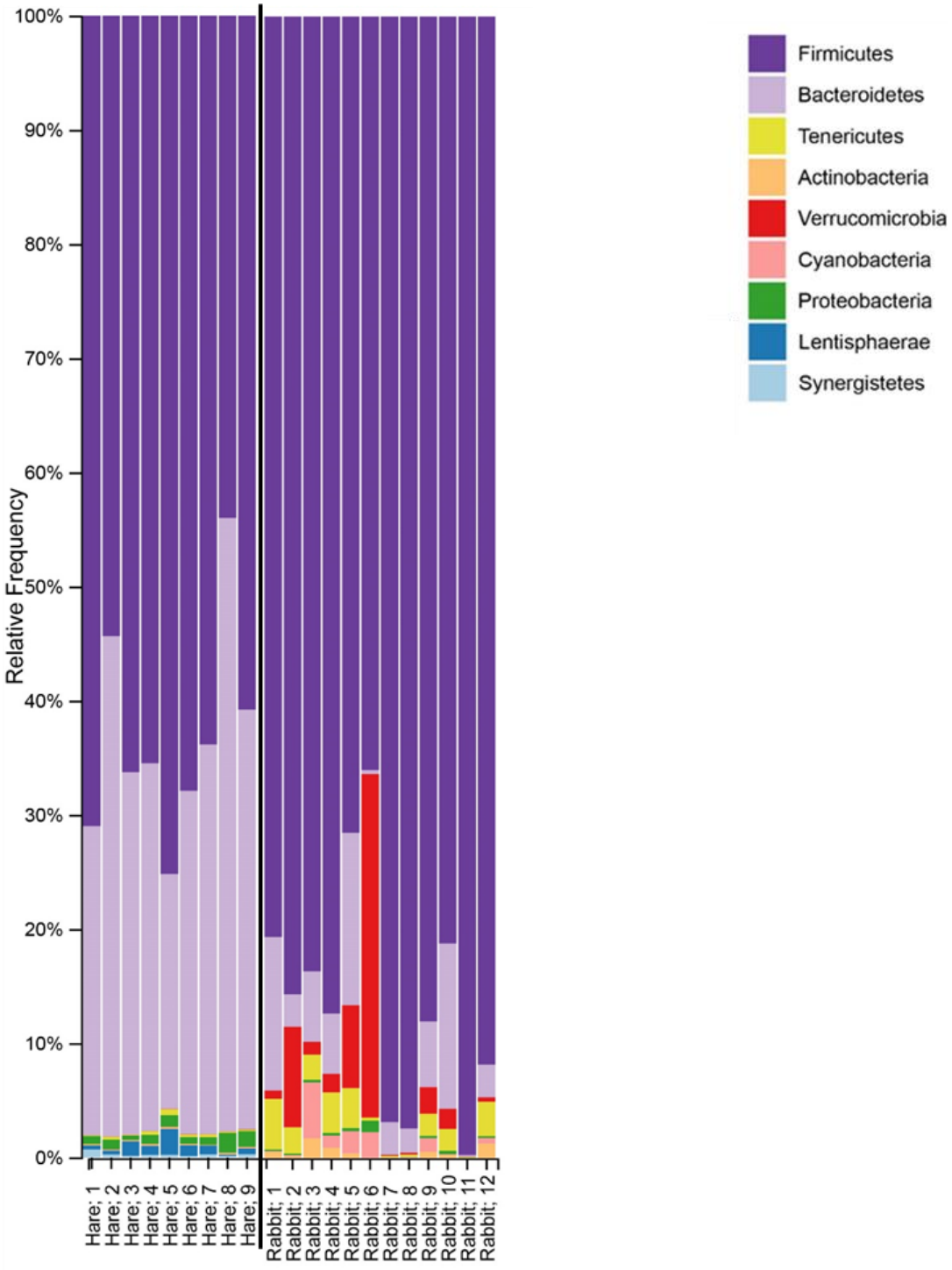


Figure 10: Difference in taxonomic classification between hare and rabbit samples at phylum rank. Taxonomic bar plot generated by analysing entire region of 16S rRNA gene depicting relative frequency of all hare and rabbit samples at a cut-off > 0.5% relative frequency.

Taxonomic classification at family and genus rank are shown in Figure 11 and Figure 12. The figure is constructed on the same basis as of Illumina MiSeq results. The most dominant genera were *Clostridium* (*Clostridiaceae*), present more than 50% relative frequency in only three samples (rabbit-7,8 and 11), followed by *Ruminococcus* (*Ruminococcaceae*) with mean relative

frequency of 15.61 ± 4.9% in hares and 2.70 ± 2.5% in rabbits. *Bacteroides* (*Bacteroidaceae*) was the next abundant genus with a mean relative frequency of 13.44 ± 4.29% in hares and 0.39 ± 0.52% in rabbits. Family *Synergistaceae* was present only in hare samples and was completely absent in rabbit samples (shown in Figure 11). Some genera were present only in few samples but not all; one of them was *Paraclostridium* (*Peptostreptococcaceae*).

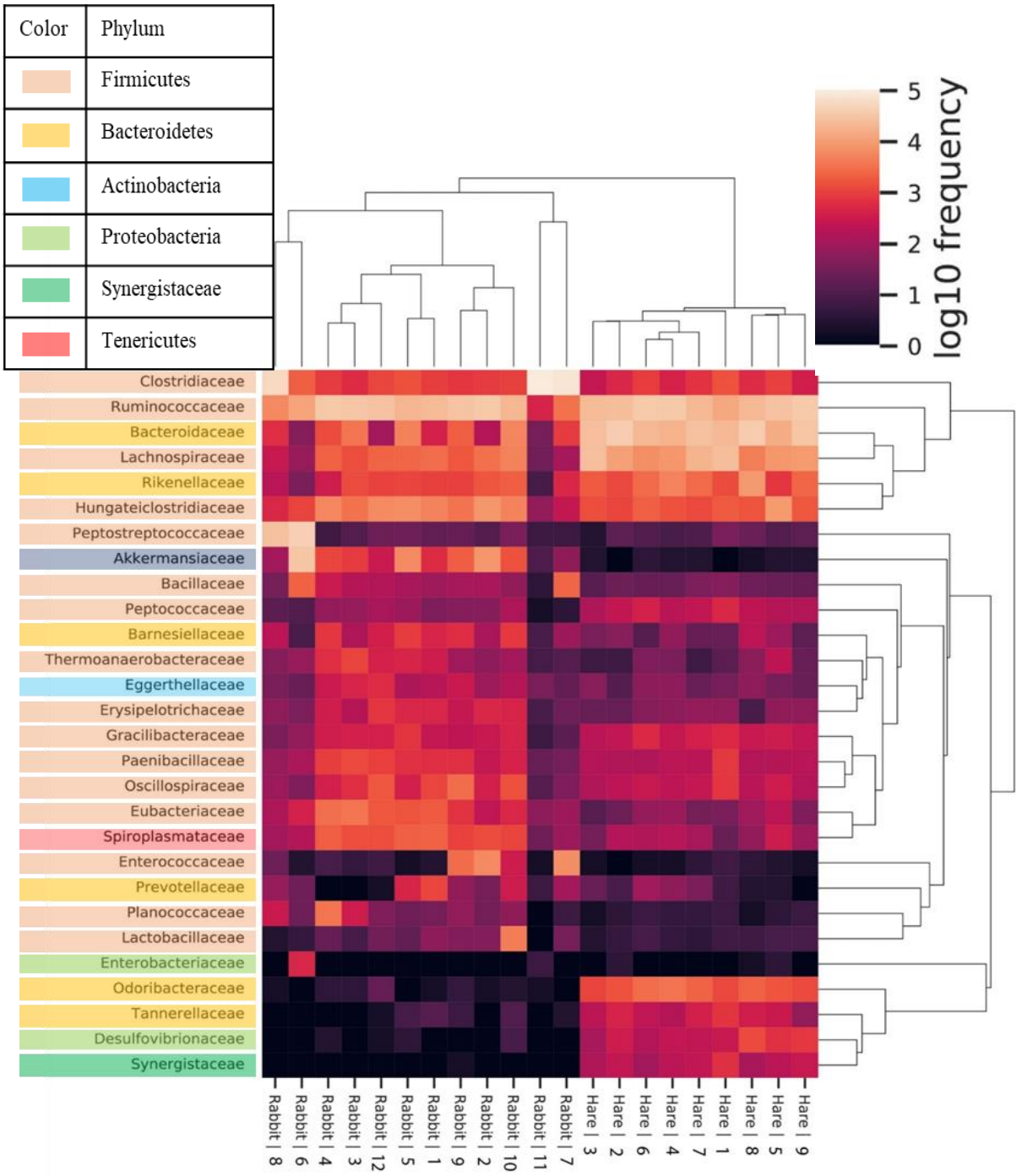


Figure 11: Difference in identified bacterial families between hare and rabbit samples. Normalized heat map clustered with respect to samples and OTUs represented by dendrograms in all hare and rabbit samples at cut-off > 0.5% relative frequency. Each color highlighted in bacterial family corresponds to the phylum represented in the adjacent table.

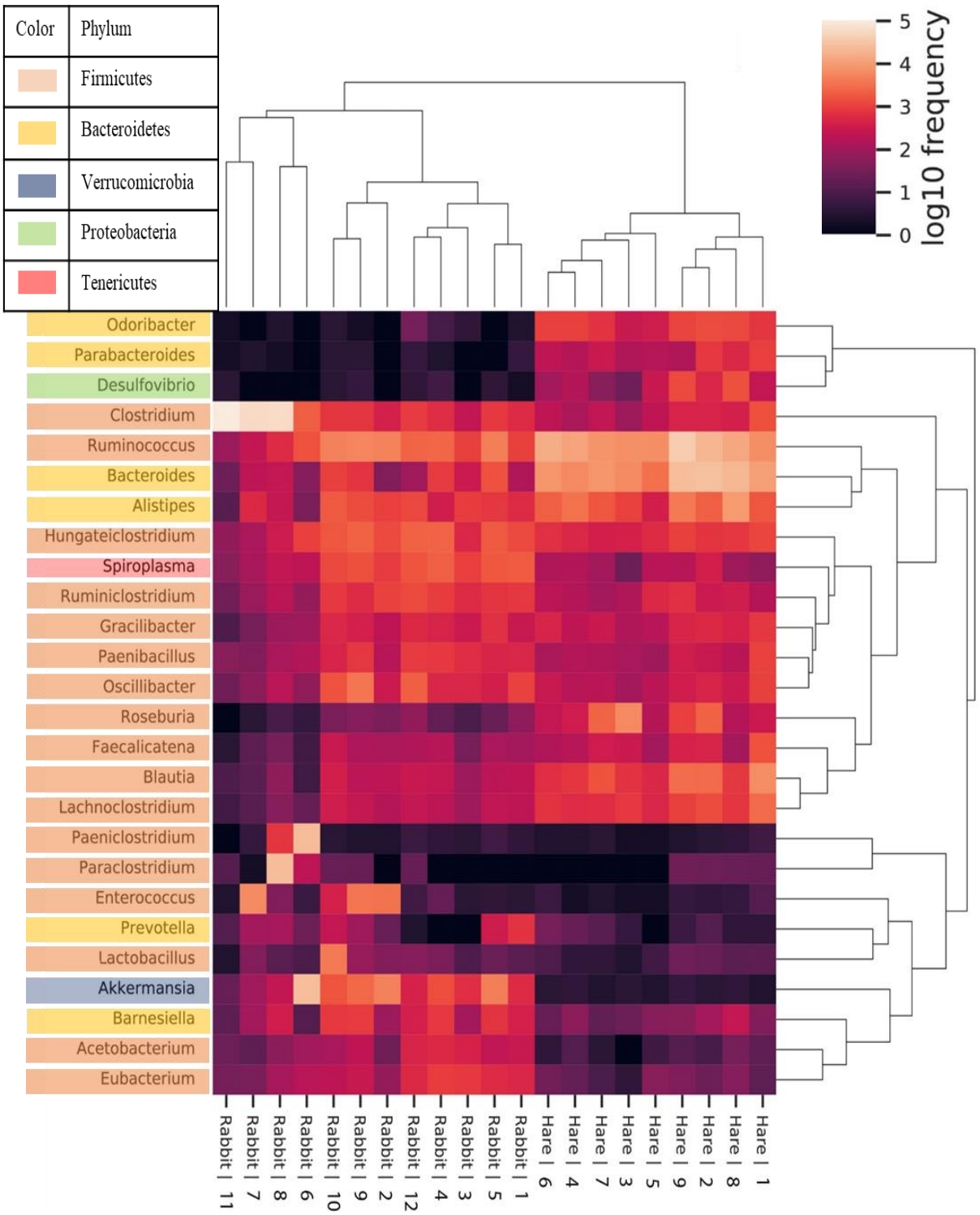


Figure 12: Difference in identified bacterial genera between hare and rabbit samples. Normalized heat map clustered with respect to samples and **O**TUs of all hare and rabbit samples represented by dendrograms at cut-off > 0.5% relative frequency. Each color highlighted in bacterial genera corresponds to the phylum represented in the adjacent table.

Overall, significant difference was observed between hare and rabbit faecal samples with respect to taxonomy (Kruskal-Wallis test; sample size: 21; $H = 4.25$; $p < 0.05$) and phylum *Firmicutes* was the most abundant.

3.3 Illumina MiSeq vs nanopore MinION for microbiome analysis

3.3.1 Difference in reads between MiSeq and MinION 16S data

The total yield was more in nanopore MinION (~16 Gb) compared to Illumina MiSeq (~8.94 Gb). Number of reads and distribution of reads per sample for both MiSeq and MinION dataset are depicted in the figure 13. The read length obtained was more in MinION (~1550 bp) than in Illumina (~ 431 bp) based on the region chosen for sequencing. The quality score chosen for MiSeq data was 30 in a Phred scale but for MinION, the threshold value chosen was 7. It cannot be compared since MinION reads do no follow Phred scale. Sequence time was similar for both MiSeq and MinION run, however, the latter can be stopped any time and produce real-time results for analysis.

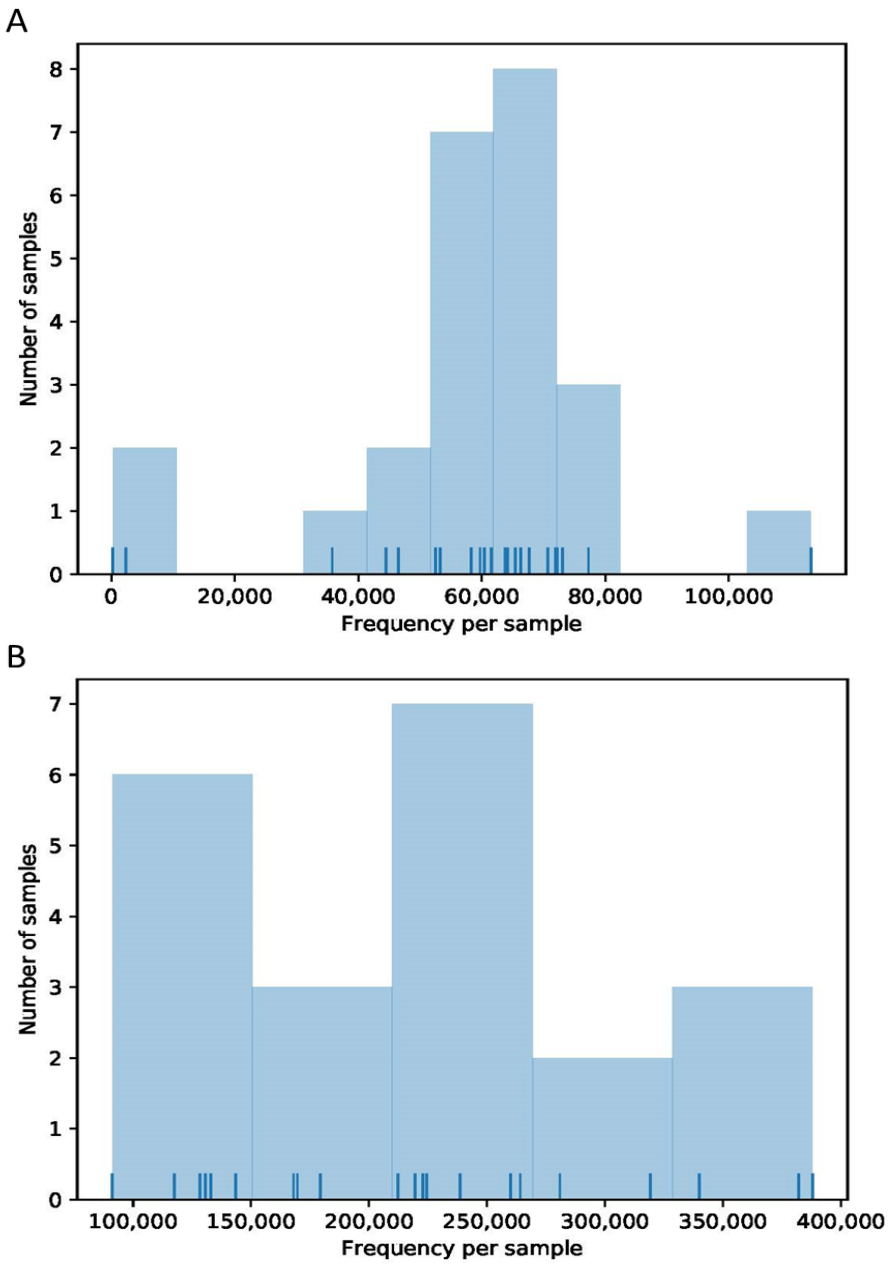


Figure 13: Frequency of reads generated per sample in MiSeq and MinION pipelines after quality control (A) 24 Illumina MiSeq samples with two samples present close to zero reads (NTC and rabbit-control) (B) 21 nanopore MinION samples with no controls. Y-axis depicts number of samples that fall under a certain frequency range (x-axis). Blue line in the x-axis indicate samples present at a specific frequency.

3.3.2 Difference in diversity analysis

I observed that both MiSeq and MinION pipelines produced similar results each showing significant difference between hare and rabbit samples. MinION data did not have phylogenetic information to perform Faith's pd and weighted UniFrac distance matrix, however, comparison was done based on Shannon diversity index and Bray-curtis dissimilarity test.

Shannon index using MiSeq data did not show significant difference between hare and rabbit samples, meaning there were no difference between them on the basis of species richness and evenness (Figure 2). However with MinION data, Shannon index showed significant difference between hare and rabbit samples (Figure 8).

Bray-curtis test in both MiSeq and MinION pipelines showed significant difference between hare and rabbit sample on the basis of bacterial composition (Figure 4 and 9). Both the pipelines showed that rabbit population to vary more than hare population.

3.3.3 Taxonomic difference between MiSeq and MinION sequenced data

Illumina (MiSeq) and nanopore (MinION) data were similar at phylum level, except for over-representation of phylum *Verrucomicrobia* in rabbit 6 of MinION data and presence of a superphylum *Patescibacteria* in MiSeq dataset which was absent in MinION dataset.

A considerable difference was observed at family level in MinION and MiSeq datasets. Families *Marinifilaceae, Victivallaceae, Veillonellaceae, Defluviitaleaceae, Atopobiaceae*, *Christensenellaceae* and *Crocinitomicaceae* were present in MiSeq data at a cut-off > 0.5% but were present at low frequency in MinION dataset. Families *Thermoanaerobacteraceae, Gracilibacteraceae, Paenibacillaceae, Oscillospiraceae, Spiroplasmataceae, Prevotellaceae, Odoribacteraceae* and *Tannerellaceae* were present at a cut-off > 0.5% in MinION dataset but was either absent (*Thermoanaerobacteraceae, Oscillospiraceae, Spiroplasmataceae* and *Odoribacteraceae)* or were present at much lower frequency in MiSeq dataset. Family *Enterobacteriaceae* was present in hare-2 and hare-3, rabbit-6 and rabbit-11 in MiSeq dataet while in MinION dataset, it was present in hare-2, hare-5 and hare-8, rabbit- 6 and rabbit-11 (Figure 6-A and 11).

Genus *Akkermansia* was present in only rabbit sample of MiSeq dataset while the MinION dataset showed presence of very few reads associated with hare samples as well. Similar difference was observed with all genera except for bacterial genera that belonged to the family *Synergistaceae.*

3.3.4 Species level resolution of MinION 16S data

Species level classification was only available from data obtained through MinION pipeline. Name of the top 30 species identified with cut-off greater than 0.5% relative frequency, percentage of accuracy in alignment and mean relative frequency are given in the Table 8. Species with less than 50 reads are classified as too low in the below table.

Table 8: Species level classification and alignment accuracy of MinION data[3]

Genus	**Genus present in MiSeq data**	**Species**	**% Accuracy**	**Mean relative frequency (%)**	
				Hare	**Rabbit**
Ruminococcus	yes	*Ruminococcus albus*	83%	11.89 ± 3.97%	1.85 ± 2.04%
		Ruminococcus champanellensis	83%	2.92 ± 2.57%	0.56 ± 0.57%
		Ruminococcus flavefaciens	84%	1.37 ± 1.25	Too low
Clostridium	yes	*Clostridium sardiniense*	87%	Too low	Present in only three samples: rabbit-7 (2.25%),

[3] Bacterial species with reads less than 50 are classified as too low

Genus	Genus present in MiSeq data	Species	% Accuracy	Mean relative frequency (%)	
					rabbit-8 (5.07%) and rabbit-11 (64.5%)
		Clostridium baratii	86%	Too low	Too low (40.33% in rabbit-7 and 7% in rabbit-11)
		Clostridium moniliforme	86%	Too low	Too low (20% in rabbit-8 and 3% in rabbit-11)
		Clostridium paraputrificum	83%	Too low	Too low (5% in rabbit-7 and 8% in rabbit-11)
		Clostridium senegalense	85%	Too low	Too low (12% in rabbit-8)
		Clostridium budayi	89%	Too low	Too low (3% in rabbit-7 and rabbit-11)
Akkermansia	yes	*Akkermansia muciniphila*	82%	Too low	0.5% (29% in rabbit-6)
Bacteroides	yes	*Bacteroides uniformis*	86%	2.67 ± 2.13%	Too low

Genus	Genus present in MiSeq data	Species	% Accuracy	Mean relative frequency (%)	
		Bacteroides sartorii	85%	1.42 ± 1.07%	Too low
		Bacteroides mediterraneensis	83%	1.31 ± 0.47%	Too low
		Bacteroides eggerthii	86%	1.44 ± 1.49%	Too low
		Bacteroides caecigallinarum	84%	1.27 ± 0.96%	Too low
		Bacteroides intestinalis	87%	0.4 ± 0.5%	Too low
		Bacteroides helcogenes	82%	0.44 ± 0.18%	Too low
		Bacteroides xylanisolvens	86%	0.76 ± 0.66%	Too low
		Bacteroides dorei	86%	0.55 ± 0.55%	Too low
Paeniclostridium	yes	*Paeniclostridium sordellii*	89%	Too low	Too low (26.8% in rabbit-6)
Paraclostridium	yes	*Paraclostridium bifermentans*	88%	Too low	Too low (12.68% in rabbit-8)
		Paraclostridium benzoelyticum	93%	absent	Too low (6.02% in rabbit-8)
Alistipes	yes	*Alistipes putredinis*	82%	1.07 ± 0.98%	0.3 ± 0.29%
		Alistipes senegalensis	82%	0.59 ± 0.57%	0.13 ± 0.12%

Genus	Genus present in MiSeq data	Species	% Accuracy	Mean relative frequency (%)	
		Alistipes onderdonkii	84%	0.48 ± 0.41%	Too low
Roseburia	yes	*Roseburia hominis*	83%	1.45 ± 2.07%	Too low
Odoribacter	yes	*Odoribacter splanchnicus*	81%	0.81 ± 0.29	Too low
Blautia	yes	*Blautia coccoides*	82%	0.57 ± 0.4	Too low
		Blautia producta	83%	0.47 ± 0.05%	Too low
Oscillibacter	yes	*Oscillibacter valericigenes*	83%	0.23 ± 0.16	0.67 ± 0.69

Based on Table 8, species from the genera *Ruminococcus*, *Bacteroides*, *Alistipes*, *Roseburia*, *Odoribacter* and *Blautia* were present more in hare samples compared to rabbit samples. Species from genera *Clostridium*, *Paraclostridium*, *Paeniclostridium*, *Akkermansia* and *Oscillibacter* were present more in rabbit samples compared to hare samples. These findings also corresponded to Illumina MiSeq results except for *Oscillibacter* which was observed to be absent in rabbit samples and were only present in hares (Figure 7-A).

Overall, significant difference was observed between hare and rabbit samples with respect to taxonomy and diversity in both Illumina MiSeq and Nanopore MinION data (Kruskal-Wallis test; sample size 21; $H = 4.25$; $p = 0.03$). Furthermore, sex, lactation, weight and month of sample collection (season) did not seem to affect bacterial diversity or abundance in this study (Kruskal-Wallis test; sample size 21; $H = 3.79$; $p > 0.1$).

Chapter 4: Discussion

4.1 Overview

The major aim of my research was to study and compare the microbiome of sympatric hares and rabbits. The results demonstrated in the previous section showed that there was a significant difference between the faecal microbiome of hares and rabbits. Furthermore, diversity analysis revealed that all hare samples had similar diversity but there was considerable difference between rabbit samples. This difference was not found to be correlated with age, sex, lactation and month of sample collection. The second part of my aim was to compare the difference observed between two sequencing pipelines. As indicated by previous studies, Illumina MiSeq platform failed to provide species-level classification (Earl et al., 2018) while nanopore MinION provided species-level classification at low alignment accuracy. I discussed these findings in the upcoming sections.

4.2 Difference in microbiomes of sympatric hares and rabbits could be associated with difference in diet

The most dominant bacteria in both European brown hares and European rabbits were from the phyla *Firmicutes* (Genus *Ruminococcus 1*) and *Bacteroidetes* (Genus *Bacteroides*). However, these bacteria are not unique to only animals from the family *Leporidae*, but rather they are common gut inhabitants of most mammals including humans (Xu and Knight, 2015). The difference lies in the composition of these bacteria present. It has been shown that the *Firmicutes/Bacteroidetes* ratio change with respect to age, body mass index (BMI) and diet in humans (Mariat et al., 2009; Koliada et al., 2017; Singh et al., 2017). In my data, I observed this ratio to be higher in rabbits than hares since rabbit samples had comparatively lower abundance of *Bacteroidetes*. Higher abundance of *Bacteroidetes* in hare samples could be associated with diet rich in fat and protein (Riaz Rajoka et al., 2017). This statement was further supported by a previous study that suggested hares to prefer a diet rich in crude fat (Schai-Braun et al., 2015). However, further research is required to support this hypothesis.

Previous study that focused on estimating hare microbiome in Europe, identified phylum *Spirochaetes* to be the third dominant phylum following *Firmicutes* and *Bacteroidetes* in faecal

sample of hares (Stalder et al., 2019). However, *Spirochaetes* was absent in my samples. This could be associated with difference in geography which affects sunlight, ultraviolet rays (UV), temperature, average rainfall and grass species present. These abiotic factors could have had an impact on diet and hence, the observed difference in this bacterial phylum.

Phylum *Verrucomicrobia* was not only absent in all of my hare faecal samples, but was also absent in previous study that estimated microbiome of hares as well (Stalder et al., 2019). However, it was present in rabbit faecal samples. In this phylum, *Akkermansia muciniphila* was the dominant species present and it is a bacterium of interest for several upcoming studies. This beneficial bacterium is a common resident in the intestine of several mammals and is known to have a mucin-degrading property that aids in healthy functioning of the gut (Belzer and de Vos, 2012). *A. muciniphila* uses mucin that lines the walls of the intestine as an energy source and protects the host from pathogens by competitive exclusion (a law that-states two species cannot exist in an environment competing for same resources). Previous study showed that increase in uptake of dietary polyphenols increased the abundance of this bacterium in the intestine of mice (Anhê et al., 2016). Hence, it is possible to hypothesise here that rabbits consume a specific diet that increases *A. muciniphila* but hares may avoid or prefer other diet, hence absence of this bacteria. Phyla *Lentisphaerae* and *Synergistetes* were present in only hare samples and not in rabbit samples (Figure 5-A and 10), the same was identified in previous studies as well (Stalder et al., 2019, Velasco-Galilea et al., 2018). There is no possible explanation why it was observed only in hares but it could be associated with diet as well. In Greengenes database, I found the dominant genus present in the phylum *Synergistetes* to be *Synergistes* which corresponded to species *jonsii.* This species is commonly associated with plants *Leucaena leucocephala* and *Mimosa pudica*. Leaves and seeds of this plant contain mimosine, a toxic amino acid, which when consumed at more than 1% in diet could cause alopecia, necrotic spots, liver congestion and edema in mammals (Fayemi et al., 2011; Mauldin and Peters-Kennedy, 2016). The bacteria *S. jonsii* breaks down toxic pyridinediols produced by mimosine inside the body, thus neutralizing the toxic effect of these plants (Allison et al., 1992). Presence of this bacteria in the faecal sample of hares could indicate inclusion of this plant species in their diet. The existence of these plant in Mulligans flat are not yet confirmed but they are invasive plant species in Australia commonly found in Queensland, northern regions and ACT as well. However when we took a short trip to Mulligans Flat during early May, we observed several plants from the clade Mimosoideae which is also the clade of above mentioned plants. Further verification is required to confirm the presence of these plant species in Mulligans Flat and then hypothesise that rabbits avoid these plants, hence absence of *S. jonsii* in their faecal sample.

This genus was not observed in SILVA database or with MinION data (NCBI database) but several unclassified reads were present in the family *Synergistaceae.*

Bacteria from the family *Enterobacteriaceae* was considerably less abundant in both hare and rabbit faecal samples. This could represent a healthy intestinal tract since most of the bacteria in this family are associated with gastrointestinal diseases (Stalder et al., 2019).

Since the hares and rabbits in this study were sympatric, they must have been exposed to same food source. However, differences in the microbiome observed suggest that they may prefer different diets. This is further supported by previous studies which mentioned that hares and rabbits prefer different food source all year round leading to minimal competition for food (Hewson, 1989; Katona et al., 2004).

Another previous study identified that hares consumed only 47 plant taxa (based on their stomach content) out of 349 plants species present in the study area and only ten among the 47 plants were positively selected throughout the year (Schai-Braun et al., 2015). This could imply that hares are selective feeders and prefer only certain plant taxa in their diet. They are forced to change their diet only when their preferred diet is unavailable which may affect hare's nutrient requirement and eventually might result in population decline (Reichlin et al., 2006). Additionally, previous studies indicated hares to have highest proportion of polyunsaturated fatty acids (PUFAs) in their skeletal muscle tissue (65.7 ± 0.3%) compared to any other mammals including rabbits (31.9 ± 8.4%) (Valencak et al., 2003; Dalle Zotte, 2014; Cobos et al., 1995). This proportion elevated further during winter by 2.3%. Increase in muscle PUFA level had been hypothesised to affect basal metabolic rate (BMR) and thermoregulation (Nowack et al., 2017; Ruf and Arnold, 2008). Since hares are fast runners and survive above ground during winter, I speculate that hares require higher energy and this energy is supplemented by choosing a PUFA rich diet since PUFA cannot be synthesised within the body. This could explain the difference in the microbiome as rabbits may require lower percentage of PUFA in their diet and consume food rich in crude fibre instead, increasing the relative abundance of *Firmicutes*. Hares are referred to as being meticulous in choice of their diet where they avoid crude fibre and select for food rich in fat and energy (Schai-Braun et al., 2015; Stalder et al., 2019). This preference explains higher relative abundance of *Bacteroidetes* in hare samples compared to rabbits as discussed above.

Though the faecal microbiome is not an indicative of the total microbiome of a host, it is still a good representation of the gut microbiome. However, the relative abundance of each bacterium present may vary depending on the organ or tissue they are confined to. Coprophagy will also

influence the bacterial diversity in the faeces as mentioned before. Besides differences in the bacterial diversity between hare and rabbit samples, I also observed difference in bacterial diversity within rabbit samples.

4.3 Alpha and beta diversity analysis reveals difference in bacterial diversity between rabbit samples but all hare samples were of similar diversity

The alpha and beta diversity analysis which I performed differentiated the microbiome of hares and rabbits. However, beta diversity analysis clustered all hare samples together indicating similar diversity while the rabbit samples were spread out indicating difference in diversity. These results suggest that hares have a defined lifestyle which does not vary among individuals whereas each individual rabbit has a varying lifestyle.

Though hares prefer solitude and rarely socialise, the observed results suggest that they are selective in diet. Hares graze together for forage as long as the area is spaced with multiple food patches (Monaghan and Metcalfe, 1985). When there is a limit in availability of food source, they display aggressive behaviour and the dominant hare chases off other hares. This was observed to be a failed approach as the chased hares generally return and pursue the same food. This pattern could further support the theory that hares are selective feeders and only have narrow ranges of diet preference leading to competition for food amongst themselves.

Rabbits are social and generally forage in groups, however their bacterial diversity varied speculating difference in diet between each individual rabbit. There are several possible explanations for this. A previous study suggested that wild rabbits can transfer their food preference from one generation to another but there were no such reports for hares (Altbäcker et al., 1995). If this is the case, each rabbit will have different preference over a specific food source which could influence their microbiome. Rabbits are also known to occasionally shift between different social groups, especially male rabbits, due to aggressiveness, territoriality, and availability of resources (DiVincenti and Rehrig, 2016). I believe that their social group also influences their diet selection. Another possible explanation I speculate is that rabbits are not particular about their diet and consume wide range of plants which could diversify their microbiomes as well.

Another interesting observation was that the bacterial diversity in both hares and rabbits did not seem to correlate with age, sex, weight, lactation and month of sample collection in this study.

Previous studies had also identified that age and sex did not influence the microbiome of hares, however no such reports were shown for rabbits (Stalder et al., 2019). It was surprising to observe that month of sample collection (season) and lactation/pregnancy did not correlate with observed difference within each population. Lactating or pregnant hares require considerably more energy since they produce milk with high fat content of over 20% and most of which originates from consuming a diet rich in fat (Hackländer et al., 2002). If this was the case, then the diversity of bacteria in pregnant and/or lactating hares should vary at least moderately from other hares with respect to certain bacterial phylum. But it is difficult to come to a conclusion due to limited sample size. No previous study identified difference in food preference of lactating wild rabbits but domestic rabbits are shown to produce milk rich in fat when consuming a fat rich diet (Hackländer et al., 2002). Like mentioned before, season also has an effect on diet since some plants may fail to thrive during colder conditions. In that case, hares may be forced to change their diet especially since they may require higher energy to cope in winter. This would affect their microbiome as well. But again, no change was observed in bacterial diversity of the samples collected between January/February (summer) and June/July (winter) in my study.

4.4 Illumina MiSeq data accurate on a per base level but nanopore MinION data provided species-level resolution

In this study, total yield obtained from bases sequenced were more in nanopore than Illumina even though only 21 samples were sequenced in the former. However, average error rate in nanopore sequencer is higher than Illumina sequencer which limits the application of the former. Furthermore, a clear comparison of quality score cannot be performed between Illumina MiSeq and nanopore MinION system since the latter does not follow Phred quality score system (Q-score) unlike the former (Laver et al., 2015). For MiSeq system, the q-score set was 30, which means that the reads are at 99.9% accuracy (probability of incorrect base call is 1 in 1000). I set the threshold quality for MinION to seven as it is the recommended setting widely used which is probably lower than Illumina threshold quality. Even with the mentioned limitation, nanopore MinION system produced longer reads making it possible to access species-level information at lower accuracy. One advantage of longer reads is that the misclassification in few random nucleotides of a sequence will not result in identification of a completely different bacterial species. However, this limitation will make MinION

comparatively less suitable for single nucleotide polymorphism (SNP) based studies but not for microbiome estimation.

Comparing number of reads generated from Illumina MiSeq platform with nanopore MinION is arbitrary in this case since MiSeq produced reads with quality score (fastq) whereas MinION produced raw fast5 reads. Due to high error rate in MinION reads, it cannot be processed in QIIME 2 using DADA2 or Deblur pipeline. Hence, I used EPI2ME desktop agent to process and align the reads against the only database supported in 16S workflow, NCBI 16S database. Though MiSeq seemed to produce slightly more reads than MinION, majority of the reads (over 70%) were removed as chimeras after DADA2 (Table 4), thereby greatly reducing the output reads produced. MinION data processing stages did not include any chimera removal step and most of the reads were retained after filtering. This is further discussed in the limitations section.

Both Illumina and MinION produced similar taxonomical classifications, both showing presence and absence of same bacteria between hare and rabbit samples especially at phylum rank. This indicates that both the techniques are suitable and adequate for amplicon sequencing of 16S rRNA gene.

However, some bacteria were identified in one technique but not in the other. This difference could be due to difference in amplicon length for sequencing. Nanopore reads contains information of the entire 16S rRNA gene (~1500 bp) whereas Illumina reads contain information of only V3-V4 (~460 bp) hypervariable region and hence can classify to different taxonomies (Shin et al., 2018). This can be confirmed by truncating the length of MinION reads to match MiSeq reads and then look for similarity between both datasets. Furthermore, each region of 16S rRNA gene is best suitable to identify certain species of bacteria; V4 region cannot identify every bacterial taxa due to higher degree of sequence conservation compared to other regions (Chakravorty et al., 2007). Since each region has its own limitation and since entire 16S rRNA region cannot be sequenced in Illumina MiSeq sequencer, I opted for sequencing V3-V4 region. Additionally, efficiency of a primer pair to bind to a bacterial sequence will vary and hence, may not always result in identification of taxa (Thijs et al., 2017). Moreover, PCR bias could also affect sequencing result since each sample has a mix of several bacteria present (multi-template samples) and thus may result in unequal amplification resulting in incorrect information on the abundance (Kanagawa, 2003; Acinas et al., 2005). Illumina MiSeq library preparation included two PCR amplification stages and hence the chances of PCR bias is more in this sequencing platform than in nanopore MinION platform. Hence, PCR-free approaches are more preferred to produce robust conclusions in sequencing based studies.

These factors could have contributed together, resulting in observed difference between the taxonomy in MiSeq and MinION dataset.

Another reason for difference in taxonomy between nanopore and Illumina platform could be due to difference in database used. Currently, QIIME 2 is only compatible with two 16S database; Greengenes and SILVA. Hence, I used SILVA 16S database for classifying MiSeq reads whereas EPI2ME 16S analysis workflow I used, classified MinION reads against NCBI 16S database. This difference in database may have classified some taxa differently. For example, difference in taxonomic nomenclature between *Ruminococcus 1* / *Ruminococcus 2* in SILVA and *Ruminococcus* in NCBI database. Also, SILVA groups certain phyla together as superphylum (phylum *Patescibacteria* in Illumina data) which is not observed in NCBI classification. Moreover, I also tried using Greengenes database to process MiSeq data which produced slight difference in taxonomy as opposed to taxa identified through SILVA database. However, for this research, I only analysed results produced from SILVA database since Greengenes database was outdated and therefore would be less effective.

4.5 Limitations

4.5.1 Sample size

One of the limitation in this study was the small sample size. I found the rabbit kitten sample (rabbit-6) to have low diversity with over-representation of bacterial families *Peptostreptococcaceae* and *Akkermansiaceae.* The former is a family comprising several pathogenic bacteria while the latter comprises beneficial bacteria. However, due to limit in sample size (kitten (n) =1), it was difficult to understand the microbiome of kitten and whether this observation correlates with diseased conditions. Also, the observed differences between diversity of each rabbit sample with respect to age, lactation and month of sample collection could have been understood better if the sample size of rabbits was larger than twelve. However even with a smaller sample size, I was able to clearly differentiate between the microbiome of sympatric hares and rabbits.

4.5.2 Sample collection and storage may have an impact on quality of DNA extracted

The samples used in this study were collected in the year 2016. When I extracted the DNA from faecal pellet and ran on agarose gel electrophoresis, it appeared to have been degraded. This could imply either improper sample collection procedure or poor preservation of animals/tissues post killing which could have degraded the DNA. This in turn could have affected the microbial diversity/composition as well. A previous research had identified that DNA degradation produced a bias in bacterial community analysis although it was only minor compared to bias due to PCR (Krehenwinkel et al., 2018). Hence, degraded DNA may not have a major impact on microbiome analysis but in future, proper sample collection and storage protocol must be followed to generate more accurate results. Another possible explanation is that hares and rabbits are hind-gut fermenters who depend on coprophagy to receive required nutrients. The bacteria in the soft faeces are hence digested again. This may result in repeated exposure of bacteria to pH of stomach and lytic enzymes leading to DNA degradation.

4.5.3 Contamination in reagents

I used reagent-only controls to estimate bacterial contamination in reagents and no-template control (NTC) to estimate contamination that may occur during amplification and sequencing. Reagent only control, especially hare control, seemed to be contaminated with DNA from the hare samples possibly due to aerosol contamination. This may not have a major impact on the results since the observed results between Illumina and nanopore sequencing systems were similar. In addition to this, there were several phyla of bacteria that were uniquely present in the reagent controls which was absent in the samples. The only phyla present in NTC was phylum *Proteobacteria* which may not have been due to contamination from sample DNA. A previous study had revealed that this phylum of bacteria was a common contaminant in DNA extraction kits and laboratory reagents (Salter et al., 2014; Laurence et al., 2014). Careful laboratory practises must be followed to prevent cross-contaminations in future especially in culture independent techniques. Careful interpretation of the result is also necessary.

4.5.4 Semi-quantitative 16S analysis approach

Amplicon sequencing such as 16S analysis is a semi-quantitative approach; it cannot discriminate between live and dead bacteria present (Kong, 2011). Dead bacteria usually interfere with the analysis as it does not give accurate representation of gut microbiome since most of the bacteria are killed by acids present in the stomach. Therefore, it is important that the dead bacteria are removed especially when processing clinical samples. One way to know whether the bacteria is alive or dead is to reconstruct their genome using 16S gene sequences. Knowing the gene sequences or proportion of metabolites detected, will help in estimating the viability of bacteria (Emerson et al., 2017).

Furthermore, 16S sequencing analysis is only limited to most bacterial and archaeal identification. Hence, it does not give an estimate of the total microbial diversity of faecal samples. To overcome this, whole genome sequencing can be done which can estimate the overall microbial composition. However, every sequencing strategy has disadvantages. Though limited to only bacterial identification, 16S sequencing is still better than culture based studies which limits identification due to varying growth preferences of bacteria.

4.5.5 Chimeric reads in Illumina sequencing platform

Over 70% of my MiSeq reads was discarded as chimeras after DADA2. Chimeric reads are formed when two or more sequences are incorrectly joined together. This must be removed prior to downstream analysis to ensure no incorrect sequences are classified as novel. There are few possible explanations why most of the reads were discarded as chimeras here. Hypervariable regions in 16S rRNA gene chosen for amplification can influence chimera formation; 16S V1,V2 and V3 region is the longest and hence longer extension time with fewer PCR cycles are required (Shin et al., 2014; Qiu et al., 2001). Since I included a V3 region in my study, perhaps longer extension time was required. Additionally, type of DNA polymerase used and the input DNA concentration for PCR could also influence formation of chimeras (Porazinska et al., 2012). It is worthwhile to modify the extension time and increase the input DNA concentration prior to amplifying longer regions such as V3 region in future studies. Though majority of the reads were eliminated as chimeras, there were sufficiently enough reads left to perform downstream analysis. No chimera removal step was included when processing

MinION data but there is a possibility that chimeras may be present in this dataset as well. Since MiSeq results mostly matched with MinION results, chimeras may not have a major influence in this study.

Additionally, I observed the cluster density of the sequenced MiSeq data to be lower (619 ± 23 K/mm^2) than expected (1200-1400 K/mm^2). This could be due to improper quantification or low input amount of DNA library. Hence, in future studies different library loading amount must be tried to achieve expected cluster density. Though the quality of data is better at low cluster density, it results in low data output. But since I had only 24 samples sequenced, the output data generated was sufficiently enough for downstream analysis.

4.5.6 Difference in database used may result in taxonomical difference

As mentioned before, difference in database resulted in difference in bacterial composition as observed between Illumina MiSeq and nanopore MinION datasets. This difference was mostly due to difference in nomenclature of some bacteria at family/genus level between two databases and hence difference in observed composition of some bacteria. Even with this limitation, the bacteria identified between both the platforms/databases were comparatively similar.

Furthermore, a previous study in the year 2017, identified SILVA 16S database to be larger than NCBI 16S database (Balvočiūtė and Huson, 2017). However both these databases were large enough to eliminate false positive results.

4.6 Troubleshooting

Initially my research was focused on estimating the microbiome of liver, duodenum and faecal pellet of 64 hare samples using both amplicon and whole genome sequencing technology. These 64 hare samples were collected from two different regions in Australia; Victoria and ACT. Since liver and duodenum are fairly sterile organs with limited bacterial diversity and composition, 16S PCR product (targeting entire ~1500 bp of 16S rRNA gene) only showed a faint band on agarose gel electrophoresis (Appendix Supplementary Figure 1-A and 1-B). However, faecal samples were rich in bacteria as observed by presence of strong band in PCR product when checked using agarose gel. Hence, I shifted my research towards studying only

the faecal microbiome of hare and rabbit samples collected from a single region in ACT. Due to time constraints and unavailability of extra MinION flow cells, I performed only amplicon sequencing in this study.

4.7 Future scope

In this study, I analysed only the 16S rRNA gene which is limited to detection of only bacteria and archaea. Whole genome sequencing (WGS) is therefore necessary to study bacteria, fungi, virus and other micro-organisms present. This approach will give us an estimate of total microbiome present in faecal samples of hares and rabbits. Rabbit haemorrhagic disease virus 2 (RHDV2) is a common biocontrol agent used to reduce rabbit population in Australia. This virus is known to infect and subsequently result in death of both hares and rabbits (Mahar et al., 2018). Hence in future, it would be interesting to study how this virus affect bacterial diversity in these lagomorphs using WGS or amplicon sequencing approach.

In this study, I identified several bacteria that were associated with pathogenic conditions such as from the phylum *Chlamydia*. Since the number of sequences associated with this phylum was low and also because it was present only in few samples, it is not possible to confirm its presence. Another bacterium that is considered a pathogen is from the species *Paeniclostridium sordelli* (*Peptostreptococcaceae*) which was present comparatively more in kitten sample. The presence of these bacteria must be confirmed by culturing and/or by designing species-specific primer. By confirming the presence of these bacteria, it is possible to know whether hares and rabbits are reservoirs to these pathogens and whether they can infect other animals and/or humans.

Additional information is necessary to understand how hares and rabbits behave in a sympatric environment. This information can further support various hypothesis about variation in diet between these two lagomorphs which leads to variation in bacterial diversity. A bigger sample size from various geographic locations in the future will aid in more accurate conclusion of how lactation, pregnancy and season affects the microbiome. This together with WGS data will further improve our understanding about the microbiome in sympatric hares and rabbits.

4.8 Conclusion

The faecal microbiome gives an estimate of the gut microbiome which is largely dependent on the type of food we eat. Mammal's diet may be influenced by several parameters such as age, sex, season and lactation which in turn influences the microbiome. This is the first study that compared the faecal microbiome of sympatric hares and rabbits in Australia. The observed results showed that there is significant difference between the microbiome of sympatric rabbits and hares indicating difference in behaviour and diet. Furthermore, diversity consistently varied within different rabbit sample which could indicate that rabbits may have preference over a wide range of plant source as opposed to hares who are selective feeders. On a broad scale, this could imply why rabbits successfully colonised and spread much faster than hares becoming a pest. However, further investigation is required to come to a robust conclusion. Future researches should therefore include larger sample sizes to confirm how various parameters affect the microbiome of hares and rabbits in a natural sympatric environment.

Additionally, I also compared two popular sequencing approach currently available; Illumina MiSeq and Nanopore MinION sequencer. The choice of sequencing platform for microbiome studies depends on the cost based on number of samples, time and desired sequencing output (short or long read sequencing). The short-read sequencing strategy provides more accurate results with limited information while long read sequencing provides more information at less accuracy. Only few years ago long read sequencing was introduced and ever since then, it has been constantly updated to improve the accuracy of results. Future improvements may fix these issues completely but until then, both Illumina and Oxford's Nanopore are suitable for amplicon based microbiome analysis each with their own advantages and disadvantages.

References

ACINAS, S. G., SARMA-RUPAVTARM, R., KLEPAC-CERAJ, V. & POLZ, M. F. 2005. PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and environmental microbiology,* 71**,** 8966-8969.

ALEKSEYEV, Y. O., FAZELI, R., YANG, S., BASRAN, R., MAHER, T., MILLER, N. S. & REMICK, D. 2018. A Next-Generation Sequencing Primer-How Does It Work and What Can It Do? *Academic pathology,* 5**,** 2374289518766521-2374289518766521.

ALLISON, M. J., MAYBERRY, W. R., MCSWEENEY, C. S. & STAHL, D. A. 1992. Synergistes jonesii, gen. nov., sp.nov.: A Rumen Bacterium That Degrades Toxic Pyridinediols. *Systematic and Applied Microbiology,* 15**,** 522-529.

ALTBÄCKER, V., HUDSON, R. & BILKÓ, Á. 1995. Rabbit-mothers' Diet Influences Pups' Later Food Choice. *Ethology,* 99**,** 107-116.

ANA POPOVIC, C. B., PAULINE W. WANG, DAVID S. GUTTMAN, WIEGER VOSKUIJ, MICHAEL E. GRIGG & PARKINSON, R. H. J. B. A. J. 2018. Design and application of a novel twoamplicon approach for defining eukaryotic microbiota.

ANHÊ, F. F., PILON, G., ROY, D., DESJARDINS, Y., LEVY, E. & MARETTE, A. 2016. Triggering Akkermansia with dietary polyphenols: A new weapon to combat the metabolic syndrome? *Gut Microbes,* 7**,** 146-153.

ARDUI, S., AMEUR, A., VERMEESCH, J. R. & HESTAND, M. S. 2018. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Research,* 46**,** 2159-2168.

ARRAZURIA, R., PÉREZ, V., MOLINA, E., JUSTE, R. A., KHAFIPOUR, E. & ELGUEZABAL, N. 2018. Diet induced changes in the microbiota and cell composition of rabbit gut associated lymphoid tissue (GALT). *Scientific Reports,* 8**,** 14103.

BAHASSI, E. M. & STAMBROOK, P. J. 2014. Next-generation sequencing technologies: breaking the sound barrier of human genetics. *Mutagenesis,* 29**,** 303-310.

BAHRNDORFF, S., ALEMU, T., ALEMNEH, T. & LUND NIELSEN, J. 2016. The Microbiome of Animals: Implications for Conservation Biology. *International Journal of Genomics,* 2016**,** 7.

BALVOČIŪTĖ, M. & HUSON, D. H. 2017. SILVA, RDP, Greengenes, NCBI and OTT - how do these taxonomies compare? *BMC genomics,* 18**,** 114-114.

BELZER, C. & DE VOS, W. M. 2012. Microbes inside—from diversity to function: the case of Akkermansia. *The Isme Journal,* 6**,** 1449.

BESSER, J., CARLETON, H. A., GERNER-SMIDT, P., LINDSEY, R. L. & TREES, E. 2018. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clinical Microbiology and Infection,* 24**,** 335-341.

BIRD, P., MUTZE, G., PEACOCK, D. & JENNINGS, S. 2012. *Damage caused by low-density exotic herbivore populations: The impact of introduced European rabbits on marsupial*

herbivores and Allocasuarina and Bursaria seedling survival in Australian coastal shrubland.

BLEIDORN, C. 2016. Third generation sequencing: technology and its potential impact on evolutionary biodiversity research. *Systematics and biodiversity,* 14**,** 1-8.

BOKULICH, N. A., KAEHLER, B. D., RIDEOUT, J. R., DILLON, M., BOLYEN, E., KNIGHT, R., HUTTLEY, G. A. & GREGORY CAPORASO, J. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome,* 6**,** 90.

BOLYEN, E., RIDEOUT, J. R., DILLON, M. R., BOKULICH, N. A., ABNET, C., AL-GHALITH, G. A., ALEXANDER, H., ALM, E. J., ARUMUGAM, M., ASNICAR, F., BAI, Y., BISANZ, J. E., BITTINGER, K., BREJNROD, A., BRISLAWN, C. J., BROWN, C. T., CALLAHAN, B. J., CARABALLO-RODRÍGUEZ, A. M., CHASE, J., COPE, E., DA SILVA, R., DORRESTEIN, P. C., DOUGLAS, G. M., DURALL, D. M., DUVALLET, C., EDWARDSON, C. F., ERNST, M., ESTAKI, M., FOUQUIER, J., GAUGLITZ, J. M., GIBSON, D. L., GONZALEZ, A., GORLICK, K., GUO, J., HILLMANN, B., HOLMES, S., HOLSTE, H., HUTTENHOWER, C., HUTTLEY, G., JANSSEN, S., JARMUSCH, A. K., JIANG, L., KAEHLER, B., KANG, K. B., KEEFE, C. R., KEIM, P., KELLEY, S. T., KNIGHTS, D., KOESTER, I., KOSCIOLEK, T., KREPS, J., LANGILLE, M. G. I., LEE, J., LEY, R., LIU, Y.-X., LOFTFIELD, E., LOZUPONE, C., MAHER, M., MAROTZ, C., MARTIN, B. D., MCDONALD, D., MCIVER, L. J., MELNIK, A. V., METCALF, J. L., MORGAN, S. C., MORTON, J., NAIMEY, A. T., NAVAS-MOLINA, J. A., NOTHIAS, L. F., ORCHANIAN, S. B., PEARSON, T., PEOPLES, S. L., PETRAS, D., PREUSS, M. L., PRUESSE, E., RASMUSSEN, L. B., RIVERS, A., ROBESON, I. I. M. S., ROSENTHAL, P., SEGATA, N., SHAFFER, M., SHIFFER, A., SINHA, R., SONG, S. J., SPEAR, J. R., SWAFFORD, A. D., THOMPSON, L. R., TORRES, P. J., TRINH, P., TRIPATHI, A., TURNBAUGH, P. J., UL-HASAN, S., VAN DER HOOFT, J. J. J., VARGAS, F., VÁZQUEZ-BAEZA, Y., VOGTMANN, E., VON HIPPEL, M., WALTERS, W., WAN, Y., WANG, M., et al. 2018. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ Preprints,* 6**,** e27295v2.

BUGLE, C. & RUBIN, H. B. 1993. Effects of a nutritional supplement on coprophagia: A study of three cases. *Research in Developmental Disabilities,* 14**,** 445-456.

BULL, M. J. & PLUMMER, N. T. 2014. Part 1: The human gut microbiome in health and disease. *Integrative Medicine: A Clinician's Journal,* 13**,** 17.

CALLAHAN, B. J., MCMURDIE, P. J. & HOLMES, S. P. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The Isme Journal,* 11**,** 2639.

CALLAHAN, B. J., MCMURDIE, P. J., ROSEN, M. J., HAN, A. W., JOHNSON, A. J. A. & HOLMES, S. P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods,* 13**,** 581-583.

CHAKRAVORTY, S., HELB, D., BURDAY, M., CONNELL, N. & ALLAND, D. 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of microbiological methods,* 69**,** 330-339.

CLARRIDGE, J. E., 3RD 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clinical microbiology reviews,* 17**,** 840-862.

COBOS, A., DE LA HOZ, L., CAMBERO, M. I. & ORDOÑEZ, J. A. 1995. Chemical and fatty acid composition of meat from Spanish wild rabbits and hares. *Zeitschrift für Lebensmittel-Untersuchung und Forschung,* 200**,** 182-185.

CONLON, M. A. & BIRD, A. R. 2014. The impact of diet and lifestyle on gut microbiota and human health. *Nutrients,* 7**,** 17-44.

COWAN, D. P. 1987. Aspects of the Social Organisation of the European Wild Rabbit (Oryctolagus cuniculus). *Ethology,* 75**,** 197-210.

DALLE ZOTTE, A. 2014. Rabbit farming for meat purposes. *Animal Frontiers,* 4**,** 62-67.

DAVID, L. A., MAURICE, C. F., CARMODY, R. N., GOOTENBERG, D. B., BUTTON, J. E., WOLFE, B. E., LING, A. V., DEVLIN, A. S., VARMA, Y., FISCHBACH, M. A., BIDDINGER, S. B., DUTTON, R. J. & TURNBAUGH, P. J. 2013. Diet rapidly and reproducibly alters the human gut microbiome. *Nature,* 505**,** 559.

DE LANNOY, C., DE RIDDER, D. & RISSE, J. 2017. The long reads ahead: de novo genome assembly using the MinION [version 2; peer review: 2 approved]. *F1000Research,* 6.

DESANTIS, T. Z., HUGENHOLTZ, P., LARSEN, N., ROJAS, M., BRODIE, E. L., KELLER, K., HUBER, T., DALEVI, D., HU, P. & ANDERSEN, G. L. 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology,* 72**,** 5069-5072.

DITCHFIELD, J., MEADS, E. & JULIAN, R. 1960. Tularemia of muskrats in Eastern Ontario. *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique,* 51**,** 474-478.

DIVINCENTI, L., JR. & REHRIG, A. N. 2016. The Social Nature of European Rabbits (Oryctolagus cuniculus). *Journal of the American Association for Laboratory Animal Science : JAALAS,* 55**,** 729-736.

EARL, J. P., ADAPPA, N. D., KROL, J., BHAT, A. S., BALASHOV, S., EHRLICH, R. L., PALMER, J. N., WORKMAN, A. D., BLASETTI, M., SEN, B., HAMMOND, J., COHEN, N. A., EHRLICH, G. D. & MELL, J. C. 2018. Species-level bacterial community profiling of the healthy sinonasal microbiome using Pacific Biosciences sequencing of full-length 16S rRNA genes. *Microbiome,* 6**,** 190-190.

EMERSON, J. B., ADAMS, R. I., ROMÁN, C. M. B., BROOKS, B., COIL, D. A., DAHLHAUSEN, K., GANZ, H. H., HARTMANN, E. M., HSU, T., JUSTICE, N. B., PAULINO-LIMA, I. G., LUONGO, J. C., LYMPEROPOULOU, D. S., GOMEZ-SILVAN, C., ROTHSCHILD-MANCINELLI, B., BALK, M., HUTTENHOWER, C., NOCKER, A., VAISHAMPAYAN, P. & ROTHSCHILD, L. J. 2017. Schrödinger's microbes: Tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome,* 5**,** 86.

ESHAR, D. & WEESE, J. S. 2014. Molecular analysis of the microbiota in hard feces from healthy rabbits (Oryctolagus cuniculus) medicated with long term oral meloxicam. *BMC veterinary research,* 10**,** 62.

FAITH, D. P. 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation,* 61**,** 1-10.

FAYEMI, P. O., ONWUKA, C., ISAH, O., JEGEDE, A., ARIGBEDE, O. & MUCHENJE, V. 2011. Effects of mimosine and tannin toxicity on rabbits fed processed Leucaena leucocephala (Lam) De Wit. leaves. *African Journal of Agricultural Research,* 6**,** 4081-4085.

GEGGEL, L. 2016. What's the Difference Between Rabbits and Hares? *Live Science*

GIBB, J. A. 1990. The European rabbit Oryctolagus cuniculus. *JA Chapman e JEC Flux (a cura di), Rabbits, Hares and Pikas-Staus Survey and Conservation Action Plan. IUCN/SSC Lagomorph Specialist Group***,** 116-120.

GIDENNE, T. 2010. Recent advances in rabbit nutrition: emphasis on fibre requirements. A review. *World Rabbit Science,* 8**,** 23-32.

GOODWIN, S., MCPHERSON, J. D. & MCCOMBIE, W. R. 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics,* 17**,** 333.

GREEN, K., DAVIS, N. E. & ROBINSON, W. A. 2017. Diet of two fossorial herbivores in a seasonally snow-covered environment. *Australian Mammalogy,* 39**,** 169-177.

HACKLÄNDER, K., TATARUCH, F. & RUF, T. 2002. The Effect of Dietary Fat Content on Lactation Energetics in the European Hare (Lepus europaeus). *Physiological and Biochemical Zoology,* 75**,** 19-28.

HALLMAIER-WACKER, L. K., LUEERT, S., ROOS, C. & KNAUF, S. 2018. The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis. *Scientific reports,* 8**,** 6292.

HAQUE, S. Z. & HAQUE, M. 2017. The ecological community of commensal, symbiotic, and pathogenic gastrointestinal microorganisms - an appraisal. *Clinical and experimental gastroenterology,* 10**,** 91-103.

HERNÁNDEZ-MARTÍNEZ, C. A., MALDONADO HERRERA, J. A., MÉNDEZ-ZAMORA, G., HERNÁNDEZ-LUNA, C. E. & GUTIÉRREZ-SOTO, G. 2017. Enzymatic extract of Trametes maxima CU1 on productive parameters and carcass yield of rabbits. *Canadian Journal of Animal Science,* 97**,** 683-688.

HEWSON, R. 1989. Grazing preferences of mountain hares on heather moorland and hill pastures. *Journal of Applied Ecology***,** 1-11.

HIRAKAWA, H. 2001. Coprophagy in leporids and other mammalian herbivores. *Mammal Review,* 31**,** 61-80.

HU, Y. & SCHWESSINGER, B. 2018. Amplicon sequencing using MinION optimized from 1D native barcoding genomic DNA.

HUERTA-CEPAS, J., SERRA, F. & BORK, P. 2016. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Molecular biology and evolution,* 33**,** 1635-1638.

ILLUMINA. 2013. *16s metagenomic sequencing library preparation* [Online]. Available: http://sapac.support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf [Accessed 05/11 2019].

JANDA, J. M. & ABBOTT, S. L. 2007. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. *Journal of Clinical Microbiology,* 45**,** 2761-2764.

JOHNSON, E. L., HEAVER, S. L., WALTERS, W. A. & LEY, R. E. 2017. Microbiome and metabolic disease: revisiting the bacterial phylum Bacteroidetes. *Journal of molecular medicine (Berlin, Germany),* 95**,** 1-8.

KANAGAWA, T. 2003. Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of Bioscience and Bioengineering,* 96**,** 317-323.

KATONA, K., BÍRÓ, Z., HAHN, I., KERTÉSZ, M. & ALTBACKER, V. 2004. Competition between European hare and European rabbit in a lowland area, Hungary: a long-term

ecological study in the period of rabbit extinction. *FOLIA ZOOLOGICA-PRAHA-,* 53**,** 255-268.

KLINDWORTH, A., PRUESSE, E., SCHWEER, T., PEPLIES, J., QUAST, C., HORN, M. & GLÖCKNER, F. O. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic acids research,* 41**,** e1-e1.

KOLIADA, A., SYZENKO, G., MOSEIKO, V., BUDOVSKA, L., PUCHKOV, K., PEREDERIY, V., GAVALKO, Y., DOROFEYEV, A., ROMANENKO, M., TKACH, S., SINEOK, L., LUSHCHAK, O. & VAISERMAN, A. 2017. Association between body mass index and Firmicutes/Bacteroidetes ratio in an adult Ukrainian population. *BMC microbiology,* 17**,** 120-120.

KONG, H. H. 2011. Skin microbiome: genomics-based insights into the diversity and role of skin microbes. *Trends in Molecular Medicine,* 17**,** 320-328.

KREHENWINKEL, H., FONG, M., KENNEDY, S., HUANG, E. G., NORIYUKI, S., CAYETANO, L. & GILLESPIE, R. 2018. The effect of DNA degradation bias in passive sampling devices on metabarcoding studies of arthropod communities and their associated microbiota. *PLOS ONE,* 13**,** e0189188.

KRUSKAL, W. H. A. W., W ALLEN 1952. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association,* 47**,** 583-621.

KUIJPER, D., VAN WIEREN, S. & BAKKER, J. 2004. Digestive strategies in two sympatrically occurring lagomorphs. *Journal of Zoology,* 264**,** 171-178.

LAURENCE, M., HATZIS, C. & BRASH, D. E. 2014. Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. *PloS one,* 9**,** e97876-e97876.

LAVER, T., HARRISON, J., O'NEILL, P. A., MOORE, K., FARBOS, A., PASZKIEWICZ, K. & STUDHOLME, D. J. 2015. Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomolecular Detection and Quantification,* 3**,** 1-8.

LEGER, A. & LEONARDI, T. 2019. pycoQC, interactive quality control for Oxford Nanopore Sequencing.

LOZUPONE, C. A., HAMADY, M., KELLEY, S. T. & KNIGHT, R. 2007. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and environmental microbiology,* 73**,** 1576-1585.

MAHAR, J. E., HALL, R. N., PEACOCK, D., KOVALISKI, J., PIPER, M., MOURANT, R., HUANG, N., CAMPBELL, S., GU, X., READ, A., URAKOVA, N., COX, T., HOLMES, E. C. & STRIVE, T. 2018. Rabbit Hemorrhagic Disease Virus 2 (RHDV2;

GI.2) Is Replacing Endemic Strains of RHDV in the Australian Landscape within 18 Months of Its Arrival. *Journal of virology,* 92**,** e01374-17.

MARIAT, D., FIRMESSE, O., LEVENEZ, F., GUIMARĂES, V., SOKOL, H., DORÉ, J., CORTHIER, G. & FURET, J. P. 2009. The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC microbiology,* 9**,** 123-123.

MAULDIN, E. A. & PETERS-KENNEDY, J. 2016. Integumentary system. *Jubb, Kennedy & Palmer's Pathology of Domestic Animals: Volume 1*. Elsevier.

MCKINNEY, W. Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference, 2010. Austin, TX, 51-56.

METZKER, M. L. 2009. Sequencing technologies — the next generation. *Nature Reviews Genetics,* 11**,** 31.

MONAGHAN, P. & METCALFE, N. B. 1985. Group foraging in wild brown hares: effects of resource distribution and social status. *Animal Behaviour,* 33**,** 993-999.

MYERS, K. 1970. The rabbit in Australia. *Proceedings of the Advanced Study Institute on 'Dynamics of Numbers in Populations', Oosterbeek, The Netherlands'.(Eds PJ den Boer and GR Gradwell.) pp***,** 478-506.

NAKATSUJI, T., CHEN, T. H., BUTCHER, A. M., TRZOSS, L. L., NAM, S.-J., SHIRAKAWA, K. T., ZHOU, W., OH, J., OTTO, M., FENICAL, W. & GALLO, R. L. 2018. A commensal strain of Staphylococcus epidermidis protects against skin neoplasia. *Science Advances,* 4**,** eaao4502.

NOWACK, J., GIROUD, S., ARNOLD, W. & RUF, T. 2017. Muscle Non-shivering Thermogenesis and Its Role in the Evolution of Endothermy. *Frontiers in Physiology,* 8.

O' DONNELL, M. M., HARRIS, H. M. B., ROSS, R. P. & O'TOOLE, P. W. 2017. Core fecal microbiota of domesticated herbivorous ruminant, hindgut fermenters, and monogastric animals. *MicrobiologyOpen,* 6**,** e00509.

OHARA, S., SATO, T. & HOMMA, M. 1974. Serological Studies on Francisella tularensis, Francisella novicida, Yersinia philomiragia, and Brucella abortus. *International Journal of Systematic and Evolutionary Microbiology,* 24**,** 191-196.

PAIRET, M., BOUYSSOU, T. & RUCKEBUSCH, Y. 1986. Colonic formation of soft feces in rabbits: a role for endogenous prostaglandins. *American Journal of Physiology-Gastrointestinal and Liver Physiology,* 250**,** G302-G308.

PEREZ, F. & GRANGER, B. E. 2007. IPython: A System for Interactive Scientific Computing. *Computing in Science & Engineering,* 9**,** 21-29.

POLLARD, M. O., GURDASANI, D., MENTZER, A. J., PORTER, T. & SANDHU, M. S. 2018. Long reads: their purpose and place. *Human Molecular Genetics,* 27**,** R234-R241.

QUAST, C., PRUESSE, E., YILMAZ, P., GERKEN, J., SCHWEER, T., YARZA, P., PEPLIES, J. & GLÖCKNER, F. O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research,* 41**,** D590-D596.

RANJAN, R., RANI, A., METWALLY, A., MCGEE, H. S. & PERKINS, D. L. 2016. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and Biophysical Research Communications,* 469**,** 967-977.

REICHLIN, T., KLANSEK, E. & HACKLÄNDER, K. 2006. *Diet selection by hares (Lepus europaeus) in arable land and its implications for habitat management*.

RIAZ RAJOKA, M. S., SHI, J., MEHWISH, H. M., ZHU, J., LI, Q., SHAO, D., HUANG, Q. & YANG, H. 2017. Interaction between diet composition and gut microbiota and its impact on gastrointestinal tract health. *Food Science and Human Wellness,* 6**,** 121-130.

ROGNES, T., FLOURI, T., NICHOLS, B., QUINCE, C. & MAHÉ, F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ,* 4**,** e2584.

RUF, T. & ARNOLD, W. 2008. Effects of polyunsaturated fatty acids on hibernation and torpor: a review and hypothesis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology,* 294**,** R1044-R1052.

SALTER, S. J., COX, M. J., TUREK, E. M., CALUS, S. T., COOKSON, W. O., MOFFATT, M. F., TURNER, P., PARKHILL, J., LOMAN, N. J. & WALKER, A. W. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology,* 12**,** 87.

SCHAI-BRAUN, S. C., REICHLIN, T. S., RUF, T., KLANSEK, E., TATARUCH, F., ARNOLD, W. & HACKLÄNDER, K. 2015. The European Hare (Lepus europaeus): A Picky Herbivore Searching for Plant Parts Rich in Fat. *PLOS ONE,* 10**,** e0134278.

SHEET, I. S. F. 2011. *European wild rabbit (Oryctolagus cuniculus)* [Online]. Department of environment and energy, Australian Government. Available: https://www.environment.gov.au/biodiversity/invasive-species/publications/factsheet-european-wild-rabbit-oryctolagus-cuniculus [Accessed 02/05 2019].

SHIN, H., LEE, E., SHIN, J., KO, S.-R., OH, H.-S., AHN, C.-Y., OH, H.-M., CHO, B.-K. & CHO, S. 2018. Elucidation of the bacterial communities associated with the harmful microalgaeAlexandrium tamarenseandCochlodinium polykrikoidesusing nanopore sequencing. *Scientific Reports,* 8**,** 5323.

SIMON BAHRNDORFF, T. A., 2 TEMESGEN ALEMNEH,2 AND JEPPE LUND NIELSEN1 2016. The Microbiome of Animals: Implications for Conservation Biology. *International Journal of Genomics,* 2016.

SINGH, R. K., CHANG, H.-W., YAN, D., LEE, K. M., UCMAK, D., WONG, K., ABROUK, M., FARAHNIK, B., NAKAMURA, M., ZHU, T. H., BHUTANI, T. & LIAO, W. 2017. Influence of diet on the gut microbiome and implications for human health. *Journal of translational medicine,* 15**,** 73-73.

SPERANSKAYA, A. S., KHAFIZOV, K., AYGININ, A. A., KRINITSINA, A. A., OMELCHENKO, D. O., NILOVA, M. V., SEVEROVA, E. E., SAMOKHINA, E. N., SHIPULIN, G. A. & LOGACHEVA, M. D. 2018. Comparative analysis of Illumina and Ion Torrent high-throughput sequencing platforms for identification of plant components in herbal teas. *Food Control,* 93**,** 315-324.

STALDER, G., PINIOR, B., ZWIRZITZ, B., LONCARIC, I., JAKUPOVIĆ, D., VETTER, S., SMITH, S., POSAUTZ, A., HOELZL, F. & WAGNER, M. 2019. Gut microbiota of the European Brown Hare (Lepus europaeus). *Scientific reports,* 9**,** 2738.

STEFANINI, I. & CAVALIERI, D. 2018. Metagenomic Approaches to Investigate the Contribution of the Vineyard Environment to the Quality of Wine Fermentation: Potentials and Difficulties. *Frontiers in Microbiology,* 9.

STOTT, P. 2003. Use of space by sympatric European hares (Lepus europaeus) and European rabbits (Oryctolagus cuniculus) in Australia. *Mammalian Biology,* 68**,** 317-327.

STOTT, P. 2008. Comparisons of digestive function between the European hare (Lepus europaeus) and the European rabbit (Oryctolagus cuniculus): Mastication, gut passage, and digestibility. *Mammalian Biology,* 73**,** 276-286.

SUBRAMANIAN, S., HUQ, S., YATSUNENKO, T., HAQUE, R., MAHFUZ, M., ALAM, M. A., BENEZRA, A., DESTEFANO, J., MEIER, M. F., MUEGGE, B. D., BARRATT, M. J., VANARENDONK, L. G., ZHANG, Q., PROVINCE, M. A., PETRI, W. A., JR., AHMED, T. & GORDON, J. I. 2014. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature,* 510**,** 417-421.

THAT'SINTERESTING, A. 2016. *How Bunny Rabbits Took Over Australia — And What The Government Is Doing About It* [Online]. ati. Available: https://allthatsinteresting.com/bunny-rabbits-australia [Accessed 03/04 2019].

THIJS, S., OP DE BEECK, M., BECKERS, B., TRUYENS, S., STEVENS, V., VAN HAMME, J. D., WEYENS, N. & VANGRONSVELD, J. 2017. Comparative Evaluation of Four Bacteria-Specific Primer Pairs for 16S rRNA Gene Surveys. *Frontiers in microbiology,* 8**,** 494-494.

VALDES, A. M., WALTER, J., SEGAL, E. & SPECTOR, T. D. 2018. Role of the gut microbiota in nutrition and health. *BMJ,* 361**,** k2179.

VALENCAK, T. G., ARNOLD, W., TATARUCH, F. & RUF, T. 2003. High content of polyunsaturated fatty acids in muscle phospholipids of a fast runner, the European brown hare (Lepus europaeus). *Journal of Comparative Physiology B,* 173**,** 695-702.

VELASCO-GALILEA, M., PILES, M., VIÑAS, M., RAFEL, O., GONZÁLEZ-RODRÍGUEZ, O., GUIVERNAU, M. & SÁNCHEZ, J. P. 2018. Rabbit Microbiota Changes Throughout the Intestinal Tract. *Frontiers in Microbiology,* 9.

WEIRATHER, J. L., DE CESARE, M., WANG, Y., PIAZZA, P., SEBASTIANO, V., WANG, X.-J., BUCK, D. & AU, K. F. 2017. Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. *F1000Research,* 6**,** 100-100.

WICK, R. R., JUDD, L. M., GORRIE, C. L. & HOLT, K. E. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microbial genomics,* 3**,** e000132-e000132.

WICK, R. R., JUDD, L. M. & HOLT, K. E. 2018. Demultiplexing barcoded Oxford Nanopore reads with deep convolutional neural networks. *bioRxiv***,** 366526.

WU, X., ZHANG, H., CHEN, J., SHANG, S., WEI, Q., YAN, J. & TU, X. 2016. Comparison of the fecal microbiota of dholes high-throughput Illumina sequencing of the V3–V4 region of the 16S rRNA gene. *Applied Microbiology and Biotechnology,* 100**,** 3577-3586.

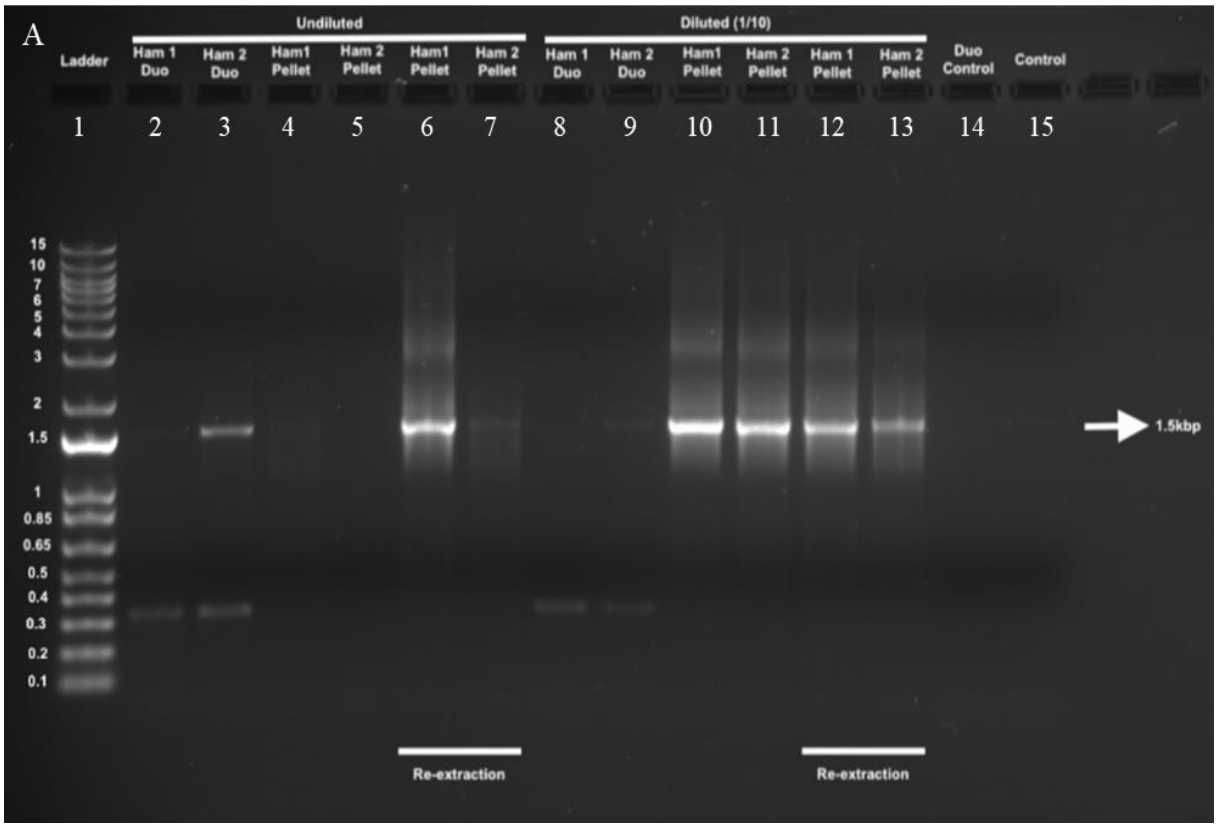
XU, Z. & KNIGHT, R. 2015. Dietary effects on human gut microbiome diversity. *British Journal of Nutrition,* 113**,** S1-S5.

ZENG, B., HAN, S., WANG, P., WEN, B., JIAN, W., GUO, W., YU, Z., DU, D., FU, X., KONG, F., YANG, M., SI, X., ZHAO, J. & LI, Y. 2015. The bacterial communities associated with fecal types and body weight of rex rabbits. *Scientific Reports,* 5**,** 9342.

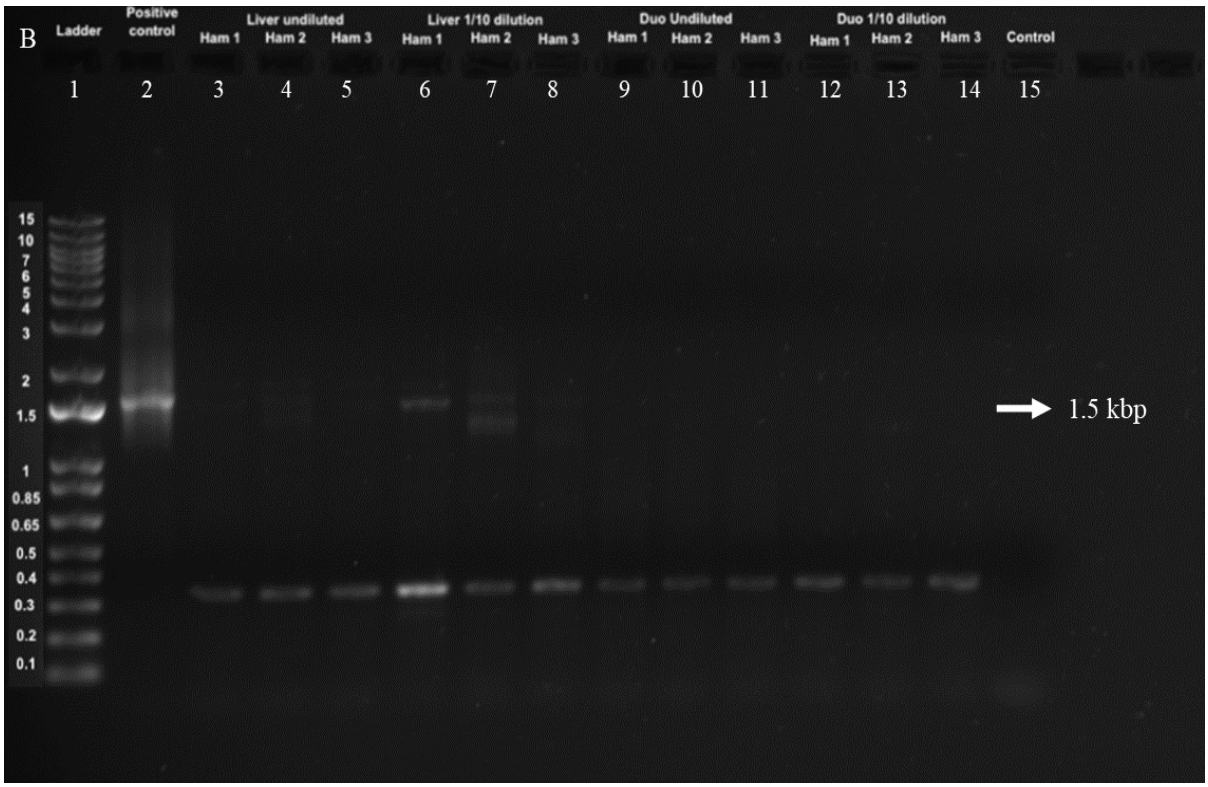
ZHANG, H., JAIN, C. & ALURU, S. 2019. A comprehensive evaluation of long read error correction methods. *BioRxiv***,** 519330.

ZHU, Y., WANG, C. & LI, F. 2015. Impact of dietary fiber/starch ratio in shaping caecal microbiota in rabbits. *Canadian journal of microbiology,* 61**,** 771-784.

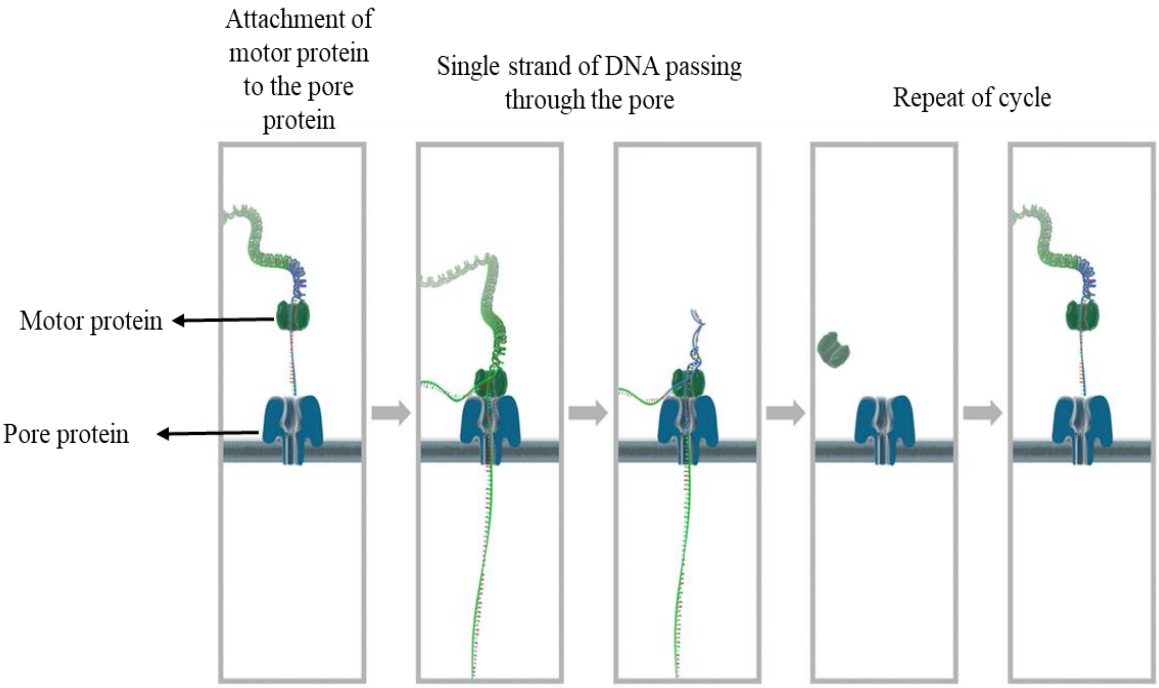
Appendix



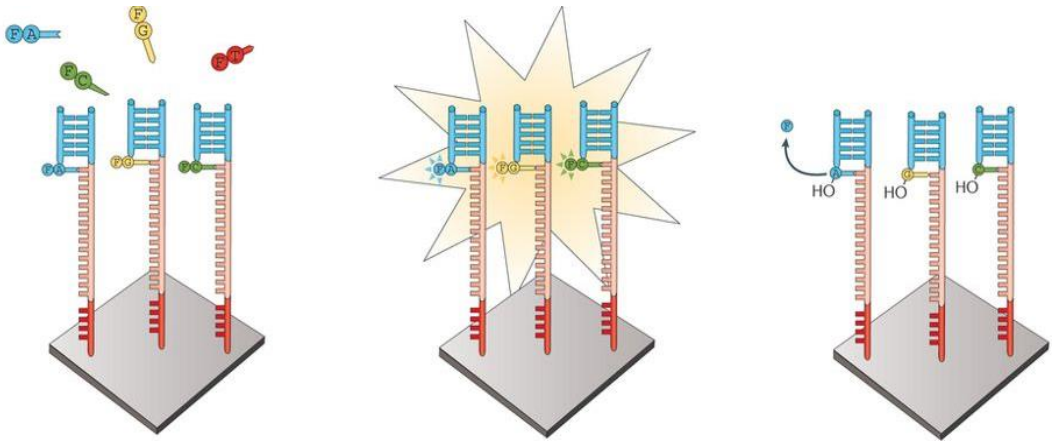
Supplementary Figure 1-A: 1% Agarose gel electrophoresis of two duodenum (duo) and two faecal (pellet) hare samples after 16S PCR (~1500 bp). Ham refers to sample from the region Hamilton in Victoria, Australia. Lane 2 to 7 were run using undiluted DNA and lane 8 to 13 were run using diluted DNA (1:10). Pellet samples in lane 4, 5, 10 and 11 contain low concentrated DNA. Lane 6, 7, 12 and 13 were run with re-extracted DNA (increasing initial tissue weight which increased DNA concentration). Only faint/no band were observed in duodenum samples (lane 2 and3; lane 8 and 9) whereas a strong band is observed in case of pellet samples (lane 6, 10, 11, 12 and 13) at 1.5 kbp. Faint bands are observed at ~300 bp. The duo control (lane 14) and control (lane 15) are negative controls made only with nuclease free water. The ladder (lane 1) used here is 1 kb Invitrogen plus ladder.



Supplementary Figure 1-B: 1% Agarose gel electrophoresis of three hare liver and duodenum samples after 16S PCR (~1500 bp). The samples are both diluted (1:10) and undiluted showing faint/no bands at 1.5 kbp and faint bands at ~300 bp. Lane 3,4 and 5 corresponds to undiluted liver sample and lane 6, 7 and 8 corresponds to diluted liver samples. Lane 9, 10 and 11 corresponds to undiluted duodenum samples and lane 12, 13 and 14 corresponds to diluted duodenum samples. The ladder used is Invitrogen 1 kb plus DNA ladder (lane 1) with faecal sample used as a positive control (lane 2) and negative control (lane 15) used was nuclease free water.



Supplementary Figure 2: Working of nanopore MinION sequencer. MinION sequencer consists of several pores which attaches to the adapter sequence of the DNA. When applied voltage, the motor protein present in the nanopore denatures the DNA into single strand which is then pulled through the pore one at a time. This movement of DNA through the pore disrupts the electric current which passes through the pore, producing a signal. These signals are measured and then decoded to retrieve the sequence of the DNA strand. Image modified from de Lannoy et al. (2017)



Attachment of labelled and blocked nucleotide to the complimentary base

Fluorescence emitted based on the base incorporated into the cluster

Detachment of labelled and blocked nucleotide from the base and beginning of new cycle

Supplementary Figure 3: Working of Illumina (sequencing by synthesis). Illumina works by sequence by synthesis approach where each terminally blocked and cleavable fluorophore labeled nucleotide present will be get incorporated into a fragment in each cluster. These clusters emit a certain color corresponding to the base incorporated when imaged by total internal reflection fluorescence (TIRF) microscopy. The nucleotide is then cleaved off and new cycle begins which will result in generation of output. Image modified from Goodwin et al. (2016)