Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.
[http://researchspace.auckland.ac.nz/feedback](http://researchspace.auckland.ac.nz/feedback)

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
Abstract

The kakapo is one of New Zealand’s iconic endemic birds, possessing an array of characteristics that are unique among avians. The kakapo is also critically endangered, and only survives through an intensive management effort by the Department of Conservation. In order to aid in conservation practices and decision making, detailed research has been conducted previously into the ecology, physiology, genetics and breeding behaviour of the kakapo. By contrast, the interactions between kakapo and the microorganisms associated with them have remained unexplored. In order to address this lack of knowledge regarding the kakapo microbiota, cultivation-based and state-of-the-art molecular techniques were employed with the overall aim of identifying the common microbes that inhabit the kakapo gastrointestinal tract and understanding the patterns that influence them.

Analysis of bacterial 16S rRNA genes revealed that the kakapo gut harbours a bacterial community of low taxonomic diversity, dominated by members of just two bacterial phyla (Proteobacteria and Firmicutes). Unexpectedly, the age of the bird played little role in shaping the kakapo microbiota, with both phylogenetic and functional data indicating no difference between age groups. By contrast, other studied factors did impact upon the microbiota, such as sampling location within the gut as well as antibiotic treatment and captivity. These analyses showed that although the kakapo microbiota is influenced by antibiotic treatment the effect is short lived, with the kakapo microbiota returning to a ‘normal’ state within weeks of release back into the wild. Genomic and metagenomic analyses reaffirmed the finding of low taxonomic diversity at broad phylogenetic levels, but showed that at fine-scale resolution (bacterial species and strains) the kakapo microbiota is much more diverse than inferred from traditional (16S rRNA-based) marker gene analysis. The functional potential of gut microbes was also examined, and was similar to that of other herbivore gut systems, despite the low taxonomic diversity.

This thesis represents the first concerted effort to study the kakapo microbiota and provides a foundation of knowledge for guiding future study and management practices.
Acknowledgements

While each data chapter of this thesis (Chapters 2 through 6) contain specific acknowledgements to the people who greatly helped in the sample acquisition, experimental design and eventual manuscripts, there are a number of people who deserve special acknowledgement. First and foremost, I would like to thank my supervisor Dr. Mike Taylor for offering me the opportunity to undertake this work. Mike has been a great supervisor, being supportive but also leaving me to independently learn and approach the thesis questions however I felt necessary (for better or worse). I would also like to thank my co-supervisor Professor Mick Clout, and advisor Dr Jacqueline Beggs who provided the contacts within the Department of Conservation that made this thesis possible.

Equally important, I would like to acknowledge Peter ‘Sneaky Pete’ Deines, who introduced me to the lab and taught me everything I needed in order to get started. Pete’s manner in the lab, wide knowledge on all subjects, and approachability provided a great environment when I was getting started. I’d also like to thank Dr Melissa Dsouza, who has been my good friend for the past 6 years through the undergraduate program, the postgraduate program, the PhD program and now in the workforce. The constant coffee breaks and friendly competition (which I usually won) made the years much more enjoyable. Rachel Simister and Kristi Biswas have also helped greatly with troubleshooting tips during some of the trickier parts of my thesis, and more coffee breaks.

There are also a number of people without whom this thesis would not exist. From the Department of Conservation I would like to acknowledge and thank Ron Moorhouse, Daryl Eason, Deidre Vercoe and Jo Ledington for their support of this project. I would also like to thank Errol Nye and Hayley Ricardo for letting me trail along behind them on Codfish Island with my trusty sampling tubes. I don’t think I slowed them down too much, but maybe they were just being kind. I would also like to thank the New Zealand Centre for Conservation Management, in particular John Potter, Richard Jakob-Hoff and Bethany Jackson for collecting and transporting samples whenever I needed them. From the University of Auckland I would like to thank Sue Skelly and Marlene Pillay, who shielded me from the majority of the administrative work that plagues us poor students wherever we go.
I would also like to acknowledge the many students and staff who have come and gone in the time I’ve been in the lab. In no particular order; Manpreet Dhami, Mike ‘Mini Mike’ Mawdsley, Sarah McRae, Anna Lau, Jack Lee, Dr Gavin Lear, Brett Wagner, Mike ‘MichaelMichael’ Hoggard, Rose Williams, Carmen Astudillo, Elahe Kia, Jieyun Wu and Dr Kelvin Lau for the good times, frequent beer and generally making the lab an awesome place to work in.

Finally, but by no means last, I’d like to thank my parents and family for their love and support during my project. Thanks must also go to my wife Siân, who has been a massive support and helped to keep me aware of the bigger world outside of the lab.
# Contents

Abstract

Acknowledgements

Contents

List of Figures and Tables

## 1 General Introduction

1.1 The kakapo .................................................. 2
  1.1.1 Ecology and diet .................................. 2
  1.1.2 Gastrointestinal anatomy and digestive strategy .... 5
  1.1.3 Conservation status ................................. 7
    1.1.3.1 Conservation islands ......................... 10
    1.1.3.2 Supplemental feeding ......................... 10
    1.1.3.3 Nest monitoring .............................. 10
    1.1.3.4 Hand-rearing chicks ......................... 11
  1.1.4 Rationale for kakapo microbiology ................. 11
 1.2 Microbial ecology and host symbiosis ................ 13
    1.2.1 Microbial symbiosis ............................ 14
    1.2.2 Bacterial symbiosis in the gastrointestinal tract 15
      1.2.2.1 Roles in digestion ......................... 15
      1.2.2.2 Roles in immune response .................. 16
    1.2.3 Bacterial symbiosis in birds ................... 17
    1.2.4 Influence of age ............................... 19
    1.2.5 Influence of captivity ........................ 19
    1.2.6 Influence of digestion strategy and gut structure 20
 1.3 Microbiological considerations of the kakapo ........ 21
 1.4 Microbiological methods .............................. 24
    1.4.1 The full-cycle rRNA approach .................. 24
    1.4.2 Genomics and metagenomics ..................... 26
    1.4.3 Cultivation-based approaches .................. 29
 1.5 Thesis objectives ...................................... 31
2 Gut Microbiome of the Critically Endangered New Zealand Parrot, the Kakapo (*Strigops habroptilus*)

2.1 Abstract .............................................. 33
2.2 Introduction ......................................... 34
2.3 Materials and Methods ............................ 37
2.3.1 Sampling ........................................... 37
2.3.2 DNA Extraction .................................... 37
2.3.3 PCR and clone library construction .......... 38
2.3.4 Phylogenetic Analysis ......................... 39
2.3.5 Determination of *Bacteroidetes* and *Archaea* sensitivity ... 39
2.4 Results .................................................. 41
2.4.1 Bacterial community composition within the kakapo gastrointestinal tract .................. 41
2.4.2 Determination of *Bacteroidetes* and *Archaea* sensitivity ... 46
2.5 Discussion ............................................ 48
2.6 Acknowledgments ................................... 52

3 Influence of Hand-rearing and Bird Age on the Faecal Microbiota of the Critically Endangered Kakapo

3.1 Abstract .............................................. 54
3.2 Introduction ......................................... 55
3.3 Materials and Methods ............................ 58
3.3.1 Sample collection .................................. 58
3.3.2 DNA extraction, PCR amplification and amplicon pyrosequencing ............................ 60
3.3.3 Bacterial community analysis .................. 61
3.4 Results .................................................. 63
3.4.1 The kakapo microbiota is of low diversity and dominated by *Firmicutes* and *Proteobacteria* .......................... 63
3.4.2 The microbial community structure does not vary between juvenile and adult individuals, but changes with time ... 68
3.4.3 Relative OTU abundance changes within the faecal microbiota ................................. 70
3.5 Discussion ............................................ 73
3.6 Acknowledgments ................................... 79

4 Quantifying the Impact of Storage Procedures for Faecal Bacteriotherapy in the Critically Endangered New Zealand Parrot, the Kakapo (*Strigops habroptilus*)

4.1 Abstract .............................................. 81
4.2 Introduction ......................................... 82
4.3 Materials and Methods ............................ 85
5 Genomic and Metagenomic Analysis of the Faecal Microbiota of the Critically Endangered Kakapo

5.1 Abstract ..................................... 95
5.2 Introduction ................................... 96
5.3 Materials and Methods ........................... 99
  5.3.1 Extraction and sequencing of genomic DNA .... 99
  5.3.2 Bioinformatic analysis ......................... 100
5.4 Results and Discussion .......................... 103
  5.4.1 The species-level diversity of the kakapo microbiota is much higher than that indicated by amplicon sequencing .... 103
  5.4.2 Functional potential of the kakapo microbiota .... 109
    5.4.2.1 The kakapo metagenome possesses cellulolytic and amylolytic potential .... 109
    5.4.2.2 Short-chain fatty acid utilisation .......... 114
    5.4.2.3 Nitrogen metabolism within the kakapo metagenome 117
  5.4.3 The kakapo faecal metagenome resembles that of a hindgut-fermenting herbivore .......... 117
  5.4.4 Challenges to interpreting the kakapo metagenome .... 120
  5.4.5 Concluding remarks ......................... 121
5.5 Acknowledgments ............................... 122

6 Characterising the Avian Gut Microbiota: Membership, Driving Influences and Potential Function ........................................ 123

6.1 Abstract ..................................... 124
6.2 Introduction ................................... 125
6.3 Materials and Methods ........................... 129
  6.3.1 Data acquisition and quality control ........... 129
  6.3.2 Correlating metadata to community structure .......... 130
  6.3.3 Functional prediction of gut microbiota .......... 131
6.4 Results and Discussion .......................... 132
  6.4.1 Taxonomic Classification of OTUs ................. 132
  6.4.2 Factors shaping the avian microbiota: study versus host ...... 132
  6.4.3 Factors shaping the avian microbiota: biological factors .... 138
    6.4.3.1 Factors shaping the kakapo microbiota: upper and lower GI .......... 140
List of Figures

1.1 An adult kakapo and six juvenile kakapo on a branch . . . . 4
1.2 Representative examples of avian gastrointestinal tract . . . 6
1.3 The avian microbiota at broad phylogenetic resolution . . . 18
1.4 An overview of the so-called 16S rRNA full-cycle approach 25

2.1 Broad taxonomic distribution of kakapo-derived 16S rRNA
gene sequences ........................................ 42
2.2 OTU-level distribution of kakapo-derived 16S rRNA gene
sequences ........................................... 43
2.3 Phylogeny of Firmicutes (lactic acid bacteria) 16S rRNA
gene sequences ....................................... 44
2.4 Phylogeny of Firmicutes (Clostridia) 16S rRNA gene se-
quences .............................................. 45
2.5 Phylogeny of Gammaproteobacteria 16S rRNA gene se-
quences ............................................. 47

3.1 Phylogenetic distribution of bacterial OTUs in the kakapo
faecal microbiota ...................................... 64
3.2 Changes in community structure in wild kakapo samples . . 69
3.3 Statistically significant changes in relative OTU abundance
during captivity ....................................... 72
3.4 Visualisation of bacterial community structure of the kakapo
and hoatzin hindgut microbiota ................... 74

4.1 Geometric mean colony counts, represented as colony-forming
units/gram of faecal material ......................... 88

5.1 Taxonomic classification of the kakapo faecal microbiota
using three analysis metrics ........................ 104
5.2 Comparisons of bacterial diversity between sampling ap-
proaches ............................................. 105
5.3 Comparison of *Escherichia* genomes from kakapo faeces with existing *Escherichia* genome sequences available in IMG ................................................................. 107
5.4 Emergent self-organising map of the differential coverage patterns in the kakapo microbiota .......................................................... 108
5.5 The propanoate metabolic pathway as represented in MEGAN with kakapo data highlighted ......................................................... 110
5.6 Abundance of COGs relevant to the carbohydrate load of the kakapo diet. ............................................................... 113
5.7 The butyrate metabolic pathway as represented in MEGAN with kakapo data highlighted ......................................................... 116
6.1 Relative proportion of clone-library OTUs represented in each study ............................................................ 136
6.2 Relative proportion of SRA OTUs represented in each study 137
6.3 Unweighted UniFrac distances for within- and between-study comparisons ................................................................. 141
6.4 Constrained Canonical Analysis of community structure based on clone-library sequence data .............................. 142
6.5 Constrained Canonical Analysis of community structure based on short-read sequence data .............................. 143
# List of Tables

1.1 Categories used to define the ‘critically endangered’ conservation status by the IUCN ........................................... 9

1.2 Average daily energy intake in a variety of parrots .......................... 22

1.3 An overview of genomic and metagenomic analysis .......................... 28

2.1 Sequences for primers used in study ............................................. 40

3.1 Sample sizes and mean individual age and additional notes taken at the time of sampling ........................................... 59

3.2 List of individuals sampled at each point in the juvenile survey .......................... 60

3.3 Common diversity and richness estimators calculated by using OTUs of ≥97% sequence similarity .......................... 66

3.4 Differences in core kakapo microbiota based on different partitioning of samples ........................................... 67

3.5 Statistically significant changes in OTU abundance between sample groups of interest ........................................... 70

4.1 Identity of each isolate, based on full-length 16S rRNA gene sequence ........................................... 89

5.1 Summary of publicly available metagenome samples used for this study ........................................... 102

5.2 The relative frequency of cellulolytic and amylolytic enzyme classes with the kakapo microbiome ........................................... 112

5.3 Results of ANOSIM testing between the functional community structure between metagenomes ........................................... 119

6.1 Published sequence data obtained from molecular analysis of avian samples ........................................... 128

6.2 Data obtained from clone-library based studies and the published study that reported the sequences ........................................... 134
6.3 Data obtained from short-read studies .................................. 135
6.4 Calculated fit of metadata factors to community distances
using PERMANOVA with linear model fitting .......................... 139
6.5 Summary of key findings in differences between predicted
metagenomes ........................................................................ 147
This thesis contains four chapters that have been published in international peer-reviewed journals, and one chapter that is in preparation for submission. These manuscripts were the result of a collaborative effort between myself and other researchers. The majority of the work was performed by myself, as per the University of Auckland regulations. Co-Authorship Forms have been submitted with the hardcopy of this thesis with the permission of the following authors;

1. **Chapter 2 co-authors**
   
   (a) Dr. Peter Deines  
   (b) Dr. Mike Taylor

2. **Chapter 3 co-authors**
   
   (a) Daryl Eason  
   (b) Dr. Mike Taylor

3. **Chapter 4 co-authors**
   
   (a) Dr. Peter Deines  
   (b) Dr. Mike Taylor

4. **Chapter 5 co-author**
   
   (a) Dr. Mike Taylor

5. **Chapter 6 co-author**
   
   (a) Dr. Mike Taylor
Chapter 1

General Introduction

The endemic New Zealand kakapo, *Strigops habroptilus*, is one of the world’s rarest bird species with only 126 remaining individuals. Recovery of the kakapo population is a major conservation priority, as the birds possess a slow breeding rate, lay small clutches of eggs and are susceptible to predation from stoats, dogs and other introduced species (Lloyd and Powlesland 1994; Elliott et al. 2001). While many aspects of kakapo biology have been studied in an effort to augment current conservation practices, a currently unexplored avenue of kakapo biology is the potential role of microorganisms in kakapo health. Microorganisms such as bacteria and archaea play many roles, both beneficial and antagonistic, in animal health and this area of biology is likely to contribute towards the continued conservation of the kakapo. This thesis aims to explore the diversity and function of bacteria associated with this iconic New Zealand parrot and to provide useful data in aiding future conservation efforts.
1.1 The kakapo

1.1.1 Ecology and diet

While once widespread across the New Zealand mainland, kakapo now live primarily on Codfish and Anchor Islands off the coast of southern New Zealand (Wood 2006), although a recent translocation program has seen a small population established on Little Barrier Island. The kakapo has a distinctive appearance, with pale green plumage and ‘whiskers’ that surround the beak (Figure 1.1). It is the world’s heaviest parrot, with males averaging slightly over 2 kg and females around 1.5 kg (Smith 1976; Powlesland et al. 1992; Eason et al. 2006). The kakapo is flightless, and the only known nocturnal parrot. At night, the kakapo forages for food, where it consumes a herbivorous diet that includes shrubs, leaves, bark, fruit and seeds from a wide variety of plants (Best 1984; Elliott et al. 2001; Atkinson and Merton 2006; Butler 2006; Wilson et al. 2006; Horrocks et al. 2008). Unlike ruminants, kakapo do not ingest a large quantity of fibrous plant material, instead using their beak to crush and strip their food in order to take as much nutritional content as possible without the fibre, which the bird discards as ‘chews’ (Oliver 1955; Butler 2006). Reproduction in the kakapo is tightly linked to mast fruiting of the native rimu (Dacrydium cupressinum) and pink pine (Halocarpus biformis). During the mast season, kakapo feed on the fruit of these plants (Cottam et al. 2006; Wilson et al. 2006), which act as a trigger for breeding (Cockrem 2006; Harper et al. 2006; Houston et al. 2007), although the precise mechanism by which breeding is triggered remains unknown (Harper et al. 2006). Supplemental feeding of kakapo does not trigger breeding, although it has proven successful in other wild bird populations (Robb et al. 2008). The case is not so clear in kakapo, as energy intake does not appear to be the limiting factor in reproduction and supplemental feeding does not increase the frequency of breeding (Atkinson and Merton 2006; Houston et al. 2007). Supplemental feeding can raise clutch size and contribute to survivorship among juveniles, although this is dependent on the exact composition of the feed provided. Differential impact of supplemental feeding on egg production has been observed, indicating that the frequency of kakapo reproduction is not dependent on energy requirement per se, but rather on the nutritional content of the diet (Elliott et al. 2001; Houston et al. 2007). Previous studies have reported a population containing up to four times as many males as females and analysis of kakapo
fossils indicates a historic ratio of two males per female bird (Trewick 1997). Speculation that the female diet influences the gender allocation of offspring has led to a new strategy being applied to the kakapo breeding program (Sutherland 2002) and appears to have been successful, with the current kakapo population evenly split between males and females (Kakapo Recovery Program 2014).

The kakapo is the only parrot to carry out lek breeding (Merton et al. 1984). Lek breeding is a mating system through which the breeding adults of a population gather and mating displays or competitions are performed in order to ‘win’ mates. This differs from other polygynous systems in that it is not a resource-based form of mate acquisition (Emlen and Oring 1977; Fiske et al. 1998). Rather than the traditional physical displays, male kakapo perform so-called ‘exploded lek’ or ‘exploded arena’ competition (Gilliard 1963; Oring 1982) in which the competition is based upon dispersed males calling for females. In order to perform this behaviour, male kakapo dig a bowl-like depression in the soil in a location that is acoustically favourable (Merton et al. 1984). During the nights, when they are most active, the male kakapo call to females with a distinctive booming call which can be heard up to five kilometres away. Interestingly, males do not feed or roost in their bowls but instead nest nearby, and will utilise the same bowl for many years, gradually creating new bowls linked together by earth tracks (Merton et al. 1984). Male kakapo maintain their bowls from October to April, with booming and mating typically occurring from January to March (Powlesland et al. 1992; Eason et al. 2006). Following mating, male kakapo take no part in the rearing of chicks, meaning females must leave the nest unattended while foraging (Powlesland et al. 1992), an aspect of kakapo behaviour that has likely contributed to chick mortality and the decline of the kakapo.
Figure 1.1: An adult kakapo (left) and six juvenile kakapo on a branch (right). Images provided courtesy of Jacqueline Beggs and Ron Moorhouse.
1.1.2 Gastrointestinal anatomy and digestive strategy

Owing to the extreme rarity of kakapo, the kakapo gastrointestinal (GI) tract has not been well studied, although the general avian GI tract is outlined in Figure 1.2. Following ingestion, food material is held in the crop, a muscular pouch analogous to the foregut found in some mammals. The crop functions as a storage pouch, holding excess food until it is ready to be digested. The crop has also been seen to act as a site of primary digestion, fermentation and nutrient liberation in some birds (Pritchard 1972; Grajal et al. 1989; Shetty et al. 1990), though this role is only considered nutritionally significant in the South American hoatzin. Food then passes through the proventriculus where it mixes with digestive fluids, then into the gizzard which grinds the ingested material into finer particles. The detailed structure of the gizzard is comprehensively described by (Svihus 2011). Some birds augment gizzard functionality by ingesting stones and grit to aid in breaking down hard food types (McLelland 1979; Feduccia 2011). Once ejected from the gizzard ingested material moves through the small and large intestines before being evacuated through the cloaca as waste. An additional organ, the ceca, is present in some birds at the junction between the small and large intestines. Although absent from parrots, the ceca are sites of intense nitrogen recycling, converting waste ammonia into amino acid precursors (Barnes 1972; Mead 1989; Vispo and Karasov 1997). The ceca also perform microbially-mediated carbohydrate fermentation and recover water from digesta (McNab 1973; Józefiak et al. 2004).
Figure 1.2: Representative examples of avian gastrointestinal tract. Ingested material passes from the crop (blue) to the gizzard (orange) via the proventriculus. Leaving the gizzard, ingested material passes through the small and large intestine. The ceca (green) lie at the border between these two organs. Figure adapted from Stevens and Hume (1998).
What little is known of the kakapo GI tract is based on dissection of opportunistically collected kakapo remains (Kirk et al. 1993). As expected for a parrot, kakapo lack ceca, which has given rise to the hypothesis that members of this species perform foregut fermentation (discussed in section 1.3). Due to their spatial proximity, it has been hypothesised that the enlargement of the crop associated with foregut fermentation would force a reduction in flight muscle mass (Morton 1978), which has been observed in the hoatzin. The kakapo crop accounts for approximately 7.5% of the bird’s body mass, compared to 14% in the hoatzin (Grajal et al. 1989), and this does not appear to have any diminishing influence on the ventral flight muscles of the kakapo (Livezey 1992). The most significant anatomical adaptation of the kakapo to its diet is the extensive development of the mandibles to enable the crushing of fibrous plant material. The kakapo beak is short and powerful, enabling the grasping and tearing of tough plant fibre. The palate and tongue are equipped with tough, keratinised surfaces to facilitate the crushing of plant material prior to ingestion, the mechanisms of which are outlined by (Kirk et al. 1993).

1.1.3 Conservation status

Despite the fact that kakapo were once spread throughout the New Zealand mainland (Wood 2006), the kakapo is currently ‘critically endangered’ according to the International Union for Conservation of Nature (IUCN) Red List (IUCN 2013b). This classification included 196 mammals, 198 avians and 164 reptiles at the last update (IUCN 2013a), and in order to be classified as critically endangered an organism must fit into any of the five categories provided in Table 1.1 (IUCN 2012). A major cause of the decimation of the kakapo population is their susceptibility to mammalian predators, which is a common theme among endemic New Zealand fauna. Due to the historical absence of these predators, endemic birds and reptiles have poor defence against carnivores such as dogs, cats and stoats (Lloyd and Powlesland 1994). While an inability to fly may have been advantageous to kakapo prior to human colonisation, it now leaves kakapo unable to escape predators and the visual camouflage that historically provided protection from the now-extinct giant eagles has no effect against stoats and dogs, which hunt largely by scent (Lloyd and Powlesland 1994; Powlesland et al. 2006).
In addition to predation, early Polynesian and European settlers hunted the kakapo for meat and feathers while simultaneously deforesting their habitats. The loss of habitat did not have the direct effect on survival that might be expected, but because settlers preferentially took fertile land for farming the kakapo were forced into areas of poor soil quality and climate meaning that the plants they rely on for food were less productive (Houston et al. 2007). Following the extreme decline in kakapo numbers, a major conservation initiative, the Kakapo Recovery Program (Department of Conservation 2012), was established with the aim of stabilising the decline of kakapo and encouraging breeding to restore population numbers (Lloyd and Powlesland 1994; Clout and Merton 1998; Elliott et al. 2001; Clout 2006; Eason and Moorhouse 2006; Eason et al. 2006). Conservation efforts can broadly be categorised into four fields, albeit with some overlap. These approaches are designed to reduce predation or otherwise improve survivorship among the population.
**Table 1.1:** The categories used to define the ‘critically endangered’ conservation status of an animal as detailed by the IUCN. The kakapo qualifies as critically endangered due to meeting categories A.1, A.2, A.3, C, and E.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
</table>
| A        | Occurring over less than 100 km² and two of:  
           |   1. Severe habitat fragmentation or existing at just one location  
           |   2. Decline in extent of occurrence, area of occupancy, area/extent/quality of habitat, number of  
               locations/subpopulations, or amount of mature individuals  
           |   3. Extreme fluctuations in extent of occurrence, area of occupancy, number of locations/subpopulations,  
               or amount of mature individuals  
| B        | As above, but less than 10 km² (used to show differing levels of severity)  
| C        | Declining population of less than 250 mature individuals and either:  
           |   1. A decline of 25% over one generation or three years  
           |   2. Extreme fluctuations, or over 90% of mature individuals in a single subpopulation, or no more than  
               50 mature individuals in any one subpopulation  
| D        | Numbers less than 50 mature individuals  
| E        | At least 50% chance of going ‘Extinct in the Wild’ over three generations or 10 years |
1.1.3.1 Conservation islands

One of the best known conservation practices used with many of New Zealand’s endangered fauna is the establishment and maintenance of predator-free islands to facilitate animal breeding. As well as protection from predators, conservation islands aim to remove competitor species such as weka and possums (Lloyd and Powlesland 1994). Translocation of the kakapo started in 1974, with the first kakapo arriving on Codfish Island in 1987 (Lloyd and Powlesland 1994). By translocating kakapo to small, remote islands birds can be easily monitored and such close observation means that sickness in an individual is quickly detected, and the bird removed to prevent spread of disease, as well as to administer treatment. In addition, issues in the breeding program such as infertile males or inbreeding can be easily detected and prevented within the island population so as to prevent undesirable mating attempts. More crucially, between 1981 and 1991, 43% of kakapo chicks hatched in the wild succumbed to predation while in the nest (Powlesland et al. 1992; Elliott et al. 2001). Maintenance of predator-free breeding islands removes the threat of predation, both preventing loss of chicks and improving survivorship among adult kakapo (Lloyd and Powlesland 1994; Clout and Merton 1998).

1.1.3.2 Supplemental feeding

As mentioned previously, supplemental feeding has some influence improving clutch sizes although it does not act as a catalyst for breeding. Supplemental feeding provides additional benefit in ensuring the health and wellbeing of the foraging parent and reducing the time of foraging expeditions (Elliott et al. 2001). Although based on small sample sizes, observation of cold damage to eggs has been cautiously attributed to prolonged exposure resulting from foraging (Elliott et al. 2001). Diet supplementation in the kakapo is an ongoing area of research, with continuous changes being made to the feed in order to better maintain bird health.

1.1.3.3 Nest monitoring

Despite the lack of predators on conservation islands, nest monitoring is still of great benefit to kakapo survivorship. Importantly, kakapo eggs (or chicks) can be
incubated by conservation workers while the mother forages, reducing the previously mentioned risk of mortality due to exposure. Monitoring also provides the opportunity for health checks on new-born chicks to ensure that they are developing within healthy guidelines. In addition, while mammalian predators are absent from the islands, petrels and other sea birds will occasionally attempt to colonise kakapo nests and kakapo fight against these invasions aggressively (Hill and Hill 1987; Higgens 1999; Powlesland et al. 2006). While these encounters are of no real danger to the adult kakapo, the struggle has been known to leave eggs damaged or juvenile kakapo injured (Daryl Eason, personal communication). Nest monitoring allows for intervention by conservation staff, either scaring away invading petrels before the adult kakapo discovers them or through proactively removing petrels from the vicinity of nesting females.

1.1.3.4 Hand-rearing chicks

Chicks rescued from the wild, or eggs raised in captivity, require a period of surrogacy to carry the chick through its parent-dependent stage, until it is ready to fledge. The first attempts at hand-rearing kakapo were unsuccessful (Sibley 1994; Elliott et al. 2001), though immense effort has since been expended to ensure that captivity mimics the natural environment as closely as possible. These efforts even go so far as to replicate the periods of cold which chicks undergo while the mother is foraging, and inoculating the chicks with faeces collected from healthy adults (Eason and Moorhouse 2006). These approaches have proven to be of great success in recent breeding years, with the 2011 breeding season producing 11 healthy juveniles despite a period of illness in each bird.

1.1.4 Rationale for kakapo microbiology

While the application of microbiology to kakapo conservation may not seem intuitive at first, there is a real need to understand the ecology of the microbes that interact with the kakapo. From a veterinary standpoint, the knowledge of kakapo pathogens is limited (Brangenberg et al. 2003; Gartrell et al. 2005) and this alone would be a reason to at least understand the typical kakapo-associated commensal bacterial communities. If it is assumed that there is a beneficial or at least neutral community of microbes associated with the kakapo, then this raises the issue of
Chapter 1. General introduction

whether or not this community can establish itself during hand-rearing/captivity. If not, a cultivated repository of kakapo-derived bacteria could be critical in establishing an alternative option to the current practice of faecal spiking of the captive kakapo diet (Eason and Moorhouse 2006), through the development of a kakapo ‘probiotic’ formula. In addition to the potential of promoting the establishment of a ‘normal’ kakapo microbiota, probiotics represent an alternative route through which veterinary concerns and infections could be treated. Probiotics have proven successful in combating medical conditions in humans (Shornikova et al. 1997; Van Niel et al. 2002; Lin et al. 2005; Fang et al. 2009) (discussed in Chapter 4) and represent a practical application of microbiological knowledge to a conservation system.

There are also sound academic reasons to study the microbiology of kakapo. Along with the kea and kaka, the kakapo represents the deepest branching parrot lineage, the *Strigopidae*. Studying the microbiota of the kakapo potentially provides a case study for determining the impact of host phylogeny on shaping the vertebrate microbiota. In addition, there has been speculation that the kakapo performs a physiological process known as foregut fermentation. Discussed in detail later in this chapter, foregut fermentation is the process of utilising a microbial community in the foregut to digest and ferment food before it passes into the gut. This process is reasonably common in mammals, for example cows and sheep. It is much rarer in avians, however, and the kakapo has been suggested as one of few birds that may perform this process (Morton 1978; Lopez-Calleja and Bozinovic 2000). As such, it is of paramount importance to study and understand the kakapo microbiota in order to address important practical and academic questions regarding the identity and function of the microbes associated with healthy members of this unique parrot species.
1.2 Microbial ecology and host symbiosis

The diversity and ubiquity of single-celled bacterial and archaeal organisms is unlike that of any other branch of life. Microbial ecology refers to the study of these microorganisms within the context of their environments and interactions with other organisms using traditional ecological paradigms. The concept of a ‘species’ is surprisingly hard to define (De Queiroz 2005), although utilising the term to define a group of individuals who are reproductively isolated from others serves as a working definition in eukaryote biology. In microbiology the use of asexual reproduction as a means to reproduce makes this definition irrelevant, so proxy measurements for species need to be applied. The traditional standard for species delineation within the microbial world has been through the use of DNA:DNA hybridisation experiments, measuring how well the DNA of one bacterial isolate pairs with that of another (Ward and Fraser 2005), or through commercially available phenotypic identification kits. A simpler manner of taxonomically classifying bacteria was first reported in 1977 by Carl Woese, when sequencing the DNA of the small subunit ribosomal RNA (16S rRNA) gene revealed the existence of a previously unknown domain of life (now known as Archaea) (Fox et al. 1977; Woese and Fox 1977). While bacterial genomes are notoriously variable even within a species, ribosomes are present in all bacteria (Ludwig and Schleifer 1994) and are subject to evolutionary pressure to remain conserved. Within the conserved structure of the ribosome exist variable loops that are open to mutation. Due to this mixture of conservation and rapid mutation, the 16S rRNA gene provides an excellent proxy for defining a bacterial species (Woese 1987; Ludwig 2007), even though the concept of species is not entirely applicable to non-eukaryotes.

In the early 1990s 16S rRNA gene sequencing was applied to bacterial communities, as it provided the ability to detect bacteria that may have been present in an environment but were not amenable to cultivation. It quickly became apparent to microbiologists that traditional cultivation approaches vastly under-represented the true species diversity found in many environments, sometimes by a factor of up to 1,000 (Fry 2000; Leadbetter 2003). Since this revelation the microbiology community has, to a degree, moved away from cultivation-based approaches to the study of marker genes, including the 16S rRNA gene, in an effort to capture the ‘true’ microbial diversity in an environment. The advent of culture-independent methods has changed the landscape of microbial ecology, with environments previously thought to be well understood now being re-examined to discover just
Chapter 1. General introduction

how diverse the microbial communities that inhabit them truly are. Molecular studies of a range of environments including soil (Rappé and Giovannoni 2003), human skin (Human Microbiome Jumpstart Reference Strains Consortium 2010), ‘extreme’ environments (Hetzer et al. 2007) and the sea (Venter et al. 2004) have led to the number of known species ballooning. In 1987 there were 12 recognised bacterial and archaeal phyla (Rappé and Giovannoni 2003), whereas current 16S rRNA gene databases recognise anywhere between 41 (Cole et al. 2009) and 88 (DeSantis et al. 2006; Quast et al. 2013).

1.2.1 Microbial symbiosis

There is debate over the exact definition of a ‘symbiosis’, but the term is usually used to define a relationship that is generally physiological, between different organisms that interact closely (Goff 1982). A well-known example of symbiosis is the interaction between pollinating insects (such as bees) and flowering plants. The insect feeds on the flower nectar and in return inadvertently pollinates the plant as it moves from flower to flower. Such an interaction has developed to the point that if one organism (for example, the bee) was to disappear overnight, the dependent flowering plant would die within a generation as it would have no mechanism to reproduce. In the case of microbial relationships with larger eukaryotic organisms, the relationship can be more intimate. Symbiosis can be broken into two distinct groupings: ectosymbiosis, the symbiotic relationship between two organisms that live in physical contact, and endosymbiosis, in which one of the organisms is internalised by the other. A well-studied example of endosymbiosis is the nitrogen-fixing root nodules found in legumes. Nitrogen-fixing bacteria are internalised to specific endosomes in the legume root tissue to provide nitrates derived from atmospheric nitrogen gas to the host plant (Fauvart and Michiels 2008). In the case of highly evolved relationships, the host organism may provide specialised living areas for the microbes upon which it relies, as seen in the legume example with endosomes. In some cases, often involving bacteria/insect symbioses, a reduction in genome size is often observed in at least one of the symbionts as duplicate, redundant genes are lost in one organism but provided for in the other (Sällström and Andersson 2005; Nakabachi et al. 2006; McCutcheon et al. 2009a,b). In less specialised cases, such as those that involve the gut or skin, the microbes may be merely adapted to the specific microenvironment that they inhabit and no special attention is paid by the host. These are known as
facultative symbiotic relationships, and arise when the presence of the symbiotic partner confers an advantage upon the organism of interest but is not essential for survival (Kikuchi 2009).

1.2.2 Bacterial symbiosis in the gastrointestinal tract

There are many beneficial roles played by microbial cells in the GI tract of animals. These include the degradation of ingested plant fibre and liberation of nutrients inaccessible to the host, prevention of pathogen colonisation, stimulation of the adaptive immune system and promotion of angiogenesis throughout the stomach and intestine.

1.2.2.1 Roles in digestion

Links between microbial community structure and increased energy harvest from food have been demonstrated for a wide range of organisms by a variety of indirect techniques, such as functional gene analysis of bacteria found in human and mouse guts (Gill et al. 2006; Turnbaugh et al. 2006), cDNA analysis of bacterial genes retrieved from termite guts (Ohkuma 2008), and correlating increased energy harvest with particular bacterial communities in chickens (Torok et al. 2008). In controlled murine models, these effects can be shown in a much more direct fashion, with the use of gnotobiotic (including germ-free) rodents as controls in experiments that demonstrate the role of bacteria in beneficially regulating gene expression in the liver (Björkholm et al. 2009; Meini et al. 2009), promoting angiogenesis in the small intestine (Stappenbeck et al. 2002), and regulating fat deposition (Bäckhed et al. 2004). Microbes in the GI tract often produce vitamins required by the host (Uphill et al. 1977; Hill 1997), and are important producers of short-chain fatty acids (SCFAs), a major source of energy in eukaryotes. In the anaerobic hindgut of all studied eukaryotes, SCFAs are produced by bacterial fermentation and these are of nutritional importance to their host, with up to 99% of bacterially-produced SCFAs being absorbed by the host (Von Englehardt et al. 1989; Scheppach 1994). SCFAs such as lactate and acetate are primary sources of energy in many vertebrates (Brady et al. 1979; Ogata et al. 1982), and carry additional benefits to the host such as preventing pathogen colonisation (Van Der Wielen et al. 2000; Maslowski et al. 2009; Fukuda et al. 2012).
The key microbes that perform these roles are highly adapted to the host environment, with the community being shaped by host-specific factors in a range of organisms (Van Der Wielen et al. 2002; Khachatryan et al. 2008; Vaishnava et al. 2011). Microbes transplanted into a new gnotobiotic host species provide significantly reduced benefits to the new host (Meinl et al. 2009), indicating a degree of specialisation to the original host conditions. Beyond mice and rats in controlled experiments, studies of the human GI tract-associated microbiota have revealed links between bacteria and increased energy harvest from food (Turnbaugh et al. 2009). These observations regarding energy harvest have been made in a range of other organisms such as termites (Ohkuma 2008), chicken (Torok et al. 2008), sheep (Chaucheyras-Durand et al. 2010), cattle (Dodd et al. 2011), and pandas (Zhu et al. 2011). In addition, certain metabolites of use to the human body are only detected in individuals colonised by particular bacterial communities (Li et al. 2008).

### 1.2.2.2 Roles in immune response

The ability of the gut microbiota to protect the host from pathogens also has a significant impact on the wellbeing of the host. Far from just out-competing pathogens in the gut, in some organisms the gut microbiota has adapted to regulate the host immune system, essentially repressing the immune response in the presence of beneficial microbes (Brisbin et al. 2008; Hill et al. 2010). Stimulation of the host immune system by GI tract microbes has also been recognised in response to both viral and bacterial challenge (Brisbin et al. 2008; Ivanov et al. 2009; Ichinohe et al. 2011). The importance of these bacteria has become so well recognised that probiotic supplements are commonly used to shape the gut microbiota of farmed chickens in order to defend against common pathogenic infections (Morishita et al. 1997; Willis and Reid 2008).

Development of gut-associated lymphoid tissues is increased in conventionally raised mice over germ-free individuals (Hudson and Luckey 1964; Hooijkaas et al. 1984), but the reaction against bacteria in the gut is comparatively weak compared to that found in individuals maintained in a sterile state (Albiger et al. 2007). This has led to speculation regarding recognition pathways between commensal bacteria and their host. Bacteria associated with the gut of the host organism can, in some cases, cause down-regulation of immune-related signalling molecules, such
as NF-κB, in order to prevent an immune response to the symbionts (O’Hara and Shanahan 2006). This type of activity is not one-sided in favour of the bacteria. It has recently been shown, using transgenic knockout mice, that particular chemicals secreted by the stomach epithelium act to ward off gut-associated bacteria from colonising the stomach surface (Vaishnava et al. 2011). Should these bacteria make contact with the stomach wall, the adaptive immune system is triggered and the host body attacks the gut microbes, but by maintaining a zone of no colonisation both organisms obtain benefits from the relationship.

### 1.2.3 Bacterial symbiosis in birds

Symbiotic relationships with bacteria have been found in all animals studied to date. Although the microbiology of birds has generally been less studied than other that of other vertebrates, certain areas, such as chicken farming, are active pockets of research in both pathogen defence and probiotic usage (Zhang et al. 2011). Increasingly, the relationship between wild birds and bacterial symbionts is attracting attention within microbial ecology, with a substantial increase in microbial community-based studies in the last 10 years. A general picture of the avian microbiota is thus starting to emerge, with the bacterial phyla *Proteobacteria, Firmicutes, Actinobacteria* and *Bacteroidetes* commonly found in a range of birds (Figure 1.3). In new-born chickens, the cell density of colonising bacteria in the cecum can reach as high as $10^{10}$ cells per gram in the first day of life (Yegani and Korver 2008), and numerous structuring factors can influence this community of cells.
### Figure 1.3: A summary of the avian microbiota at broad phylogenetic resolution.

Findings reported cover a range of methods (cultivation, 16S rRNA gene sequencing and microarray) and therefore should not be taken as authoritative. For the hoatzin, the total number of detected phyla is currently over 40, although the rare tail is omitted from this table. For all other birds, all known bacterial taxa are displayed (Damere et al. 1979; Pacheco et al. 2004; Blanco et al. 2006; Gibbs et al. 2007; Gong et al. 2007; Janiga et al. 2007; Godoy-Vitorino et al. 2008; Lu et al. 2008; Lu and Domingo 2008; Matsui et al. 2010; Xenoulis et al. 2010; Wienemann et al. 2011).

<table>
<thead>
<tr>
<th></th>
<th><strong>Proteobacteria</strong></th>
<th><strong>Firmicutes</strong></th>
<th><strong>Actinobacteria</strong></th>
<th><strong>Bacteroidetes</strong></th>
<th><strong>Spirochaetes</strong></th>
<th><strong>Fusobacteria</strong></th>
<th><strong>Verrucomicrobia</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α</td>
<td>β</td>
<td>γ</td>
<td>δ</td>
<td>ε</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capercaillie</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chicken</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gull</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ostrich</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hoatzin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green-Rumped Parrotlet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellow-Headed Blackbird</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpine Accentors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mealy Parrot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue-Yellow Parrot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red-Green Parrot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red Kite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1.2.4 Influence of age

Although study of this area has been restricted to a small number of organisms thus far, a factor that consistently influences the microbiota of an animal is the age of the organism. Differences in the microbiota of avians (Van Der Wielen et al. 2002; Scupham 2007; Gong et al. 2008; Godoy-Vitorino et al. 2010; Van Dongen et al. 2013), mice and humans (Vaahtovuo et al. 2001; Lozupone et al. 2012; Schloss et al. 2012; Yatsunenko et al. 2012) have been experimentally demonstrated. While the juvenile microbiota is variable compared to that of the adults, it has been shown that the microbiota of juveniles eventually stabilises and converges with that of the adult (Palmer et al. 2007; Godoy-Vitorino et al. 2010; Yatsunenko et al. 2012; Van Dongen et al. 2013), although the time this takes varies by host and lifestyle. The differences between the juvenile and adult occur in a manner similar to ecological succession, with rapid colonisers eventually being displaced by the more ‘standard’ adult microbial community profile (Lu et al. 2003; Palmer et al. 2007). Due to the fact that the juvenile microbiota membership and succession pattern differ by host (Lu et al. 2003; Scupham 2007), and captivity programs influence this succession (Stanley et al. 2013), it is difficult to apply general rules to the establishment of the GI microbiota in a particular host species, other than to say that age is a factor likely to play a role in any study system.

1.2.5 Influence of captivity

In light of the significant human influence on the natural kakapo life cycle, it is important to consider whether or not there exists the potential for displacement or modification of the gut microbiota of the kakapo as a result of human intervention. Most literature to date pertaining to other birds consists of studying either wild birds or farmed birds, but with no overlap between the two. A recent paper published by Xenoulis et al. (2010) aimed to specifically track the differences between wild and captive species of parrots. Common molecular microbiology techniques were used to characterise cloacal samples taken from three species of parrots, collected from both the wild and from breeders. The results of this study indicated two things: that human influence modifies the gut community composition, and that in general there is an increase in diversity as a result of this influence (Xenoulis et al. 2010). Significant changes in the avian microbiota due to captivity have also been reported for the capercaillie (Wienemann et al. 2011) and
Chapter 1. General introduction

20

the turkey (Scupham et al. 2008). Interestingly, although analysis of the turkey microbiota showed changes in community structure due to human influence, the overall diversity of the communities did not differ. The effect of captivity is highly variable, with not only differences between wild and captive populations reported, but also differences between captive populations (Stanley et al. 2013).

1.2.6 Influence of digestion strategy and gut structure

Like any environment, the GI tract is not homogeneous, but consists of regions of varying physiological parameters, such as oxygen availability, surface association and pH. These changes in local environment mean that for microbes the gut provides a gradient of differing niches to occupy. Within chickens and turkeys, the bacterial communities that reside in the GI tract are known to change between regions of the gut (Bjerrum et al. 2006; Gong et al. 2007; Torok et al. 2008). Equally important in shaping the microbiota of the gut is the digestion strategy employed by the host organism (Muegge et al. 2011; Godoy-Vitorino et al. 2012). As previously mentioned, many herbivores employ a process known as foregut fermentation, through which ingested plant material is held in an enlarged organ (the crop in birds, or rumen in mammals) where bacterial activity degrades the fibrous plant content and performs the SCFA fermentations usually seen in the hindgut. This practice is common in herbivorous mammals (Muegge et al. 2011) but essentially unique among birds to the South American hoatzin (Opisthocomus hoazin) (Grajal et al. 1989). Recent molecular comparisons between the microbes that inhabit the GI tract of hoatzin and that of cows have shown that the digestion strategy of the animal has a greater influence in shaping the microbiota than the host animal itself (Godoy-Vitorino et al. 2012). As the gut strategy of the kakapo is unknown, it is essential that if bacterial analysis of the kakapo is to be performed, the influence of gut location must be considered.
1.3 Microbiological considerations of the kakapo

The interplay between a host and its microbiota is a facet of kakapo research that has previously gone unnoticed, and many aspects of the kakapo make it an interesting animal to study for its microbiology alone. As previously mentioned, degradation of cellulose in the gut is difficult for most vertebrates and in all studied animals is aided by a community of microbes in the gut. Due to this globally common association it is reasonable to assume that the kakapo will also rely on a community of microbes to assist with the digestion of plant sugars. In the case of a kakapo, a further complicating matter is the uncertainty regarding which part of the gut houses the anaerobic fermentation process essential for extracting energy from food. In hindgut-fermenting birds, the primary site of fermentation is in finger-like appendages that sprout from the colon (Clench and Mathias 1995), however, these organs are absent in psittaciformes (parrots), the phylogenetic order that includes kakapo. The kakapo has occasionally been speculated to possess foregut fermentation due to its diet and high body mass (Clench and Mathias 1995; Lopez-Calleja and Bozinovic 2000). However, although larger than normal, kakapo crops are not significantly different from those of the kaka or kea, the kakapo’s two closest relatives (Kirk et al. 1993), and there are behavioural differences between the kakapo and the hoatzin. For example, the hoatzin became a focus of study due to the fact that it was able to fly on a folivorous diet that should not have been able to sustain flight. This implied that there was some mechanism in play that allowed the hoatzin to harvest more energy from the volume of ingested material than would be expected in a normal bird. In contrast, kakapo daily energy expenditure is extremely low, requiring as little as a quarter of that required by flying birds, and about half that of their flying relative, the kea (Table 1.2). In addition, kakapo chew their food excessively, swallowing juices and severely masticated soft plant tissue before spitting out the fibrous content in characteristic ‘chews’ (Oliver 1955; Horrocks et al. 2008). As such, while the hoatzin and ruminants may require intense microbial activity to perform the level of degradation necessary to liberate nutrients from their fibrous food, the kakapo is selectively feeding on a relatively high amount of sugar and nutrients compared to what is found in its food sources.
Table 1.2: **Average daily energy intake (DEI) in a variety of parrots.** DEI is given in raw values (based on averages obtained from their respective references), and normalised to body mass. For kakapo and kea, values represent daily energy expenditure. The study by Wolf et al. (2007) noted no change in body mass during the study, indicating that the DEI of studied avians is likely close to the expenditure.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Average Body Weight (g)</th>
<th>Daily Energy Intake (kJ / day)</th>
<th>DEI / g</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lorikeet</td>
<td>45</td>
<td>82.1</td>
<td>1.8</td>
<td>Wolf et al. (2007)</td>
</tr>
<tr>
<td>Lovebird</td>
<td>45</td>
<td>66.4</td>
<td>1.5</td>
<td>Wolf et al. (2007)</td>
</tr>
<tr>
<td>Rainbow Lorikeet</td>
<td>130</td>
<td>140.7</td>
<td>1.1</td>
<td>Wolf et al. (2007)</td>
</tr>
<tr>
<td>Cockatiel</td>
<td>90</td>
<td>120</td>
<td>1.3</td>
<td>Wolf et al. (2007)</td>
</tr>
<tr>
<td>Goldie’s Finch</td>
<td>47</td>
<td>86.8</td>
<td>1.8</td>
<td>Wolf et al. (2007)</td>
</tr>
<tr>
<td>Kea</td>
<td>1000</td>
<td>848</td>
<td>0.8</td>
<td>Bryant (2006)</td>
</tr>
<tr>
<td>Kakapo</td>
<td>2000</td>
<td>799</td>
<td>0.4</td>
<td>Bryant (2006)</td>
</tr>
</tbody>
</table>
Ultimately, the location within the gut is an important factor that must be considered when studying the microbiota of an animal. Given the scarcity of kakapo, sampling via dissection and removal of the GI tract is not feasible, which essentially limits sampling to faecal collections. For hindgut fermenters, or studies that track an individual through time, faeces are an accepted proxy for direct gut sampling, but with uncertainty surrounding the kakapo fermentation strategy faeces alone may not be sufficient for analysis. Swabbing of a surface is a commonly applied technique for sampling bacteria from a range of settings (Moreno et al. 2003; Blanco et al. 2006; Shawkey et al. 2007; Klomp et al. 2008; Xenoulis et al. 2010) and, despite its potential biases, is the only viable option for obtaining upper-GI tract samples from kakapo.

In addition to gut strategy and diet, the development of the juvenile microbiota is strongly influenced by interaction with the parent (Palmer et al. 2007; Dominguez-Bello et al. 2010; Godoy-Vitorino et al. 2010; Yatsunenko et al. 2012; Indrio et al. 2013; Penders et al. 2013; Khodayar-Pardo et al. 2014). For kakapo chicks that undergo hand rearing and captivity, the major avian transfer mechanism of regurgitation-based feeding is obviously not possible. In an attempt to introduce the appropriate complement of microbes to hand-reared chicks, the Department of Conservation supplements their food with faeces collected from adult kakapo (Eason and Moorhouse 2006). There is a range of potential problems with this approach, most critically the fact that the efficacy of this technique is unknown. Faecal samples collected from wild kakapo will have spent an unknown amount of time outside of the body, where they are subjected to a change in temperature (symbionts are most likely adapted to life at the body temperature of the bird, about 37°C) and may be contaminated by bacteria and fungi not associated with the kakapo prior to collection. This problem provides the biggest problem facing kakapo microbiology - without knowing what microbes comprise the healthy kakapo microbiota, it is impossible to know what microbes need to be screened out or tested for during routine operations. In addition, since faecal samples must be obtained opportunistically, once screened for pathogens they are frozen for later use. This process has an unknown effect on the microbial community. These samples are also not standardised, making evaluation of this technique impossible to objectively assess due to the variation in faecal composition and microbiota between samples.
1.4 Microbiological methods

1.4.1 The full-cycle rRNA approach

Culture-independent analyses have become the backbone of microbial ecology, and as such represent the first port of call for many studies. The full-cycle 16S rRNA approach is a term used to describe a basic workflow investigating the microbial community of a new environment (Figure 1.4). Following the extraction of genomic DNA from an environmental sample, a marker gene is amplified using the polymerase chain reaction (PCR). This marker gene is frequently the 16S rRNA gene, hence the name of the system, although others have been proposed, such as the RNA polymerase β-subunit (Mollet et al. 1997; Dahllöf et al. 2000; Case et al. 2007), 23S rRNA gene (Christensen et al. 1998) and DNA gyrase B subunit (Fukushima et al. 2002). The amplified gene of choice is then sequenced to determine its exact genetic code. The chemistry of DNA sequencing varies among the technologies used, with the traditional dideoxy chain-termination method (Sanger and Coulson 1975; Sanger et al. 1977) being supplanted by so-called ‘next-generation’ sequencing (NGS) technologies (Metzker 2010) in recent years. These newer sequencing methods provide orders of magnitude more sequence data for a fraction of the cost, but are generally limited to shorter sequence lengths compared to Sanger sequencing. For example, Sanger sequencing can provide read lengths of 1,000 basepairs (bp) while NGS methods advertise between 75 and 700 bp depending on the platform used (AllSeq 2014; Illumina Inc. 2014; Roche Diagnostics 2014; Life Technologies 2014).
Figure 1.4: An overview of the so-called 16S rRNA full-cycle approach. This approach provides a mechanism to iteratively refine the understanding of a microbial community through analysis of a particular environment at more targeted levels each time the cycle is repeated.
Once representative sequences are obtained from a microbial community the gene sequences can be analysed to infer the phylogeny of the individual microorganisms. The pooled sequences can be further analysed as a community, using traditional ecological diversity measures to identify similarities or differences in the community structure, with the aim of linking these differences back to environmental factors. Recent studies have employed sophisticated techniques to show that the physiological processes of an individual bacterium are dependent on the community in which it resides, not just its genetic potential (Berry et al. 2013). Although 16S rRNA gene-based approaches are limited to a strictly ‘who is there’ style of analysis, under some conditions functional prediction of bacterial lineages can be made from 16S rRNA gene sequence (Langille et al. 2013). Further, the results of 16S rRNA gene surveys can be iteratively refined with successive experiments to provide information about how communities change through time or between differing environments, or combined with in situ functional probing techniques (Andreasen and Nielsen 1997; Dahllöf 2002; Okabe et al. 2004; Ginige 2005).

1.4.2 Genomics and metagenomics

The term ‘genome’ was first coined by Hans Winkler in 1920 to refer to a haploid set of chromosomes (Winkler 1920), and by extension the coding and non-coding information contained within them. While this definition essentially holds true today, it has been unofficially expanded to refer to the genes and non-coding nucleic acid sequence of an organism including the non-chromosomal plasmids of bacteria and the genetic content of RNA-based viruses. In practice, the field of genomics aims to sequence and reconstruct the entire genome of an organism of interest and to then identify and annotate the genes and regulator sequences; in contrast to the single-gene sequencing techniques discussed previously which probe for the presence of a single, pre-defined gene.

At the turn of the century, genome projects were an expensive and laborious task, with only a handful of high profile projects being performed, such as the Human Genome Project (International Human Genome Sequencing Consortium 2001, 2004). As a biological technique, genome sequencing has undergone a massive increase in popularity due largely to the applicability of NGS to genome projects and the number of completed genome projects has drastically increased in the last decade, proportional to the decrease in sequencing cost facilitated by
the prevalence of NGS technology (Pagani et al. 2012; National Human Genome Research Institute 2013). In response to this increase in sequenced genomes the National Centre for Biotechnology Information (NCBI) now curates a database specifically for annotated genomes (Pruitt et al. 2005) in addition to its traditional gene/protein/nucleotide databases. This drastic increase in the amounts of sequenced DNA has driven the development of software capable of performing robust draft assemblies with minimal user intervention (Swain et al. 2012), and online analysis pipelines to centralise the computational resources required for processing (Aziz et al. 2008; Markowitz et al. 2012). Genomics has proven a valuable tool for microbiology as a method of generating hypotheses regarding the functional capability of isolates, which can swiftly be followed up with validation using classical techniques.
**Table 1.3:** An overview of the steps required for genomic and metagenomic analysis, as well as some of the current bioinformatic tools used for analysis. A major challenge in this field is that these tools do not usually work with each other natively, so the process of converting the output from a particular piece of software to the input of the next must be devised by the researcher.

<table>
<thead>
<tr>
<th>Step</th>
<th>Genome</th>
<th>Metagenome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Binning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CARMA (Krause et al. 2008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PhymmBL (Brady and Salzberg 2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sort-ITEMS (Monzoorul Hague et al. 2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PhylopythiaS (McHardy et al. 2007; Patil et al. 2012)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Assembly (de novo)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SOAPdenovo (Luo et al. 2012)</td>
<td>Meta-IDBA (Peng et al. 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CLC Workbench (CLC Genomics 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Assembly (reference)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AMOS (AMOS Consortium 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Post-assembly improvement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>iCORN (Otto et al. 2010)</td>
<td>IMAGE (Tsai et al. 2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSPACE (Boetzer et al. 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gene prediction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prodigal (Hyatt et al. 2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FragGeneScan (Rho et al. 2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MetaGeneMark (Zhu et al. 2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MetaGeneAnnotator (Noguchi et al. 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Functional classification</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(algorithms)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BLAST (Altschul et al. 1990)</td>
<td>USEARCH (Edgar 2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Functional classification</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(databases)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MEGAN (Huson et al. 2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SEED (Overbeek et al. 2005)</td>
</tr>
</tbody>
</table>
Chapter 1. General introduction

Just as the analysis of a single marker gene from a bacterium has been expanded to an environmental bacterial community, whole-genome sequencing in microbiology has been expanded into the field of metagenomics. The concept of metagenomics is simple, essentially to extract bulk DNA from an environmental sample, rather than an isolate, and to sequence the DNA of the community. From a practical perspective this approach has always been difficult to implement, partly due to the difficulty in achieving adequate sequencing throughput to yield sufficient coverage of all genomes in the sample. DNA extracted from a source is biased by the relative abundance of the microbes that inhabit the source and, due to the random nature of a PCR-free sequencing approach, a large number of sequences (frequently on the order of gigabases worth of DNA) is required in order to provide sufficient depth to completely cover the genomes of rarer organisms. Once sufficient DNA sampling is achieved, a greater challenge is how to interpret these data, and what to do with the outcome. The seminal work in this area was published in 2004, where researchers were able to separate and reconstruct the individual genomes comprising a simple metagenomic sample (Tyson et al. 2004). From these newly assembled genomes, functional predictions were made, making it possible to subsequently cultivate a previously uncultured bacterium (Tyson et al. 2005). From this point, metagenomics became a common technique for analysing microbial communities of increasing complexity. Although the technical challenges surrounding metagenomic analysis are greater than those surrounding simple genomic analysis, development in this area has progressed swiftly with a variety of tools developed to address the common issues of metagenomics (Table 1.3), with web-based platforms also freely accessible (Meyer et al. 2008; Markowitz et al. 2012).

1.4.3 Cultivation-based approaches

Although molecular techniques are a better tool for capturing total diversity in a system for which there is no prior knowledge, the importance of traditional cultivation techniques cannot be overlooked. Cultivation refers to the practice of growing microorganisms under controlled conditions, providing nutrient and mineral sources through either a solid or liquid growth, as well as allowing the control of other factors such as oxygen and acidity levels. This approach is an iconic technique in the public perception of science, using petri dishes to grow bacteria and fungi for research purposes. Cultivation generally takes place to
either obtain a pure strain of a particular microbe of interest, or to study the impact of a particular bioactive product on the growth patterns of organisms of interest (for example, the impact of an antibiotic on a pathogen). By strictly controlling the growth conditions of the microbes in culture, researchers can use bacterial cultivation to control all variables in the system except for the one of interest, allowing for more direct response to a particular change in conditions.

While 16S rRNA analysis may reveal greater diversity than a cultivation approach, it reveals little regarding the physiological capabilities of the bacteria detected other than inferring function from closely related species. While it is possible to probe bacterial communities in situ for specific functional capabilities (discussed previously) this can be a labour-intensive task as each potential compound must be tested independently. More general approaches exist, in the forms of genetic probing techniques (Zhang et al. 1992; Gelfand et al. 1996; Ahmed et al. 2007; MacAulay and Voet 2014) and analysis of metabolic profiles (Mashego et al. 2007). Such approaches are not without their own limitations, however. Any approach based on the presence of genes will only yield information about the potential abilities of the microbes present, as expression levels of the gene(s) remain unknown. Moreover, some genes that have only ever been identified in silico may have a different function from that which is attributed to them, or they may perform multiple functions. Metabolite analysis, as well as suffering from its own extraction biases, will be limited to the current flux state of the sample (which may be disturbed during the process of sample collection) and provide complex readouts that do not associate a particular metabolite to a particular organism.

Any system that is to have a practical application, be it industrial (Third et al. 2001; Ahn 2006), medicinal (Shornikova et al. 1997; Lin et al. 2005; Fang et al. 2009), or food preparation (Trivedi et al. 1986; Erten and Tanguler 2010), requires a physical sample that can be utilised. Although it is unlikely to be achieved during this thesis, the long-term goal of developing a probiotic for kakapo must also start with the cultivation of potentially desirable bacteria. In order to accurately assess the functional capabilities of potential kakapo symbionts, it is therefore essential to cultivate as many microbial species as possible, so that the functional capabilities can be tested directly.
1.5 Thesis objectives

The overall aim of this study is to perform the first comprehensive analysis of the kakapo microbiota, in terms of microbial community membership, function and the potential factors that influence the development of the juvenile kakapo microbiota. To achieve these broad aims, this thesis has been broken down into the following five specific objectives:

- Document the variation of the kakapo microbiota throughout the gastrointestinal tract and establish a baseline of microbes that are associated with healthy kakapo (Chapter 2).

- Test the influence of factors relating to the captivity of juvenile kakapo on the microbiota, and establish overall patterns within the development of the kakapo microbiota (Chapter 3).

- Evaluate the survivorship of faecal microbes during freezing, and test the impact of a cryo-protectant agent as a means of improving the current juvenile seeding practices (Chapter 4).

- Utilise a (meta)genomic approach to analyse the functional components of the kakapo microbiota and to compare these functions to those of animals with a known digestion strategy (Chapter 5).

- Compare the microbiota of kakapo to that of other avians, and attempt to understand the broad patterns that shape avian microbiota (Chapter 6).
Chapter 2

Gut Microbiome of the Critically Endangered New Zealand Parrot, the Kakapo (Strigops habroptilus)

David W. Waite¹, Peter Deines¹,², Michael W. Taylor¹

¹. Centre for Microbial Innovation, School of Biological Sciences, The University of Auckland, Auckland, New Zealand
². Institute of Natural Sciences, Massey University, Auckland, New Zealand

Published in the Peer-Reviewed Journal PLoS ONE
April 18th, 2012

PLoS ONE, Volume 7(4), Page e35803
DOI: 10.1371/journal.pone.0035803
2.1 Abstract

The kakapo, a parrot endemic to New Zealand, is currently the focus of intense research and conservation efforts with the aim of boosting its population above the current ‘critically endangered’ status. While virtually nothing is known about the microbiology of the kakapo, given the acknowledged importance of gut-associated microbes in vertebrate nutrition and pathogen defense, it should be of great conservation value to analyze the microbes associated with kakapo. Here we describe the first study of the bacterial communities that reside within the gastrointestinal tract (GIT) of both juvenile and adult kakapo. Samples from along the GIT, taken from the choana (≈throat), crop and faeces, were subjected to 16S rRNA gene library analysis. Phylogenetic analysis of >1000 16S rRNA gene clones, derived from six birds, revealed low phylum-level diversity, consisting almost exclusively of Firmicutes (including lactic acid bacteria) and Gammaproteobacteria. The relative proportions of Firmicutes and Gammaproteobacteria were highly consistent among individual juveniles, irrespective of sampling location, but differed markedly among adult birds. Diversity at a finer phylogenetic resolution (i.e. operational taxonomic units (OTUs) of 99% sequence identity) was also low in all samples, with only one or two OTUs dominating each sample. These data represent the first analysis of the bacterial communities associated with the kakapo GIT, providing a baseline for further microbiological study, and facilitating conservation efforts for this unique bird.
2.2 Introduction

The kakapo (*Strigops habroptilus*) is one of the world’s rarest bird species, with only 126 individuals remaining on two predator-free islands off New Zealand’s south coast. Endemic to New Zealand, the kakapo possesses a range of behaviours and physiological characteristics that make it unique: it is the world’s heaviest parrot, the only flightless parrot and the only parrot to carry out lek breeding (Merton et al. 1984). Due to a combination of infrequent mating, low clutch numbers, and poor defense against mammalian predators the kakapo has been pushed to the verge of extinction (Lloyd and Powlesland 1994; Houston et al. 2007), though intensive conservation efforts by New Zealand’s Department of Conservation have recently reversed the decline in numbers. Research programs into kakapo ecology, nutrition and genetics are well established and a management program has been enacted with the aim of restoring the kakapo population in New Zealand. Such practices as confining birds to predator-free islands, supplementary feeding, breeding programs and constant human supervision of both newborn chicks and adults have had a marked effect on the kakapo population – from just 62 remaining individuals in 1991 (Elliott et al. 2001) to the current level. By contrast, the potentially important roles of symbiotic microorganisms in kakapo nutrition and pathogen defense remain unstudied, although positive bacterial influence on the gastrointestinal tract (GIT) was first observed in vertebrates almost 50 years ago (Dubos and Schaedle 1964).

The interactions between hosts and GIT-associated bacterial communities have been the subject of intense study in mammals, particularly humans (Ley et al. 2008b; Zoetendal et al. 2008), with murine models often used to demonstrate causal links between microbes and aspects of host health (Yi and Li 2012). Among avians, microbial research has mainly focused on either pathogen detection, or effects on weight gain in broiler chickens (Yegani and Korver 2008). In the last decade, the study of microbial communities in the GIT of birds has become commonplace, with cultivation-dependent and -independent methods used to examine microbial presence and activity within avian gastrointestinal environments. The microbial communities associated with commercially farmed species such as turkey (Lu and Domingo 2008) and ostrich (Matsui et al. 2010) have been investigated, as well as a range of wild birds, including parrots (Pacheco et al. 2004; Xenoulis et al. 2010) and the South American hoatzin (Godoy-Vitorino et al. 2008, 2010, 2012), and their roles in bird fitness extend far beyond involvement in digestion and
nutrient uptake. For example, studies on the effect of feather-degrading bacteria on mate selection and breeding fitness have revealed novel mechanisms through which bacteria can influence the lifecycle of their host (Shawkey et al. 2007; Burtt et al. 2011).

Links between microbial community structure and increased energy harvest from food have been demonstrated for a wide range of organisms by a variety of indirect techniques (Gill et al. 2006; Turnbaugh et al. 2006; Ohkuma 2008; Torok et al. 2008). In controlled murine models, these effects can be shown at a much more direct level, with gnotobiotic (germ-free) rodents used as controls in experiments that demonstrate the role of bacteria in regulating gene pathways in a range of organs (Stappenbeck et al. 2002; Bäckhed et al. 2004; Meinl et al. 2009; Björkholm et al. 2009). Microbes isolated from a particular host gut have been shown to be highly adapted to the host environment with the community being shaped by host-specific factors in a range of organisms (Van Der Wielen et al. 2002; Khachatryan et al. 2008; Vaishnava et al. 2011). Microbes transplanted into a new gnotobiotic host provide significantly reduced benefits to the new host (Meinl et al. 2009).

Stimulation of the host immune system by GIT microbes has also been recognized in response to both viral and bacterial challenge (Brisbin et al. 2008; Ivanov et al. 2009; Ichinohe et al. 2011), and development of gut-associated lymphoid tissues is increased in conventionally raised mice compared to their germ-free counterparts (Hudson and Luckey 1964; Hooijkaas et al. 1984).

With such varied and important roles being influenced by microbes, the lack of an accurate baseline description of kakapo-associated microbes represents a major gap in our knowledge of kakapo biology. Identification of the indigenous microbial community would be of great value to conservation efforts by enabling identification of allochthonous – potentially pathogenic – microbes. The existing literature surrounding kakapo-associated bacteria has so far focused on detecting and responding to pathogen outbreaks. Such an event occurred in 2004, when three kakapo died from erysipelas within 72 hours of translocation. The birds had been checked for known pathogens (Brangenberg et al. 2003), and erysipelas had not previously been observed in kakapo (Gartrell et al. 2005). While attacks from previously unidentified pathogens are unavoidable, this highlights an area in which molecular microbiology could play a key role in aiding kakapo recovery efforts, through the use of specific, high-sensitivity molecular probing techniques to detect pathogens before their numbers expand to levels that affect the bird.
Human interaction with wild birds can influence the composition of the GIT community (Xenoulis et al. 2010; Janiga et al. 2007), and the potential for human impact on the kakapo GIT community is great (although unavoidable). In times of sickness, wild kakapo are taken into captivity and frequently treated with broad-spectrum antibiotics to combat pathogens. In captivity kakapo are fed a diet supplemented with fruit and pellets not available in the wild and hand-reared chicks are fed on bird formula exclusively until approximately 30 days of age (Eason and Moorhouse 2006). A better understanding of kakapo microbiology carries clear potential for aiding conservation of this endangered bird, yet there are also sound academic reasons for researching this area. The kakapo diet consists mainly of shoots and leaves, and there has been speculation that kakapo may utilize microbes in the foregut to ferment ingested plant material (Morton 1978). While this process is common in ruminants (e.g. cattle and sheep) it is almost unknown among avians, with only the hoatzin known to use the foregut to facilitate fermentation (Grajal et al. 1989). The hoatzin, sole member of the family *Opisthocomidae*, exploits a diverse microbial community in its enlarged crop to aid in digestion, utilising up to 40 bacterial phyla as well as archaea to ferment plant material in the crop (Godoy-Vitorino et al. 2010). The kakapo has been suggested as a possible candidate for foregut fermentation due to its lack of a cecum, which is the primary site of hindgut fermentation (Clench and Mathias 1995), and its similar diet to the hoatzin.

The key aim of this study was to document the microbial community of the kakapo digestive tract in both newly hatched chicks and adults, using samples derived from both the fore- and hindgut to ensure maximum coverage of the GIT. 16S rRNA gene analysis was used to identify bacteria at each sampling site, and the samples were compared to test for changes in community structure along the GIT. This study represents the first step in a wider investigation of the kakapo microbiome, with the ultimate goal of aiding conservation and management of this critically endangered bird.
2.3 Materials and Methods

2.3.1 Sampling

Samples were obtained from four kakapo on Codfish Island (46°47′S 167°38′E), off the coast of Stewart Island, New Zealand, during the nesting season, between 17th and 23rd April 2011. Two additional faecal samples had previously been collected from adult birds when they were brought to Auckland Zoo. A total of 13 samples were used in this analysis, collected from three unfledged chicks, and three adults. Samples of the upper GIT were taken by Department of Conservation staff using sterile rayon-tipped swabs (Copan, #170KS01), and samples were taken from the choana and crop of chicks, and choana of one adult. The choana is an opening in the roof of the mouth that joins to the nasal passage. Due to difficulties in restraining adults, crop samples could not be taken from adult birds. A faecal sample was collected from all six birds at the time swabbing was performed. As interference with nesting mothers can cause them to abandon the nest, samples from chick parents could not be obtained. Swabs were stored in sterile polypropylene tubes and kept on ice until they were frozen at the ranger hut on Codfish Island. Samples were shipped to The University of Auckland on dry ice, and stored at -20°C upon arrival.

2.3.2 DNA Extraction

Despite considerable efforts to standardize the DNA extraction procedure, it was necessary to use a different approach for extracting DNA from swab versus faecal samples due to unreliable DNA retrieval from hard-to-obtain swab samples and unwillingness to risk valuable samples. Genomic DNA was extracted from swabs using heat lysis. Swab tips were placed in a 1.5 mL cryotube containing 1 mL extraction buffer (20 mM EDTA, 0.1 M Tris (pH 8.0), 1% CTAB, 56 mM NaCl), briefly vortexed, then incubated at 94°C for 15 min in order to lyse both Gram-negative and Gram-positive cells (Sadeghi et al. 2008; Petersen et al. 2010). The solution was cooled briefly on ice, 300 μL of chloroform/isoamyl alcohol (24:1 ratio) was added and the solution mixed by inversion, then centrifuged at 13,000 rpm for 5 min at room temperature. The supernatant was transferred to a 2 mL microcentrifuge tube, to which 0.6 volume (vol) isopropanol and 0.1 vol 3 M
sodium acetate (pH 5.2) were added. Samples were incubated overnight at -20°C then centrifuged at 13,000 rpm at 4°C for 30 min. The supernatant was discarded and the pellet washed twice with ice-cold 70% ethanol followed by centrifugation at 13,000 rpm at 4°C for 10 min. Samples were dried and suspended in 50 μL UltraPure water (Invitrogen).

Genomic DNA was extracted from kakapo faecal samples using a variation on an extraction protocol previously described (Costa et al. 2004). 100 mg of faeces were suspended in 1 mL of 70% ethanol with 200 mg of 0.1 mm zirconia/silica beads in a 1.5 mL cryotube. Samples were agitated using a FastPrep FP120 bead beater, at 5.5 ms⁻¹ for 30 s, followed by centrifugation at 13,000 rpm for 5 min and removal of supernatant. 1 mL of extraction buffer was added to each tube in addition to 30 mg of polyvinylpolypyrrolidone (PVPP), before being agitated using the previous settings. Samples were then incubated at 65°C for 30 min, with gentle mixing every 10 min. Samples were centrifuged at 13,000 rpm for 5 min and the supernatant was transferred to a fresh 1.5 mL microcentrifuge tube containing 0.5 mL of chloroform/isoamyl alcohol solution (24:1 ratio) and inverted to mix. Samples were centrifuged at 13,000 rpm for 5 min, before the supernatant (approximately 1 mL) was transferred to a 2 mL microcentrifuge tube containing 0.6 vol isopropanol and 0.1 vol 3M sodium acetate (pH 5.2). Samples were mixed then incubated on ice for 1 h, then centrifuged at 13,000 rpm at 4°C for 1 min to remove any remaining sediment (presumed to be leftover SDS). The supernatant was transferred to a new microcentrifuge tube and centrifuged under the same conditions for 30 min. The supernatant was removed and the pellet washed twice using ice-cold 70% ethanol followed by 10 min centrifugation at 13,000 rpm, 4°C. The pellet was dried and resuspended in 50 μL UltraPure water (Invitrogen).

2.3.3 PCR and clone library construction

PCR was performed using the forward primer 616V and reverse primer 1492R, targeting *Escherichia coli* positions 8 - 27 and 1492 - 1513 respectively, to amplify bacterial 16S rRNA genes, and 21F/958R for archaeal 16S rRNA genes (Table 2.1). Reactions were conducted in 25 μL volumes, containing 20 mM Tris-HCl, 50 mM KCl (buffer), 1.5 mM MgCl₂, 25 μM of each dNTP, 2.5 μM of each primer, 0.5 units Taq polymerase and 1.0 μL of extracted DNA template. Cycling conditions for the 616V/1492R primer pair were as follows: initial denaturing at 94°C for
Chapter 2. Gut microbiome of the kakapo

5 min, 30 cycles of 94°C for 45 s, 57°C for 45 s and 72°C for 1.5 min, followed by a final elongation step at 72°C for 7 min. Cycling conditions for 21F/958R were described previously (Webster et al. 2004). In order to successfully amplify from faecal samples, the addition of 2% bovine serum albumin per tube was required (Wintzingerode et al. 1997). Cloning was performed using the P-GemT Easy Vector kit (Promega, Inc, Madison WI, USA) following the manufacturer’s instructions. Approximately 96 clones from each of the 13 clone libraries were selected for sequencing (Macrogen Inc, Seoul, South Korea).

2.3.4 Phylogenetic Analysis

Sequences were analyzed using the Geneious software package (Drummond et al. 2010) and low-quality data from the ends of each sequence removed. Chimeras were identified with the Pintail algorithm using the Mallard software package (Ashelford et al. 2006) and subsequently removed from the data set. Sequences were aligned via the SINA web aligner (Pruesse et al. 2007) and imported into ARB using the SILVA 108 database (Ludwig et al. 2004). Sequence data were divided into operational taxonomic units (OTU) of 99% sequence identity using mothur (Schloss et al. 2009) and one sequence to represent each OTU per sample was used in tree construction. Sequences representing each OTU were submitted to the DDBJ/EMBL/GenBank databases under accession numbers JQ283115 - JQ283245, JQ302756, and JQ302757. Phylogenetic trees were constructed in ARB using the maximum likelihood method RAxML. Bootstrap values were calculated using 5,000 parsimony replications. Unweighted UniFrac analyses were performed in mothur to statistically compare bacterial community composition among different sample types.

2.3.5 Determination of Bacteroidetes and Archaea sensitivity

A pure culture of Chryseobacterium formosense (phylum Bacteroidetes), originally isolated from wastewater, was obtained from a colleague and cultivated at the original isolation conditions (R2A broth, 28°C for 48 h, C. Brown, personal communication). C. formosense cells were added to samples of kakapo chick faeces
at proportions down to 0.15% total bacterial cell load, calculated through enumeration of *C. formosense* via plating counts, and DAPI staining of faecal samples, then subjected to both extraction methods detailed previously. A fragment of the 16S rRNA gene was amplified using the 341-GC/518R primer pair. Cycling conditions consisted of an initial denaturing step at 94°C for 5 min, followed by 25 cycles at 94°C for 1 min, 60°C for 1 min, and 72°C for 30 s, then a final elongation step at 72°C for 5 min. The product was analyzed using denaturing gradient gel electrophoresis (DGGE) with a denaturing gradient of 40 - 70%. A positive control of pure *C. formosense* DNA was used as an indicator of a *Bacteroidetes* band in the gel pattern. A pure culture of *Methanosarcina acetivorans* (domain Archaea, strain DS2834) was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ) and added to samples of kakapo chick faeces at proportions down to 0.4% total cell load. A fragment of the archaeal 16S rRNA gene was amplified using the 21F/958R primer pair. Cycling conditions were described previously (Webster et al. 2004). Samples were visualised on a 1% agarose gel and analysed using the BioRad Gel Doc imaging system.
2.4 Results

2.4.1 Bacterial community composition within the kakapo gastrointestinal tract

Bacterial 16S rRNA gene amplification was successful for all samples, whereas no archaea were amplified from any samples. A total of 1,007 clones yielded high-quality sequence that passed chimera checking. The phyla Gammaproteobacteria and Firmicutes were present in all libraries, with only slight representation from Fusobacteria in a single chick choana sample (Figure 2.1). When sequence data were dereplicated into 99% OTUs it was revealed that most of the sequences belonged to only a few key OTUs, such as Haemophilus felis and Streptococcus pasteurianus (Figure 2.2). A Chao1 diversity estimator for each clone library was calculated at the 99% OTU level, and in almost all cases the expected number of OTUs per library was close to the observed number. The remainder of the diversity in each library was split among several low-abundance OTUs. Phylogenetic trees of kakapo-associated Firmicutes (Figures 2.3 and 2.4) and Gammaproteobacteria (Figure 2.5) are shown for comparison.

The extent of differences in bacterial community composition between samples was tested at the OTU level using an unweighted UniFrac analysis. Sequences obtained from kakapo chick, and adult faecal, samples were pooled according to sample type. Sequence data obtained from the single adult choana swab were not included in statistical comparison. Pairwise comparisons were made between sample types to test the null hypothesis that the bacterial community is homogeneous throughout the GIT. Significant differences in community structure were observed between the chick choana/crop (p < 0.001) and crop/faeces (p = 0.002), but not between the choana/faeces. Between chick and adult faecal samples, no significant difference was seen. Given that Sass and Sirocco had been subject to considerable human intervention prior to sampling, their faecal samples were compared separately to the chick samples and wild adult sample (Millie). The faecal communities of Sass/Sirocco were significantly different from those of the wild adult (p = 0.002), but did not differ significantly from the chick samples. There was no significant difference between the wild adult and chick faecal samples.
Figure 2.1: Phylum-level distribution of kakapo-derived 16S rRNA gene sequences. Phylum-level affiliation of 16S rRNA gene sequences obtained from the kakapo GIT. Samples to the left of the dotted line represent clone libraries derived from juveniles, and samples on the right represent adult-derived sequences.
Chapter 2: Gut microbiome of the Kakapo

### Nearest BLAST

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Species name</th>
<th>Sample</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>AY362886</td>
<td>Actinobacillus capsulatus</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>U65583</td>
<td>Actinobacillus porcinus</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY362895</td>
<td>Actinobacillus rossii</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AF530475</td>
<td>Escherichia fergusonii</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AF224292</td>
<td>Haemophilus felis</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY362907</td>
<td>H. haemoglobinophilus</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY362908</td>
<td>H. parainfluenzae</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY362909</td>
<td>H. parasuis</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AF129440</td>
<td>Klebsiella oxytoca</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AF053893</td>
<td>Mannheimia varigena</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>U78182</td>
<td>Psychrobacter arenosus</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>X96966</td>
<td>Raoultella ornithinolytica</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>X96963</td>
<td>Shigella dysenteriae</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY362910</td>
<td>S. flexneri</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY216868</td>
<td>Volucribacter psittacicida</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>X68174</td>
<td>Leptotrichia goodfellowii</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY438672</td>
<td>Clostridium baratii</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY438673</td>
<td>C. bartlettii</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>X76748</td>
<td>C. colinum</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>Y18176</td>
<td>C. disporicum</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>X76750</td>
<td>C. glycolicum</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY485602</td>
<td>Streptococcus lutetiensis</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AY485603</td>
<td>S. oralis</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>DQ232528</td>
<td>S. pasteurianus</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AF349724</td>
<td>S. suis</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AF009477</td>
<td>S. suis</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
<tr>
<td>AF349725</td>
<td>S. suis</td>
<td>Hakatere</td>
<td>Tutoko</td>
</tr>
</tbody>
</table>

**OTUs**: 11, 10, 7, 9, 12, 8, 9, 22, 7, 2, 7, 11, 17

**Chao1 Estimate**: 13.5, 16, 10, 16.5, 19, 9.5, 9.6, 40.5, 13, 2, 10, 15, 35

**Legend**

- >40%
- >15%
- >5%
- >4%
- >3%
- >2%
- >1%

**Figure 2.2**: OTU-level distribution of kakapo-derived 16S rRNA gene sequences. OTU-level affiliation of 16S rRNA gene sequences obtained from the kakapo GIT. Values in the heatmap are scaled as a proportion of the total number of sequences per clone library. Observed numbers of OTUs at 99% sequence similarity are provided below the figure, as well as the estimated diversity for each library using the Chao1 estimator.
Figure 2.3: Phylogeny of Firmicutes (lactic acid bacteria) 16S rRNA gene sequences derived from kakapo. 16S rRNA gene-based phylogenetic analysis of lactic acid bacteria recovered from kakapo samples. Solid junctions represent >90% bootstrap support, and hollow junctions >75%. Colouring denotes faecal clones (red), crop clones (blue) and choana clones (purple). Dashed lines indicate sequence length <1200 bp. Scale bar, 10% sequence divergence.
Figure 2.4: Phylogeny of Firmicutes (Clostridia) 16S rRNA gene sequences derived from kakapo. 16S rRNA gene-based phylogenetic analysis of Clostridia recovered from kakapo samples. Solid junctions represent >90% bootstrap support, and hollow junctions >75%. Colouring denotes faecal clones (red), crop clones (blue) and choana clones (purple). Dashed lines indicate sequence length <1200 bp. Scale bar, 10% sequence divergence.
2.4.2 Determination of *Bacteroidetes* and *Archaea* sensitivity

*Bacteroidetes* are common GIT-associated bacteria in many vertebrates, but were not detected in any of the kakapo samples. As certain DNA extraction techniques can lead to under-representation of *Bacteroidetes* in a sample (Boom et al. 1990), we tested whether our DNA extraction methods are able to detect the presence of *Bacteroidetes* within faecal and swab samples. *Bacteroidetes* DNA was successfully detected in all spiked faecal and swab samples (data not shown), down to approximately 0.15% of bacterial cell load, indicating that *Bacteroidetes* were not excluded by our DNA extraction protocols. Similarly, while *Archaea* were not detected in un-spiked kakapo faeces, the archaeal 16S rRNA gene could be detected when swab and faecal samples were spiked with archaeal cells, down to approximately 0.4% of cell load.
Figure 2.5: Phylogeny of Gammaproteobacteria-associated 16S rRNA gene sequences derived from kakapo. 16S rRNA gene-based phylogenetic analysis of Gammaproteobacteria found within kakapo samples. Solid junctions represent >90% bootstrap support, and hollow junctions >75%. Colouring denotes faecal clones (red), crop clones (blue) and choana clones (purple). Dashed lines indicate sequence length <1200 bp. Scale bar, 10% sequence divergence.
2.5 Discussion

This paper describes the first molecular examination of the bacterial communities within the kakapo GIT, and provides evidence that qualitative differences exist between sites sampled throughout the GIT. The kakapo GIT appears to harbor a low-diversity community of microbes, with essentially only two phyla detected, *Gammaproteobacteria* and *Firmicutes*. Microbes in the kakapo GIT are abundant, with both cultivation-based measurements and DAPI cell counts indicating a microbial cell density in the order of $10^{10}$ cells per gram of faecal material (data not shown), yet each sample is dominated by only a few genera, typically *Haemophilus*, *Streptococcus*, and *Clostridium*. As the *Fusobacteria* discovered were only in a single sample, and found in low abundance, it is possible that their presence represents some form of contamination which occurred during sampling or DNA extraction. The *Fusobacteria*-associated sequences were similar to isolates and clones of the genus *Leptotrichia*, a bacterium commonly found in the human oral cavity (Eribe and Olsen 2008).

At the phylum level, bacterial diversity is well conserved among all chicks sampled, but within the adults there is large variation in terms of relative abundance of each phylum. This may be explained by a range of factors regarding the adults, such as the frequent handling of Sirocco and, to a lesser extent, Sass, or the age difference between Sass and Millie/Sirocco. The bird Sass died several weeks after the collection of faecal samples, but not due to pathogen-related illness, and had not been treated with antibiotics prior to sampling (which can disrupt the GIT community (O’Hara and Shanahan 2006; Dethlefsen et al. 2008)). Subject age has been linked to a shift in the bacterial gut community in humans (Hopkins et al. 2001; Rajilić-Stojanović et al. 2009) and mice (Vaahtovuo et al. 2001), so it is conceivable that such a community change may be a natural phenomenon. While functional roles of the bacteria detected in this study can only be speculated upon, those bacteria encountered in the kakapo GIT correspond to genera commonly observed in other herbivores. In a study of the gut microbiota of deer it was recognized that *Streptococcus* played a role in degrading tannins ingested by the host animal, and many of the *Streptococcus* detected in the kakapo clone libraries were closely related to this species (Figure 2.3, *Streptococcus gallolyticus* sub. *macedonicus*, AB563237) (Hiura et al. 2010). Most of the bacterial genera detected throughout the kakapo GIT are known anaerobic fermenters, capable of converting sugars such as glucose and cellulose into acids such as acetate, which...
are utilized by the host. Members of the genus *Clostridium* are frequently identified as cellulolytic (Shoham et al. 1999; Warnick et al. 2002; Varel and Pond 1992; Sabathé et al. 2002), and have been found to increase in proportion within the herbivore gut in the absence of starch (Laure et al. 2005). The inability to detect *Bacteroidetes* in a parrot, using either 16S rRNA-based techniques or cultivation, is not unique to our study (Pacheco et al. 2004; Xenoulis et al. 2010). In addition to playing roles in butyrate production and bile acid metabolism (Kim and Milner 2007; Thomas et al. 2011), *Bacteroidetes* are well-characterized degraders of starch and cellulose in the gut (Dongowski et al. 2000; Chassard et al. 2007; Martens et al. 2009; Bolam and Sonnenburg 2011). Historically the kakapo have relied on a low-starch diet (Horrocks et al. 2008), which may have selected against *Bacteroidetes* colonization, as diet has been identified as one of the factors that shape gut microbiota (Finegold et al. 1974; Hehemann et al. 2010; Martínez et al. 2010; Shanks et al. 2011). While the DNA extraction method utilized in this study is capable of extracting detectable levels of DNA from *Bacteroidetes* comprising less than 1% of the community, it is conceivable that the inability to detect *Bacteroidetes* stems from low sequence counts compared to those obtained using next-generation sequencing technologies.

Given the endangered status of the kakapo, destructive sampling (via dissection) is not possible. As such, our analyses were limited to swab and faecal samples rather than direct tissue and gut content samples. Although surface swabs may not give a perfect representation of the local bacterial community, they have been previously applied in a range of avian study systems (Xenoulis et al. 2010; Shanks et al. 2011; Moreno et al. 2003; Blanco et al. 2006; Klomp et al. 2008) where dissection of the target organism was not feasible. There still exists the potential that mucosa-associated bacteria of the crop may not be recovered through the swabbing of live animals, indicating a potential blind spot in sampling. Nevertheless, it appears that swabbing of the kakapo crop is adequate for differentiating between microbial communities of the choana and crop, despite the fact that any probe into the crop risks potential contamination as the swab passes the choana. The use of faecal samples as a proxy for hindgut bacterial communities has been used extensively in a range of vertebrates, including humans and birds. While several studies have highlighted differences between bacterial recovery from mucosal biopsies and faecal samples (Zoetendal et al. 2008; Eckburg et al. 2005), this appears to be due to faecal samples containing not only mucosa-associated bacteria that have been shed into the faeces, but also bacteria that colonize the faecal substrate.
directly. Community data taken from faecal samples contains a reasonable representation of microbes within the hindgut, and differences in faecal microbiota (both at presence/absence and functional levels) have been shown to reflect differences in the intestinal tract of the host (Vaahtovuo et al. 2001; Savage 1977), although it must be stressed that they do not provide an exact representation of microbial community structure and function within the intestine itself. Based on unweighted UniFrac analysis it appears that the faecal bacterial communities of adults and chick kakapo are not significantly different, which may indicate a vector for inoculation of kakapo chicks with their parents’ microbiota. The lack of significant difference between choana and faecal communities in the chicks is not surprising considering the lifestyle of unflended kakapo chicks. Essentially immobile, the chicks are unable to distance themselves from their own faeces. The chicks studied in this project have since been fledged and given the low population of kakapo and constant attention to the birds, these present an excellent opportunity for longitudinal studies throughout the lifespan of the birds.

Given the low-energy diet of the kakapo and its lack of cecum, it has been speculated that the kakapo may utilize microbially-mediated foregut fermentation to derive additional energy from its food. While this study was not targeted at confirming or rejecting the notion of kakapo foregut fermentation, the possibility that key microorganisms may be resident in the crop rather than hindgut was taken into consideration when planning this study. There are several frequently found bacterial phyla in the microbial community of foregut-fermenting mammals and the hoatzin, predominantly Firmicutes and Bacteroidetes, with representation from Verrucomicrobia, Actinobacteria and Spirochaetes commonly observed (Godoy-Vitorino et al. 2010; Ley et al. 2008a). Methanogenic archaea are also commonly found in the rumen or crop of foregut fermenters (Godoy-Vitorino et al. 2008; Tokura et al. 1999; Tajima et al. 2001; Irbis and Ushida 2004; Shin et al. 2004; Yu et al. 2008). With the exception of Firmicutes, none of the above-mentioned taxa were detected in the kakapo samples. In the hoatzin it has been shown that the microbial community of the foregut is similar to that of ruminants (Godoy-Vitorino et al. 2012), but given the apparent absence of so many bacterial phyla in the kakapo crop it is unlikely that the kakapo shares this community structure and gut adaptation. Foregut fermentation is an adaptation to a diet rich in cellulososes that the host cannot digest (Russell and Rychlik 2001), but kakapo do not retain and digest cellulose in the manner seen in ruminants and the hoatzin, typically spitting away masticated plant material after extracting juices from the flesh.
(Horrocks et al. 2008; Oliver 1955). Although merely speculative at this stage, it thus seems unlikely that kakapo perform foregut fermentation in the traditional manner.

One observation from this study that may prove to be of future concern is the high number of Pasteurellaceae-like sequences within the choana and crop swabs. Many of the sequences were clustered with bacterial genera such as Haemophilus, or with several non-cultivated clades commonly detected in the avian respiratory tract, which are frequently found as respiratory pathogens in vertebrates (Christensen et al. 2003). It has been noted that certain Pasteurellaceae which were present in our libraries (Bisgaard Taxon 34, Bisgaard Taxon 44, Figure 2.5) are frequently associated with respiratory disease in psittacine birds. Although not all bacterial species in these clades are causative agents of disease, their presence should be considered a warning, as they are often found in sick birds (Gregersen et al. 2010). During the 2011 breeding season several chicks were removed from the nest due to respiratory problems, although this did not cause long-term health issues in the birds (D. Eason, personal communication). While there is no data to imply a causal link between the observed Pasteurellaceae and the illness, pathogens do appear to have been introduced to the kakapo population previously through avian vectors (Gartrell et al. 2005).

In summary, we performed the first 16S rRNA-based microbial analysis of the bacteria that inhabit the kakapo GIT. We have shown that the GIT is inhabited by a few key organisms, and that the community composition changes throughout the GIT. Our results also provided preliminary evidence that the human influence on kakapo lifestyle appears to cause a shift in these bacterial communities, although whether this has a positive, negative, or neutral effect on the bird remains unknown.
2.6 Acknowledgments

We gratefully thank Daryl Eason and Jo Ledington (Department of Conservation), and John Potter (Auckland Zoo) for provision of all kakapo samples, and Caroline Brown for supplying the required *Bacteroidetes* strain. We would also like to thank Ron Moorhouse and Deidre Vercoe (Department of Conservation), and Richard Jakob-Hoff (Auckland Zoo) for their endorsement and support of this project, plus Jacqueline Beggs and Mick Clout (University of Auckland) for useful discussions on this topic.
Chapter 3

Influence of Hand-rearing and Bird Age on the Faecal Microbiota of the Critically Endangered Kakapo

David W. Waite¹, Daryl K. Eason², Michael W. Taylor¹

¹. Centre for Microbial Innovation, School of Biological Sciences, The University of Auckland, Auckland, New Zealand
². Research Development and Improvement Division, Department of Conservation, Nelson, New Zealand

Published in the Peer-Reviewed Journal Applied and Environmental Microbiology

May 16th, 2014

Applied and Environmental Microbiology, Volume 80(15), Pages 4650 - 4658
DOI: 10.1128/AEM.00975-14
3.1 Abstract

The critically endangered New Zealand parrot, the kakapo, is subject to an intensive management regime aiming to maintain bird health and boost population size. Newly hatched kakapo chicks are subject to human intervention and are frequently placed in captivity throughout their formative months. Hand-rearing greatly reduces mortality among juveniles, but the potential long-term impact on the kakapo gut microbiota is uncertain. To track development of the kakapo gut microbiota, faecal samples from healthy, pre-fledged juvenile kakapo, as well as unrelated adults, were analysed using 16S rRNA gene amplicon pyrosequencing. Following the original sampling, juvenile kakapo underwent a period of captivity, so further sampling during and post-captivity aimed to elucidate the impact of captivity on the juvenile gut microbiota. Variation in the faecal microbiota over a year was also investigated, with resampling of the original juvenile population. Amplicon pyrosequencing revealed a juvenile faecal microbiota enriched with particular lactic acid bacteria when compared to the adults, although overall community structure did not differ significantly among kakapo of different ages. Abundance of key OTUs was correlated with antibiotic treatment and captivity, although the importance of these factors could not be proven unequivocally within the bounds of this study. Finally, the microbial community structure of juvenile and adult kakapo changed over time, reinforcing the need for continual monitoring of the microbiota as part of regular health screening.
3.2 Introduction

The kakapo (Strigops habroptilus) is a flightless, nocturnal parrot endemic to New Zealand. It possesses few defences against introduced mammalian predators (Lloyd and Powlesland 1994), has a slow reproductive rate and usually lays only a single egg. Although kakapo were once common throughout New Zealand (Wood 2006), the population has declined to 126 individuals at the time of publishing, confined to three predator-free islands off the coast of New Zealand. Since the 1980s, kakapo have been subject to an intensive management program focused on preserving the species and eventually restoring the population to self-sustaining levels (Clout and Merton 1998; Powlesland et al. 2006; Eason et al. 2006). In an effort to optimize management practices, the New Zealand Department of Conservation has collaborated with researchers from a wide range of biological disciplines, including behavioural ecology, physiology, genetics, nutrition and, recently, microbiology (Merton et al. 1984; Moorhouse and Powlesland 1991; McNab and Salisbury 1995; Atkinson and Merton 2006; Cockrem 2006; Cottam et al. 2006; Waite et al. 2012).

The kakapo has an array of biological characteristics that make it an unusual animal to study. Apart from being the world’s heaviest parrot, the only flightless parrot and the only parrot to carry out lek breeding (Merton et al. 1984), the kakapo has been identified as a potential foregut fermenter due to its herbivorous diet and lack of ceca (Clench and Mathias 1995). Microbially-mediated foregut fermentation is a common trait in mammals but is rare among avians, with only the South American hoatzin known to perform this process (Grajal et al. 1989). Notable differences in feeding strategy do exist between the kakapo and hoatzin, however, with kakapo rarely ingesting fibrous plant material, instead extracting the juices from shoots and leaves and discarding the undigested ‘chews’ (Oliver 1955; Horrocks et al. 2008). There is currently no empirical evidence to support or dispute the occurrence of foregut fermentation in kakapo. Intriguingly, the reproductive cycle of the kakapo is linked to the fruiting of particular native New Zealand trees. The rimu tree undergoes a mast season every three or four years, during which time kakapo mate and rear young. While the link between kakapo and rimu is well established, the causal link between these two phenomena is unclear and is not simply a matter of additional dietary energy enabling reproduction (Houston et al. 2007).
Chapter 3. *Influence of hand-rearing*

The role of microbial symbionts in the gastrointestinal (GI) tract of vertebrates is well documented, and a range of mechanisms through which microbes contribute to the nutrition (Hill 1997; Turnbaugh et al. 2006; Torok et al. 2011; Stanley et al. 2012) and development of the host gut (Stappenbeck et al. 2002; Rahimi et al. 2009; Cao et al. 2012) have been identified. Our previous research into the GI-associated bacteria of the kakapo revealed a community dominated by only a handful of operational taxonomic units (OTUs), mainly from the phyla *Proteobacteria* and *Firmicutes* and apparently lacking in archaea (Waite et al. 2012, 2013). The kakapo GI tract appears to have a low phylum-level diversity of bacteria compared to other birds (Kohl 2012; Waite et al. 2013) and this has led to speculation of a population bottleneck on the gut microbiota. The kakapo microbiota is not well understood, but it is likely that current management practices have an impact on the microbial community as kakapo are removed from the wild and given veterinary care at the first sign of sickness (Eason and Moorhouse 2006). This captivity results in a change in diet and often includes antibiotic treatment, both of which are frequently linked to shifts in microbial community structure (Rahimi et al. 2009; Tannock 1997; Dethlefsen et al. 2008; Hammons et al. 2010; Hill et al. 2010).

Of major relevance to kakapo microbiology is the fact that the developmental pattern of the kakapo gut microbiota is completely unknown, thereby making it difficult to address questions regarding the effect of diet or captivity on the gut microbiota. Previous research into the development of the gut microbiota in other host species has reported that the microbiota of juveniles differs significantly from that of adults in both avians and mammals (Gong et al. 2008; Godoy-Vitorino et al. 2010; Lozupone et al. 2012; Yatsunenko et al. 2012; Van Dongen et al. 2013), although it is not clear that this pattern is reflected in the kakapo (Waite et al. 2012). In many avian species the juvenile microbiota is a dynamically changing community (Van Der Wielen et al. 2002; Lu et al. 2003; Scupham 2007) that gradually develops towards the adult community structure (Godoy-Vitorino et al. 2010; Van Dongen et al. 2013), but the changes in microbiota as the subject ages vary by host. For example, chickens are enriched in *Lactobacillaceae* in the first week of life (Lu et al. 2003), while juvenile turkeys appear to harbour a large proportion of *Clostridiales* until around 10 weeks of age (Scupham 2007). Moreover, even genetically related individuals undergo different developmental patterns when geographically isolated (Stanley et al. 2013). In order to better understand the temporal dynamics of the kakapo microbiota, and potentially gain insights into the impact of human intervention, samples were collected from juvenile kakapo.
born during the 2011 breeding season, spanning four time points throughout the first year and a half of life. The aims of this study were to compare differences in the juvenile and adult faecal microbiota, to understand the time required for juvenile kakapo to develop a ‘normal’ adult gut microbiota, and to investigate the potential effects of captivity on the juvenile microbiota.
3.3 Materials and Methods

3.3.1 Sample collection

In the 2011 breeding season, 11 kakapo chicks were hatched on Codfish Island, off the coast of New Zealand (46°47′S 167°38′E). Samples were collected from these juveniles at four time points, which are summarised in Table 3.1. During the nesting period each juvenile required a period of captivity, either due to low weight gain or sickness. While undergoing the period of captivity the juveniles were fed a diet of fruit, Lactated Ringer’s solution and the proprietary parrot hand-rearing formula Exact (Kaytee Products Inc, Chilton, WI).

Eight of the captive juveniles were also treated with the commercially available antibiotics Augmentin and Clavulox, which combine a \( \beta \)-lactam (amoxicillin) and \( \beta \)-lactamase inhibitor (clavulanic acid) (Tables 3.1 and 3.2).

Samples collected during this period were taken following five days of captivity. Following release from captivity, juveniles were given a ‘recovery’ period of two weeks, after which additional faecal samples were collected. A final sample was collected approximately one year later, at which point the ‘juveniles’ had fledged from the nest and were now independent adults.

Fresh faecal samples were also collected from adult kakapo in the same manner. Samples were taken from adults with no history of sickness, and no history of captivity other than the original translocation to Codfish Island, at time points coinciding with the Juvenile_First and Juvenile_Fourth samplings. The adult age data are incomplete as some birds were born wild and only introduced to Codfish Island later in life. However, the youngest adults sampled were born on Codfish Island in 2005, making the ‘adult’ kakapo representative of a substantially older sub-population than those individuals representing the ‘juvenile’ dataset, even though at the Juvenile_Fourth time point the original juvenile kakapo were mature kakapo.

Fresh faeces were collected aseptically during routine health inspections of the juveniles and stored in sterile polypropylene tubes on ice until they were able to be frozen, a period of less than one hour following collection. Sample sizes vary throughout the study due to difficulty with capture and recapture of birds.
Table 3.1: Sample sizes and mean individual age and additional notes taken at the time of sampling. For ‘adult’ kakapo, age data is incomplete but the youngest adults were $\geq 6$ years old in 2011.

<table>
<thead>
<tr>
<th>Sample time point</th>
<th>Number of individuals</th>
<th>Mean age (days)</th>
<th>Sample year</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juvenile_First</td>
<td>8</td>
<td>16</td>
<td>2011</td>
<td>Fecal sample from wild juvenile kakapo</td>
</tr>
<tr>
<td>Juvenile_Second</td>
<td>10</td>
<td>51</td>
<td>2011</td>
<td>Fecal sample from captive juvenile kakapo fed an artificial diet of lactated Ringer’s solution and fresh fruit; 8 juveniles were also treated with antibiotics during captivity</td>
</tr>
<tr>
<td>Juvenile_Third</td>
<td>5</td>
<td>69</td>
<td>2011</td>
<td>Fecal sample from juvenile kakapo 2.5 wk following release</td>
</tr>
<tr>
<td>Juvenile_Fourth</td>
<td>8</td>
<td>569</td>
<td>2012</td>
<td>Fecal sample from juvenile kakapo during the next round of health screening; at this point, individuals were mature, independent birds</td>
</tr>
<tr>
<td>Adult_First</td>
<td>10</td>
<td></td>
<td>2011</td>
<td>Fecal sample from wild adult kakapo collected at the same time as Juvenile_First</td>
</tr>
<tr>
<td>Adult_Second</td>
<td>7</td>
<td></td>
<td>2012</td>
<td>Fecal sample from wild adult kakapo collected at the same time as Juvenile_Fourth</td>
</tr>
</tbody>
</table>
Table 3.2: **List of individuals sampled at each point in the juvenile survey.** Individuals marked with an asterisk (*) received antibiotic treatment with the commercially available antibiotic amoxicillin-clavulanic acid formulas.

<table>
<thead>
<tr>
<th>Juvenile_First</th>
<th>Juvenile_Second</th>
<th>Juvenile_Third</th>
<th>Juvenile_Fourth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hakatere</td>
<td>Atareta</td>
<td>Atareta</td>
<td>Hakatere</td>
</tr>
<tr>
<td>Ian</td>
<td>Hakatere*</td>
<td>Ian</td>
<td>Ian</td>
</tr>
<tr>
<td>Ihi</td>
<td>Ian*</td>
<td>Taonga</td>
<td>Ihi</td>
</tr>
<tr>
<td>Stella</td>
<td>Ihi*</td>
<td>Tutoko</td>
<td>Stella</td>
</tr>
<tr>
<td>Taonga</td>
<td>Stella*</td>
<td>Waikawa</td>
<td>Tia</td>
</tr>
<tr>
<td>Tia</td>
<td>Taonga*</td>
<td></td>
<td>Tutoko</td>
</tr>
<tr>
<td>Tutoko</td>
<td>Tia*</td>
<td></td>
<td>Waa</td>
</tr>
<tr>
<td>Waa</td>
<td>Tutoko*</td>
<td></td>
<td>Waikawa</td>
</tr>
<tr>
<td></td>
<td>Waa*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Waikawa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.3.2 DNA extraction, PCR amplification and amplicon pyrosequencing

DNA extraction was performed on samples using a previously described bead-beating method (Waite et al. 2012). PCR was performed using 16S rRNA gene-specific primers targeting Bacteria: 533f (5′-GTG CCA GCA GCY GCG GTM A-3′) and 907R (5′-CCG TCA ATT MMY TTG AGT TT-3′) (Simister et al. 2012). Each primer was synthesised with an FLX Titanium adaptor sequence (533f: CCA TCT CAT CCC TGC GTG TCT CCG AC, 907R: CCT ATC CCC TGT GTG CCT TGG CAG TC) and a 10 bp multiplex identifier barcode (MID) was attached to the forward primer using commercially available barcode sequences (Roche Diagnostics Corporation, Branford, CT). For each DNA extraction, three 25 µL PCR reactions were performed with a specific barcoded primer. Reactions contained 20 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl₂, 100 µM dNTP mixture, 2.5 µM of forward and reverse primer, 2% bovine serum albumin, 0.5 units Taq polymerase and 10 ng of template DNA. Cycling conditions were as follows: initial denaturation of 94°C for 5 min, 20 cycles of touchdown PCR (94°C for 30 s, 60°C for 30 s and 72°C for 45 s, with a 0.5°C decrease in annealing temperature per
cycle), 10 cycles of 94°C for 30 s, 50°C for 30 s and 72°C for 45 s, and a final elongation step of 72°C for 10 min, with one negative control run for each primer pair. Following amplification, PCR products from each sample were pooled and purified using the Agencourt AMPure XP bead system (Agencourt, Beckman Coulter, MA, USA), according to the manufacturer’s instructions. Product size was measured using the Agilent DNA 1000 kit (Agilent Technologies, Waldbronn, Germany) run on the Agilent 2100 Bioanalyzer platform. Purified product was quantified using the Qubit Quant-iT DNA high-sensitivity assay and combined with fragment size to equalise the number of molecules pooled per sequencing run. Samples were randomised across sequencing runs and pyrosequencing was performed on a Roche GS-FLX Titanium platform (Roche, NJ, USA) by Macrogen Inc (Seoul, South Korea). Sequence data were submitted to the NCBI Sequence Read Archive under the accession numbers SAMN02369276 - 329 and SAMN02420182 - 88.

3.3.3 Bacterial community analysis

Sequence reads were processed using mothur version 1.31.2 (Schloss et al. 2009). Briefly, pyrosequencing flowgrams were denoised using the mothur implementation of AmpliconNoise (Quince et al. 2011). Sequences with length <200 bp or homopolymers >8 bp were removed, as were sequences with more than one MID barcode mismatch or two primer mismatches. Sequences were then aligned against a reference database (Schloss 2014a) and sequences that could not be aligned were removed. Chimeras identified using the UCHIME algorithm (Edgar et al. 2011) were removed from the data set and the remaining sequences classified using a previously reported method (Simister et al. 2012; Schmitt et al. 2012). Sequences classified as chloroplast were removed from the dataset, as were sequences that could not be classified even to domain level. From a starting total of 216,759 sequences, the above-mentioned procedures led to a final total of 208,121 high-quality sequence reads for further analysis.

Following these initial quality control steps, sequence data were binned into genus- and species-approximating OTUs using definitions of ≥95% and ≥97% sequence similarity (OTU_{0.95} and OTU_{0.97}), respectively. Yue-Clayton theta, Jaccard index (Yue and Clayton 2005) and phylogeny-based UniFrac distances (Lozupone and Knight 2005) were calculated between each group using each OTU definition. Average distances between communities were calculated by randomly subsampling
1,400 sequences per sample and calculating distance measures 10,000 times. Apparent changes in community structure were tested using analysis of molecular variance (AMOVA) (Excoffier et al. 1992; Anderson 2001). Changes in OTU\textsubscript{0.97} abundance between sample groups were tested using metastats (White et al. 2009), but are only reported if the OTU of interest had a relative abundance of $\geq 1\%$ of the obtained sequences across any sample, as rarer OTUs were frequently only observed in a single kakapo individual and therefore were not informative when surveying the overall microbiota. Correlations between OTUs that differed significantly between captive and wild individuals were measured using the Point-Biserial correlation coefficient ($r_{pb}$) in the R software environment version 2.15.2 (R Core Team 2012).

Coverage of sequencing was estimated by calculating Good’s Coverage index after subsampling the OTU\textsubscript{0.97} table to a depth of 1,400 OTUs per sample. Richness and diversity estimators were also calculated on the subsampled data using Shannon’s diversity and evenness, Simpson diversity index, Chao1 and ACE estimators. In order to test how well previous clone-library data represented the full faecal microbiota, representative OTU\textsubscript{0.97} sequences were mapped against full-length clone sequences using usearch (Edgar 2010) with a minimum similarity value of 0.97. The results of the usearch mapping were recorded both in terms of the total proportion of reads that were mapped against the clone-library data, and as a proportion of the total number of OTU\textsubscript{0.97} generated during analysis.

The core microbiota was calculated by scoring the presence/absence of each OTU in each sample then calculating presence as a proportion of all samples. An overall core was calculated using all available wild samples (excluding Juvenile\_Second), as well as separate cores consisting of exclusively adult kakapo (Adult\_First, Adult\_Second, Juvenile\_Fourth) and exclusively healthy juvenile kakapo (Juvenile\_First, Juvenile\_Third). A final core was calculated from the captive, antibiotic-treated kakapo ($n = 8$) but not from the untreated pair of birds, due to the low sample size. The core microbiota was categorised as the ‘core’ microbes present in $> 90\%$ of individuals surveyed and the ‘variable’ microbes present in $> 60\%$ of individuals surveyed. All other OTUs were considered either transient or individual specific.
3.4 Results

3.4.1 The kakapo microbiota is of low diversity and dominated by *Firmicutes* and *Proteobacteria*

Amplicon pyrosequencing yielded an average of 4648 sequences per sample, compared to 77 obtained via clone libraries in our previous paper (Waite et al. 2012). Good’s Coverage values show that amplicon pyrosequencing was able to describe essentially the entire microbiota of each sample group, while ecological diversity measures indicate an uneven bacterial community of low diversity (Table 3.3). This unevenness was reflected in the low Shannon evenness calculated for each sample. The Shannon evenness index is a ratio of the Shannon diversity index of a sample compared to the maximum possible Shannon index in the sample, with a value of 1 being perfectly even. The evenness of all samples was low according to this metric (Table 3.3) and this was apparent visually (Figure 3.1). Community richness (Chao1 and ACE) were lowest during the juvenile captivity period, although this was not reflected clearly in the diversity estimators. Similarly, when amplicon sequences were mapped to previous kakapo gut microbiota clone-library data, the diversity within the clone-library accounted for a comparatively large proportion of the total reads sequenced, but a much lower proportion of the total OTUs (Table 3.3). Similar to previous findings, the phylum-level membership consisted predominantly of *Proteobacteria* and *Firmicutes*. In contrast to the clone-library data, *Bacteroidetes* and *Actinobacteria* were also frequently, though not universally, detected (in 70% and 33% of samples, respectively).
Figure 3.1: **Phylogenetic distribution of bacterial OTUs in the kakapo faecal microbiota.** High-level taxonomic information is provided as per classification. OTUs are defined as groups of 16S rRNA gene sequences that share ≥97% similarity and are ordered by phylum, then sub-ordered by class. For clarity only the 50 most abundant OTUs are plotted, representing 98.0% of the total reads, with 477 OTUs comprising the remainder of the reads following removal of singletons. OTU abundances are scaled as a proportion of all sequences in the respective sample. OTU02 (mentioned in the manuscript) is noted with an asterisk (\(\ast\))
In order to create a baseline for future kakapo microbiology research, we defined core and variable communities of OTUs that were observed in the kakapo faecal microbiota (Table 3.4). The ‘core’ microbiota consisted of OTUs present in >90% of individuals sampled and the ‘variable’ microbiota in >60% of individuals. These groupings were further stratified by age (adult core, juvenile core), and treatment core was calculated for juvenile kakapo under the influence of antibiotics (treated core). In all core calculations of wild birds, two OTUs were classified as ‘core’, taxonomically assigned as Escherichia and Streptococcus. The ‘variable’ microbiota differed between juvenile and adult birds with more OTUs recruiting to the variable microbiota of juveniles, consistent with the higher diversity and richness indices reported in the juvenile population (Table 3.3). The treated core microbiota differed from that of the juvenile core, and accounted for more of the microbiota than in wild samples. It is important to note that the classification scheme used was only based on presence/absence of OTUs and did not account for their relative abundance in the sample, although abundance carries biological significance in the interpretation of these data. With a single exception, core OTUs accounted for >5% of the amplicon reads in their respective grouping (Table 3.4). Some variable OTUs, including OTU02 in the antibiotic-treated microbiota, OTU04 in the adult microbiota and OTU06 in the wild juvenile microbiota accounted for a large proportion of the microbiota in some individuals, but were not commonly observed amongst different individuals.
Table 3.3: **Common diversity and richness estimators calculated by using OTUs of ≥97% sequence similarity.** The median value for each sample group is reported. Reads mapped refers to the proportion of amplicon sequences that could be mapped to preexisting clone library data. OTUs mapped refers to the number of representative OTUs that could be mapped to preexisting clone library data.

<table>
<thead>
<tr>
<th>Sample group</th>
<th>Good's coverage</th>
<th>Shannon diversity index</th>
<th>Shannon evenness index</th>
<th>Simpson diversity index</th>
<th>Chao1 estimator</th>
<th>ACE estimator</th>
<th>Reads mapped (%)</th>
<th>OTUs mapped (%)</th>
</tr>
</thead>
</table>
Table 3.4: Differences in core kakapo microbiota based on different partitioning of samples. OTU labels are provided to allow consistency for comparisons with other tables that report OTU abundances and changes. Values report the mean relative abundance (percent) of an OTU in the overall microbiota for its grouping. Values in boldface type denote core OTUs in their group, while values in lightface type indicate that an OTU was variable. The final row reports the total proportion of the microbiota that is accounted for by these OTUs (percent).

<table>
<thead>
<tr>
<th>OTU</th>
<th>OTU taxonomy</th>
<th>Mean relative abundance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Overall</td>
</tr>
<tr>
<td>OTU01</td>
<td><em>Escherichia</em></td>
<td>34.45</td>
</tr>
<tr>
<td>OTU02</td>
<td>Unclassified <em>Proteobacteria</em></td>
<td></td>
</tr>
<tr>
<td>OTU03</td>
<td><em>Streptococcus</em></td>
<td>16.33</td>
</tr>
<tr>
<td>OTU04</td>
<td><em>Clostridium</em></td>
<td>5.73</td>
</tr>
<tr>
<td>OTU05</td>
<td><em>Enterococcus</em></td>
<td>1.04</td>
</tr>
<tr>
<td>OTU06</td>
<td><em>Lactobacillus</em></td>
<td></td>
</tr>
<tr>
<td>OTU07</td>
<td><em>Clostridium</em></td>
<td></td>
</tr>
<tr>
<td>OTU10</td>
<td><em>Pseudomonas</em></td>
<td></td>
</tr>
<tr>
<td>OTU11</td>
<td><em>Lactobacillus</em></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>57.55</td>
</tr>
</tbody>
</table>
3.4.2 The microbial community structure does not vary between juvenile and adult individuals, but changes with time

Adult and juvenile community structures were not significantly different from each other under any distance/OTU combination when juvenile and adult samples from matching time points were compared. Comparing the juvenile samples sequentially revealed no differences in community structure (AMOVA, all distances, all OTUs). The community structures of the first and last juvenile samples were significantly different (p < 0.001, all distances, all OTUs) and this difference was also observed for the adult samples (p < 0.001, all distances, all OTUs). The progression of the community structure through time is visualised in Figure 3.2.
**Figure 3.2:** Changes in community structure in wild kakapo samples. Non-metric multidimensional scaling of the weighted UniFrac distances between individual samples obtained from wild kakapo. Left: distances calculated based on OTU$_{0.97}$ (stress = 0.15, $r^2$ = 0.90). Right: distances calculated based on OTU$_{0.95}$ (stress = 0.15, $r^2$ = 0.91).
Table 3.5: Statistically significant changes in OTU abundance between sample groups of interest. Statistical testing and q value corrections were performed across the entire OTU table, but only OTUs that accounted for \( \geq 1\% \) of the bacterial community are reported. OTUs marked with an asterisk were observed in 2 individuals within the Juvenile.First cohort and thus may reflect random interindividual variation rather than a cohort-related difference. OTU labels are provided to allow consistency for comparisons with other tables that report OTU abundances and changes.

<table>
<thead>
<tr>
<th>OTU</th>
<th>OTU Taxonomy</th>
<th>Mean relative abundance (%)</th>
<th>q value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTU06</td>
<td><em>Lactobacillus</em></td>
<td>11.16</td>
<td>( &lt;0.001 )</td>
</tr>
<tr>
<td>OTU07</td>
<td><em>Clostridium</em></td>
<td>4.49</td>
<td>( &lt;0.001 )</td>
</tr>
<tr>
<td>OTU10</td>
<td><em>Pseudomonas</em></td>
<td>1.41</td>
<td>( &lt;0.001 )</td>
</tr>
<tr>
<td>OTU16</td>
<td><em>Lactobacillus</em></td>
<td>1.91</td>
<td>( &lt;0.001 )</td>
</tr>
<tr>
<td>OTU19</td>
<td><em>Lactobacillus</em></td>
<td>2.58</td>
<td>( &lt;0.001 )</td>
</tr>
<tr>
<td>OTU50</td>
<td><em>Clostridium</em></td>
<td>1.78</td>
<td>( &lt;0.001 )</td>
</tr>
</tbody>
</table>

3.4.3 Relative OTU abundance changes within the faecal microbiota

Although the community structure overall did not vary significantly between early juvenile and adult sample groups, statistically significant changes in the relative abundances of OTUs were detected. Six OTUs, generally classified as lactic acid bacteria, were present at significantly higher levels in the Juvenile.First microbiota than in the time-equivalent Adult.First (Table 3.5). These OTUs reduced in abundance in the juvenile microbiota and there were no significantly different OTU abundances between the Adult.First and Juvenile.Third communities. We conclude that the apparent increase in abundance of these OTUs is associated with the young age of the Juvenile.First cohort, with the exception of two OTUs (Table 3.5, asterisk) that were observed in only two of the individuals comprising the Juvenile.First cohort and that may represent individual-specific variation in the microbiota.

The potentially confounding effects of captivity on the microbiota of Juvenile.Second and its subsequent development was a cause for concern within this study, although given the nature of kakapo conservation this was unavoidable. Due to the
differential treatment of juveniles in captivity, and the rich wild-type data from the Adult_First, Juvenile_First and Juvenile_Third groups, some inference can be made from the data, although pending additional testing with an adequate control group these findings remain speculative. OTUs that fluctuated significantly throughout the captivity period are reported in Figure 3.3 and are summarised as follows: OTU03 (Streptococcus) showed a strong, negative correlation with the β-lactam antibiotic treatment kakapo group when compared to any non-treated group ($r_{pb} = -0.45$, $p = 0.009$) and a weaker, statistically insignificant correlation with captivity overall ($r_{pb} = -0.16$, $p = 0.37$). Conversely, OTU05 (Enterococcus) was enriched in the antibiotic-treated ($r_{pb} = 0.7$, $p < 0.0001$) and captive kakapo ($r_{pb} = 0.59$, $p = 0.0003$) when compared to non-treated groups. Samples obtained from the captivity without antibiotic treatment showed no differences that could not be explained by the observed differences between Juvenile_First and Juvenile_Third or the lack of difference between Juvenile_Third and Adult_First (i.e. some OTUs were enriched in the youngest juvenile samples, but not in the later samples or adult samples, implying that individual age was driving the decline in abundance).
Figure 3.3: Statistically significant changes in relative OTU abundance during captivity. Groups Captive+AB and Captive−AB refer to the cohort Juvenile_Second split by whether or not antibiotics were administered. Excluding changes in OTU abundance that could be attributed to bird age, only OTU03 and OTU05 were significantly different in antibiotic-treated samples compared to all other kakapo samples (including captive kakapo without antibiotic treatment).
3.5 Discussion

The gut microbiota of the kakapo is likely to contribute greatly to the health and well-being of the bird, however until recently (Waite et al. 2012, 2013) it remained virtually unstudied. Here we document, for the first time, temporal changes in the bacterial communities within the kakapo GI tract, but do not find consistent differences between juvenile and adult birds at a community-wide level. Amplicon pyrosequencing confirms our previous conclusion that the kakapo microbiota is an uneven, low-diversity community. Although differences in experimental factors and sampling depth mean that between-study comparisons must be treated with caution, the diversity and richness estimators calculated in this study are nonetheless low compared to those obtained from other avians. A recent analysis of the emu hindgut reported a mean Shannon diversity index of 3.4 (kakapo = 0.95) and Chao1 richness of 624 (kakapo = 60.5) (Bennett et al. 2013), while clone-library based analysis of the chicken cecum reported a Chao1 value of 121 (Gong et al. 2007). Molecular analysis of the hoatzin cecum (Godoy-Vitorino et al. 2012) reported a median inverse Simpson diversity of $>400$ (kakapo = 1.78). Kakapo and hoatzin are frequently compared, with the kakapo occasionally mentioned as a potential candidate for avian foregut fermentation. Despite the marked difference in microbial diversity between the kakapo and hoatzin hindgut, the taxonomic differences between these microbiota are not great. Hierarchical clustering of phylotyped data obtained from hoatzin hindgut analysis (Godoy-Vitorino et al. 2012) reveals no separation between the kakapo juvenile and adult faecal microbiota and the hoatzin ceca microbiota (Figure 3.4). While the microbiota of the adult kakapo crop is unknown it is interesting to observe a degree of convergence in the hindgut microbiota of these two geographically isolated birds.
Figure 3.4: **Visualisation of bacterial community structure of the kakapo and hoatzin hindgut microbiota.** Unweighted Pair Group Method with Arithmetic Mean clustering of kakapo and hoatzin hindgut microbiota based on phylotyping community data at the class (left) and genus (right) level. Distances were calculated using the Jaccard index using data obtained from (Godoy-Vitorino et al. 2012). Hoatzin sample is bolded and appeared to cluster with the kakapo samples, rather than clustering away from kakapo communities.
Chapter 3. *Influence of hand-rearing*

The greater sequencing depth afforded by amplicon pyrosequencing has allowed us to examine the microbiota in greater detail than our previous effort (Waite et al. 2012). Indeed, while the core microbiota in the current analysis consisted of only *Gammaproteobacteria* and *Firmicutes*, members of additional phyla were also frequently detected in the birds tested. Representatives of the *Bacteroidetes* were detected in approximately 70% of samples, although often at levels lower than our previous methodology was capable of detecting (Waite et al. 2012). This finding is somewhat at odds with our previous conclusions, and emphasises the value of increased sequencing depth, even in apparently low-diversity environments. *Bacteroidetes*-associated OTUs occurred in an individual-specific manner so although *Bacteroidetes* were observed in many samples, no *Bacteroidetes*-OTUs classified into the kakapo core microbiota. In other avian systems, the most abundant bacterial phyla detected with molecular methods are the *Firmicutes* and *Bacteroidetes*, followed by *Actinobacteria* and *Proteobacteria* (Kohl 2012), which is notably different to the bacterial community profile of the kakapo. The functional implications of this differing community composition compared to that of other herbivorous birds is a matter for future investigation, and knowledge of the core bacteria that comprise the kakapo microbiota will be of benefit for potential future kakapo bacteriotherapy and probiotic development (Waite et al. 2013).

The lack of differentiation between juvenile and adult community structures is an interesting finding, as strong differences have been reported in other avian systems in which the juveniles were of a similar age to those studied here (Gong et al. 2008; Van Dongen et al. 2013), or older (Godoy-Vitorino et al. 2010). Despite the lack of difference in community structure, some OTUs varied significantly in abundance between younger and older birds, likely reflecting at least some role of bird age in shaping the microbiota. Due to concerns regarding the handling of nesting adults, adults sampled in this study were not the parents of the studied juveniles and to the authors’ knowledge there was no contact between these juvenile and adult birds. We speculate that this apparently homogenous bacterial community may be due to the small size of Codfish Island, with all adult kakapo utilising an almost identical diet. It has been observed that kakapo on Codfish Island share overlapping home ranges, with individuals often sharing feeding stations when supplemental feed is provided. Further study into the development of the juvenile microbiota may better resolve this pattern by using a finer time-scale to investigate the community structure much sooner after hatching. Although age did not appear to influence community structure, the relative abundance of
particular OTUs (including those within the core microbiota) varied with age and time, consistent with the wider avian literature (Gong et al. 2008; Lu et al. 2003; Scupham 2007). In general, microbial OTUs that were classified as part of the overall core microbiota were present at high levels within the faecal samples and in each sample grouping the core and variable OTUs represented over half of the microbiota (Table 3.4), indicating that these microbes are likely of biological significance to the kakapo. It was interesting to note that the core microbiota of the juvenile samples accounted for more of the total microbiota and comprised more OTUs than that of the adults. The core microbiota was even more conserved in the antibiotic-treated kakapo, which may be a consequence of a reduction in available niches brought on by a controlled diet and antibiotic pressure.

The need to maintain the health of juvenile kakapo is clearly paramount from a conservation standpoint, and accordingly it was not possible to maintain a wild juvenile control group when individuals needed to be taken into captivity. With the entire juvenile kakapo population in captivity, there was no adequate control group for comparison other than the adult group, but we nevertheless attempted to identify significant changes in the microbiota that correlated with this period. Studying kakapo that were captive but not treated with antibiotics allowed us to tease apart the influence of diet and antibiotic treatment, albeit with a smaller sample size than is ideal. The reported changes in the microbiota were more strongly correlated with antibiotic treatment (OTU03 $r_{pb} = -0.45$, OTU05 $r_{pb} = 0.7$) than captivity in general (OTU03 $r_{pb} = -0.16$, OTU05 $r_{pb} = 0.59$), implying that antibiotic treatment was the primary factor driving the observed differences during captivity. We also observed changes in the microbiota of captive, antibiotic-treated individuals, which adds further evidence that captivity alters the kakapo microbiota. The lack of a juvenile control group makes these findings tentative, but there appears to be no long-term impact of these changes as the bacterial community structures of treated juveniles and untreated adults converged in later samples and no differences in relative OTU abundance were detected between adults and juveniles following release from captivity. In future breeding seasons it would be desirable to attempt to repeat this experiment with finer time scales and more rigid control groups to validate these data (bird health permitting).

A previous meta-analysis (Shade et al. 2013) has reported that microbial communities vary through time across a range of natural ecosystems. Consistent with this observation, the kakapo faecal microbiota varied over time, with juvenile and
adult samples differing after approximately one year between samples. It is possible that this change is related to the change in diet, as reproduction in the kakapo correlates with the fruiting of specific trees native to New Zealand. In addition, supplementary feeding is a common practice during breeding seasons and these changes in kakapo ecology provide a mechanism that may account for the apparent changes in gut microbiota over time. When comparing differences between the early and later samples, a Proteobacteria-associated OTU (Figure 3.3, OTU02) was only sporadically observed in the early samples but was one of the most abundant OTUs at the final time point (Figure 3.1). This OTU could not be classified below the phylum level, with results generally alternating between the classes Betaproteobacteria and Gammaproteobacteria. Additional chimera testing was performed using both UCHIME and the chimera.slayer command in mothur against the SILVA gold database (Schloss 2014a), but this OTU was not determined to be chimeric in either test. The presence and identity of this OTU will be important in future analysis of the kakapo microbiota as its identity may be better resolved with longer sequence data or a different amplicon region. Changes in the microbiota over time carry significance for the wider field of disease ecology, as they emphasise the need for up to date knowledge of animal-associated microbiota in order to better enable pathogen detection (Belden and Harris 2007).

Cut-off values of 95% and 97% sequence similarity are commonly reported in the literature, although there is concern over what value is appropriate (Keswani and Whitman 2001; Stackebrandt and Ebers 2006; Koeppel and Wu 2013). In order to account for potential biases introduced by the level of OTU similarity, sequence data were binned into two different OTU definitions (OTU0.95, OTU0.97) and statistical testing performed between each sample group using each OTU cut-off. The overall robustness of a finding was not only based on the statistical significance but also whether a finding was consistent across different OTU definitions and distance calculations. Statistically significant findings were preserved across OTU definition, giving confidence that the conclusions regarding community structure were not merely artefacts of the bioinformatic approach.

Overall our findings suggest that the faecal community structure of kakapo is not influenced by host age per se, but does change through time. That age is not a significant factor in shaping the community structure of the kakapo microbiota contrasts with patterns seen in other avian hosts, although age-related differences in OTU abundance indicate that age still plays a role in shaping at least a subset
of the kakapo microbiota, although possibly on a shorter timescale than normally observed. By comparing our juvenile data to data obtained from adult kakapo, we have teased apart age- and time-based differences and established a working core set of bacteria for further study in kakapo microbiology. Finally, in confirming that the microbiota changes over time we have reaffirmed the need for continuous sampling of the microbiota of this endangered bird in order to ensure that knowledge of the ‘healthy’ microbiota is accurate. The described data - in combination with ongoing ecological and genetic studies - should contribute to the management and ultimately the survival of this ancient parrot.
3.6 Acknowledgments

This work was supported by funding from The University of Auckland Faculty Research Development Fund (grant 9841 3626187). D.W.W. was supported by a University of Auckland Doctoral Scholarship.

We gratefully thank Ron Moorhouse, Deidre Vercoe and Jo Ledington (Department of Conservation) for their support with this project. We also thank Peter Deines and Siân Morgan-Waite for helpful comments on this manuscript.
Chapter 4

Quantifying the Impact of Storage Procedures for Faecal Bacteriotherapy in the Critically Endangered New Zealand Parrot, the Kakapo (*Strigops habroptilus*)

David W. Waite\(^1\), Peter Deines\(^2\), and Michael W. Taylor\(^1\)

**1.** Centre for Microbial Innovation, School of Biological Sciences, The University of Auckland, Auckland, New Zealand

**2.** Institute of Natural & Mathematical Sciences, Massey University, Auckland, New Zealand

Published in the Peer-Reviewed Journal *Zoo Biology*  
September 9\(^{th}\), 2013

*Zoo Biology*, Volume 32, Pages 620 - 625  
DOI: 10.1002/zoo.21098
4.1 Abstract

The endemic New Zealand kakapo is classified as ‘critically endangered’ and, in an effort to prevent extinction and restore the kakapo population, intensive handling of rare kakapo chicks is often utilised to reduce mortality and improve health outcomes among juveniles. Due to concerns that hand-reared chicks may not receive a full bacterial complement in their gut in the absence of regurgitated food from their mother, conservation workers feed a suspension of frozen adult faeces to captive chicks. However, the efficacy of this practice is unknown, with no information about the viability of these bacteria, or whether certain bacterial taxa are selected for or against as a consequence of freezing. In this study we experimentally determined the effects of freezing and reanimation on bacterial cell viability and diversity, using a faecal sample obtained from a healthy adult kakapo. Freezing reduced the number of viable bacterial cells (estimated by colony-forming units) by 99.86%, although addition of a cryoprotectant prior to freezing resulted in recovery of bacterial cells equivalent to that of non-frozen controls. Bacterial taxonomic diversity was reduced by freezing, irrespective of the presence of a cryoprotectant. While this study did not address the efficacy of faecal supplementation \textit{per se}, the obtained data do suggest that faecal bacteriotherapy using frozen faeces (with a cryoprotectant) from healthy adult birds warrants further consideration as a conservation strategy for intensively managed species.
4.2 Introduction

The kakapo (*Strigops habroptilus*) is a critically endangered parrot endemic to New Zealand. Due to their poor defence against mammalian predators and slow reproductive rate, kakapo numbers have been decimated (Lloyd and Powlesland 1994). While kakapo were once abundant throughout New Zealand, their population is now limited to three predator-free islands off the New Zealand coast. Kakapo reproduction is tied to the mast fruiting cycle of native rimu trees which occurs roughly once every two to six years (Houston et al. 2007). Combined with small clutch sizes, this makes the recovery of the kakapo population an extremely slow process, with every newborn chick subject to intense handling and veterinary scrutiny to ensure survival to sexual maturity. Due to health complications that can arise from rapid or slow weight gains, juvenile kakapo are often taken into captivity and fed a controlled diet to keep their weight gain within healthy parameters. If kakapo are displaying signs of illness, they are prophylactically treated with broad-spectrum antibiotics in an attempt to prevent the establishment of pathogenic microorganisms.

Faecal microbiota transplantation, often referred to as faecal bacteriotherapy, is a procedure developed to combat chronic infections in humans, most notably chronic *Clostridium difficile* infections (Eiseman et al. 1958; Brandt 2012; Lemon et al. 2012; Weissman and Coyle 2012). It is conceptually simple, utilizing the transplantation of faecal material from a healthy donor into the patient’s gastrointestinal (GI) system via a range of procedures (Duplessis et al. 2012; Kassam et al. 2012; Kelly et al. 2012), as an alternative to conventional antibiotic treatment (Pépin et al. 2005; Zar et al. 2007). As a treatment of *Clostridium difficile* infections, success rates vary based on method of administration, with upper GI administration achieving a 76% rate of resolution, and lower GI over 90% (Gough et al. 2011). The rationale for such treatment is that while treatment with antibiotics can weaken the infection, with no way of promoting the healthy gut microbiota to establish, there is no guarantee of recovery. Re-introducing the ‘normal’ gut microbiota to the patient and allowing competition between the pathogen and ‘normal’ community provides a dual mechanism for both inhibition of the pathogen, and re-establishment of the healthy community. A recent study (Van Nood et al. 2013) found the transplantation of faecal material with a success rate of 94% to be 3-4 times more effective than the antibiotics treatment in curing gut infections with *Clostridium difficile*. An alternative approach to such treatment, which relies on
the same principle, is the use of oral probiotics. Such treatment has been clinically tested in humans for a variety of infections (Shornikova et al. 1997; Van Niel et al. 2002; Lin et al. 2005; Fang et al. 2009), and is also being investigated in poultry farming for similar reasons to its consideration in the kakapo recovery program (Morishita et al. 1997; Willis and Reid 2008; Cao et al. 2012).

It is possible that such a treatment could be utilized in the kakapo management program to combat the effects of hand-rearing and antibiotic treatment, and indeed a similar practice has been applied in a previous breeding season (Eason and Moorhouse 2006). In this event, fresh faeces were collected from healthy adult kakapo during routine inspections, screened for known pathogens and frozen, to later be fed to juveniles. The impact of such freezing on the bacterial cell load is unknown and may result in the death of important bacterial species within the faeces. On the other hand, if the faeces are not frozen then it is likely that the faecal community will change, due to changes in temperature, oxygen availability, depletion of nutrient content, and potential establishment of airborne or human-associated bacteria during handling. It is impractical to collect fresh faeces every time a chick needs to be fed, especially when the need for pathogen screening is considered.

A potential alternative is to freeze pathogen-free faeces with a cryoprotectant in order to suspend microbial activity without damage to the cells. The use of non-toxic cryoprotectant agents is a common practice, with chemicals such as glycerol, skim milk, and sugar compounds frequently cited either as stand-alone reagents (Bonten et al. 1997; Dan et al. 1989; Cody et al. 2008; Savini et al. 2010; Kanmani et al. 2011) or in combination (O’Brien et al. 2001; Kanmani et al. 2011; Navarta et al. 2011). The relative efficacies of these reagents is highly variable, although as stand-alone reagents glycerol and skim milk additives are consistently more effective than either sugar or other commercial additives (Bonten et al. 1997; Kanmani et al. 2011; Navarta et al. 2011). Between skim milk and glycerol additives, there are varying reports of comparative efficacy, although these are often confounded by additional factors such as the bacterial strains tested, quantities used and freezing regime (Cody et al. 2008; Savini et al. 2010; Kanmani et al. 2011; Navarta et al. 2011). For this experiment, glycerol was selected as a cryoprotectant as it is routinely used in our lab when creating long-term stocks of bacterial cultures, has low toxicity to mammals and birds (Cerrate et al. 2006), and is an ingredient in common commercial parrot feed. Although a freezing
regime of -70°C or liquid nitrogen is often recommended for maximum recovery (Dan et al. 1989; Siberry et al. 2001), the technical practicalities of working on Codfish Island prevent these being a viable option, so a freezing temperature of -20°C was utilised.

Previous molecular analysis of the bacterial community of the kakapo GI tract revealed a low diversity system, dominated by several abundant bacterial species, generally members of the genera *Escherichia/Shigella* and *Streptococcus*, as well as a range of other lactic acid bacteria (Waite et al. 2012). In addition, the faecal communities of juvenile and wild adult kakapo are not different in terms of membership so it is possible that feeding adult faeces to juvenile kakapo may seed the juvenile with the ‘correct’ bacteria and promote the establishment of the ‘normal’ gut microbiota. While this approach has the potential to work, fundamental questions regarding the bacterial survival rates from freezing must first be addressed. The primary aim of this research was to investigate the effects of freezing on bacterial survivorship, and to determine the efficacy of glycerol as a cryoprotectant for faecal samples, to preserve both bacterial community density and diversity.
4.3 Materials and Methods

4.3.1 Sampling

Fresh kakapo faeces were collected aseptically on Codfish Island from a healthy adult male kakapo (named Bonus) and immediately stored in a sterile polypropylene tube on ice. The sample was not frozen, in order to prevent potential cell death due to freeze/thaw cycling. The sample was shipped overnight to The University of Auckland and processed immediately.

4.3.2 Cultivation and identification of bacteria

Upon arrival, the sample was separated into three portions: the first was plated immediately (hereafter referred to as ‘fresh’ sample), while the second faecal portion was frozen at -20°C for 35 days (‘frozen’ sample). The final portion was mixed with sterile glycerol to a final concentration of 15%, briefly vortexed, then frozen at -20°C for 35 days (‘glycerol + frozen’ sample). For bacterial cultivation, 100 mg of faecal sample was suspended in sterile phosphate-buffered saline (PBS) and serially diluted (from $10^{-2}$ to $10^{-7}$). Fifty microliters of each dilution was then spread-plated in triplicate onto agar plates containing Luria Broth (LB), Brain-Heart Infusion (BHI), M17 and MRS agar, chosen for their ability to select for bacteria previously observed in kakapo faeces (Waite et al. 2012). LB and BHI plates were incubated both aerobically and anaerobically. M17 and MRS plates were incubated anaerobically only, giving a total of six growth conditions. For anaerobic conditions, cells were grown in an anaerobic chamber (Coy Labs, MI, USA). All growth plates were incubated at 37°C for 24 h, then colonies were enumerated (colony-forming units (CFU)), and 24 random colonies picked from each of the six growth conditions and streaked for purity. Pure isolates were grown in the respective liquid broth and stored at -80°C with 15% glycerol.

To identify the isolated bacteria, isolates were cultivated in liquid media and genomic DNA extracted by heating a portion of culture broth to 94°C for 30 min, then centrifuging at 13,000 rpm for 15 min to separate suspended DNA from lysed cellular material. The 16S rRNA gene was amplified by PCR, using the forward primer 616V (5'-AGA GTT TGA TYM TGG CTC AG-3') (Spring et al. 1998) and
reverse primer 1492R (5′-GGT TAC CTT GTT ACG ACT T-3′) (Polz and Cavanaugh 1998), to amplify a ≈1,500 bp region of the gene. Twenty-five microliter reactions were performed containing 20 mM Tris-HCl, 50 mM KCl (buffer), 1.5 mM MgCl2, 100 µM nucleotide solution (equal parts A, T, G, C), 2.5 µM of each primer, 0.5 units of Taq polymerase and 10 ng template DNA. Cycling conditions were as follows: initial denaturation at 94°C for 5 min, 30 cycles of 94°C for 45 s, 57°C for 45 s, 72°C for 90 s, with a final elongation at 72°C for 7 min. Redundant sequences were identified by performing restriction fragment length polymorphism (RFLP) with each PCR product using the enzyme HaeIII (Invitrogen) according to the manufacturer’s specifications. One representative isolate of each banding pattern, for each growth condition, was sequenced (Macrogen Inc., Seoul, South Korea). Sequences were uploaded to DDBJ/EMBL/GenBank under the accession numbers KC986845 - KC986859.

4.3.3 Total bacterial cell counts from faeces

In order to determine the total density of bacteria in kakapo faeces, we performed a cultivation-independent fluorescent staining of bacterial cells. Three samples of faecal material (100 mg each) were suspended in 1 mL PBS and fixed in 2% (v/v) formalin. Fixed samples were serially diluted and bacterial cells were then filtered onto polycarbonate membrane filters (pore size 0.22 µm, diameter 47 mm, Millipore Ltd.). Filters were cut into sections (≈10 mm chord length), stained with DAPI (4′,6-diamidino-2-phenylindole), washed twice with distilled water and once with 96% ethanol before air drying. Filter pieces were mounted on glass slides and analysed by epifluorescence microscopy. For the enumeration of bacterial numbers at least 500 cells were counted per replicate.

4.3.4 Culture-independent bacterial community analysis via 16S rRNA gene amplicon pyrosequencing

In order to determine whether the faecal microbiota of Bonus was representative of a healthy adult kakapo, DNA was extracted from frozen faecal material using a previously established method (Waite et al. 2012). PCR amplification of bacterial 16S rRNA genes, pyrosequencing and bioinformatics analyses were as reported previously (Simister et al. 2012).
4.4 Results and Discussion

The practice of feeding a frozen faecal suspension to captive kakapo chicks has previously been reported in the scientific literature, although the efficacy of this practice is unknown (Eason and Moorhouse 2006). While the new research described here does not ascertain whether or not faecal bacteriotherapy has any impact (beneficial or not) on the kakapo GI microbiota, it represents a starting point to understand whether there is the potential for such treatment to work. Such an approach could in principle be applied not only to kakapo, but to any intensively managed species.

Our first step was to quantify the rate of bacterial survivorship following freezing of kakapo faeces (Figure 4.1). As a percentage of initial plate counts, averaged across all cultivation conditions, the proportion of surviving cells in the frozen samples was 0.14% that of the fresh samples. This effect was similarly pronounced irrespective of the growth conditions employed. Addition of glycerol prior to freezing resulted in the overall plate count being marginally (18.6%) higher compared with the fresh sample, although this difference was not statistically significant (paired t-test, $t = 2.58$, $p = 0.66$). In order to test statistical significance of differences within a given growth condition, counts obtained from each treatment (fresh, frozen, glycerol + frozen) were compared using a one-way ANOVA testing the null hypothesis of equal means. If a significant difference was observed, pairwise comparisons were made using Tukey’s HSD (Figure 4.1). Under all growth conditions, CFU counts from frozen samples were significantly ($p < 0.01$) lower than for the corresponding fresh sample, whereas frozen + glycerol samples were variable in their response (Figure 4.1).
Figure 4.1: Geometric mean colony counts, represented as colony-forming units/gram of faecal material. Error bars represent 2 standard deviations. Asterisk (*) between two bars represents a statistically significant difference (Tukey’s HSD, *p < 0.01).
**Table 4.1: Identity of each isolate, based on full-length 16S rRNA gene sequence.** Where two matches are given, both hits gave an equal BLAST score

<table>
<thead>
<tr>
<th>Media</th>
<th>Banding pattern</th>
<th>Closest BLAST match</th>
<th>Sequence identity (%)</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>LB (aerobic)</td>
<td>1</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><em>Escherichia coli</em></td>
<td>99</td>
<td>NR_074891.1</td>
</tr>
<tr>
<td>BHI (aerobic)</td>
<td>2</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td>LB (anaerobic)</td>
<td>2</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td>BHI (anaerobic)</td>
<td>2</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td>M17 (anaerobic)</td>
<td>2</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td>MRS (anaerobic)</td>
<td>2</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><em>Escherichia fergusonii/Shigella flexneri</em></td>
<td>99</td>
<td>NR_074902.1/NR_026331.1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td><em>Streptococcus gallolyticus</em></td>
<td>99</td>
<td>NR_074849.1</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td><em>Streptococcus gallolyticus</em></td>
<td>99</td>
<td>NR_074849.1</td>
</tr>
</tbody>
</table>
While the aforementioned data point to a drastic reduction in the number of viable bacterial cells following freezing of kakapo faeces, it appears that the addition of glycerol prior to freezing does ameliorate this effect. However, it is important to consider the identities of those bacteria that survive (or do not survive) freezing, to determine whether the faecal samples used for bacteriotherapy contain the full complement of intestinal bacteria. We therefore sought to identify the bacteria which were cultivated following each treatment. A total of five RFLP patterns, each one interpreted as being representative of an individual bacterial type, were seen across 432 screened isolates. When separated by growth condition, a total of 15 pattern/condition combinations were observed. Representative sequences of each were identified using BLAST (Table 4.1). Consistent with previous molecular characterisation of the kakapo GI microbiota (Waite et al. 2012), these new cultivation-based data identify *Escherichia fergusonii* and *Streptococcus gallolyticus* as the two most abundant microbes within the faeces/hindgut. Our 16S rRNA gene pyrosequencing of the Bonus faecal microbiome supported this finding. From 3503 high-quality sequence reads, *E. fergussoni* comprised 92% of the reads and *S. gallolyticus* 2%, making it the second most abundant bacterium (data not shown). The remaining 6% of pyrosequencing reads were affiliated primarily with members of the phyla *Proteobacteria* and *Firmicutes*, with a single sequence assigned to the phyla *Actinobacteria*. The pyrosequencing data also suggest that the faecal microbiome of Bonus is representative of healthy kakapo; in a comparison with previously obtained faecal samples from eight wild kakapo adults, the faecal bacteria of Bonus only differed significantly from one of those individuals (pairwise unweighted UniFrac test, data not shown).

Following freezing, banding patterns corresponding to *Streptococcus* bacteria were not observed. In the unfrozen faecal cultivation, Streptococcus-affiliated banding patterns accounted for 13 of the 24 isolates cultivated on MRS agar. Assuming that *Streptococcus* is equally sensitive to freezing as *Escherichia*, the probability of failing to sample at least one *Streptococcus* from the frozen samples is $7.4 \times 10^{-9}$, based on the observed frequency prior to freezing. It is thus more likely that the absence of *Streptococcus* is due to a differential impact of freezing on the Gram-positive Streptococcus compared with the Gram-negative *Escherichia* and *Shigella* strains. We conclude that although glycerol treatment is capable of preventing the loss of viable cells, it does not protect all members equally and thus does not protect from a loss of viable bacterial diversity.
Culture-independent enumeration of bacteria revealed a mean bacterial cell count in fresh Bonus faeces of $4.03 \times 10^{11}$ cells per gram of faecal material (S.E. = $4.08 \times 10^{10}$). Although it is difficult to compare to cultivated and uncultivated cell counts due to biases inherent in cultivation this was roughly an order of magnitude higher than the cell counts obtained through culture. Combined with our pyrosequencing data it is clear that there were bacteria within the faecal bacterial community that resisted cultivation. Whether this was due to the cultivation conditions not supporting growth of these microbes or due to competition effects between the relatively fast-growing \textit{E. fergussoni} and these additional bacteria is unclear.

While it is often possible to isolate rare microbes from a wide range of environmental samples using basic cultivation techniques (Sibley et al. 2011; Oakley et al. 2012; Shade et al. 2012), our ability to obtain more unique isolates was likely inhibited by the rapid growth of \textit{Escherichia}-like isolates under all cultivation conditions. While \textit{E. coli} 16S rRNA genes (and presumably cell abundance) are relatively rare in the human gut (Suau et al. 1999; Eckburg et al. 2005; Mariat et al. 2009), the rapid growth of this bacterium under cultivation conditions has led it to be one of the most studied bacteria to date. Given that an \textit{Escherichia}-like OTU represents almost the entire bacterial diversity with pyrosequencing, it is not surprising that this genus was dominant in cultivation efforts. While RFLP analyses were useful to reduce sequence redundancy, they did not map well against the BLAST classification of full-length sequences. This is not surprising when considering a commonly cited limitation of the 16S rRNA gene – its conserved nature can prove problematic when trying to distinguish species- or strain-level differences. This problem has long been known when differentiating between \textit{E. coli} and \textit{Shigella flexneri} (Christensen et al. 1998; Fukushima et al. 2002), and it should be noted that for each of the 13 \textit{Escherichia}-like isolates reported, \textit{E. coli}, \textit{E. fergussoni}, and \textit{Shigella flexneri} were all top matches, with only minor differences in overall BLAST score. The exact phylogenetic relationship of these isolates, as well as their functional roles, is a matter for future analysis.

Due to practical considerations in the management of kakapo chicks, it is not possible to directly test the effects of faecal supplementation on juvenile kakapo at this stage. With the last breeding season occurring in early 2011, another is not expected for several years. In addition, the small clutch sizes make a controlled study difficult due to low sample sizes, and the need to partition these low numbers across multiple test and control groups. The data generated by this study indicate
that freezing drastically reduces overall viable bacterial cell numbers, and also that the rate of survivorship between bacterial species is not uniform. Based on these findings it is unlikely that faecal supplementation in its current form will be sufficient to inoculate juvenile kakapo with the full microbial complement, although it may have some success with the dominant *E. fergussoni* strain. With such rare opportunities to conduct robust field trials of faecal supplementation on juvenile kakapo, we recommend further analysis of faecal handling and storage practices in order to develop an optimal storage and handling practices prior to deployment in the next generation of kakapo. Overall, we believe that faecal bacteriotherapy deserves further attention as a possible tool in the management of endangered species.
4.5 Acknowledgments

We gratefully thank Bethany Jackson (Auckland Zoo) for transporting the required fresh sample material to Auckland. We would like to also thank Daryl Eason, Ron Moorhouse, and Deidre Vercoe (Department of Conservation) for their support of this project. We are grateful to the Department of Conservation for funding, as well as The University of Auckland Faculty Research Development Fund (grant 9841 3626187). DWW was supported by a University of Auckland Doctoral Scholarship.
Chapter 5

Genomic and Metagenomic Analysis of the Faecal Microbiota of the Critically Endangered Kakapo

David W. Waite, Michael W. Taylor

Centre for Microbial Innovation, School of Biological Sciences, The University of Auckland, Auckland, New Zealand

Manuscript in preparation for submission to Environmental Microbiology
Chapter 5. *Metagenomic microbiota of kakapo*

5.1 Abstract

The kakapo is a critically endangered, herbivorous parrot endemic to New Zealand. The kakapo hindgut houses a dense microbial community of low taxonomic diversity. The functional roles of the bacteria that reside in the kakapo hindgut are currently unknown, with only limited cultivation and molecular analysis performed on the microbiota. In order to better understand the kakapo hindgut microbiota we employed metagenomic shotgun sequencing of the faeces of adult and juvenile kakapo, as well as genomic shotgun sequencing of selected kakapo bacterial isolates. Metagenomic analysis revealed a bacterial community of higher diversity and richness than that inferred from amplicon pyrosequencing of the 16S rRNA gene, but seemingly absent of archaea. Notably, functional genes associated with *Bacteroidetes* were detected in all metagenome samples, despite the corresponding 16S rRNA genes being absent from the majority of amplicon data. Detailed functional reconstruction of the metabolic pathways within the bacteria showed that although the taxonomic diversity of the kakapo hindgut is limited, it contains the potential to perform a range of roles which are beneficial to the host bird. Finally, we compared the functional profile of the kakapo metagenome to those obtained from other herbivorous animals. Statistical testing of the community structures showed that the kakapo hindgut most closely resembles that of a hindgut fermenting herbivore. This research represents the first effort to understand the functional roles of the kakapo microbiota, and to address the digestive strategy employed by this parrot.
5.2 Introduction

The critically endangered kakapo (*Strigops habroptilus*) is an endemic New Zealand parrot, known for its unusual diet, nocturnal habits and lack of flight. The kakapo is classified as ‘critically endangered’ and the population reached a low point of 51 birds in 1995 (Clout and Merton 1998; Powlesland et al. 2006). The kakapo population has since recovered to 126 with the assistance of the New Zealand Department of Conservation, with all individuals confined to three predator-free islands off the coast of New Zealand. In order to better understand this parrot, and to aid in conservation efforts, a body of scientific study has been conducted with the aim of better understanding kakapo behaviour and ecology, physiology, nutritional requirements and microbiology. There are several attributes of the kakapo that make it an interesting animal to study: it is the only parrot known to perform lek breeding, it is the world’s heaviest parrot and the only flightless parrot (Merton et al. 1984). The diet of the kakapo is also something of a peculiarity - although kakapo are nominally herbivorous they do not ingest large volumes of plant material, instead preferring to crush plant fibres in their beaks, extracting the juices and discarding the plant fibre (Oliver 1955; Horrocks et al. 2008). The diet of the kakapo generally revolves around shoots and leaves, with the exception of the infrequent breeding seasons during which kakapo feed extensively on the fruit of the native rimu tree. The kakapo breeding cycle is entrained to the rimu mast fruiting, although the exact mechanism that links the two is unclear (Houston et al. 2007).

While the genetics and behavioural ecology of kakapo are well understood, and frequently manipulated to ensure optimal genetic diversity among offspring, kakapo microbiology remains largely unstudied. With the exception of limited earlier study of specific pathogens (Brangenberg et al. 2003; Gartrell et al. 2005), the kakapo microbiota has only recently been investigated, and even this has so far been limited to molecular techniques focusing on the 16S rRNA gene as a marker for microbial species identity (Waite et al. 2012, 2013, 2014). Based on these molecular techniques, it is evident that the kakapo gastrointestinal (GI) tract contains low diversity, but distinct, bacterial communities within the crop (foregut) and hindgut (Waite et al. 2012). Unusually for a vertebrate gut community, the kakapo microbiota is dominated by *Gammaproteobacteria* and *Firmicutes*, with *Actinobacteria* and *Bacteroidetes* observed only at low abundances in most individuals (Waite et al. 2014).
GI tract-associated bacteria have been of interest for over half a century (Dubos and Schaedle 1964) and recent work in a range of organisms has linked the functional activity of the gut microbiota to improved energy harvest from food sources (Hill 1997; Turnbaugh et al. 2006; Torok et al. 2008; Stanley et al. 2012), vitamin and nutrient synthesis (Hill 1997; Uphill et al. 1977) and gut development (Stappenbeck et al. 2002; Björkholm et al. 2009). Such relationships are likely to exist in the kakapo as well, but while the 16S rRNA gene is a well-known and accepted measure of species identity it provides only limited insights into the functional capabilities of microbial OTUs (Langille et al. 2013) and is only a starting point for detailed analysis. With massive reductions in the cost of DNA sequencing in recent years, the use of metagenomics has become much more common as a standard laboratory technique, and it allows for detailed analysis of poorly understood microbial communities without the need for cultivation or targeted amplification of genes of interest or cultivation (Schloss and Handelsman 2005), as well as providing a platform for novel virus discovery (Kristensen et al. 2010; Mokili et al. 2012). Recent applications of metagenomic community analysis have revealed the importance of strain-level differences in community membership that are critical to clinical or functional outcomes, yet overlooked by 16S rRNA gene amplicon sequencing (Morowitz et al. 2011; Sharon et al. 2013).

Even ignoring the microbial ecology aspects of strain diversity within the kakapo GI tract, the basic bacterial roles in the kakapo gut are a black box, with the roles and functions of the microbiota inferred from cultivation and 16S rRNA-based species identification. Microbial fermentation in the gut is a key area of interest, with the production of short chain fatty acids (SCFAs) of major importance to the host vertebrate. Several SCFAs are considered critical products of the gut microbiota: acetate, butyrate, and propionate (Cummings and Macfarlane 1991), as well as lactate, which is of special importance in the avian gut (Brady et al. 1979; Ogata et al. 1982). Although metabolism of these bacterial products by the host is a major source of energy for many animals (Parrett and Edwards 1997; Morrison et al. 2006; Donohoe et al. 2011), there is emerging evidence that these products also have secondary health benefits to the host. Acetate can protect the gut of chickens and mice from pathogen colonisation (Van Der Wielen et al. 2000; Maslowski et al. 2009; Fukuda et al. 2012), while butyrate has roles in reduction of inflammation, inhibition of NF-κB and histone deacetylation (Lührs et al. 2002; Hamer et al. 2008), and propionate may reduce lipogenesis and serum cholesterol (Hosseini et al. 2011). Most of the bacteria identified in human faeces are known
SCFA producers (Chassard et al. 2008), and it has long been known that up to 99% of bacterially fermented SCFAs are absorbed through the host gut epithelium (Von Englehardt et al. 1989; Scheppach 1994). In order to address the functional aspects of the kakapo microbiota and to better resolve the community diversity we employed genomic and metagenomic sequencing of kakapo faecal isolates and faecal material obtained from healthy juvenile and adult individuals.
5.3 Materials and Methods

5.3.1 Extraction and sequencing of genomic DNA

Eight previously obtained kakapo faecal isolates (Waite et al. 2013) were grown to stationary phase and genomic DNA was extracted using a standard enzymatic digest protocol (Feil et al. 2012) with modifications. Isolates were grown in either Luria or M17 broth for 24 h then stored at 4°C for 2 h to allow DNA replication to finish. Cells were then removed from broth by centrifugation and suspended in TE buffer. 740 µL of bacterial suspension was transferred to a fresh 1.5 mL microcentrifuge tube. 40 µL lysozyme (100 mg/mL) (Sigma-Aldrich, St Louis MO, USA), 40 µL 10% SDS and 16 µL proteinase K (10 mg/mL) (Sigma-Aldrich, St Louis MO, USA) were added, and tubes were incubated overnight at 37°C with gentle mixing. 100 µL NaCl (5 M) and 100 µL CTAB/NaCl mixture (described in original protocol) were added, and samples were incubated at 65°C for 10 min. Following incubation, 500 µL chloroform:isoamyl alcohol (24:1) was added and tubes gently mixed by inversion then centrifuged for 15 min at 13,000 rpm. 1 mL supernatant was transferred to a fresh tube and the process repeated with 500 µL phenol:chloroform:isoamyl alcohol (25:24:1). 1 mL supernatant was transferred to a fresh 2 mL microcentrifuge tube to which 0.6 vol isopropanol and 0.1 vol sodium acetate (3 M, pH 5.2) were added. Samples were mixed by inversion and incubated at room temperature for 1 h, followed by centrifugation at 16,000 g at 4°C for 30 min. Supernatant was removed and the pellet washed twice with ice-cold 70% ethanol followed by centrifugation at 16,000 g at 4°C for 10 min. Samples were dried and suspended in 20 µL TE buffer with RNAse A (QIAGEN, Germantown, MS, USA). Samples were incubated for 20 min at 37°C then stored at -20°C until further needed. DNA concentration was calculated using a QuBit Quant-iT DNA high-sensitivity assay and DNA was electrophoresed on a 2% agarose gel to assess shearing.

For the metagenomic analyses, frozen faeces were obtained from three healthy adult (named Bonus, Rakiura and Ellie) and juvenile kakapo (Tia, Waa and Tutoko), including the same sample from which isolates were cultured (Bonus). Prior to extraction, faeces were fractionated using a previously established method (Peris-Bondia et al. 2011). One gram of faecal material was suspended in 5 mL PBS and vortexed for 2 min, followed by centrifugation at 800 g for 2 min. The
upper fraction was collected and centrifuged at 7,500 g for 7 min, then the supernatant removed and pelleted biomass was washed twice with 1 mL PBS. DNA extraction was performed as above, with the following modification: following addition of CTAB/NaCl, samples were incubated at 94°C for 30 min, then briefly cooled on ice. Library preparation and sequencing were performed under the auspices of New Zealand Genomics Limited. Raw DNA was prepared using the Nextera XT kit and samples pooled for sequencing. Sequencing was performed in three runs, with all 14 samples sequenced in each run, allowing for adjustment of template ratios between runs to correct for over-/under-represented samples. Sequencing was performed using the Illumina MiSeq platform using 2 x 250 bp paired-end reads. In addition, metagenomic DNA was extracted from each sample using a standard bead-beating technique (Waite et al. 2012) and subjected to 16S rRNA gene amplicon pyrosequencing in order to compare the amplicon- vs metagenome-derived diversity, using previously reported PCR and library preparation techniques (Simister et al. 2012; Waite et al. 2013).

5.3.2 Bioinformatic analysis

Raw genomic reads were filtered and assembled using the PAGIT toolkit (Swain et al. 2012). Assembled genomes were uploaded to the IMG pipeline (Markowitz et al. 2012) for annotation and functional analysis. Comparisons between amplicon and metagenomic diversity were performed using a combination of mothur, PhymmBL and custom python scripts (Brady and Salzberg 2009; Schloss et al. 2009). Amplicon data and metagenomic sequences were classified into taxonomic bins and an OTU table built using a phylotype approach, grouping each read by its taxonomic origin using the program PhymmBL. In addition, 16S rRNA gene fragments were identified from the metagenomic samples using the communityM prediction software (Parks 2014) with a training set built upon the SILVA SSU 111 database (Quast et al. 2013). Putative 16S rRNA gene sequences were then classified using the naïve Bayesian method against the RDP taxonomic database (Wang et al. 2007; Schloss 2014b). Due to the fragmented state of these 16S rRNA sequences, distance-based OTUs could not be generated and therefore OTU diversity could not be obtained. Diversity estimation was performed using the commonly reported Shannon and Simpson diversity estimators after rarefying the OTU tables to the lowest sample depth obtained from amplicon sequencing.
Metagenomic functional analysis was performed in two ways. First, in order to investigate in detail the functional potential of the kakapo microbiota, reads from all samples were pooled then filtered and assembled as per the initial steps of the PAGIT toolkit, followed by additional metagenomic assembly using MetaVelvet (Zerbino and Birney 2008; Namiki et al. 2012). Gene prediction was performed using Prodigal (Hyatt et al. 2010) and predicted genes identified using BLAST against the NCBI non-redundant (nr) protein database. BLAST output was loaded into MEGAN (Huson et al. 2007) for functional pathway analysis. Differential coverage scores were calculated by exporting gene sequences associated with the phyla Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria from MEGAN and mapping the unassembled, quality filtered sequence reads from each metagenomic sample against these reference genes using UPARSE (Edgar 2013). The resulting mapping file was then used to calculate coverage scores for each gene sequence per metagenomic sample and combined with taxonomic origins using custom python scripts. The resulting coverage table was inputted into the Databionic ESOM Tools package (Ultsch and Moerchen 2005) and an emergent self-organising map generated to demonstrate the co-occurrence patterns of each functional gene, using default settings.

Comparative metagenomic analysis was performed by identifying publicly available faecal metagenomes from IMG/M and MG-RAST for comparison with those of the six kakapo sampled. Samples collected are reported in Table 5.1, and were chosen for being derived from faecal samples obtained from either a hindgut-fermenting herbivore, foregut-fermenting herbivore, or hindgut-fermenting omnivore (human). These categories were chosen due to the unconfirmed nature of the kakapo digestive strategy (foregut vs hindgut fermentation), and also to test for functional similarity to humans as, although nominally wild, kakapo are subject to frequent human handling. Quality-filtered DNA from each sample was downloaded from the respective repository, and gene prediction performed using Prodigal. Predicted genes were then classified against a local copy of the COG database (Tatusov et al. 2000, 2003). COGs were scored as present/absent in each sample, then pair-wise distances calculated using the Jaccard index, and the degree of structuring tested using ANOSIM (Clarke 1993). Specific queries regarding the relative abundance of particular COG were performed by extracting subsets of the gene annotation from the main data set and statistical analysis performed in R (R Core Team 2012).
Table 5.1: **Summary of publicly available metagenome samples used for this study.** Column ‘Data ID’ refers to either the MG-RAST sample ID (numeric) or the NCBI Gold Project ID (prefaced with ‘Gm’). For human metagenome samples, samples from six individuals were selected (Gold Sample IDs Gs0001812, Gs0001825, Gs0001833, Gs0001865, Gs0002025, Gs0002246) which referred to the first sample collected from each individual.

<table>
<thead>
<tr>
<th>Fermentation Strategy</th>
<th>Diet</th>
<th>Organism</th>
<th>Data Source</th>
<th>Data ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foregut fermentation</td>
<td>Herbivore</td>
<td>BigHorn Sheep</td>
<td>MG-RAST</td>
<td>4461345.3, 4461346.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Colobus</td>
<td>MG-RAST</td>
<td>4461355.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gazelle</td>
<td>MG-RAST</td>
<td>4461358.3, 4461360.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Giraffe</td>
<td>MG-RAST</td>
<td>4461358.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rock Hyrax</td>
<td>MG-RAST</td>
<td>4461363.3, 4461364.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kangaroo</td>
<td>MG-RAST</td>
<td>4461365.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Okapi</td>
<td>MG-RAST</td>
<td>4461368.3, 4461369.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Springbok</td>
<td>MG-RAST</td>
<td>4461377.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urial Sheep</td>
<td>MG-RAST</td>
<td>4461379.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Warty Pig</td>
<td>MG-RAST</td>
<td>4461380.3</td>
</tr>
<tr>
<td>Hindgut fermentation</td>
<td>Herbivore</td>
<td>Black Rhino</td>
<td>MG-RAST</td>
<td>4461349.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Capybara</td>
<td>MG-RAST</td>
<td>4461352.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elephant</td>
<td>MG-RAST</td>
<td>4461341.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gorilla</td>
<td>MG-RAST</td>
<td>4461360.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Horse</td>
<td>MG-RAST</td>
<td>4461361.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Orangutan</td>
<td>MG-RAST</td>
<td>4461370.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rabbit</td>
<td>MG-RAST</td>
<td>4461372.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zebra</td>
<td>MG-RAST</td>
<td>4461383.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chicken</td>
<td>MG-RAST</td>
<td>Gm00130</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Panda</td>
<td>IMG/M</td>
<td>Gm00363</td>
</tr>
<tr>
<td>Hindgut fermentation</td>
<td>Omnivore</td>
<td>Human</td>
<td>IMG/M</td>
<td>Gm00341</td>
</tr>
<tr>
<td>Unknown</td>
<td>Herbivore</td>
<td>Kakapo</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
5.4 Results and Discussion

5.4.1 The species-level diversity of the kakapo microbiota is much higher than that indicated by amplicon sequencing

Consistent with our previous studies based on 16S rRNA gene amplification and sequencing (Waite et al. 2012, 2013, 2014), the phylum-level diversity of the metagenomic samples was limited to a pair of dominant phyla, the Proteobacteria (specifically Gammaproteobacteria) and Firmicutes (Figure 5.1). Actinobacteria were also observed in all samples, although at lower proportions, and consistent with our previous data archaea were absent. 16S rRNA genes classifying as Bacteroidetes were only observed in two individuals, but functional genes from Bacteroidetes were detected in all metagenomes. Despite similarities in overall taxonomic classification, there was a marked increase in phylogenetic diversity within the metagenomic shotgun data compared to that of the amplicon data at finer taxonomic resolution (Figure 5.2). Although the 16S rRNA gene is a well-established proxy for bacterial phylogeny and functional speculation (Langille et al. 2013), it possesses limited capability to differentiate at species and strain level. Biologically, this finding has two possible interpretations. The kakapo microbiota could contain a high level of strain diversity which is masked by traditional 16S rRNA-based analysis methods, and therefore not observed at finer taxonomic levels. Alternatively, the microbiota may have a low level of species/strain diversity but these microbes may have an expanded genomic capacity, having obtained an array of genes from environmental sources via mechanisms such as horizontal gene transfer which would not be apparent in a classification-based binning approach.
Figure 5.1: **Taxonomic classification of the kakapo faecal microbiota using three analysis metrics.** Left: taxonomic classification of 16S rRNA gene amplicons analysed by pyrosequencing. Middle: classification of 16S rRNA gene fragments extracted from the metagenomic shotgun sequence data using communityM. Right: classification of all DNA present in the metagenomic shotgun sequencing. Symbols represent the individual kakapo Bonus (A), Rakiura (B), Ellie (C), Tia (D), Waa (E) and Tutoko (F).
**Figure 5.2:** Comparisons of bacterial diversity between sampling approaches. All samples were rarefied to the lowest sequencing depth obtained during sequencing (amplicon data). Data are reported in pairs of grouping (gray: genus-level phylotyping, white: species-level phylotyping). For amplicon OTU data, clustering values of $\geq 95\%$ 16S rRNA gene sequence similarity (genus) and $\geq 97\%$ similarity (species) were used. Top: Shannon diversity estimator, bottom: Simpson diversity estimator.
Previous cultivation of kakapo faecal samples yielded a number of putative *Escherichia* isolates (Waite et al. 2013), from which we randomly selected in order to investigate the genomes of individual microbes. Assembled *Escherichia* genomes were compared to existing *Escherichia* genomes available on IMG, and the results are visualised in Figure 5.3. Consistent with an interpretation of high strain-level diversity, kakapo faeces-derived genomes did not cluster together but were instead split across the tree. In contrast to the expanded-genome interpretation, genomes were not differentiated from previously examined *Escherichia* genomes, which would be apparent as unique branch length belonging to the genomes. Statistical testing of the tree structure shown in Figure 5.3 using unweighted UniFrac (Lozupone and Knight 2005) revealed no statistically significant branch length unique to the kakapo isolates (UniFrac distance = 0.64, p = 0.46). We concluded that the kakapo microbiota contains a diverse community of *Escherichia*-like strains.

Despite only a sporadic occurrence of *Bacteroidetes*-associated 16S rRNA gene sequences (Figure 5.1, left and middle panels), functional gene sequences of supposed Bacteroidetes origin were observed in all metagenomic samples. Extracting 16S rRNA gene fragments from metagenomic data yielded *Bacteroidetes* sequences in three of the samples (Bonus, Rakiura, Tutoko, data not shown). Classification of the extracted 16S rRNA gene sequences using difference taxonomic databases (Greengenes, SILVA) was performed but did not change this finding. Reconstruction of 16S rRNA genes using EMIRGE (Miller et al. 2011, 2013) was also performed but yielded no *Bacteroidetes* sequences. In lieu of adequate marker genes we investigated the possibility that these genes were derived from a different bacterial origin, either through horizontal gene transfer or as a result of poor annotation, by analysing their differential coverage profiles (Albertsen et al. 2013). Observing the co-occurrence patterns of functional genes revealed that all *Bacteroidetes* functional genes were occurring in tight synchrony with functional genes obtained from *Actinobacteria* (Figure 5.4), which may suggest that they are occurring as part of *Actinobacteria* genomes. Supporting this observation is the fact that the metagenome richest in *Bacteroidetes* genes (Ellie, 0.5%) contained no corresponding 16S rRNA genes in either amplicon or metagenome-extracted 16S data. Regardless, without cultivating these organisms this hypothesis cannot be confirmed, but it does provides an example of 16S rRNA amplicon data not adequately detailing the true functional diversity of a microbial system.
**Figure 5.3:** Comparison of sequenced *Escherichia* genomes from this study with existing *Escherichia* genome sequences available in IMG. Genome differences are calculated based on COG profiles and the tree constructed using hierarchical clustering with an outgroup of 10 *Bacteroidetes* genomes. Kakapo-derived genomes are highlighted in bold type, and contain the taxonomic classification of the 16S rRNA gene sequence in brackets.
Figure 5.4: Emergent self-organising map of the differential coverage patterns for each functional gene detected in the kakapo microbiota. Points are coloured as red (Proteobacteria), blue (Firmicutes), green (Actinobacteria), yellow (Bacteroidetes), other phyla (cyan). Similar background colour represents neurons of similar height.
5.4.2 Functional potential of the kakapo microbiota

De novo assembly, gene prediction and annotation of the kakapo metagenome yielded a total of 235,693 genes, which were predominantly bacterial in origin (98.95%). Only 14 genes of archaeal origin were detected, which is consistent with our previous finding that archaea are extremely rare, if not absent, within the kakapo microbiota (Waite et al. 2012). The remainder of genes were either of eukaryotic (avian and fungal) or viral origin, but were so rare that complete pathway analysis was not feasible. Viral genes were exclusively derived from bacteriophages belonging to the groups Enterobacteria phage N15, Salmonella phage Vi II-E1, podoviridae and the Mu-like virus family.

5.4.2.1 The kakapo metagenome possesses cellulolytic and amylolytic potential

Unsurprisingly for a bacterial community living in the gut of a herbivorous vertebrate, genes involved in the degradation of plant fibres were commonly observed. Genes responsible for the degradation of cellulose to cellubiose and \( \beta \)-glucans were exclusively detected in the Gammaproteobacteria genomes, as were genes involved in the conversion of pectin to pectate (Figure 5.5). However, pathways for converting both cellubiose and \( \beta \)-glucans to glucose were encoded by Gammaproteobacteria, Firmicutes, Actinobacteria, and Bacteroidetes lineages within the metagenome, hinting at functional synergies within the microbial community. Extracellular sucrose could be imported and converted to glucose by members of the Gammaproteobacteria, Firmicutes and Actinobacteria. The metabolism of extracellular trehalose was encoded by two pathways, either through import into the bacterial cell followed by conversion to glucose (Gammaproteobacteria and Firmicutes), or by converting extracellular trehalose to maltose (Gammaproteobacteria and Actinobacteria). Trehalose could be created from the degradation of starch, provided by the Gammaproteobacteria and Firmicutes. While this amylolysis pathway could not be traced to a species-level resolution within the metagenomic data, this pathway has been frequently documented in Escherichia coli and members of the Lactobacilli so was likely being provided by these abundant members of the kakapo microbiota.
Figure 5.5: The propanoate metabolic pathway (KO00640) as represented in MEGAN (version 4.70.4) with kakapo data highlighted. KEGG entries are coloured by taxonomic origin: *Gammaproteobacteria* (red), *Firmicutes* (blue), *Actinobacteria* (green), *Bacteroidetes* (yellow).
Although metagenomic sequencing provides no indication of expression and localisation of bacterial proteins, the wide range of sources through which the Gammaproteobacteria were capable of acquiring glucose may explain why they are generally the most abundant microbes observed in the kakapo hindgut. The presence of starch degradation genes in the kakapo metagenome is an interesting finding, as the natural kakapo diet is extremely starch poor. The historic kakapo diet appears to contain no starch (Horrocks et al. 2008), with only a single analysed sample containing starch in trace amounts. More recently, Butler (2006) found that up to 50% of the plant mass ingested by the kakapo consists of complex carbohydrates but no detectable levels of starch. Analysis of the nutrient contents of the kakapo chick crop, and that of fruits found on Codfish Island, suggests that starch comprises less than 5% of the kakapo diet compared to \( \approx 75\% \) of the diet consisting of cellulose and lignin (Cottam et al. 2006). Despite this, COGs associated with the digestion of starch were observed in all metagenomic samples (Table 5.2) although they were less common and less abundant than COGs for cellulolytic COGs (Table 5.2, Figure 5.6). The supplemental feeding used by the Department of Conservation contains rich sources of starch, including corn, wheat, maize and oats (Products 2015). The addition of these starch compounds to the kakapo diet may have lead to the abundance of Escherichia strains in the kakapo GI tract, where their ability to utilise an abundant nutrient not normally associated with the kakapo diet gives them a competitive edge over the ‘normal’ bacteria.
Table 5.2: The relative frequency of cellulolytic and amylolytic enzyme classes with the kakapo microbiome. COGs are recorded presence/absence and scored for the proportion (%) of faecal metagenomes they are detected in. For genomic data, COGs are recorded as present in all cultivated isolates, or *Escherichia* only.

<table>
<thead>
<tr>
<th>Carbohydrate</th>
<th>Enzymatic activity</th>
<th>COG Identifier</th>
<th>Representation in metagenomes</th>
<th>Representation in genomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellulose</td>
<td>β-Glucosidase</td>
<td>1486</td>
<td>100.0</td>
<td><em>Escherichia</em> isolates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2723</td>
<td>100.0</td>
<td>All isolates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1472</td>
<td>100.0</td>
<td><em>Escherichia</em> isolates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4354</td>
<td>83.3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>β-Xylosidase</td>
<td>3507</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3664</td>
<td>50.0</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3940</td>
<td>50.0</td>
<td>-</td>
</tr>
<tr>
<td>Xylanase</td>
<td></td>
<td>0726</td>
<td>100.0</td>
<td>All isolates</td>
</tr>
<tr>
<td>β-1, 4- Endoxylanase</td>
<td></td>
<td>3693</td>
<td>66.7</td>
<td>-</td>
</tr>
<tr>
<td>Endo-1, 4-β Glucanase</td>
<td></td>
<td>1363</td>
<td>100.0</td>
<td>All isolates</td>
</tr>
<tr>
<td>Starch</td>
<td>Glycoside hydrolases</td>
<td>1449</td>
<td>50.0</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3387</td>
<td>83.3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4945</td>
<td>33.3</td>
<td>-</td>
</tr>
<tr>
<td>Glycosyltransferase</td>
<td></td>
<td>0438</td>
<td>100.0</td>
<td>All isolates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1216</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5581</td>
<td>100.0</td>
<td><em>Escherichia</em> isolates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2943</td>
<td>100.0</td>
<td><em>Escherichia</em> isolates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1807</td>
<td>100.0</td>
<td><em>Escherichia</em> isolates</td>
</tr>
<tr>
<td>Pectin lyase</td>
<td></td>
<td>3386</td>
<td>16.7</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 5.6: Abundance of COGs relevant to the carbohydrate load of the kakapo diet. Data reflect the total abundance of carbohydrate degrading COG families in the kakapo metagenome (top) and cultivated bacteria (bottom). Cellulolytic COGs (left) are as follows; β-glucosidase (red), β-xylosidase (green), xylanase (COG0726), β-1,4-endoxylanase (COG3696) and endo-1,4-β glucanase (COG1363). Amylolytic COGs (right) are as follows; glycoside hydrolases (purple), glycosyltransferases (teal) and pectin lyase (yellow).
Chapter 5. *Metagenomic microbiota of kakapo*

Although, historically, bacterial cellulolysis in the hindgut was thought to be rare (Barnes 1972; McNab 1973; Mead 1989), those studies and reviews were based on cultivation, which is now known to drastically underestimate bacterial diversity and hence the community functional capability. Even before molecular microbiology became commonplace, it was observed that pure cultures obtained from the avian digestive tract could not degrade cellulose, but mixed cultures from the same source could (Domínguez-Bello et al. 1993). Our study is certainly not the first metagenomic analysis of a herbivore gut that encountered the potential for cellulolytic metabolism (Zhu et al. 2011; Pope et al. 2012; Sergeant et al. 2014) and it is therefore likely that the cellulolytic genes observed in the kakapo hindgut are not simply a passive part of the bacterial genomes, but are actively used by the bacteria that reside in the kakapo hindgut.

### 5.4.2.2 Short-chain fatty acid utilisation

Consistent with historic research demonstrating bacterial SCFA production in the avian cecum (Annison et al. 1968), and more recent vertebrate gut microbiology in general (Cummings and Macfarlane 1991; Macfarlane and Macfarlane 2012), a range of SCFA-producing pathways were identified, coding for the production of lactate, butyrate, propionate and acetate. The functional importance (or lack thereof) of such products is unclear in the kakapo but in other systems, both mammalian and avian, their roles have been elucidated. One of the best studied SCFAs in avian physiology is lactate, in which it is a glucose precursor (Brady et al. 1979; Ogata et al. 1982; Watford 1985). Although microbially produced lactate may not be utilised exclusively by the host vertebrate (Scanes and Braun 2013), a review of the microbial metabolism of the avian hindgut concluded that bacterial lactose utilisation is greater than that of lactate, with approximately 50% of anaerobic isolates capable of metabolising lactose, but fewer than 20% metabolising lactate (Mead 1989). Such an environment would suggest that there is an excess produced in the hindgut that goes unused by the microbial community, and this is corroborated by experimental observations that elevating microbially produced lactate does not lead to an increase in production of downstream SCFAs such as acetate and butyrate (Józefiak et al. 2006).

Acetate has been reported as the most abundant SCFA in the avian cecum (Annison et al. 1968; Goldstein 1989) and has been of interest in the chicken, where it
selectively inhibit pathogens but not autochthonous *Lactobacillus* (Van Der Wielten et al. 2000). Acetate can also be an inhibitor of pathogen colonisation in mice (Maslowski et al. 2009; Fukuda et al. 2012), with similar effects also reported for other SCFAs (Marounek et al. 1999; Józefiak et al. 2000). Acetate and lactate production pathways within the kakapo metagenome were evenly distributed among bacterial phyla detected in the faeces, but both butyrate and propionate production appeared biased towards *Proteobacteria* origins. Microbial production of propionate has been linked to *Bacteroidetes, Firmicutes* and *Actinobacteria* (Macfarlane and Macfarlane 2012) or *Escherichia coli* (Hosseini et al. 2011). Within the kakapo, a majority of the common synthesis pathways were encoded by members of the *Gammaproteobacteria* and *Actinobacteria* (Figure 5.5). Butyrate production is usually associated with species from the *Firmicutes* (Macfarlane and Macfarlane 2012; Russell et al. 2013), and indeed *Firmicutes* pathways were observed in the kakapo as well as *Gammaproteobacteria*. It was interesting to note that the method of conversion from butanoyl-CoA to butyrate was different for the *Gammaproteobacteria* and *Firmicutes*, even though they shared the same pathway for converting both pyruvate and acetoacetyl-CoA to butanoyl-CoA (Figure 5.7). It appears that, despite the taxonomic simplicity of the kakapo faecal microbiota, the microbes in the kakapo hindgut are, collectively, functionally equivalent to those of other vertebrates, and have filled the ‘required’ pathways through the *Gammaproteobacteria*. Moreover, functional redundancy exists in these pathways, with different microbial taxa achieving the same end product through unique means.
Figure 5.7: The butyrate metabolic pathway (KO00650) as represented in MEGAN (version 4.70.4) with kakapo data highlighted. KEGG entries are coloured by taxonomic origin: Gammaproteobacteria (red), Firmicutes (blue), Actinobacteria (green), Bacteroidetes (yellow).
5.4.2.3 Nitrogen metabolism within the kakapo metagenome

Bacterial metabolism of uric acid has been suspected since early observations that many faecal isolates were capable of utilising urate as a nitrogen source (Barnes 1972; Mead 1989; Vispo and Karasov 1997), and the observation that in cockerels almost 80% of uric acid is decomposed within one hour of entering the ceca (Karasawa et al. 1988). Stable isotope labelling of urea has shown that nitrogen absorbed from the cecum is primarily in the form of protein, amino acids and urea (Karasawa and Maeda 1995). While kakapo lack ceca, bacteria with the capability for uric acid decomposition have been isolated from the hindgut of birds without ceca (Preest et al. 2003), and as uric acid is the major waste product from avians (Wright 1995) it is an excellent resource for microbial metabolism. Genes responsible for the reduction of urate to ammonia (Actinobacteria and Gammaproteobacteria), nitrate/nitrite to ammonia (Actinobacteria, Firmicutes and Gammaproteobacteria), and nitrogen to ammonia (Gammaproteobacteria), were detected within the faecal microbiota. These three phyla, as well as Bacteroidetes, contained pathways for converting this ammonia to a range of amino acids, which are likely of benefit to the host bird, although further analysis is required to confirm the role of bacteria in uric acid recycling, and rates of activity.

5.4.3 The kakapo faecal metagenome resembles that of a hindgut-fermenting herbivore

The bacterial community structure of the kakapo GI tract is somewhat unusual, with high-throughput amplicon sequencing revealing a surprisingly simple bacterial community, often lacking in common gut bacteria of the phylum Bacteroidetes. Previous analysis of the kakapo microbiota has revealed no statistically significant difference in bacterial community structure between adults and juveniles (Waite et al. 2014) and this lack of overall differentiation was observed within the metagenomic analysis as well (ANOSIM, R = 0.04, p = 0.27). The kakapo has occasionally been suggested as a candidate for foregut fermentation (Morton 1978; Lopez-Calleja and Bozinovic 2000), a practice that is common among mammals but essentially unique to the hoatzin among avians (Grajal et al. 1989). Kakapo have a comparatively low daily energy expenditure; approximately half that of their flying relative, the kea (Bryant 2006), and when normalised for body weight this can be as little as 20% of the intake of other avians (Bryant 2006;
Wolf et al. 2007). Previous research comparing the GI microbiota of vertebrates with differing gut fermentation strategy has demonstrated that the function of the gut provides greater structuring influence than the host vertebrae (Muegge et al. 2011; Godoy-Vitorino et al. 2012), which allowed for speculation about the fermentation strategy of the kakapo gut. Functional annotations of publicly available faecal metagenomic samples from a range of hosts (Table 5.1) were compared.

The overall COG composition of each metagenome was compared, as well as COG categories relevant to gut function such as carbohydrate metabolism, amino acid metabolism and energy metabolism. Due to the fact that the relative abundance of a COG in a sample does not reflect the activity of the COG (i.e. factors such as expression level and localisation of the protein heavily influence the activity), COGs were only scored in terms of presence/absence. The overall metagenomic structure was different between the kakapo and the metagenomes obtained from herbivorous foregut/hindgut fermenters and humans (omnivorous hindgut fermenters) (Table 5.3). However, when the functional annotations were split into roles relevant to gut function, the differences were much less between that of kakapo and herbivores, with the differences between kakapo and herbivorous hindgut fermenters frequently non-significant (both statistically, and in terms of interpreting the ANOSIM Global R value). We concluded that the kakapo faecal metagenome is functionally similar to that of hindgut-fermenting herbivores, although without detailed knowledge of the function of the crop microbiota we cannot state conclusively whether the kakapo performs foregut fermentation.
Table 5.3: Results of ANOSIM testing between the functional community structure of the kakapo faecal metagenome to metagenomes derived from animals with known fermentation strategy. Columns report the Global R value between the kakapo and that listed. Asterisk (*) signifies a statistically significant Global R value (p < 0.005).

<table>
<thead>
<tr>
<th></th>
<th>Herbivore foregut fermentator</th>
<th>Herbivore hindgut fermentator</th>
<th>Omnivore hindgut fermentator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>1.00*</td>
<td>0.80*</td>
<td>1.00*</td>
</tr>
<tr>
<td>Amino acid transport and metabolism</td>
<td>0.57*</td>
<td>0.11</td>
<td>0.75*</td>
</tr>
<tr>
<td>Carbohydrate transport and metabolism</td>
<td>0.59*</td>
<td>0.21</td>
<td>0.89*</td>
</tr>
<tr>
<td>Energy production and conversion</td>
<td>0.76*</td>
<td>0.34</td>
<td>0.93*</td>
</tr>
</tbody>
</table>
5.4.4 Challenges to interpreting the kakapo metagenome

While metagenomics is a powerful tool for rapidly elucidating the potential functionality and interactions within a microbial community, interpretations of results must be performed cautiously as there are assumptions that are frequently made in metagenomic analysis that must be accounted for. The assumption that all genes observed in the metagenome are being expressed, and that when these pathways combine the activity of several different microbes that these microbes are naturally occurring in close proximity, is not always correct. For example, while lactate production within gut communities is performed by *Lactobacilli*, it only occurs during certain parts of the cell growth cycle (Pritchard 1972; Chen et al. 2013) and is therefore not universally expressed. It has also been frequently shown that supplementation of feed with enzymes such as xylanase and glucanase (Onderci et al. 2008; Cowieson et al. 2010; Mathlouthi et al. 2011; Ghahri et al. 2012; Ribeiro et al. 2012) improves animal weight gain, even though bacterial xylanase and glucanase are already present in the chicken cecum (Beckmann et al. 2006; Sergeant et al. 2014), indicating that the bacterial functional potential is not indicative of actual function, at least in terms of efficiency.

More challenging to the biological interpretation of metagenomic data is the fact that it is not immediately clear which products of the metagenome are simple intermediates in the microbial system and which are of benefit or harm to the host. A key example of this issue in the kakapo metagenome (and likely others) is the presence of acetoacetate as a fermentation product. Acetoacetate can influence the regulation of protein synthesis and turnover in avian skeletal muscle (Wu and Thompson 1990; Thompson and Wu 1991) and thus may be of importance to kakapo physiology. Equally likely, it may simply be an intermediate produced during bacterial fermentation of butyrate (Figure 5.7). Likewise, while fumarate has been identified as an inhibitor of methanogenesis in ruminants (Ungerfield et al. 2007), fumarate and malate are also intermediates produced during the Krebs cycle, and so the bacterial fermentation of fumarate in pathways other than the Krebs cycle may simply lead to the production of new intermediates for this crucial, intracellular pathway. As long as caution is taken in interpreting the data, however, metagenomics provides a powerful tool for providing rapid insights into microbial communities like that of the kakapo.
5.4.5 Concluding remarks

To summarise, we have performed the first detailed investigation of the functional diversity of the kakapo faecal microbiota. Our data show that, despite superficial simplicity at the 16S rRNA gene level, the microbiota is diverse and performs functional capabilities similar to that of other herbivores, and reinforce our previous findings that the microbial community structure does not change between juvenile and adult kakapo (Waite et al. 2012, 2014). We have also shown that the functional structure is closer to that of a hindgut-fermenting herbivore than a foregut fermenter. Although this is not proof of the digestive strategy of the bird, it is the first reported evidence (to our knowledge) that addresses this issue. Finally, we have shown that despite an absence of 16S rRNA genes that indicate Bacteroidetes presence, their functional ability is maintained in the microbiota, although the mechanism by which this occurs is unclear.
5.5 Acknowledgments

We are grateful to Philip Hugenholtz for hosting D.W.W at the Australian Centre for Ecogenomics as part of this project, and to Yuji Sekiguchi for his helpful insights into the kakapo metagenome project. This work was supported by funding from The University of Auckland Faculty Research Development Fund (grant 9841 3626187). D.W.W. was supported by a University of Auckland Doctoral Scholarship.
Chapter 6

Characterising the Avian Gut Microbiota:
Membership, Driving Influences and Potential Function

David W. Waite, Michael W. Taylor

Centre for Microbial Innovation, School of Biological Sciences, The University of Auckland, Auckland, New Zealand

Published in the Peer-Reviewed Journal Frontiers in Microbiology
May 16th, 2014

Frontiers in Microbiology, Volume 5, Article 223
DOI: 10.3389/fmicb.2014.00223
6.1 Abstract

Birds represent a diverse and evolutionarily successful lineage, occupying a wide range of niches throughout the world. Like all vertebrates, avians harbour diverse communities of microorganisms within their guts, which collectively fulfil important roles in providing the host with nutrition and protection from pathogens. Although many studies have investigated the role of particular microbes in the guts of avian species, there has been no attempt to unify the results of previous, sequence-based studies to examine the factors that shape the avian gut microbiota as a whole. In this study, we present the first meta-analysis of the avian gut microbiota, using 16S rRNA gene sequences obtained from a range of publicly available clone-library and amplicon pyrosequencing data. We investigate community membership and structure, as well as probe the roles of some of the key biological factors that influence the gut microbiota of other vertebrates, such as host phylogeny, location within the gut, diet and association with humans. Our results indicate that, across avian studies, the microbiota demonstrates a similar phylum-level composition to that of mammals. Host bird species is the most important factor in determining community composition, although sampling site, diet and captivity status also contribute. These analyses provide a first integrated look at the composition of the avian microbiota, and serve as a foundation for future studies in this area.
6.2 Introduction

The role of the gut microbiota in shaping the health and physiology of vertebrate hosts is a well-established, highly exciting area in microbiology. The diversity and function of microbes in the gastrointestinal (GI) tract is an area of ongoing research, with recognised roles for the vertebrate microbiota in nutrition (Jin et al. 1998; Preest et al. 2003; Turnbaugh et al. 2006; Angelakis and Raoult 2010; Stanley et al. 2012), gut development (Stappenbeck et al. 2002; Rahimi et al. 2009; Zhang et al. 2011; Cao et al. 2012) and regulation of host physiology (Bäckhed et al. 2004; Björkholm et al. 2009; Meini et al. 2009). 16S rRNA gene sequencing has been employed in a range of studies to assess the diversity and phylogenetic relationships of gut microbes and this has proven to be a powerful tool for understanding the factors that shape microbial communities, due to both its informative and predictive potential. A secondary benefit of the 16S rRNA gene is that, in addition to reporting the results of findings in scientific journals, it is customary to deposit the primary sequence data into publicly available databases which allow for a second wave of meta-study. By aggregating data from a variety of sources or environments, researchers have been able to discern large-scale patterns in microbial ecology, analysing the bacterial communities of mammalian (Ley et al. 2008a) and fish (Sullam et al. 2012) guts, as well as across other non-biological factors (Lozupone and Knight 2007; Chu et al. 2010; Shade et al. 2013). One area that has arguably not undergone such a revolution is that of the avian microbiota. While several notable exceptions exist, such as commercially farmed broiler chickens and turkeys as well as the South American hoatzin, the majority of avian systems have not been studied outside of immediate pathogenic concerns.

Similar to other vertebrates, the GI tract of birds is colonised by a community of microbes, with a density as high as $10^{11}$ c.f.u / g in the hindgut (Barnes 1972). The role of microbes in the avian gut has long been a topic of study, with groundbreaking research throughout the 1960’s identifying the role of bacteria in starch degradation and volatile fatty-acid production within the bird gut (Bolton 1965; Annison et al. 1968; Pritchard 1972; Pacheco et al. 2004) and microbially mediated fermentation of lactate (Bolton 1962; Pritchard 1972; Moore et al. 2004). Cellulolytic microbes have occasionally been observed in avian
crops (Shetty et al. 1990; Domínguez-Bello et al. 1993), but significant bacterial cellulolysis has only been reported in the hoatzin (Grajal et al. 1989; Domínguez-Bello et al. 1993), with only low levels of cellulose fermentation reported for other birds (Clemens et al. 1975; Cutler et al. 2005). The ceca are the sites of recycling of urea (Barnes 1972; Mead 1989; Vispo and Karasov 1997; Preest et al. 2003), retention of water (McNab 1973) and fermentation of carbohydrates (Józefiak et al. 2004). It has been observed that a cellulose-rich diet leads to increased size of the ceca (Leopold 1953; McNab 1973; Miller 1976; Duke et al. 1984; Redig 1989; Stevens and Hume 1998), but there is contradictory evidence for the direct utilisation of cellulose in the avian hindgut (Barnes 1972; McNab 1973; Mead 1989).

With the rise of 16S rRNA gene sequencing a large portion of avian microbiology has shifted from microbial physiology to the diversity and phylogeny of avian gut microbes. Specific studies have addressed areas of avian microbial ecology, such as the variation in microbial diversity along the GI tract (Bjerrum et al. 2006; Gong et al. 2007; Torok et al. 2008; Waite et al. 2012), the influence of diet (Rubio et al. 1998; Blanco et al. 2006; Torok et al. 2008; Janczyk et al. 2009; Hammons et al. 2010), age (Van Der Wielen et al. 2002; Godoy-Vitorino et al. 2010; Van Dongen et al. 2013) or other host-specific factors (Zhu et al. 2002; Lucas and Heeb 2005; Banks et al. 2009; Benskin et al. 2010; Wienemann et al. 2011). While there is extensive evidence that microbial colonisation of the GI tract brings benefits to the host bird (Jin et al. 1998; Torok et al. 2008; Angelakis and Raoult 2010; Torok et al. 2011; Zhang et al. 2011; Cao et al. 2012; Stanley et al. 2012), there are also pathways through which the normal colonisation of microbes can be of detriment to the host (Ford and Coates 1971; Potti et al. 2002; Cao et al. 2012; Singh et al. 2013). Although there are many published studies exploring aspects of the avian microbiota, it has evidently been uncommon for authors to publish their sequence data to an archive, somewhat limiting the potential for avian metastudies. As an example of this, in their 2008 meta-analysis of the vertebrate microbiota Ley et al. had access to rich clone-library data from insects (19 studies), humans (20 studies) and other vertebrate species (23 studies, including five from birds) (Ley et al. 2008b). In 2012, Sullam et al. identified for analysis 24 pre-existing clone-libraries derived from fish guts (Sullam et al. 2012). By contrast, in the same year Kohl only identified eight avian libraries with any significant microbiota data (Kohl 2012). A survey of the recent literature has shown that the picture of the avian microbiota has since improved significantly, with the continued usage of
clone-libraries and incorporation of amplicon pyrosequencing into existing study systems (Table 6.1).

In order to gain new insights into the avian gut microbiota, we sought to amalgamate the existing knowledge and determine whether patterns detected in individual studies were consistent across avians as a whole. To achieve this goal we collected publicly available data from NCBI GenBank and MG-RAST and reanalysed the data using established bioinformatics pipelines.
Table 6.1: Published sequence data obtained from molecular analysis of avian samples. Asterisk (*) denotes a study that analysed the bacterial communities associated with multiple species of birds, but with common phylogenetic or geographic grouping. For 16S rRNA gene amplicon data, reference names are the last name of submitter where available. Short-read data with an unknown reference refers to data which could not be tracked back to a published paper.

<table>
<thead>
<tr>
<th>16S rRNA gene clone data</th>
<th>16S rRNA gene amplicon data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host</td>
<td>Site sampled</td>
</tr>
<tr>
<td>Adelie penguin</td>
<td>Faecal</td>
</tr>
<tr>
<td>Capercaillie</td>
<td>Cecum</td>
</tr>
<tr>
<td>Chicken</td>
<td>Cecum</td>
</tr>
<tr>
<td>Ileum/Cecum</td>
<td>Lu et al. (2003)</td>
</tr>
<tr>
<td>Crop/Cecum</td>
<td>Gong et al. (2007)</td>
</tr>
<tr>
<td>Aggregate</td>
<td>Wei et al. (2013)</td>
</tr>
<tr>
<td>Crane</td>
<td>Faecal</td>
</tr>
<tr>
<td>Crop</td>
<td>Wright et al. (2009)</td>
</tr>
<tr>
<td>Crop</td>
<td>Godoy-Vitorino et al. (2010)</td>
</tr>
<tr>
<td>Kakapo</td>
<td>Crop/Faecal</td>
</tr>
<tr>
<td>Shorebirds*</td>
<td>Cloaca</td>
</tr>
<tr>
<td>Gull</td>
<td>Faecal</td>
</tr>
<tr>
<td>Parrot*</td>
<td>Cloaca</td>
</tr>
<tr>
<td>Ostrich</td>
<td>Cecum</td>
</tr>
<tr>
<td>Stork</td>
<td>Feathers</td>
</tr>
<tr>
<td>Turkey</td>
<td>Cecum</td>
</tr>
<tr>
<td>Faecal</td>
<td>Lu and Domingo (2008)</td>
</tr>
<tr>
<td>Cecum</td>
<td>Scupham et al. (2008)</td>
</tr>
<tr>
<td>Aggregate</td>
<td>Wei et al. (2013)</td>
</tr>
</tbody>
</table>
6.3 Materials and Methods

6.3.1 Data acquisition and quality control

Clone-library data were obtained from GenBank through a comprehensive literature survey, followed by the retrieval of clone-library sequence data of interest. Short amplicon data from next-generation sequencing studies were obtained from MG-RAST and the NCBI Sequence Read Archive (hereafter referred to as short-read data) by browsing for the publicly available data sets. Data sources are as reported in Table 6.1, with the exception of the database provided by Wei and colleagues (Wei et al. 2013), which was excluded from analysis as their data overlapped significantly with sequences obtained from original studies.

All downloaded data were re-analysed using mothur version 1.32.1 (Schloss et al. 2009). For short-read data, flowgrams were trimmed to a single length then denoised. Where flowgrams were not available, sequences were trimmed using the trim.seqs command, removing the barcode and primer sequences and discarding sequences with an average quality score of less than 25, or sequences with a homopolymer run of greater than eight bases. All sequence data were then aligned, screened for chimeras with uchime (Edgar et al. 2011) and classified against the Greengenes taxonomy using the naïve Bayesian method (DeSantis et al. 2006; Wang et al. 2007). Sequences that could not be classified to domain level, or were classified as Cyanobacteria, were removed from the dataset as they likely represent ingested plant material. Chimeric sequences and sequences that could not be aligned were also removed from the data set.

For data obtained from clone libraries it is common practice to simply upload representative sequences to GenBank, rather than the complete dataset. In order to account for the loss of abundance information from the original clone libraries, taxonomic classification was reported by calculating operational taxonomic units (OTUs) of 97% sequence similarity for each sample and assigning taxonomy using the classify.otu command in mothur. Although short-read data does contain the data from the complete sequencing run, studies did not always utilise the same 16S rRNA gene region and so could not be directly compared. In lieu of OTU generation, genus-level phylotypes were constructed using the sequence classification. For short-read data, the phylotype table was rarefied to a depth of 1,500 data points and Shannon and Simpson diversity indices calculated.
6.3.2 Correlating metadata to community structure

For clone data, sequences were trimmed to an 800 bp overlapping region and a phylogenetic tree constructed using the clearcut neighbour-joining algorithm (Evans et al. 2006) for UniFrac analysis. Sequences less than 800 bp in length were discarded, resulting in the loss of three avian samples compared with the previous classification. Due to the potential bias in relative abundance incurred by the selective uploading of data, only unweighted UniFrac distance was calculated. For short-read data there was no contiguous region of sequence common to all samples, so analysis was performed by constructing genus-level phylotypes of the classified data. Community differences were calculated using Jaccard (presence/absence) and Yue-Clayton theta (abundance) distance by randomly subsampling each community to 1,500 sequences 20,000 times and averaging the community distances across iterations.

Metadata regarding the host, sample type, animal diet and captivity status were recorded and their impact on community differences compared using the vegan package (version 2.0-8) (Oksanen et al. 2013) in the R software environment (R Core Team 2012). Samples were grouped according to the following categories: host animal, diet and captivity status. Diet consisted of three categories - carnivore, herbivore and grain-fed – that reflected a ‘typical’ diet of the host. When dividing animals based on diet, the distinction was made between an herbivorous diet (leaves and green plant material, such as eaten by the kakapo and hoatzin) and grain-fed diet (pelleted feed, such as found in farmed chickens) due to the different nutrient content and availability in these diets. Captivity status consisted of simply dividing samples into those animals that are wild or farm-raised. For short-read data the study that provided the data was also used as a test for how much the dynamics of the study itself shaped the data. This factor could not be applied to the clone-library data as not every original study uploaded sequences with sufficient information to recapture biological replication with the sequence data.

Permutational multivariate analysis of variance (PERMANOVA) with linear model fitting was performed (Anderson 2001; McArdle and Anderson 2001) in R. Samples were grouped according to each metadata factor and tested for how well the grouping accounted for the variation between samples using the ‘adonis’ function of the vegan package (Oksanen et al. 2013), measured as $R^2$. A significance value
Chapter 6. Characterising the avian gut microbiota

(p-value) was generated by comparing the obtained $R^2$ to that obtained from 1,000 random permutations of the data. For factors with a statistically significant fit, constrained canonical analysis (CCA) was performed (Ter Braak 1986) using the factor as the constraining variable to isolate the contribution of that factor to the microbial community.

6.3.3 Functional prediction of gut microbiota

Following quality control of short-read data, sequences were mapped to OTUs using closed-reference OTU picking in QIIME 1.80 (Caporaso et al. 2010). 16S rRNA gene abundance levels were then normalised against the known gene copy number for that OTU and function predictions made based on OTU membership using PICRUSt (Langille et al. 2013). Functional predictions were categorised into KEGG pathways and statistical analysis performed using STAMP v 2.0 (Parks and Beiko 2010). Data were partitioned by metadata factors and differences in relative abundance tested using ANOVA, followed by post-hoc Games-Howell test with the Benjamini-Hochberg FDR used as a multiple testing correction (Benjamini and Hochberg 1995). For testing the presence of genes involved in cellulose digestion, KEGG data were screened for pathways that mapped to COGs involved in cellulolysis and data extracted. Pair-wise comparisons were performed using Welch’s t-test (Welch 1947) with the Benjamini-Hochberg FDR.
6.4 Results and Discussion

6.4.1 Taxonomic Classification of OTUs

Quality-control of sequence data yielded a high number of high-quality sequences, of varying length and sequenced gene region, from a subset of the studies reported in Table 6.1 (Table 6.2, 6.3). Consistent with the microbiota of vertebrates in general, the avian gut microbiota appears to harbour mostly OTUs belonging to *Bacteroidetes*, *Firmicutes* and *Proteobacteria* (Figure 6.1, 6.2). Members of the phylum *Firmicutes* were present in all samples analysed, while *Proteobacteria* and *Bacteroidetes* were also widespread (*Proteobacteria*: 90% of clone samples, 100% of short-read samples; *Bacteroidetes*: 80% of clone samples, 87% of short-read samples). These three phyla are commonly observed within gut environments, and specific lineages of these phyla are frequently studied for their symbiotic roles, for example *Bacteroides thetaiotaomicron* starch degradation in humans (Don- gowski et al. 2000; Xu et al. 2003; Sears 2005), and *Lactobacilli*-associated bile salt hydrolase activity in mice and chickens (Tannock et al. 1989; Tanaka et al. 1999; Knarreborg et al. 2002). To a lesser extent, *Actinobacteria* (65% of clone samples, 89% short-read samples) and *Tenericutes* (65% of clone samples, 58% short-read samples) were also reasonably common throughout the data. Within the short-read data, a higher proportion of unclassified OTUs was observed, which may be due to a lack of phylogenetic resolution due to shorter read length. Alternatively, it has been shown that the use of the adapter/barcode construct in a single-step PCR, as is commonplace in pyrosequencing studies, can negatively affect taxonomic classification (Berry et al. 2011).

6.4.2 Factors shaping the avian microbiota: study versus host

PERMANOVA testing of the short-read data set revealed that the largest factor contributing to the shaping of the microbiota was the study itself (Table 6.4). This finding may be a real result, as most studies focused on a single bird geographically isolated from other studies (i.e. the ‘study’ variable is the product of host and location), or may be an artefact resulting from the specific DNA extraction and PCR techniques involved (Boom et al. 1990; Suzuki and Giovannoni 1996;
Martin-Laurent et al. 2001; Sipos et al. 2007; Berry et al. 2011; Kennedy et al. 2014). In order to resolve this issue, we hypothesized that if the host species was truly driving the differences observed between studies, then the phylogenetic differences between taxonomically similar bacterial lineages within each study would be smaller between studies with a closely related host bird. Alternatively, a study that investigated a range of host birds would have greater within-study variation than a study that investigated a single host.
Table 6.2: Data obtained from clone-library based studies and the published study that reported the sequences. Number of reads used, OTUs generated and average sequence length for clone library data utilised in the study. Asterisk (*) denotes a study that analysed the bacterial communities associated with multiple species of birds, but with common phylogenetic or geographic grouping.

<table>
<thead>
<tr>
<th>Host</th>
<th>Site sampled</th>
<th>Number Sequences</th>
<th>Number OTUs</th>
<th>Median Sequence Length (bp)</th>
<th>Figure 6.1 Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adelie penguin</td>
<td>Faecal</td>
<td>48</td>
<td>44</td>
<td>846</td>
<td>Banks2009</td>
</tr>
<tr>
<td>Capercaillie</td>
<td>Cecum</td>
<td>114</td>
<td>43</td>
<td>1476</td>
<td>Wienemann2011</td>
</tr>
<tr>
<td>Chicken</td>
<td>Cecum</td>
<td>329</td>
<td>213</td>
<td>433</td>
<td>Zhu2002</td>
</tr>
<tr>
<td></td>
<td>Illeum/Cecum</td>
<td>99</td>
<td>72</td>
<td>644</td>
<td>Lu2003</td>
</tr>
<tr>
<td></td>
<td>Cecum</td>
<td>74</td>
<td>52</td>
<td>1404</td>
<td>Bjerrum2006</td>
</tr>
<tr>
<td></td>
<td>Crop/Cecum</td>
<td>39</td>
<td>27</td>
<td>850</td>
<td>Gong2007</td>
</tr>
<tr>
<td></td>
<td>Cecum</td>
<td>627</td>
<td>137</td>
<td>301</td>
<td>Torok2011</td>
</tr>
<tr>
<td>Crane</td>
<td>Faecal</td>
<td>16</td>
<td>7</td>
<td>817</td>
<td>Ryu2012</td>
</tr>
<tr>
<td>Hoatzin</td>
<td>Crop</td>
<td>1235</td>
<td>376</td>
<td>1365</td>
<td>Godoy2008</td>
</tr>
<tr>
<td></td>
<td>Crop</td>
<td>2123</td>
<td>267</td>
<td>1338</td>
<td>Godoy2010</td>
</tr>
<tr>
<td>Kakapo</td>
<td>Crop</td>
<td>29</td>
<td>6</td>
<td>728</td>
<td>Waite2012C</td>
</tr>
<tr>
<td></td>
<td>Faecal</td>
<td>73</td>
<td>17</td>
<td>740</td>
<td>Waite2012F</td>
</tr>
<tr>
<td>Shorebirds*</td>
<td>Cloaca</td>
<td>64</td>
<td>34</td>
<td>192</td>
<td>Santos2012</td>
</tr>
<tr>
<td>Gull</td>
<td>Faecal</td>
<td>117</td>
<td>85</td>
<td>780</td>
<td>Lu2008b</td>
</tr>
<tr>
<td>Parrot*</td>
<td>Cloaca</td>
<td>49</td>
<td>39</td>
<td>684</td>
<td>Xenoulis2010</td>
</tr>
<tr>
<td>Ostrich</td>
<td>Cecum</td>
<td>310</td>
<td>98</td>
<td>889</td>
<td>Matsui2010</td>
</tr>
<tr>
<td>Turkey</td>
<td>Cecum</td>
<td>657</td>
<td>139</td>
<td>1450</td>
<td>Scupham2007</td>
</tr>
<tr>
<td></td>
<td>Faecal</td>
<td>688</td>
<td>423</td>
<td>472</td>
<td>Lu2008a</td>
</tr>
<tr>
<td></td>
<td>Cecum</td>
<td>104</td>
<td>67</td>
<td>1454</td>
<td>Scupham2008</td>
</tr>
</tbody>
</table>
Table 6.3: **Data obtained from short-read studies.** Number of reads used, phylotypes generated and average sequence length for clone library data utilised in the study. Asterisk (*) denotes a study that analysed the bacterial communities associated with multiple species of birds, but with common phylogenetic or geographic grouping. Reported regions sequenced are only approximate and do not accurately reflect the start/stop positions of the amplicons. Ecological diversity estimators were calculated by rarefying phylotype table to 1,500 phylotypes/sample prior to calculation and median values are reported. Shannon Evenness is calculated by dividing the Shannon Diversity by the maximum Shannon Diversity value for the depth of sampling. A value of 1 represents complete evenness.

<table>
<thead>
<tr>
<th>Host</th>
<th>Individuals Sampled</th>
<th>Number of Sequences</th>
<th>Region Sequenced</th>
<th>Number of Phylotypes</th>
<th>Shannon Diversity</th>
<th>Shannon Evenness</th>
<th>Simpson Diversity</th>
<th>Figure 6.2 Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkey</td>
<td>38</td>
<td>910,992</td>
<td>V3</td>
<td>60</td>
<td>1.27</td>
<td>0.17</td>
<td>0.48</td>
<td>Danzeisen, 2013</td>
</tr>
<tr>
<td>Duck</td>
<td>1</td>
<td>6,742</td>
<td>V1 - V3</td>
<td>105</td>
<td>1.73</td>
<td>0.24</td>
<td>0.33</td>
<td>Unno, 2010</td>
</tr>
<tr>
<td>Goose</td>
<td>1</td>
<td>7,825</td>
<td>V1 - V3</td>
<td>232</td>
<td>3.4</td>
<td>0.46</td>
<td>0.08</td>
<td>Unno, 2010</td>
</tr>
<tr>
<td>Chicken</td>
<td>1</td>
<td>6,416</td>
<td>V1 - V3</td>
<td>112</td>
<td>2.9</td>
<td>0.4</td>
<td>0.1</td>
<td>Unno, 2010</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>74,678</td>
<td>V1 - V2</td>
<td>20</td>
<td>1.6</td>
<td>0.22</td>
<td>0.3</td>
<td>Stanley, 2013</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>16,990</td>
<td>V2</td>
<td>24</td>
<td>0.56</td>
<td>0.08</td>
<td>0.72</td>
<td>PRJEB1467</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>13,243</td>
<td>V2</td>
<td>31</td>
<td>2.07</td>
<td>0.28</td>
<td>0.17</td>
<td>PRJNIA193217</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>22,384</td>
<td>V3</td>
<td>204</td>
<td>3.37</td>
<td>0.46</td>
<td>0.08</td>
<td>PRJNA169064</td>
</tr>
<tr>
<td>Emu</td>
<td>4</td>
<td>96,549</td>
<td>V2</td>
<td>39</td>
<td>1.44</td>
<td>0.2</td>
<td>0.34</td>
<td>Bennet, 2013</td>
</tr>
<tr>
<td>Kakapo</td>
<td>30</td>
<td>128,021</td>
<td>V3 - V4</td>
<td>28</td>
<td>0.83</td>
<td>0.11</td>
<td>0.56</td>
<td>PRJNA222380</td>
</tr>
<tr>
<td>Little Penguin</td>
<td>4</td>
<td>68,280</td>
<td>V2</td>
<td>53</td>
<td>0.86</td>
<td>0.12</td>
<td>0.56</td>
<td>PRJEB3384</td>
</tr>
<tr>
<td>King Penguin</td>
<td>8</td>
<td>116,937</td>
<td>V2</td>
<td>50</td>
<td>1.98</td>
<td>0.27</td>
<td>0.22</td>
<td>Dewar, 2013</td>
</tr>
<tr>
<td>Misc. Penguins*</td>
<td>3</td>
<td>18,216</td>
<td>V1 - V3</td>
<td>120</td>
<td>2.95</td>
<td>0.4</td>
<td>0.1</td>
<td>Dewar, 2013</td>
</tr>
<tr>
<td>Petrel / Prion</td>
<td>2</td>
<td>17,335</td>
<td>V2</td>
<td>107</td>
<td>2.63</td>
<td>0.36</td>
<td>0.18</td>
<td>PRJEB1549</td>
</tr>
</tbody>
</table>
Figure 6.1: Relative proportion of clone-library OTUs represented in each study. Taxonomic classification for each OTU was derived from a consensus taxonomic classification of each sequence assigned to the OTU. Samples from clone-library data.
Chapter 6: Characterising the avian gut microbiota

Figure 6.2: Relative proportion of SRA OTUs represented in each study. Top labels identify the study from which sequences were downloaded; bottom labels identify the host bird. Top letters denote studies PRJEB3384 (A), PRJEB1467 (B), PRJNA169064 (C) PRJNA193217 (D), Unno2010 (E), Bennet2013 (F) and PRJEB1549 (G). Bottom letters denote host organisms duck (H), goose (I), fairy prion (J) and petrel (K).
We identified three studies that sequenced overlapping regions of the bacterial 16S rRNA gene (Table 6.1, Unno2010, Dewar2013 and PRJEB3384) and observed that two bacterial genera were conserved across all three studies, namely *Bacteroides* and *Clostridium*. Sequences associated with these taxa were extracted from the main dataset and unweighted UniFrac distances were calculated between each biological replicate. The within- and between-study UniFrac distances are reported in Figure 6.3 and, consistent with our prediction, the within-study and between-study difference was similar when the data originated from a closely related host (Figure 6.3, Dewar2013, LittlePenguin and Dewar2013.LittlePenguin). By contrast, the differences between Dewar2013 and Unno2010, and LittlePenguin and Unno2010, were higher than the within-group difference for *Clostridium* and elevated compared to the penguin/penguin comparisons for *Bacteroides*. The within-group differences were higher for Unno2010-*Bacteroides* than for other groups, but this may be a result of the Unno2010 study itself analysing several different birds. Although the different methodologies employed in the various studies are likely to have some impact on the results, we concluded that this was overshadowed by the impact of the host organism and proceeded to analyse other metadata factors.

### 6.4.3 Factors shaping the avian microbiota: biological factors

Standard ecological diversity indices revealed varying degrees of microbial diversity among the birds studied (Table 6.3). In agreement with our previous observations of low microbial diversity within the kakapo hindgut (Waite et al. 2012, 2013), the diversity estimators for kakapo were among the lowest observed. Consistent with previously reported mammalian findings (Ley et al. 2008a), and with more targeted avian studies (Zhu et al. 2002; Lucas and Heeb 2005; Banks et al. 2009; Benskin et al. 2010), the host organism was the strongest driver of community structure in the clone-library data and second strongest in the short-read data (Table 6.4). Other factors were still significantly associated with shaping the gut community but their fit to the data was lower. The fit for any particular factor across the data was quite low (Table 6.4), which is likely a result of compounding variables from the individual studies, rather than a real lack of influence of these factors. In order to account for this variation, CCA was used to visualise patterns in the data that could be accounted for by the factor of interest. Results showed
Table 6.4: Calculated fit of metadata factors to community distances using PERMANOVA with linear model fitting. For both data types the sample collection method was tested but did not show any meaningful correlation with community structure.

<table>
<thead>
<tr>
<th>Clone-library (Unweighted UniFrac)</th>
<th>Fit ($R^2$)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host</td>
<td>0.68</td>
<td>0.001</td>
</tr>
<tr>
<td>Sample Site</td>
<td>0.25</td>
<td>0.001</td>
</tr>
<tr>
<td>Diet</td>
<td>0.17</td>
<td>0.002</td>
</tr>
<tr>
<td>Captivity</td>
<td>0.09</td>
<td>0.004</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Short-read Amplicon (Jaccard Distance)</th>
<th>Fit ($R^2$)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>0.4</td>
<td>0.001</td>
</tr>
<tr>
<td>Host</td>
<td>0.35</td>
<td>0.001</td>
</tr>
<tr>
<td>Sample Site</td>
<td>0.27</td>
<td>0.001</td>
</tr>
<tr>
<td>Diet</td>
<td>0.18</td>
<td>0.001</td>
</tr>
<tr>
<td>Captivity</td>
<td>0.13</td>
<td>0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Short-read Amplicon (Yue-Clayton theta)</th>
<th>Fit ($R^2$)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>0.41</td>
<td>0.001</td>
</tr>
<tr>
<td>Host</td>
<td>0.36</td>
<td>0.001</td>
</tr>
<tr>
<td>Sample Site</td>
<td>0.31</td>
<td>0.001</td>
</tr>
<tr>
<td>Diet</td>
<td>0.21</td>
<td>0.001</td>
</tr>
<tr>
<td>Captivity</td>
<td>0.15</td>
<td>0.001</td>
</tr>
</tbody>
</table>

clear clustering of data for clone samples (Figure 6.4), but weak clustering for short-read data (Figure 6.5). This lack of resolution within the short-read data is likely due to the loss of OTU phylogenetic information due to non-overlapping 16S rRNA gene regions between studies. Due to the lack of phylogenetic relationship between OTUs, each OTU is considered equally different from every other OTU (Lozupone and Knight 2005) and hence evolutionary information is lost.
6.4.3.1 Factors shaping the kakapo microbiota: upper and lower GI

Amplicon pyrosequencing samples obtained from the crop of juvenile birds and used in this study were compared to faecal samples obtained from the same individuals. Crop and faecal samples were extracted from the main data set and analysed using the community structure analysis reported in Chapter 3. Testing of community structure using OTU and phylogenetic distance measures showed that the crop and faecal microbiota of juvenile kakapo differed under all tested conditions ($p < 0.001$, all distances, all OTUs), reaffirming the original findings of Chapter 2, that the microbiota of the upper and lower GI tract differ in juvenile kakapo.
Figure 6.3: **Unweighted UniFrac distances for within- and between-study comparisons.** Distances were calculated by extracting reads classified as *Clostridium* (top) and *Bacteroidetes* (bottom) from each sample. Differences between each pair of samples were categorised as being the distance between samples from the same study or from different studies and plotted accordingly (blue = within study, orange = between study).
Chapter 6. Characterising the avian gut microbiota

Figure 6.4: Constrained Canonical Analysis of community structure based on fitting of metadata factors to the clone-library sequence data. Images represent host (top left), sample site (top right), diet (bottom left) and captivity status (bottom right).
Figure 6.5: Constrained Canonical Analysis of community structure based on fitting of metadata factors to the short-read sequence data. Images represent host (top left), sample site (top right), diet (bottom left) and captivity status (bottom right). Note that the ‘herbivore’ grouping represents exclusively kakapo.
6.4.4 Functional prediction of the gut microbiota

Ultimately, the study of microbial communities is of little biological value unless the functional potential of the community, or individual members, is considered. Statistical testing revealed differences in many predicted functional pathways when data were partitioned by host, but this finding was ignored as it is a likely side-effect of 16S rRNA prediction (i.e., if the 16S rRNA-defined communities differ between hosts, the metagenomic prediction based on 16S rRNA community is also likely to differ). Metagenomes were instead partitioned by diet, captivity and gut location sampled and these categorisations of data revealed interesting differences in functional capability (Table 6.5). Captive birds were predicted to have a microbiota with enhanced capability for carbohydrate metabolism and a lower rate of microbial genes associated with infectious disease. When comparing predicted metagenomes by diet, the microbiota of carnivores was predicted to have a greater capability for amino acid and energy metabolism when compared to herbivores, a finding previously reported in mammals (Muegge et al. 2011). The grain-fed microbiota was predicted to have a higher capability for carbohydrate metabolism than that of herbivores. Genes involved in lactate production were predicted in all samples, which is not surprising as lactate is a known by-product of microbial activity in the ceca and is a major metabolic precursor for glucose in avians (Brady et al. 1979; Ogata et al. 1982; Franson et al. 1985). These findings provide support for the fitting of metadata categories to the samples, as the factors that contribute to shaping the microbiota were also supported by known functional roles of these microorganisms. Partitioning of data by sample site revealed several key influences on the predicted functionality of the microbiota. For example, genes grouping into the KEGG grouping ‘signalling molecules and interaction’ were lowest in faecal samples. This grouping includes an array of genes involved in cell adhesion molecules and cytokine receptors and is likely to be involved in host/bacteria interactions. Genes involved in carbohydrate metabolism were at their lowest in foregut samples from kakapo, and elevated in the hindgut, consistent with the fact that most birds utilise their hindgut/cecum for carbohydrate fermentation (McNab 1973; Mead 1989).

Interestingly, the influence of diet did not match differences in the predicted ability of the microbiota to degrade cellulose. Between the three diet groupings, β-1,4-endoxylanase was more abundant in carnivorous birds than herbivorous birds. β-xylosidase activity was predicted to be higher in grain-fed birds than strictly
herbivorous birds, while xylanase was higher in herbivorous birds than grain-fed (Table 6.5). When taken as a proportion of the total cellulolytic potential, the microbiota of carnivorous birds had a higher predicted occurrence of $\beta$-xylosidase than that of herbivorous birds, and a higher occurrence of Cellulase M than grain-fed birds. Between the non-carnivorous birds, Cellulase M and xylanase accounted for a higher proportion of cellulolytic potential in the herbivorous birds, and $\beta$-glucosidase and $\beta$-xylosidase in grain-fed birds. These genes were detected in a range of bacterial phyla within the avian gut, but particular bacterial families were enriched in the gut microbiota, likely contributing to these differences in relative gene abundance. Of the PICRUSt OTUs that carried cellulolytic potential, members of the Bifidobacteriaceae, Bacteroidaceae and Lactobacillaceae were highly represented in metagenomes which exhibited elevated $\beta$-xylosidase and $\beta$-glucosidase levels. Leuconostocaceae were enriched in predicted metagenomes with elevated Cellulase M and $\beta$-xylosidase. Interestingly, higher abundance of xylanase genes was predominantly associated with abundance of the Enterobacteriaceae, which may reflect the influence of the Proteobacteria-rich kakapo microbiota. When normalised to a proportion of the total cellulolytic gene abundance, predicted proportions of $\beta$-1,4-endoxylanase were not significantly different between dietary groupings. Although not necessarily intuitive, these findings are supported by previous observations that the cellulolytic potential of the avian hindgut is minimal (Barnes 1972; McNab 1973; Mead 1989), and correlates with the observation that cellulolytic pre-digestion of feed boosts energy harvest and weight gain (Józefiak et al. 2006; Yu et al. 2008; Cowieson et al. 2010; Mathlouthi et al. 2011; Ghahri et al. 2012; Ribeiro et al. 2012) in farmed broiler chickens. Caution must be taken in interpreting these predictions, as a recent study has shown that the functional capabilities of the gut microbiota are dependent on community membership as well as genetic potential (Berry et al. 2013). Furthermore, the PICRUSt prediction framework can only account for sequences that can be accurately mapped to the existing database, with no provision for sequences representing novel, or unstudied, bacterial lineages. Nevertheless, the framework provided high-level predictions that were consistent with the known state of avian microbiology and therefore represents an excellent pathway for generation of novel hypotheses and for general annotation of 16S rRNA gene amplicon studies.

In summary, we have conducted a comprehensive meta-analysis of publicly available avian microbiota sequences and tested whether, despite notable differences in physiology between avians and mammals, the factors that drive community
structure are the same. We show that the avian host species is the strongest factor in determining community composition and decoupled this effect from potential study bias where the data allowed. Finally, we have analysed the potential functional profiles of 16S rRNA gene amplicon data and found that the genomic potential predicted of the communities fits well with the existing literature, and is therefore an excellent platform to leverage these data into new hypotheses and lines of inquiry.
**Table 6.5: Summary of key findings in differences between predicted metagenomes.** Comparisons are reported as column Sample 2 compared to the last entry in Sample 1. Gene abundances are reported as a relative proportion of the total predicted metagenomic content.

<table>
<thead>
<tr>
<th>Functional Group</th>
<th>Sample 1 Proportion of genes</th>
<th>Sample 2 Proportion of genes</th>
<th>p-value (corrected)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Carbohydrate Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Captive</td>
<td>11.28%</td>
<td>Wild</td>
<td>10.49%</td>
</tr>
<tr>
<td>Grain-fed</td>
<td>11.51%</td>
<td>Carnivore</td>
<td>10.85%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Herbivore</td>
<td>10.68%</td>
</tr>
<tr>
<td><strong>Infectious Disease</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild</td>
<td>0.50%</td>
<td>Captive</td>
<td>0.43%</td>
</tr>
<tr>
<td><strong>Amino Acid Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carnivore</td>
<td>10.86%</td>
<td>Herbivore</td>
<td>8.52%</td>
</tr>
<tr>
<td>Grain-fed</td>
<td>8.81%</td>
<td></td>
<td>0.026</td>
</tr>
<tr>
<td><strong>Signalling Molecules and Interaction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fecal</td>
<td>0.16%</td>
<td>Crop</td>
<td>0.20%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cecum</td>
<td>0.25%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ileum</td>
<td>0.23%</td>
</tr>
<tr>
<td><strong>β-1,4-endoxylanase</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carnivore</td>
<td>0.02%</td>
<td>Herbivore</td>
<td>0.01%</td>
</tr>
<tr>
<td><strong>β-xylosidase</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grain-fed</td>
<td>0.02%</td>
<td>Herbivore</td>
<td>0.01%</td>
</tr>
<tr>
<td><strong>Xylanase</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herbivore</td>
<td>0.01%</td>
<td>Grain-fed</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
6.5 Acknowledgments

This work was supported by funding from The University of Auckland Faculty Research Development Fund (grant 9841 3626187). D.W.W. was supported by a University of Auckland Doctoral Scholarship.
Chapter 7

General Discussion

Due to a near-complete absence of information regarding the kakapo microbiota prior to undertaking this thesis, the primary goal of the research performed was to document the microbes that inhabit the kakapo gastrointestinal (GI) tract and to attempt to understand the factors that influence them. The recent worldwide focus on the human microbiome (Human Microbiome Project 2014) has led to the development of an array of new techniques, analysis pipelines which have revealed general patterns in microbiota development and succession that can be used as a framework for understanding any previously undescribed microbial community. In this thesis, a variety of molecular methods were used to address questions relating to microbial membership in the kakapo GI tract, and the findings from these analyses were used to direct subsequent cultivation efforts. When identifying the potential factors that could influence the kakapo microbiota, factors of both academic interest (age, location in the GI tract) and those considered relevant to the objectives of the Kakapo Recovery Program (artificial diet, antibiotic treatment) were selected.
Chapter 7. General discussion

7.1 First insights into the kakapo microbiome

The flow between the chapters of this thesis reflects the progressive nature of the overall thesis goals. Chapter 2 documented the diversity and variation within the microbiota along the GI tract, using upper- and lower-GI tract samples to compare the microbiota throughout the gut environment. Care was taken to identify microbial lineages that were common to the majority of individual samples in an attempt to establish a baseline of kakapo-associated microbes for future reference. Chapter 3 extended upon this work by employing the higher sampling throughput offered by amplicon pyrosequencing, seeking to reaffirm the initial observations of low taxonomic diversity and an apparent lack of difference between the bacterial communities obtained from juvenile kakapo and a single, wild adult. I sought to use a higher throughput sequencing platform, as well as more individuals, as both of these tentative findings from Chapter 2 could potentially be explained by a lack of biological replication. Chapter 3 also attempted to reveal human actions that influenced microbial community composition, in particular the change of diet and frequent antibiotic treatment to which kakapo are subjected. While other factors also played roles in shaping the kakapo GI microbiota, the human-associated factors are directly influenced by the Department of Conservation as part of the Kakapo Recovery Program. A final observation that the microbiota varies through time is one that will need to be re-visited in future studies, as it presents the possibility that our baseline data cannot be relied on in years to come.

Chapter 4 represented our first attempt to cultivate kakapo-associated microbes and the experimental parameters of this chapter were driven by the discoveries of Chapters 2 and 3, with predominantly anaerobic conditions and growth media that favoured fermentation reactions chosen. This chapter also attempted to quantify the potential efficacy of the faecal inoculation applied in previous breeding years (Eason and Moorhouse 2006). My findings here showed that freezing did impact the viability of the microbiota, with both a lower diversity and overall cell load being recovered following freezing. As there was no breeding season concurrent with this experiment it was not possible to test this finding in greater detail, but the basic message reported from this chapter was that the storage procedures do indeed influence the viability of the microbiota, and warrant further scrutiny. Chapter 5 then utilised these cultivated isolates, as well as the overall faecal metagenome, to evaluate functional capability. Furthermore, this chapter also sought to address the issue of community composition, in terms of both its taxonomic simplicity
and structure compared to the hindguts of other herbivores. Briefly, it appeared that the lack of 16S rRNA gene diversity in the faecal samples did not translate into a lack of functional diversity, with most of the ‘expected’ bacterial metabolic pathways still being accounted for in the bacterial metagenome. This chapter also continued to interrogate the data for age-related differences in the microbiota, but once again found no community-level differences between the microbiota of adults and juveniles.

Finally, in Chapters 5 and 6 I consider my findings about the kakapo microbiota in a wider context, by comparing with the microbiota of other animals. In terms of the bacterial function related to digestive roles (e.g. carbohydrate fermentation, energy metabolism), the kakapo hindgut microbiota is most similar to that of a hindgut-fermenting herbivore. More broadly, the findings from these chapters compared previous statements regarding low diversity and low taxonomic richness to next-generation sequencing data obtained from a wider variety of sources, analysing these data through a single and methodologically consistent bioinformatics pipeline. The findings of this research are important in several aspects, most notably the added knowledge for future conservation practices and the many ways in which the kakapo microbiota defies what was originally predicted of it.

7.1.1 The kakapo microbiota defies original expectations

Like the kakapo itself, the microbiota of the bird defies many common expectations; indeed a repeated finding of this thesis is essentially a restatement of a common theme in ecology - that there is always an exception to the rule. Prior to undertaking this thesis I hypothesised that the kakapo microbiota would contain a rich community of bacteria, archaea and fungi. Such communities are typical of herbivorous animals (Ley et al. 2008a; Pope et al. 2010; Li et al. 2012; Pope et al. 2012; Kittelmann et al. 2013), including the hoatzin (Godoy-Vitorino et al. 2008, 2010, 2012), and this seemed a reasonable working hypothesis. The kakapo appears to harbour one of the lowest-diversity bacterial communities seen among avians (although certain studies with captive chickens have reported lower, see Chapter 6, Table 6.3). The data obtained throughout this thesis show that the composition of the kakapo microbiota is dominated by Proteobacteria, which is in itself an unusual finding as vertebrate guts are typically dominated by members of the Bacteroidetes and Firmicutes. The kakapo microbiota appears to be absent
of archaea, although intense metagenomic sequencing detected traces of archaea at very low abundances. Fungi were also only detected sporadically in metagenomic shotgun sequencing, although this aspect of the kakapo microbiota was not addressed in detail.

An additional way in which the kakapo microbiota defies expectation is in the fact that the microbiota of the adult and juveniles does not differ in any meaningful way. A major finding to emerge from the recent interest in the human microbiome (Human Microbiome Project 2014) was the degree to which the human microbiota changes following birth (Palmer et al. 2007; Dominguez-Bello et al. 2010; Yatsunenko et al. 2012). These changes are so distinctive that they overpower significant inter-study biases in human data (Lozupone et al. 2013), and a strong influence of age has been shown in a range of other host animals (detailed in Chapter 3). While a few differences were identified between adult and juvenile kakapo with respect to particular bacterial taxa (Chapter 3, Table 4), age is usually a major factor in structuring the gut microbiota of vertebrates and community composition has been shown to differ drastically between young and mature individuals in a number of host species (detailed in Chapter 3). The potential for differences in the microbiota between adults and juveniles was addressed three times in this thesis (Chapters 2, 3 and 5) but in each case no differences were observed between age groups, with an exception being adult birds that had undergone extensive handling prior to sampling (Chapter 2). While it is possible that there are differences in the microbiota at very early stages of life, the nature of sampling from a wild animal generally precludes such a study. While differences may exist at the very early stages of life, the fact that convergence between the adults and juveniles occurs as quickly as these differences is in itself a novel finding.

7.2 Challenges of researching a critically endangered species

A major consideration during this research was the need to prioritise the well-being and health of kakapo over the study design for each data chapter. This is not unreasonable, as research regarding the kakapo microbiota should clearly not impede the preservation of the bird. The impact of this major consideration
on study design is noticeable throughout the thesis and is particularly evident in Chapter 3. The need to hand-rear all juvenile kakapo during the 2011 breeding season left no control group available for a proper evaluation of the impact of the dietary changes and antibiotic supplementation that occur during captivity. Meaningful data were still generated within the scope of this chapter, identifying a strong influence of antibiotics on the abundance of a number of OTUs. Importantly for kakapo conservation, this trend appeared to be short-lived, with all signs of antibiotic treatment disappearing by one month after return to the wild. This is a key finding, as the potential for long-term effects of antibiotic treatment is an aspect of microbiology currently under intense scrutiny within the field of human medicine. Showing that these effects are only short-lived in the kakapo is reassuring in that there should not be complications arising (with respect to the microbiota) from the necessary handling of juvenile kakapo during the early stages of life.

Another major challenge, albeit not unique to working with kakapo, is that of differentiating autochthonous microorganisms from allochthonous or transient microbes. The high sensitivity of PCR-based marker gene surveys means that there is a good chance of detecting microbes that may simply be passing through the gut. Microbes present in food, drink, or even in the atmosphere can easily make their way into the gut, where they will leave DNA that is detectable through molecular methods. The potential for contamination of the gut with non-resident microbes is a concern that is often raised within microbiology but can largely be addressed through repetitive sampling of many individuals in an attempt to observe patterns of occurrence. A specific example from the kakapo microbiota is the presence of the bacteria *Escherichia fergusonii* and *Streptococcus gallolyticus* within the kakapo faecal samples. The consistent presence of these microbes in healthy individuals, combined with their high abundance within the microbial community, is a strong indicator that their presence in the gut is not random, but rather a sign that they occupy this environment permanently.

### 7.3 Evaluation of methods applied

The field of molecular biology, including microbiology, has recently undergone a massive (and ongoing) paradigm shift in the way that DNA sequence data are generated and analysed. Even at the time of submitting the first-year literature
review for this thesis (submitted in 2012), the provisional goals detailed the use of DNA fingerprinting techniques such as denaturing gradient gel electrophoresis (Muyzer et al. 1993; Muyzer and Smalla 1998) for addressing the research questions raised in Chapter 3, while Chapter 5 was not even considered feasible at that time. As much as possible, I have attempted to keep up with the pace of the field and many aspects of the final thesis have been re-worked to better reflect the contemporary approaches of microbiology, in some cases discarding the methods described in the original thesis proposal. Despite this, there are always areas that could have been improved by additional analyses that were considered impractical, or not considered at all, at the time the work was done. Notwithstanding the obvious improvements that can be made to any study (e.g. more replicates, longer timespans) and the previously mentioned issues regarding kakapo handling, there are several aspects of this thesis that could have been performed differently. After the initial lack of detection of \textit{Bacteroidetes} (Chapter 2), more thorough testing could have been performed, using a more sensitive technique such as catalysed reporter deposition fluorescence in situ hybridisation (CARD-FISH) (Lebaron et al. 1997; Schönhuber et al. 1997) or quantitative PCR in an attempt to identify \textit{Bacteroidetes} in the kakapo gut. Although I was careful to mention that the spike-and-recovery experiments did not preclude the possibility of \textit{Bacteroidetes} being present in the kakapo faeces (Chapter 2), a more sensitive approach would likely have revealed \textit{Bacteroidetes} in the kakapo samples, as Chapters 3 and 5 detected rare \textit{Bacteroidetes} in over half of all kakapo surveyed.

Another aspect of this study that could have been undertaken differently was the cultivation performed in Chapter 4. Although a range of cultivation conditions was used in order to isolate non-\textit{Escherichia} species, care should have been taken to employ cultivation strategies that either inhibited \textit{Escherichia} growth or created an obvious differentiation between bacterial colonies. Despite \textit{Escherichia coli} accounting for less than 1% of the human gut microbiota (Suau et al. 1999; Eckburg et al. 2005; Mariat et al. 2009) its rapid growth in many cultivation media have led to it being the model organism for much of microbiology (Lederberg and Tatum 1946; Benzer 1961; Blattner et al. 1997). In a system such as the kakapo, where over 90% of the microbiota may well be the close relative \textit{E. fergusonii}, its dominance in cultivation conditions should perhaps have been foreseen. There exist a range of cultivation media that allow visual discrimination between bacterial colonies, for example MacConkey, eosin-methylene blue and hektoen agar. Utilisation of such media would remove the randomness from isolate selection and
would likely have led to a greater range of novel isolates being selected from the growth plates.

### 7.4 Future directions and conservation implications

The major question that is left unanswered at the end of this thesis relates to the identity and community structure of the kakapo crop microbiota. While the resident microbiota of the crop was investigated in juveniles in Chapters 2 and 6, whether or not this is representative of the adult is unclear. Given the lack of differentiation between the faecal microbiota of the adult and juvenile birds, it could be inferred that the juvenile crop microbiota reflects that of the adult but this is only speculation. Crop samples from adult kakapo are difficult to come by, generally only available when a bird is seriously sick or has died. In both cases the samples obtained are likely not representative of a healthy individual and are therefore only of limited use in addressing fundamental questions surrounding the microbial ecology of the kakapo.

From a conservation biology perspective, the dominance of *Escherichia* in the kakapo GI tract is both intriguing and carries with it potential health risks. Given the lack of starch in the ‘natural’ kakapo diet, *Escherichia fergussoni*-like bacteria would not be expected to comprise a major part of the kakapo microbiome (Touchon et al. 2009). Despite this, this bacterial lineage dominates the microbiota of most individuals. Although mere speculation at this stage, the dominance of *Escherichia* in the kakapo gut is likely a symptom of the starch-rich supplemental feeding provided regularly to kakapo during breeding seasons. We noted a decline in *Escherichia* abundance after the breeding season had ended (Chapter 3), which may have been related to the reduced availability of supplemental feed to kakapo at this time. While the dominance of *Escherichia* in the kakapo GI tract does not appear to harm the birds, it does pose risks in that it drives an uneven bacterial community composition. The ecological relevance of evenness in bacterial communities is that a high community evenness provides resilience to stress and perturbation (Wittebolle et al. 2009). In the GI context this primarily translates into resilience against the establishment of disease, which is of great value to the Kakapo Recovery Program. As outlined in Chapter 1, the pathogens of kakapo
are not well understood and blanket antibiotic treatment remains the preferred approach for both prevention and treatment of sickness. Indiscriminate usage of antibiotics is a global issue and one that will not remain viable forever. Encouraging the development of a diverse bacterial community within the kakapo gut may help to reduce the occurrence of disease, while remaining neutral for direct bird wellbeing.

An additional question raised by this research is the nature of the faecal *Bacteroidetes*. Based on the co-occurrence pattern of *Actinobacteria* and *Bacteroidetes* I speculated that there may be an important ecological relationship between these two phyla, with members of the *Actinobacteria* potentially creating a niche in which *Bacteroidetes* can establish in the kakapo gut, or alternatively the *Actinobacteria* harbouring *Bacteroidetes*-like functional genes. Molecular methods such as FISH can be used to identify spatial relationships in microbial communities, but the sampling regime utilised in this research was not amenable to such a research question. While faecal samples are generally accepted as a proxy for the hindgut microbiota membership, they do not adequately reflect the spatial structuring that occurs in the gut prior to the shedding of microbial cells into faeces. As such, while it was technically possible to address this question through the course of this thesis, this aspect of kakapo microbiology was left untouched due to the invasive sampling procedures required in order to address a mostly academic question. This aspect is still of interest to the study of kakapo and, if a situation in which it can be addressed arises, it will be an excellent avenue for future study.

Another aspect of kakapo microbiology that remains to be studied is the role of both viruses and eukaryotic microorganisms such as fungi and protozoans. With decreases in the cost of sequencing, the use of metagenomics as a tool for viral ecology has become commonplace (Delwart 2007; Kristensen et al. 2010; Mokili et al. 2012) and is likely to be an informative and practical aspect of future kakapo microbiology. Some initial work was performed during this thesis to uncover the identity of common kakapo-associated eukaryotes but was not pursued for reasons of time and lack of a biologically meaningful question. Also, the original planning of this thesis sought to investigate the role of trichomonads in a recently documented disease (exudative cloacitis) of kakapo. However, additional analysis by the New Zealand Centre for Conservation Medicine reported conflicting results regarding the putatively pathogenic nature of this microbe. Due to the changing state of knowledge with regards to this eukaryote’s role in the disease it was
not investigated as part of this thesis. The causative agent of this disease is now thought to be a virus (White et al. 2014), although only preliminary data are available. Although a small number of eukaryotic genes were annotated in the kakapo metagenome (Chapter 5), the separation technique employed as part of the DNA extraction protocol appears to have purified the bacterial and archaeal cells while to a large extent excluding the larger kakapo and fungal cells. While this was desirable for the nature of this thesis, it leaves another branch of kakapo microbiology unstudied. Knowledge about the potential roles of eukaryotic microorganisms in both the ‘normal’ microbiota and in disease ecology would be an excellent complement to the data generated in this thesis.

The final area for future study of the kakapo microbiota revolves around the practicality of developing a kakapo probiotic for sick and captive birds. This is an aspect of study that must lean heavily on previous research, both for identifying the ‘core’ gut microbes that should comprise such a solution and to actually obtain them via cultivation. As such, this was not a viable goal within the time frame of this thesis, but this research has made steps towards this goal. The cultivation and analysis of *Escherichia* and *Streptococcus* strains (Chapters 4 and 5) has yielded several bacterial isolates that can be used in such an endeavour, although they are unlikely to constitute the full microbial complement, with *Lactobacilli* and members of the *Actinobacteria* frequently being detected in healthy birds (Chapters 2, 3, 4 and 5). In addition, the work presented in Chapter 5 indicates that although the kakapo harbours a single *Escherichia* (16S rRNA gene-defined) OTU, this grouping contains a range of functionally diverse bacterial strains, which will need to be independently evaluated for use in such a probiotic. Nevertheless, this thesis has conducted the initial groundwork necessary for such an undertaking and the development of a probiotic would be greatly advantageous for the Kakapo Recovery Program, potentially as a substitute for antibiotic treatment during captivity.

### 7.5 Concluding remarks

In summary, the research described within this thesis represents the first comprehensive study of the kakapo microbiota, and sheds light on at least some of the interactions between the host bird and resident bacteria. Important ecological questions surrounding bacterial community diversity and membership throughout the GI tract, as well as differences between young and old birds, have been
addressed. Practical knowledge regarding the impact of captivity and handling of birds has also been addressed and provided useful information for the future conservation of this parrot. For example, feeding frozen faeces is probably not beneficial to the development of the microbiota in juveniles, although it is unlikely to be harmful either and a probiotic substitute is not practical at this stage. In addition, although the utilisation of antibiotics on captive juveniles has a marked impact on the microbiota this appears to be transient and the microbial community reverts to a ‘normal’ state within weeks of release. At a more general level, a major take-home message from this work should be that there are no rules in gut microbiology, only guidelines. Taxonomic diversity of the gut microbiota is generally attributed to an herbivorous diet and age is considered a major factor in shaping the vertebrate microbiota. Despite repeated testing, the kakapo does not follow these patterns and shows that such patterns are not universal in microbiology.


Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, A.,
Buchner, A., Lai, T., Steppi, S., Jacob, G., Förster, W., Brettske, I., Gerber,
S., Ginhart, A., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss,
T., Lüßbmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis,
A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., and Schleifer,

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q.,
Y., Han, C., Cheung, D., Yiu, S., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li,
pirically improved memory-efficient short-read de novo assembler. *Gigascience*,

Lührs, H., Gerke, T., J.G., M., Melcher, R., Schauer, J., Boxberge, F., Schepp-
pach, W., and Menzel, T. (2002). Butyrate inhibits NF-κB activation in lamina
propria macrophages of patients with ulcerative colitis. *Scandinavian Journal


Macfarlane, G. and Macfarlane, S. (2012). Bacteria, colonic fermentation, and

Mariat, D., Firmesse, O., Levenez, F., Guimarães, V., Sokol, H., Doré, J., Corthier,

Markowitz, V., Chen, I., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Rat-
ter, A., Jacob, B., Huang, J., Williams, P., Huntemann, M., Anderson, I.,
microbial genomes database and comparative analysis system. *Nucleic Acids

antibiotics on in vitro production of volatile fatty acids and methane in caecal

catabolism by the human gut microbiota: The Bacteroidetes Sus-like paradigm.

Martin-Laurent, F., Philippot, L., Hallet, S., Chaussod, R., Germon, J., Soulas,
microbial diversity analysis methods. *Applied and Environmental Microbiology*,


