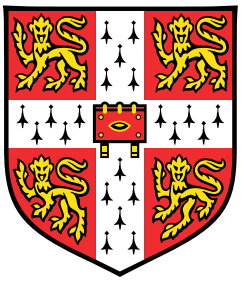

Potential of faeces for non-invasive assessment of Juan Fernandez fur seals and their environment

Constanza Valentina Toro Valdivieso

Supervisor: Dr. Barbara Blacklaws

Advisor: Dr. Cinzia Cantacessi

Department of Veterinary Medicine
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Newnham College

December 2021

In special memory of my friend Daniel Esteban Torres Castillo DVM, and all those who lost their battle against COVID 19.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text It does not exceed the prescribed word limit (60.000 words) for the relevant Degree Committee.

..... Constanza Valentina Toro Valdivieso September 2021

Constanza Valentina Toro Valdivieso
December 2021

Acknowledgements

"The oceans deserve our respect and care, but you have to know something before you can care about it." - Silvia Earle

This project has had a profound impact on who I am as a researcher and as a human. I have had the privilege to work side by side with the most amazing people one could find. This project would have never even started without the trust and support of my supervisor Dr Barbara Blacklaws; I will forever be grateful for being given a chance that I never even dreamed of. Thank you so much for your advice, patience, constructive criticism, transparency, encouragement, and for allowing me to conduct myself with intellectual freedom. Considering we did not have much funding, I think we did pretty well?! Regarding the lack of funding, I am deeply grateful for my amazing collaborators, without whom none of this analysis would have been possible. Many thanks to Dr Eduardo Castro-Nallar, Dr Frederick Toro, Prof. Jonathan Powell, Dr Ravin Jugdaohsingh, Prof. Mark Holmes, Dr Iain Kean, Dr Lajos Kalmar and last but not least, my very good friend Dr Sam Stubbs. I would also like to thank Dr Jaume Fourcada, Capt. Charles Moore, Prof. Joseph Hoffman, and the Juan Fernandez locals Pablo Rivadeneria and Monserrat Gonzalez for kindly providing prey and non-JFFS bone samples.

I would also like to show my deepest gratitude to the Department of Veterinary Medicine and Newnham College for their kind contributions to my fieldwork and The National Research and Development Agency of Chile (ANID) and Cambridge Trust for covering my scholarship. Additionally, I would like to thank the more than 60 people that donated to my research via the Crowdfunding platform. These donations marked the starting point of this research. Thank you so much for your trust.

My fieldwork would have been totally unachievable without the logistic support from Aerocardal, who sponsored my flights to the Juan Fernández Archipelago, the National Fisheries and Aquaculture Service National (SERNAPESCA), Fundacion Endemica and OIKONOS, especially Héctor Gutierrez and Pablo Marríquez Angulo, for always giving a hand when needed.

Finally, I would like to give a special thank you to the Chilean National Forestry Commission (CONAF), in particular to Guillermo Araya (Director of the Juan Fernandez National Park) and the rangers Ángela García, Ramón Schiller and Danilo Arredondo for the crucial support, teaching and friendship provided during fieldwork.

Additionally, I want to show my appreciation to Prof. Daniel Torres Navarro, one of the first scientists studying these amazing pinnipeds. His passion for them remains till these days,

decades after his last visit to the archipelago. Muchas gracias por su vocación, entusiasmo, apoyo y dedicada trayectoria.

This acknowledgement would not be complete without thanking my family and friends for their support and encouragement throughout these four years, especially to my partner Harry.

This thesis is dedicated to my family, the Juan Fernandez local community, and especially my partner Harry.

Abstract

Marine environments are at risk. Pollution, climate change, disruption of the food network and pathogen dissemination are a few examples of problems that are currently affecting the health of the oceans. Identification and investigation of useful environmental sentinel species such as marine mammals can contribute to better understand the deterioration of ocean health. However, to effectively use wild populations as sentinels, it is first necessary to establish a baseline.

The Juan Fernandez fur seal (JFFS) is a marine mammal endemic to the Juan Fernandez Archipelago in the Pacific Ocean. The archipelago is a UNESCO Biosphere Reserve and has been identified as one of the eleven irreplaceable priority sites for marine conservation worldwide. As a result of overhunting between the 17, 18 and 19th centuries, the JFFS was severely reduced. Furthermore, it was presumed extinct by the end of the 19th century. Since its "rediscovery" in the early 60s up to current times, the JFFS population have evidenced a steady recovery. Today, it is an icon for local tourism and a great example of population recovery.

Since 1995 the Chilean government has decreed a 30-year hunting ban which enabled such an impressive recovery. However, the ban on hunting will only last until 2025, and it is unclear what conservation measures will be put in place after this date. Besides intermittent basic censuses, no further monitoring has been done on this species in the last two decades, to the best of my knowledge. Complicated logistics, inaccessibility, lack of funding, and human resources may partially explain the lack of research. However, in the context of the hunting ban coming to an end in less than five years, there is an urge to build as much knowledge as possible in a relatively short period to inform policymakers when deciding on the future protection measures for this species.

To establish a baseline for the study and monitoring of the JFFS, I explored the potential of faecal samples as a non-invasive method to obtain diverse information while lowering sampling cost and logistic complexities. Here, I focused on only three of the multiple topics that can be studied from faecal samples. First, I used 16S rRNA amplicon sequencing to characterise the faecal microbiome (Chapter 3). This first faecal microbiome characterisation evidenced a clear separation of the samples into two clusters. Due to the little information available on the species, it was not possible to provide a clear explanation of the pattern observed here. However, diet and sex (associated with prey selection and therefore to diet) could be considered possible explanations. On the other hand, the phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt) inferred pathways associated with pathogenesis were enriched in cluster 2, which contained only 22 % of the samples. This first insight is an

important contribution to understanding the natural microbial diversity in free-free ranging pinnipeds.

Next, I used inductively coupled plasma mass spectrometry to evaluate heavy metal exposure (Chapter 4). The results evidenced high levels of Cd and Hg in the JFFS when compared to the Antarctic fur seal (AFS), suggesting high exposure. Diet is the most likely source of contamination. Motivated by these results, I analysed Cd in bone samples, evidencing Cd absorption. These samples, however, did not evidence any of the changes usually associated with Cd intoxication in bone, suggesting some degree of adaptation to high levels of this toxic heavy metal. Furthermore, Si levels, an ultra-trace element related to bone health, could be an interesting target for future studies on Cd tolerance in JFFS. Human studies on this topic may also benefit.

Finally, in Chapter 5, I focused on optimising a method for collection, storage and, host DNA extraction and amplification of faecal samples. For this optimisation, I targeted the mtDNA control region, five different microsatellite loci and two loci commonly targeted for molecular sex identification. Swabbing the faecal surface was usually associated with less specific PCR products. However, by performing nested PCRs, the specificity of the amplification dramatically improved in samples with poor DNA. On the other hand, when using more sensitive assays such as real-time PCR, which was used for molecular sex identification, using a nested PCR approach should only be considered when direct amplification fails to avoid sample contamination. This study showed working with faecal samples for investigating population genetics requires a lot of optimisation. However, once methods become optimised, the difficulty of processing these samples reduce while the probability of success increases. Even though not yet complete, this study is a significant contribution that will enable more rigorous monitoring of the JFFS.

The information generated from my research is an essential contribution to the knowledge about this species which is urgently needed to inform policymakers for future conservation policies. Here, I have shown that working with faecal samples can be an accessible alternative to studying various aspects of a species.

Table of contents

List of figures

List of tables

Chapter 1

Literature Review

1.1 Introduction

Marine environments provide vital services without which life on earth would not be possible. Furthermore, humans also benefit from marine environments for food, economic gain and even drug discovery (Von Schuckmann *et al.*, 2020). These are complex and highly interconnected systems subject to various environmental impacts (Borja, 2014; Von Schuckmann *et al.*, 2020) Pollution, climate change, disruption of the food network, and pathogen dissemination are a few examples of the problems currently affecting ocean integrity and function (Halpern *et al.*, 2019b). Remote insular environments are particularly sensitive to changes in the environment (Veron *et al.*, 2019). For clarity, the term "insular environment" will be used to refer to those ecosystems associated with oceanic islands. Despite their isolated location, these environments can be severely affected by ocean pollution dispersed through ocean currents or via the trophic network. The South Pacific Gyre (a gyre is a large system of rotating ocean currents), for instance, is known to be an important concentrator and distributor of plastic debris and associated toxic chemicals (Markic *et al.*, 2018).

Due to the ocean interconnectivity, assessing local environmental risk and impacts can be challenging. Thus, integrated approaches at the macro-and micro-ecological levels are needed to properly understand and manage ecological threats in these kinds of complex systems. Identifying biological targets capable of providing a large diversity of information in combination with large scale collaborative work is vital. In this context, pinnipeds and other marine mammals are known to be prominent sentinels of marine ecosystems (Bossart, 2011a; Moore, 2008).

However, to effectively use wild populations as sentinels, it is first necessary to establish a baseline dataset against which to compare for change in future studies. My research targeted a poorly studied pinniped, the Juan Fernandez fur seal (*Arctocephalus philippii philippii*, JFFS), endemic to an isolated archipelago in the middle of the South Pacific. This species is the only native marine mammal in this location and, thus, a potential bioindicator. Additionally, the species legal protective status is due to be reviewed in less than five years. I used faecal samples as a non-invasive and accessible method to study different aspects of these animals and their

environment. This work is the first study on this species in the last two decades, to the best of my knowledge.

1.2 The Juan Fernandez archipelago

The Juan Fernandez Archipelago (JFA) is a group of islands located in the south pacific about 700 km off the Chilean continental coast (coordinates: 33°38'29" S 78°50'28" W). The archipelago consist of three islands: Robinson Crusoe (RC), Santa Clara (SC) and Alejandro Selkirk (AS). From the three islands, RC is the only one permanently populated. Juan Bautista is the only town in RC and has a population of approximately 1000 people. The community is highly dependant on local marine products such a fish, octopus and lobsters. This last one is the most important economic resource (Porobic *et al.*, 2019).

AS, SC and approximately 90 % of RC are included in the "Archipielago de Juan Fernandez" national park. The archipelago was designated a UNESCO Biosphere Reserve in 1977 and is one of the ten priority sites for biodiversity conservation of the Valparaiso region. Additionally, the archipelago is considered one of the 11 irreplaceable priority sites for marine conservation worldwide (Pompa *et al.*, 2011). Furthermore, a survey on marine biodiversity published in 2016 described the JFA as a hotspot for endemism (Friedlander *et al.*, 2016) and, in 2017, was declared a Marine Protected Area (MPA) called "Mar de Juan Fernadez". The MPA includes the coastal marine parks "Loberia de Selkirk", "El Arenal" (EA), "El Palillo", and "Tierras Blanca" (TB). Thus, the study and monitoring of the JFFS may provide a valuable contribution to the monitoring and management of this protected marine ecosystem. Furthermore, two of the sampling locations included in this study were within the parks EA and TB.

1.3 Marine mammals: Fur seals

Marine mammals are mammals that feed entirely or predominantly from marine environments. There are five surviving lineages divided into three orders: Cetartiodactyla includes cetaceans (whales, dolphins, and porpoises); Sirenia includes manatees and dugongs; and Carnivora includes pinnipeds (walruses, sea lions, and seals), sea otters, and polar bears *Ursus maritimus*. Sea otters and the polar bear are the most recent members of the marine mammal group and remain closer to terrestrial mammals than marine ones (Uhen, 2007; Yuan *et al.*, 2021). Cetacean, sirenians and pinnipeds are the oldest surviving lineages. The first two are fully aquatic species, while pinnipeds still need to return to land for reproductive purposes. However, species from this group spend most of their time in the water. The amphibious lifestyle makes them a good link between oceanic and coastal environments. When on land, changes in the

population dynamics, at least during the reproductive season, are more visible and, thus, easier to monitor.

The family Otariidae, also known as eared seals, is a monophiletic clade represented by 16 species (Berta and Churchill, 2012; Berta *et al.*, 2018). However, the species number varies among proposed taxonomic classifications over which there is still no consensus. Pinnipeds originated in the northern hemisphere and later migrated to the south where the largest otariid species diversity can be found in current times. Some of the most visible difference between otariids and other pinnipeds are the presence of a pinnae (external ear flap), from which their name derived (the Greek word "otarion" means little ears), and their quadrupedal locomotion when on land, which is facilitated by their long front flippers and the ability to rotate their hind limbs forward (this last characteristic is also shared with walruses). These species mostly inhabit cold temperate waters. All otariids are polygynous and demonstrate a strong sexual dimorphism.

Even though the relationships within the otariid clade remain under debate, two subfamilies are largely recognised: sea lions (Otariinae) and fur seals (Arctocephalinae). The main differences between these two are the fur and the body size. Fur seals are characterised by small body sizes and a thick fur composed by outer guard hair and a dense layer of short, fine and waterproof underfur. Sea lions on the other hand are larger in size and only have one fur layer, relying more on their blubber for thermoregulation.

The fur seal clade is grouped in two genera: *Callorhinus*, which only includes one species, (*Callorhinus ursinus*) and *Arctocephalus*. The genus *Arctocephalus* has been a matter of debate in terms of taxonomic classification. As a result, various taxonomic groupings have been proposed. Initial morphological comparisons suggested the JFFS and the Guadalupe fur seal (GFS) were separate species. In 1954, Sivertsen (1954) suggested these two species should be reclassified into a different genus named *Arctophoca*. Three years later, Scheffer (1958) suggested these species were, instead, subspecies of *Arctocephalus philippii*. Since then, scientists have debated whether these should indeed be considered separate species. Wynen *et al.* (2001) performed the first and only molecular analysis, including all pinniped species. He targeted the mitochondrial DNA control region. This phylogenetic analysis inferred a close relationship with an extremely low interspecific divergence between JFFS and GFS. However, the study only included two unique sequences from GFS and five from JFFS.

The close relationship between these two fur seals is fascinating. First, the geographical separation between the two is more than seven thousand kilometres. For context, the GFS is the only remaining fur seal in the northern hemisphere. Additionally, the Galapagos fur seal's habitat is located between these two. Wynen *et al.* (2001), for instance, very loosely speculated that avoiding the strong Humboldt current that flows north along the western coast of South

America may have facilitated the migration towards the south. Unfortunately, this relationship between JFFS and GFS and its significance on the colonisation of the Southern Hemisphere remain unstudied. This may be in part due to poor research development around the JFFS. Thus, the study of the JFFS may contribute to a better understanding of the historical migration patterns that led fur seals to colonise the south.

1.4 The Juan Fernandez fur seal

As mentioned earlier, and like other marine mammals, the JFFS was severely hunted. Furthermore, it has been hypothesised that the pre-sealing population was as numerous as 4,000,000 individuals, and at least 3,870,170 fur seals were estimated to have been slaughtered (Hubbs and Norris, 1971). After being thought extinct for over 100 years, a group of 200 individuals were sighted by Bahamonde (1966) in AS. Later censuses consistently reported increasing numbers (Torres, 1987; Aguayo and Maturana, 1970; Aguayo *et al.*, 1970; Osman and Moreno, 2017). Moreover, 84,827 individuals (pups and adults) were counted during a census carried out in 2018 in RC and SC. An official communication sharing the summarised results can be found on CONAF's official website (www.conaf.cl).

However, it is fair to highlight that censuses have often been impaired by complicated logistics, inaccessibility to some of the colonies, and lack of human and financial resources. As a result, these animal counts were often incomplete (e.g. did not include all colonies) or inconsistent (e.g. annual censuses carried out at different stages of the reproductive season). Nevertheless, the population recovery has been evident to the local community. Furthermore, some local fishermen have reported conflicts with the JFFS, such as stealing baits, and have expressed concerns about what they perceive as a fur seal overpopulation. A hunting ban decreed in 1995 is perhaps the only and most important conservation tool to protect the JFFS. However, this ban is not indefinite and the future of the JFFS conservation needs to be urgently discussed as this legal tool is only valid until November 2025. It is expected there will be some level of support towards reopening the hunting of the JFFS as a method to limit the population expansion. It is likely that supporters will rely on the population growth evidenced by the rather limited censuses and their personal experiences of conflict. Even though, these arguments may be considered legitimate, the species remains to a large extent poorly understood.

1.4.1 What do we know so far

Very few studies focused on topics other than population abundance. Almost all of them were carried out between the 80s and 90s. Furthermore, the JFFS was the only fur seal species not

included in the recently published book on Latin American pinnipeds, reflecting the lack of research interest from the scientific community (Heckel and Schramm, 2021). In this section, I have listed all the studies I could find and their main results. Three of theses studies focused on parasite infections. *Uncinaria hamiltoni*, *Ogmogster heptalineatus*, *Phocanema decipiens*, *Anisakis* sp., and *Diphyllobothrium* sp. have been reported in the JFFS (Sepúlveda and Alcaíno, 1993; Sepúlveda, 1998; Cattan *et al.*, 1980).

Boness and Francis (1991) looked at how thermoregulatory requirements affected social behaviour in reproductive JFFS males and females. Their results showed that rapid increases in solar radiation resulted in females moving from their pupping/resting sites into the water to cool down. On the other hand, males would remain inland, protecting their territories and only cooling down in the water for a maximum of 45 min when female density was very low. However, the authors evidenced that some males would take advantage of the females' daily trips to the water. Instead of land territories, these males would hold aquatic territories to achieve as many copulation as the other males.

In 1995, Ochoa Acuna and Francis (1995) published a study on the JFFS diet with a focus on spring and summer prey. The diet was mostly small fish (80 % were myctophids) and cephalopods (mainly from the family Onychoteuthidae). This study also showed diet differed among sexes and ages, and variation was also observed between years. The authors concluded that the small range of prey is more likely to result from prey availability than specialisation.

Another study that looked at the foraging behaviour and maternal attendance of JFFS females (Francis *et al.*, 1998) showed that JFFS females undertake some of the most extended foraging trips among otariids (500 kilometres). The authors suggested that such long trips were explained by prey distribution. Further evidence showed that fat content in milk was among the highest recorded of otariids which is likely to be associated with the long intersuckling intervals due to the long foraging trips. Finally, two independent studies looking at postnatal population rate concluded there was no significant difference in growth rates between males and females pups (Osman *et al.*, 2010; Ochoa-Acuña *et al.*, 1998). However, overall differences between growth rate were observed between years (Ochoa-Acuña *et al.*, 1998). Additionally, males are, on average, heavier than females at birth (Osman *et al.*, 2010).

Additionally, Sepúlveda *et al.* (1997) published a study on heavy metal concentrations found in kidneys and liver tissues of JFFS found dead at AS. Briefly, the study showed high levels of Hg and Cd in JFFS pups. Hg concentration found in newborn and older pups suggested transplacental and milk transfer, where milk was likely to be the main contamination route in pups. The authors speculated on possible sources of contamination. Feeding behaviour, for instance, was the suggested source for Cd. At the same time, the high concentrations of Hg could be related to geothermal activities characteristic of areas with high tectonic and volcanic

activity, such as the Juan Fernandez ridge, which is where the JFA is located. In my study, I revisited the JFFS exposure to heavy metals. See Chapter 4 for more details.

Goldsworthy *et al.* (2000), published the first and only study on the topic of populations genetics. Here, the authors looked at a hypervariable segment of the mtDNA control region to explore the genetic consequence of the strong population decline that resulted in their presumed extinction. Contrary to their expectations, they found that the genetic diversity did not reflect such a severe population decline as originally thought. Nevertheless, there are reasons why more genetic studies are urgently needed. During my fieldwork, for instance, I saw several leucistic individuals (partial loss of pigmentation), which may suggest inbreeding (Figure 1.1). I also observed the presence of individuals of two vagrant fur seal species: the subAntarctic fur seal (*Arctocephalus tropicalis*, SAFS) and the Antarctic fur seal (*Arctocephalus gazella*, AFS) (Figure 1.2). Over the three fieldwork seasons I conducted, I saw three SAFS, a male and two females, and several AFS juvenile females. AFS, in particular, were observed in the middle of various reproductive colonies in RC but none in SC. There is only one previous report on the presence of vagrant fur seal species published by Torres *et al.* (1984). Thus, hybridisation events may be taking place in the JFA. Another important reason to study and monitor the JFFS genetics is to identify and measure possible selection pressures due to the rapid environmental changes affecting marine environments. A more extensive discussion can be found in Chapter 5.

Between 2003 and 2005, Osman and Moreno (2017) looked at population trends and distribution. The study showed at least 60 % of the annual production was contributed by AS, 25 % by RC and only 12 % by SC. Furthermore, the authors evidenced that SC and RC had reached an equilibrium while AS was still expanding. The study hypothesed the limited availability of adequate grounds for reproduction may explain the stagnation observed in SC and RC.

Finally, a more recent publication that looked at microplastics in faecal samples collected from various South American otariid species showed that the JFFS was the species most exposed to this contaminant (Perez-Venegas *et al.*, 2018). The authors argued that diet and their higher exposure to the plastic debris concentrated in the South Pacific gyre might explain these results.

1.4.2 Threats

There are several good reviews on global risks to pinnipeds and other marine mammals (Bester, 2014; Kovacs *et al.*, 2012; Simmonds and Isaac, 2007), However, in this section, I will mostly focus on what I consider the most relevant threats to the survival of the JFFS. Similar to other pinnipeds, climate change is perhaps one of the biggest threats to the JFFS (Bester,

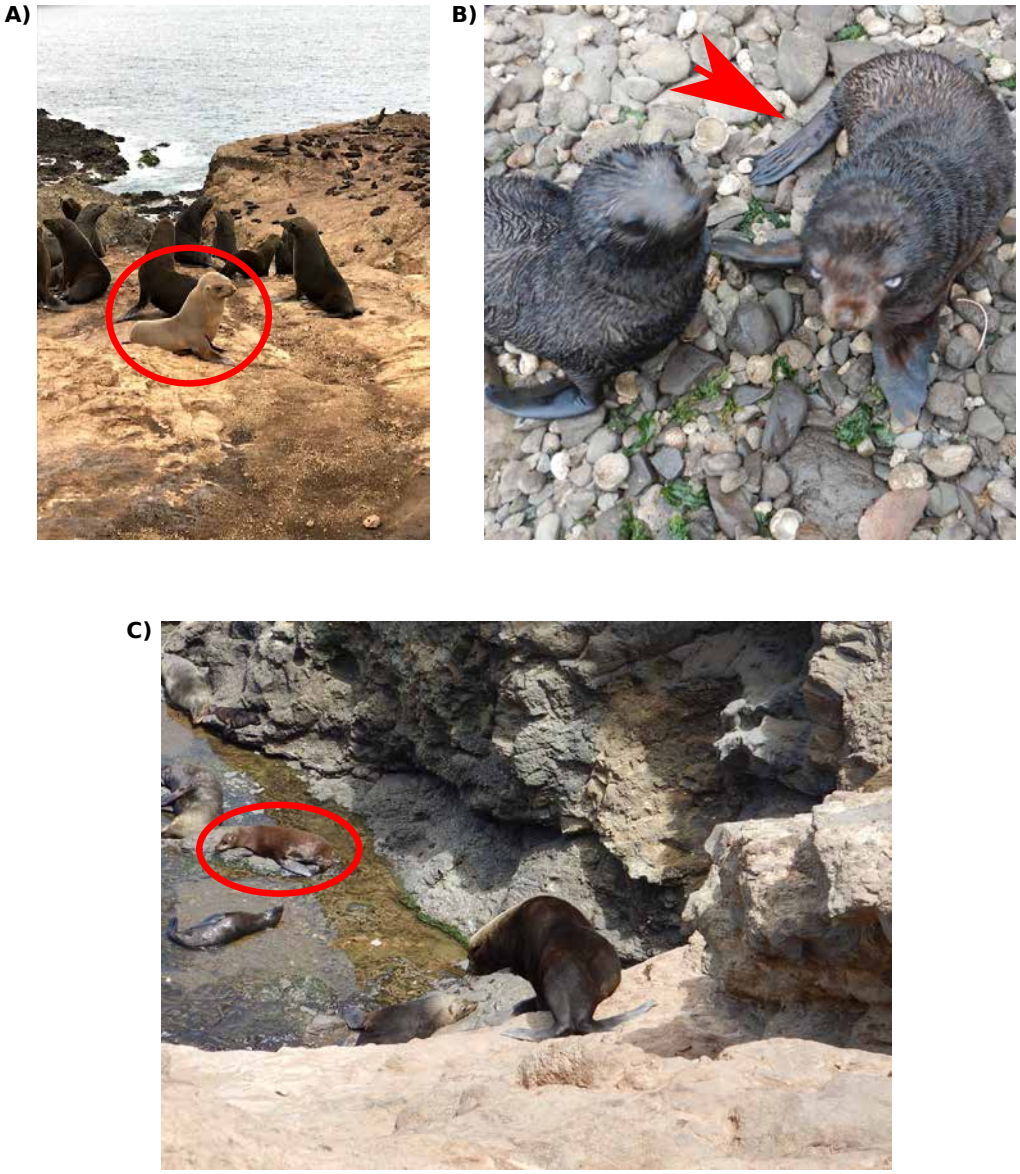

Figure 1.1. Leucistic Juan Fernandez fur seals.
Leucistic individuals are encircled in red or indicated with a red arrow. A) Female, B) Pup, C) Male.

Figure 1.2. Vagrant fur seal species registered in the Juan Fernandez archipelago. A) and B) show juvenile females of Antarctic fur seal (*Arctocephalus gazella*). C) Female Subantarctic fur seal (*Arctocephalus tropicalis*). D) Male Subantarctic fur seal (*Arctocephalus tropicalis*).

2014; Kovacs *et al.*, 2012; Allen *et al.*, 2011; Simmonds and Isaac, 2007). Changes in prey abundance and distribution are among the most documented impacts of climate change on marine environments. These impacts may result from the increase in water temperature itself or other indirect effects such as the incremented frequencies of natural circulation fluctuations such as the Southern Ocean Oscillation, also known as El Niño (Freund *et al.*, 2019). Furthermore, temperate marine ecosystems such as the JFA may be particularly vulnerable to climate change (Edwards and Richardson, 2004; Caccavo *et al.*, 2021). Additionally, evidence has shown that ocean surface temperature has dramatically influenced the distribution of otariids (Churchill *et al.*, 2014) and cooler surface temperatures have facilitated the colonisation of the Southern hemisphere by fur seals and sea lions. Thus, warmer water temperatures will likely result in a reduction in pinnipeds range and population size.

Evidence generated from other better-studied pinnipeds have already reported detrimental effects on their populations due to the impacts on prey abundance and availability resulting from climate change (Páez-Rosas *et al.*, 2021; Forcada and Hoffman, 2014a; Robinson *et al.*, 2018). Due to the characteristics of the JFFS maternal attendance, changes in prey will probably impact pup survival. Furthermore, impacts on prey are not the only possible detrimental effects of climate change on JFFS fur seal populations. Sepúlveda *et al.* (2020) reported an increment in stranding rates of newborn sea lions due to an increase in the frequency and intensity of coastal storms. The habitat characteristics of most of the JFFS colonies located in SC and RC (Osman and Moreno, 2017) are particularly exposed to these events. Thus, an increment in pup stranding due to coastal storms should be expected. It would not be surprising if these populations are already being affected, but due to the lack of adequate monitoring, variations in the population can not yet be perceived.

Marine pollution is another critical threat to the JFFS population. Microplastic and heavy metal exposure has already been evidenced in this species (Perez-Venegas *et al.*, 2018; Sepúlveda *et al.*, 1997). Furthermore, high trophic position, longevity and fat content make pinnipeds highly susceptible to bioaccumulation and biomagnification of environmental pollutants (Hazen *et al.*, 2019; Gray, 2002). Additionally, prey selection in combination with high exposure to the anthropogenic debris associated with the South Pacific gyre may increase the vulnerability of the JFFS to this threat.

Finally, JFFS harvesting by man may become a real threat if Governmental decisions are taken without the necessary information.

1.5 Gut microbiome

The microbiome is the entire community of microbes inhabiting a particular environment. However, in this dissertation, I will only focus on bacteria communities.

When inhabiting a living organism, the microbial community is known to play critical roles in host fitness, survival and disease (Daskin and Alford, 2012; Stappenbeck and Virgin, 2016). The gastrointestinal tract, especially the colon, is recognised as one of the largest microbial reservoirs in an animal host (Ley *et al.*, 2008a). This community is acquired at birth, during lactation and later from the environment and is modified by factors such as age, sex, circadian changes and fluctuations in the environment (Stappenbeck and Virgin, 2016; Ley *et al.*, 2008a).

The gastrointestinal microbiome fulfils essential functions in gut maturation, digestion/host nutrition, metabolic activity and plays a leading role in local and systemic immune system function by influencing innate and adaptive immune defences, preventing infectious pathogen colonisation and contributing to the maturation/modulation of the immune system (O'Hara and Shanahan, 2006; Woodhams *et al.*, 2020). Additionally, the composition of the gut microbiome community differs according to species, diet, habitat characteristics, individual host genotypes and even location in the gut, among other factors (Bik *et al.*, 2016; Groussin *et al.*, 2017; Ley *et al.*, 2008a).

Until recently, the study of host-microbial interaction has been focused on human, animal models and domestic animals. However, the strong influence of the microbial community on host physiology and fitness demonstrated by the human-centric studies motivated scientists to explore the host-microbial interactions in wild populations. Studies looking at the microbiome of wild animals have shown changes in the microbiome associated with anthropogenic impacts such as habitat degradation, captivity and environmental pollution (Amato *et al.*, 2013; Delport *et al.*, 2016; Alfano *et al.*, 2015; Suzuki *et al.*, 2021; Ley *et al.*, 2008a; Fackelmann and Sommer, 2019; Marangi *et al.*, 2021a). Similarly, microbial patterns are associated with physiological processes such as reproductive state (Dietrich *et al.*, 2018). Furthermore, studies in sexually dimorphic species such as the Northern elephant seal evidenced differences in the faecal microbiome composition between female and male pups.

Despite the potential of using the faecal microbiome to study wild populations, unlike controlled experiments, it is not easy to isolate different variables. Thus, there is higher risks to be directed to misleading conclusions due to confounding variables. Additionally, sampling and analytical methods often differ between studies which means they are difficult to compare. Therefore, it is critical that before asking more complex questions an adequate characterisation of the microbial composition is carried out.

1.6 Conservation genetics

In the last two decades the use of molecular techniques to study wild populations has become more feasible and widely spread. In combination with traditional information such as geography, population abundance, behaviour, environmental and historical processes the study of wildlife genetics can be a powerful tool to investigate topics including genetic diversity, hybridization, heterozygosity and patterns of historical population increments or declines (DeYoung and Honeycutt, 2005; Amos and Hoelzel, 1992).

In marine mammals, the analysis of mitochondrial DNA (mtDNA) is the most used approach to study genetic diversity and phylogeny. Due to the number of copies, this is a relatively easy target when using samples containing poor DNA quantity and quality such as fossils, faeces or hair. Due to the high degree of nucleotide polymorphism, the mtDNA control region has been widely used in population genetic studies and phylogeographic analysis (Rosel *et al.*, 2017). As a matter of fact, this is the only target that has been looked at in all pinniped species and thus, has a major contribution in the current taxonomic classification (Rosel *et al.*, 2017; Slade *et al.*, 1994; Berta *et al.*, 2018). One of the reasons for its popularity is its exceptionally high evolution rate (Tatarenkov and Avise, 2007; Stoneking, 2000). Additionally, mtDNA is maternally inherited and haploid which makes it a reliable target to study the historic maternal lineage within a population as well as the impacts of population bottlenecks on genetic diversity (Amos and Hoelzel, 1992; DeYoung and Honeycutt, 2005). However, maternal inheritance is both an advantage and a disadvantage as it only reflects the the evolutionary history of the maternal line (Rosel *et al.*, 2017). Additionally, introgressive hybridisation and heteroplasmy are common occurrences in mtDNA which may obscure its correct interpretation (Rubinoff *et al.*, 2006; Keller *et al.*, 2010).

Microsatellite analysis is another popular tool in conservation genetics. These markers are simple sequence tandem repeats (SSTRs). The repeat units usually involve between two and five nucleotides (Vieira *et al.*, 2016). The polymorphism observed in microsatellites results from the addition or deletion of entire repeat motifs cause by polymerase strand-slippage or by recombination errors. These errors results in variations of repeat numbers between individuals (Vieira *et al.*, 2016). Individuals have two copies at each locus, one inherited from each parent. Thus, the level of homozygosity within these loci can be measured and used as an indication of inbreeding. They also have a variety of uses in identifying individuals and their source, population geographic spread, historical population expansion or bottlenecks events and can be linked to disease resistance (DeYoung and Honeycutt, 2005; Hoffman, 2009). No one has studied any nuclear genetic marker in JFFS to date.

1.7 Heavy metals

The mechanisms by which heavy metals cause toxicity have been widely documented and reviewed (Jaishankar *et al.*, 2014; Balali-Mood *et al.*, 2021). Mercury (Hg) toxicity and bioavailability varies depending on its form: metallic element, inorganic salts or organic compounds. Methylmercury, for instance, is known to cause microtubules and mitochondria destruction, increased lipid peroxidation and accumulation of neurotoxic molecules (Patrick, 2002). According to the Agency for Toxic Substances and Disease Registry (ATSDR), cadmium (Cd) is amongst the most toxic heavy metal. This element affects gene expression, inhibits DNA repair, interferes with apoptosis and autophagy, induces oxidative stress and interacts with bioelements such calcium (Jaishankar *et al.*, 2014; Balali-Mood *et al.*, 2021; Buha *et al.*, 2019).

There is plenty of evidence that marine mammals are highly exposed to heavy metals, especially those who prey on cephalopods. Furthermore, these species are known for their capacity to bioaccumulate and biomagnify contaminants such as toxic heavy metals (Gray, 2002). Nevertheless, these species have evolved various adaptations that make them more resilient to these pollutants. For instance, marine mammals have high numbers of metallothioneins in relevant organs such as the liver and kidney to detoxify the heavy metal (Wang *et al.*, 2014). Despite evolving decontamination methods, marine mammals may still suffer from heavy metal toxicity (Kakuschke *et al.*, 2011; Schaefer *et al.*, 2011) as it is not yet clear what a marine mammal's resilience threshold is to heavy metals. Indeed, Desforges *et al.* (2016) suggested that the Cd adverse effect concentration (ppm) for lymphocyte proliferation suppression varies among marine mammal species.

Processes such as ocean acidification may be increasing the bioavailability of heavy metals in marine environments (Shi *et al.*, 2016). Additionally, evidence shows microplastic debris can work as a heavy metal vector (Bradney *et al.*, 2019; Brennecke *et al.*, 2016). We already know the JFFS is highly exposed to plastic contamination. Thus, it is necessary to monitor the exposure and the possible consequences on their health.

1.8 Thesis aims

In summary, the absolute numbers of the JFFS have been increasing since their rediscovery. However, information on this species is limited and outdated. Nevertheless, from published studies combined with field observations, this species is vulnerable to climate change and ocean pollution. Therefore, understanding the current status of this species is vital for its conservation. The fast generation of information about the JFFS is critical as soon they may no

longer benefit from legal protection. Therefore, identifying and optimising methods to access as much information as possible will enormously contribute to the conservation of this species and may enable the use of this species as a marine bioindicator for the JFA ecosystem.

Within this context, this PhD research aims to use a non-invasive method of sampling (faecal samples) to monitor a variety of biological parameters in the JFFS:

1. To determine the faecal bacteriome,
2. To determine heavy metal exposure of the JFFS from faecal heavy metal content,
3. To develop genetic tests for the JFFS from faecal material.
    (a) microsatellite analysis to estimate in-breeding,
    (b) sexing of samples,
    (c) mitochondrial DNA haplotype analysis to estimate population structure.

Chapter 2

Materials and Methods

2.1 Study Area

This research took place in the Juan Fernandez archipelago (JFA). Samples were collected from different locations covering two of the three islands that make up the archipelago as shown in Fig. 2.1.

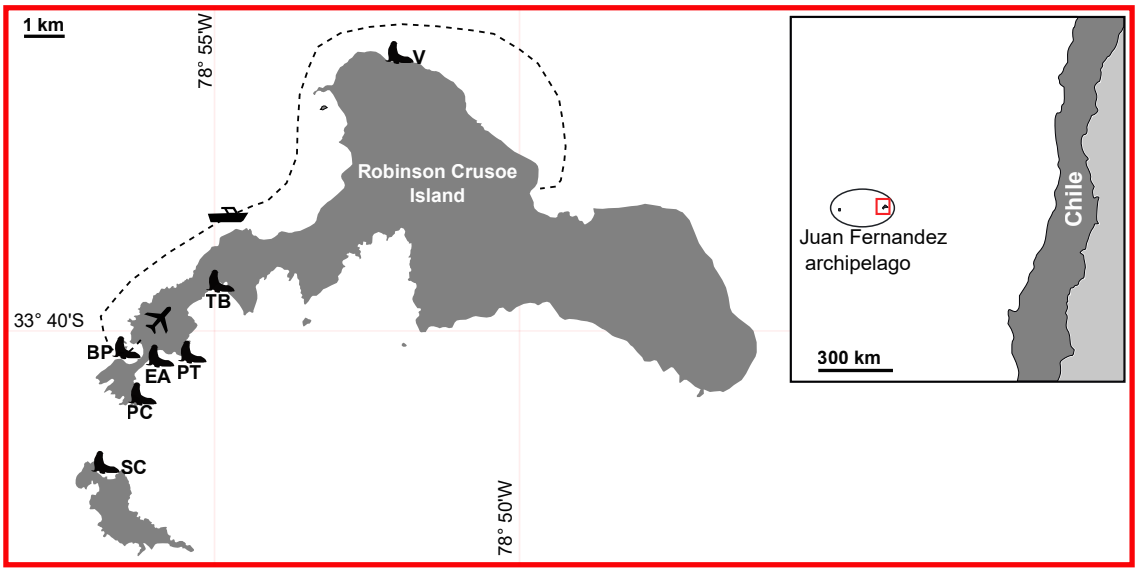

Figure 2.1. Simplified map of Robinson Crusoe and Santa Clara islands.
The plane indicates the airfield and the dotted line the access route from the airfield to San Juan Bautista Village (the only settlement on the island). Fur seal icons show the sampling locations. El Arenal (EA), Bahia El Padre (BP), Piedra Carvajal (PC), Punta Trueno (PT), Santa Clara (SC), Tierras Blancas (TB) and Vaqueria (V).

Robinson Crusoe Island (RC) (33°38'00"S 78°51'00"W) is the only permanently populated island and the main access to the JFA. The town, San Juan Bautista, is located on the north coast of the island. The island can be accessed by plane or by boat. The Juan Fernandez airfield is located on the south-western coast of RC. From here, passengers take a boat from Bahia El Padre (BP) directly to the town. During the fur seal reproductive season (specifically during January and February), BP is the location with the largest numbers of Juan Fernandez fur

seals (JFFS) and, therefore, the most exposed to direct anthropogenic pressure (water traffic, oil, sewage, tourism). Based on the recent 2018 JFFS census carried by CONAF (data not published), most of the JFFS population in Robinson Crusoe is located in the south-west peninsula, including the colonies from El Arenal (EA), the only sandy beach occasionally visited by locals and few tourists, and Tierras Blancas (TB). TB is characterised by being an open natural platform a couple of meters above the sea, protecting the resident colony from the effects of sea storms and rockfalls. Santa Clara Island (SC) (33°42'07"S 79°00'05"W) is the smallest of the three islands. This uninhabited island is one kilometre south-west of RC. The access to this island is restricted, only occasional visitors for educational, research and monitoring purposes are allowed.

2.2 Ethics

All JFFS faecal samples were collected directly from the ground in a non-invasive manner. In addition, JFFS bones and tissue samples were collected post-mortem. Disturbance of the colonies was kept to a minimum, and no animal was handled or harmed in the process. Two Chilean Institutions gave permits for the collection of JFFS samples: Comision Nacional Forestal (CONAF) (Certificate 009217) and SERNAPESCA (R.E.X.N 43 and R.E.X.N 445/201). Finally, import licenses from the Department for Environment, Food & Rural Affairs (DEFRA) were obtained (ITIMP16.1158 and ITIMP19.0401).

2.3 Sample collection

2.3.1 Faecal Samples

With the logistic support of CONAF, I collected 75 JFFS faecal samples during three consecutive reproductive seasons: 2016–2017, 2017–2018 and 2018–2019. Twenty-five samples were collected from BP, 19 from TB, 17 from SC and 10 from EA. Opportunistically, I collected single samples from Vaqueria (V), Punta Trueno (PT), Piedra Carvajal (PC) and El Pangal (EP) Supp. Tab. A.1. Not all the samples were adequate for every analysis. For instance, some of the samples were too old to be considered for any DNA-related study but were used for heavy metal analysis. When possible, I collected three different sub-samples from each specimen. I collected the first by swabbing the external surface of the sample using a sterile Copan FLOQSwab. This first swab was aimed to collect host epithelial cells shed in the outside layer of the faeces to carry out host genetic studies (Ramón-Laca *et al.*, 2015). See more details on sample collection, storage methods and the selection process for the host genetic study in

section 2.6. To study the faecal microbiome, I used a disposable sterile spatula to expose the core of the faeces and collected a second sub-sample by carefully avoiding the material in direct contact with the surrounding elements (Vlčková *et al.*, 2012; Blekhman *et al.*, 2016). Both swabs were immediately stored in independent 1.5 ml Eppendorf Tubes containing RNAlater (Sigma-Aldrich). Finally, I placed the remaining faeces in Nalgene 5005-0015 Specimen Cryogenic Vials for analysis of heavy metal content.

Collection of samples took place before noon to limit sun exposure. The samples were collected based on consistency and colour to reduce the variability between samples, introduced by the delay between defecation and collection. Host information such as sex or age was not possible to distinguish at the time of collection. I used visual cues and GPS location to decrease the risk of collecting specimens from the same individual. Samples were stored at -20°C within 32 h of collection and for 1–2 months until arrival transfer to the laboratory, where they were stored at -80°C in the case of the samples used for microbial analysis (Castro lab, Center for Bioinformatics and Integrative Biology, Universidad Andres Bello), and at -30°C for the other sub-samples (Dept. of Veterinary Medicine, University of Cambridge).

Dr. Jaume Fourcada provided an additional ten Antarctic fur seal faecal samples from the British Antarctic Survey (BAS). These samples were collected from Bird Island, South Georgia, Antarctica (54°00'S 38°03'O) during the reproductive season 2017–2018, as part of their long-term monitoring program. The faecal material was collected directly from the rectum and stored in RNAlater at -80°C until further analysis.

2.3.2 Bone samples

To look at heavy metal concentrations in JFFS bones, I collected ten old bone fragments directly from the ground during the reproductive season 2018–2019. All the bones were lower mandibles from adults (n = 5) and pups (n = 5) of undetermined sex or exact age.

For comparison, two archived bones samples were donated by Prof Joseph Hoffman from Bielefeld University, Germany. Both samples were made up of skull fragments. One of the samples belonged to an adult male grey seal collected in the Orkney Islands, Scotland, in 2003. The other sample corresponded to an adult male Antarctic fur seal collected from Bird Island, South Georgia, in 2019.

2.3.3 Non-pinniped samples

Based on the heavy metal concentration in faeces, some prey samples were also analysed for heavy metals. Eleven myctophid fish samples were provided by the Algalita Marine Research and Education Foundation. These samples were found dead, floating in different points of the

South Pacific gyre between Easter Island and the JFA. Additionally, internal organs of three octopuses, one lobster (hepatopancreas, kidneys and gills) and two whole sea cucumbers were donated by members of the local community. Finally, I collected three water samples and four soil samples from BP and TB.

2.4 Gut Microbiome

2.4.1 16S rRNA sequencing

Fifty-seven JFFS faecal samples collected during two consecutive reproductive seasons (2016–2017 and 2017–2018) from seven different locations were included in this study. Six of the seven colonies were located on Robinson Crusoe Island: El Arenal (n = 9), Bahia El Padre (n = 23), Piedra Carvajal (n = 1), Punta Trueno (n = 1), Tierras Blancas (n = 12) and Vaquería (n = 1). One colony was located on Santa Clara island (n = 12).

DNA extraction and sequencing

Samples were processed as soon as possible after the completion of each fieldwork season. Due to the possible batch effect introduced by processing samples in different years, comparisons between years of collection will not be explored in this study.

I thawed the samples on ice and centrifuged at $10,000 \times g$ for 15 min to pellet the sample out of RNAlater. I extracted the genomic DNA from each pelleted sample (approx. 180 micrograms) using the MO BIO PowerSoil DNA Isolation kit (QIAGEN) accordance with the manufacturer's instructions. Isolated DNA was quantified by Dr Eduardo Castro-Nallar's group (Castro Lab, Center for Bioinformatics and Integrative Biology, Universidad Andres Bello, Chile) on a Qubit 3.0 fluorometer (Invitrogen, Waltham, MA), using the Qubit dsDNA HS Assay Kit (Invitrogen, Waltham, MA).

The bacterial 16S rRNA gene was PCR (Polymerase Chain Reaction) amplified targeting a 250 bp region covering the V4 variable region. PCR amplification, barcode tagging, and library preparation was performed according to Kozich *et al.* (2013). Libraries were constructed using the TrueSeq DNA kit and sequenced on a MiSeq platform (Illumina). The read length target changed between the two sampling years. Sequencing was performed using v2 chemistry producing 2×250 bp paired-end reads in the 2017 samples while the 2018 sequences were 2×150 bp paired-end reads. The PCR assays and library preparation were conducted at The Microbial Systems Molecular Biology Laboratory (MSMBL), University of Michigan, USA, in collaboration with the Castro Lab.

Sequence data analysis and taxonomic classification

I manually assessed the raw sequence quality with FastQC v. 0.11.5 (Andrews, 2010). All 57 samples contained reads of consistent length (respective to the sequencing year) and the average read quality score was above 30. A drop in base quality was observed at the ends of reads (4–5 and 8–10 respectively). Then, I imported the demultiplexed raw sequences into QIIME2-2019.10 (Bolyen *et al.*, 2019) where quality control, de-replication, read truncation and paired read merging was performed using the DADA2 (Divisive Amplicon Denoising Algorithm) qiime2 plugin (Callahan *et al.*, 2016). Instead of generating operational taxonomic units (OTUs) by clustering sequences based on similarity, the final output of DADA2 is a table with exact sequence variants also known as amplicon sequence variants (ASVs), which are generated by modelling and correcting Illumina sequencing errors. To avoid the batch effect, I carried out this step separately according to the year of collection. However, to normalise between datasets, I truncated the 250 bp reads produced from 2017 samples so that the paired reads matched the length of the paired reads from 2018 samples. To confirm consistency in paired read lengths between the two years, I aligned the representative sequences generated from both years in Geneious Prime 2020.0.5 (2020) using MAFFT (Multiple Alignment using the Fast Fourier Transform) plug-in with default settings and which I then assessed by eye (Katoh and Standley, 2013).

Next, I generated a mid-point rooted, approximately-maximum-likelihood phylogenetic tree for diversity analysis using the qiime2 phylogeny plug-in which uses the MAFFT and the FastTree program (Price *et al.*, 2010). Finally, taxonomies were assigned to the ASVs using a 16S-V4-specific classifier trained against the Silva132 database clustered at 99 % sequence similarity (Quast *et al.*, 2013).

Data processing and statistical analysis

I performed statistical analysis in duplicates, once using all available data and then only with data corresponding to the core microbiome. The core microbiome is defined here as all the ASVs present in at least 50 percent of the samples.

Data processing and statistical analysis were carried out in R version 3.6.0 (R Core Team, 2019). I applied multiple filtering steps using the Phyloseq package version 1.36.0 to prepare the data (McMurdie and Holmes, 2013).

1) I manually inspected unassigned ASVs at the Kingdom level, with the Basic Local Alignment Search Tool (BLAST) before filtering based on both BLAST results (those with non-bacterial matches) and prevalence (ambiguous taxonomy at the phylum level with a prevalence of 1 and total abundance less than 5 reads) (Altschul *et al.*, 1990).

2) Based on the rarefaction curve (Fig. 3.1), I identified 3 samples having insufficient depth of sampling and these were therefore removed from the statistical analysis. A threshold of 13,980 reads was used as a cut-off. Removed samples were identified as 17JFFS16 (BP, 4463 counts), 17JFFS23 (TB, 2602 counts) and 17JFFS23 (EA, 2042 counts). Despite having reasonable read counts, the three excluded samples did not reach the plateau in the rarefaction curve, suggesting that these samples are likely to have a poor representation of the microbial community.

3) I also removed possible contamination signals by running a correlation analysis and comparing clusters with a list of previously identified reagent contaminants (Salter *et al.*, 2014).

4) Finally, the data was rarefied using the same threshold used for filtering samples (McKnight *et al.*, 2019). The effects of each filtering step are summarised in Tab. A.2.

I characterised the overall microbial composition by summing the non-normalized read counts and obtaining the relative abundance at different taxonomic levels.

Alpha diversity

I calculated the estimates of within-sample diversity (alpha diversity) using the Phyloseq package version 1.36.0. Three indices were included: a richness estimator, which estimates the total number of species in each sample (Chao1) and two different diversity estimators (Shannon-Weiner and Simpson index) (Fisher *et al.*, 1943; Shannon, 1948; Simpson, 1949; Morris *et al.*, 2014). The latter two approaches consider richness as well as abundance. However, the effect of richness and rare species strongly impact the Shannon-Weiner index, whereas the Simpson index is mainly influenced by evenness and common species. Additionally, I decided to explore the alpha diversity by looking at the effective number of species (true diversity), defined by Jost (2006) as "the number of equally-common species required to give a particular value". I calculated the diversity of order q = 1 which represents the true diversity associated the Shannon-Weiner index. This transformation to effective numbers is done by applying an exponential function to the original Shannon-Weiner index. From now on, I will refer to this effective number as Q1.

I used non-rarefied data to explore the alpha diversity. To compare locations, I performed a one-way analysis of variance test (ANOVA) or a non-parametric Kruskall Wallis test for each estimate. I tested the ANOVA assumptions by visualising the data and performing statistical testing. I also used a Shapiro-Wilk test to confirm normality and a Levene's test for heteroscedasticity. When exploring Shannon-Weiner and Simpson indices, I identified sample

18JFFS23 (SC) as an outlier (standard residual > 3) which I removed for these indices only. Q1 also showed an outlier which I removed for its respective analysis (17JFFS15). Finally, data visualisation suggested samples collected from TB differed from the other locations. Therefore, I performed a post-hoc analysis with Dunnett's or the non-parametric Dunn's test to compare each location to TB. Samples from PC, PT and V were not included in the location comparison due to their limited sample size (n = 1).

Beta diversity

To investigate the variation between samples (beta diversity) I calculated two different distances for both rarefied full, as well as, the core datasets. Bray-Curtis dissimilarity distance looks at the differences between samples based on the ASV abundances. Weighted UniFraq distance explores the phylogenetic divergence between ASVs by also taking into account the abundance of these (with an emphasis on dominant ASVs). I used the R packages Vegan version 2.4-2 to calculate Bray-Curtis dissimilarity distances and Rbiom version 1.0.2 to calculate UniFraq distances. I then ran and plotted a principal coordinate analysis (PCoA) on each distance matrix.

To further explore the clustering of samples (Cluster 1 versus Cluster 2) observed in the Bray-Curtis PCoA, I computed a permutational multivariate ANOVA (PERMANOVA) with 999 permutations to test for significant differences between the clusters. Finally, I ran a Similarity Percentages breakdown analysis (SIMPER) between the clusters to identify the genera that contributed most to the difference between clusters. I further explored the genera that contributed highly to dissimilarities between groups with the non-parametric Mann-Whitney U test.

I used a Spearman rank correlation coefficient (ρ) to explore any possible associations between the different taxa and also between the first two components of the Bray-Curtis ordination analysis. I built a correlation matrix to visualise significantly and strongly correlated ($\rho \geq |0.6|$) genera. For this method, only the core microbiome dataset was used at the genus level.

Finally, I performed a Redundancy Analysis (RDA) to explore the variation explained by location. Briefly, RDA, a constrained ordination analysis, looks for the axes (components) within a multidimensional analysis that maximises its association with the explanatory variable of interest, in this case, location. Due to the zero-rich dataset, a Hellinger tranformation was applied before the analysis (Legendre and Gallagher, 2001).

Predictive functional analysis

I performed a predictive functional analysis on the unrarefied ASV abundance table using PICRUSt2 (phylogenetic investigation of communities by reconstructing unobserved states) version 2.3.0b (Douglas *et al.*, 2020). This bioinformatics tool is based on placing the ASV sequences into a reference tree to predict the hidden-state gene families. After normalising the gene copy numbers, PICRUSt2 can then infer meta-genomes and make pathway-level predictions.

In QIIME2, I applied a couple of preliminary filtering steps on the feature table containing the ASV abundances before running the PICRUSt2 pipeline. To reduce the noise in the analysis, I removed all the ASVs with a total abundance of less than five. I also filtered the three samples excluded from the 16S rRNA taxonomy analysis. I followed the official PICRUSt2 tutorial pipeline (https://github.com/picrust/picrust2/wiki/PICRUSt2-Tutorial-(v2.4.1)) to predict the 16S gene copy numbers, the Kyoto Encyclopedia of Genes and Genomes (KEGG) orthologs and the Enzyme Comission (EC) numbers (Kanehisa and Goto, 2000; Krieger *et al.*, 2004). To infer the KO and EC metagenomes, the ASV abundances are normalised according to the predicted gene copy numbers. The final step of the pipeline is the pathway-level prediction. Based on the EC number abundances, PICRUSt2 can then infer the MetaCyc pathway abundances at the low hierarchical level by default. To infer top pathway levels, I used the mapping files curated by Jiung-Wen Chen publicly available on GitHub (https://github.com/Jiung-Wen). The inference of KEGG pathway abundances (based on KO's) is not included in the software default settings. Thus, I ran this step in R following the Rscript suggested in the FAQ section of the software repository's wiki (https://github.com/picrust/picrust2/wiki).

I performed the functional prediction statistical analysis in R version 4.1.0 for both KEGG and MetaCyc pathways. I started by looking at the Nearest Taxon Index (NSTI) distribution to investigate the quality of the predictions. Well characterised communities will show lower NSTIs (e.g. Human Microbiome Project NSTI = 0.11 *pm* 0.49) than less characterised ones (e.g. ocean dataset NSTI = 0.51 *pm* 2.06) (Langille *et al.*, 2013; Douglas *et al.*, 2019). ASVs with NSTI > 2 were automatically removed by PICRUSt2. Next, I used descriptive statistics and data visualisation tools to explore the overall pathway composition at the higher pathway levels (levels one and two). Finally, I looked for significant low-level pathways that differed between the two clusters identified in the taxonomic analysis. PICRUSt2 outputs are characterised by being compositional data; therefore, I used the ALDEx2 (ANOVA-Like Differential Expression) version 1.24.0 for this analysis. When running the ALDEx2 analysis, the data is first transformed into centred log-ratio values (CLR). This transformation is done by generating Monte-Carlo (MC) instances which, in this case, I set to 1000. The final CLR value for each feature results from calculating the geometric mean from MC instances. Next, I applied

an interquartile log-ratio correction (iqlr) to account for the asymmetry of the variance of the clr values when calculating the geometric mean. This correction only uses as denominator, features between the first and third quartiles across all samples. I then performed Welch's t and Wilcoxon rank tests on the CLR-transformed data. I only considered pathways with an effect size threshold -1 < or > 1 and an expected Benjamini-Hochberg corrected p-value < 0.05 in both tests to be distinct between clusters. For KEGG pathway predictions, I also considered median log2(fold change) differences larger than one as an additional criterion.

2.5 Heavy metals

Fourteen of the JFFS faecal samples collected during the 2017 – 2018 reproductive season and five of the AFS samples collected by the BAS team, were measured for heavy metal content to determine fur seal exposure to heavy metals. Bone samples from cadavers, were also analysed to determine the extent of gastrointestinal absorption and retention of heavy metals by the animals. To determine the sources of heavy metals in the fur seal food chain, ten whole myctophids and the internal organs of three octopuses were also investigated for heavy metals. Finally, sea water and soil samples collected at JFFS faecal sampling locations were also analysed to inform upon possible environmental contamination of the samples.

2.5.1 Trace element analysis

First, all faecal samples were air dried in a clean class II microbiological safety cabinet to eliminate any variation in water content between samples. Aliquots (0.1 g) of the dried samples were then digested with ultra-pure (UHP) nitric acid and hydrogen peroxide at room temperature, and the digest diluted with UHP water. Sample blanks were similarly prepared. The diluted samples and blanks were then measured for the concentrations of 53 different elements with a triple quadrupole Inductively Coupled Plasma Mass Spectrometer (8900 ICP-MS/MS, Agilent Technologies Inc, CA, USA), using appropriately prepared multi-element calibration standards. Food chain samples were similarly prepared for ICP-MS/MS analysis, except the samples were not dried prior to acid digestion. Water samples were diluted in 1 % UHP nitric acid prior to analysis, whilst soil samples were digested in an acid digestion microwave (UltraWave, Milestone SRL), diluted and analysed with the ICP-MS/MS. Bone samples were first thoroughly cleaned of external surface contaminants by sonicating several times in ultra-pure water in an ultrasonic bath. Cleaned bones were then dried and aliquots digested in the acid digestion microwave. The acid digests were diluted in UHP water and analysed with the ICP-MS/MS. Cadmium and silicon levels in the diluted faecal and bone

digests by ICP-MS/MS, were confirmed with analysis repeated on an ICP-Optical Emission Spectrometer (Ultima 2C ICP-OES; Horiba Jobin-Yvon). Sample digestions and ICP-MS/MS and ICP-OES analyses were carried out by Dr Ravin Jugdaohsingh (Prof Jonathan Powell's group), at the Department of Veterinary Medicine, University of Cambridge.

2.5.2 Statistical analysis

All statistical analyses were performed in R version 4.1.0 (R Core Team, 2019). In this part of the analysis, I only focused on exploring the faecal and soil samples. First, I reduced noise in the data by identifying elements detected at high concentrations in the faecal samples but which were low in the soil samples. I detected these elements by running a principal component analysis (PCA) and a cluster analysis. The six elements with the highest concentrations in the faecal samples were selected for further investigation. These were phosphorus (P), Copper (Cu), Zinc (Zn), Selenium (Se), Cadmium (Cd) and, Mercury (Hg). The low concentrations of these elements in soil meant their high concentration in the faecal samples was unlikely to be the result of cross-contamination with soil at the time of collection. Then, I ran a Spearman rank correlation to look for possible statistical relationships between the selected elements. Finally, I was interested in comparing the selected element concentrations between fur seal species. Due to the uneven samples size, I performed a non-parametric Mann-Whitney U-Test at the 0.05 significance level.

To look at elements in JFFS bones, I applied a similar statistical approach to the faecal samples. Including both bone and soil samples, I started by running a PCA and a cluster analysis to identify possible overrepresented elements due to cross-contamination with soil, followed by a non-parametric Mann-Whitney U-Test at the 0.05 significance level to compare the elements concentrations between age groups. Finally, I ran a Spearman correlation analysis, this time including the two non-JFFS bone samples.

I performed a final PCA analysis including the octopus, lobster, sea cucumber, myctophid and faecal samples to see if any non-fur seal samples had similar trends to the faecal samples to inform future work.

2.6 Host genetics

Seventy-two faecal samples (external swabs) collected throughout the three reproductive seasons were used. External swabs were used to increase the proportion of host epithelial cells in each sample, increasing the chance of DNA amplification.

2.6.1 Testing sample storage and DNA extraction kits

Dog samples were used to test the effect of different storage strategies (Freezing, RNAlater and 96 % ethanol) as well as different DNA extraction methods; QIAmp DNA stool mini kit (Qiagen), GeneAll Exgene Stool mini kit, Quick DNA Fecal/Soil Microbe MiniPrep Kit (Zymo Research) and TRIzol (Invitrogen). I collected three bag-disposed dog faeces found in the School of Veterinary Medicine surroundings, University of Cambridge. I independently swabbed the outside layer of each faecal sample three times. Each of the three swabs was placed in a separate 1.5 ml Eppendorf tube. The first tube was empty, the second tube containing 1 ml of ethanol and the third one, 1 ml of RNAlater. Tubes number two and three were left at room temperature, while tube number one was kept at – 20°C. DNA extraction took place four months later, following the protocol of each manufacturer. Samples in ethanol or RNAlater were centrifuged at 5000 $\times$ g for ten minutes to remove the storage medium prior to the extraction.

From the experiment described above (results are disclosed in Chapter 5) and considering cost-benefits, I decided to use the Qiamp kit to perform the DNA extraction from the JFFS faecal samples. I only modified step two of the manufacturer's protocol; after removing the RNAlater and adding the 1.4 ml of ASL buffer, I incubated the samples overnight at 56°C. I was able to successfully extract 24 samples before running out of QIAmp DNA stool mini kit reagents. Unfortunately, Qiagen no longer manufacture this kit, and the QIAmp alternative, QIAamp Fast DNA Stool Mini Kit, did not work well on the JFFS faecal samples. Therefore, I extracted the remaining 48 samples with TRIzol.

Additionally, I extracted DNA from tissue samples using the DNeasy Blood and tissue kit (Qiagen) following the manufacturer's protocol to optimise PCR reactions. These samples (N = 4) were collected opportunistically from dead carcasses. Due to the decomposition stage, all tissues collected were taken from the ear of each individual.

2.6.2 Primer selection

Even though swabbing the outside layer of the faeces can increase the number of host cells per sample, it does not ensure that there will be enough DNA, if any. This is a challenge that particularly affects the amplification of nuclear DNA as each cell contains only one copy of this type. On the other hand, there are many mitochondria per cell, each a mitochondrial DNA (mtDNA) copy ie multiple mtDNA copies per cell. For this reason, I decided to start by amplifying mtDNA to identify samples with a higher potential of containing enough host nuclear DNA (Reed *et al.*, 1997b). I used primers previously used in JFFS and other pinnipeds that specifically targeted the full length of the mtDNA control region; 5'-TTCCCCGGTCTTGTAAACC-3' (T-Thr)

and 5'-ATTTTCAGTGTCTTGCTTT-3' (T-Phe) with an additional internal primer; SCR-1, 5'-CCTGAAGTAAGAACCAGATG- 3', which I used for subsequently running semi-nested PCRs (Hoelzel and Green, 1992; Hoelzel *et al.*, 1993; Goldsworthy *et al.*, 2000). Due to the background noise observed when the T-Thr primer was used for sequencing the amplicons, I developed a new forward primer (T-Thr2) using primer3 (https://primer3.ut.ee/). A full list of primers used in this study can be found in Supp. Tab. A.3. Once I confirmed the presence of mtDNA, I then worked on optimising the PCR amplification of five microsatellite loci using primers previously designed for AFS samples (Hoffman, 2009, 2011). One of the problems of amplifying host DNA from faecal samples is the large proportion of non-host DNA that can interact with the primers. This contamination results in false positives and sequencing background noise. For this reason, I designed external primer sets using the published sequence derived when the PCRs were designed and using the published amplicon as a guide. I then ran this external nested/semi-nested PCR and used the product with the published primers. Apart from the SRY primer set, all primer sets used for running internal PCRs were previously published. Supp. Tab. A.3 shows a description of the primers including the amplicon fragment size, the optimised annealing temperature and the reference.

Finally, I was interested in determining the sex of the samples by performing a molecular sexing protocol. I started by using the SRY and the PinZFY primer sets previously developed for pinnipeds (Robertson *et al.*, 2018). Nucleotide BLAST with Californian sea lion sequence for ZFX (DQ811094) and gray seal sequence for SRY (AY424660) was used to identify pinnipeds ZFX/Y and SRY sequences in Genbank which were then aligned with the pinniped sequences used in the design of the original primers (Robertson *et al.*, 2018).

2.6.3 Amplification of host DNA

All the DNA amplifications (tissue and faecal samples) were performed in either 20 or 50 μl reactions using Phusion High-Fidelity DNA Polymerase (New England Biolabs, Thermo Scientific). Reactions were performed in 1X Phusion HF buffer, 200 μM dNTPs and 0.02 U polymerase/μl with primers at a final concentration of 500 nM. Two or four microliters of template were added depending on the size of the reaction.

Before amplifying host DNA from the JFFS faecal samples, I tested and optimised the PCR reactions on tissue samples using a DNA concentration of 10 ng/μl. The optimisation consisted of running a temperature gradient to identify the optimal annealing temperature. Five temperatures at steps of 2°C from highest to lowest were tested, ensuring that the midpoint was centred at the temperature calculated on the online NEB Tm calculator for the primers (https://tmcalculator.neb.com). The final cycling protocol had some small variations depending on the template characteristics and the primers annealing temperature. In general,the initial

denaturation was carried out for 2 minutes or 30 seconds (Faecal extract and PCR product respectively) at 98 ° C. To increase the PCR specificity, a set of initial touchdown cycles were incorporated to the protocol. Each of these cycles started with 15 seconds at 98°C, 15 seconds at different annealing temperatures starting at 72°C, lowering one degree per cycle until the optimised annealing temperature was reached (Korbie and Mattick, 2008) and a extension of 30 seconds at 72°C followed by 45 cycles of amplification (30 if the template was a PCR product) consisting of 15 seconds at 98°C, 15 seconds at the optimised annealing temperature, and 30 seconds at 72°C. The amplification was finalised with 10 minutes extension at 72°C.

To deal with leftover PCR reagents post amplification, a PCR purification step was perfomed between the external and the nested PCR using the Wizard SV Gel and PCR Clean-Up.

To confirm the efficacy of the primers, I checked the PCR products resulting from the tissue samples using 2% agar gel electrophoresis. The band of interest was cut out, extracted from the gel (Wizard SV Gel and PCR Clean-Up kit was used) and sent for Sanger Sequencing at Source BioScience (Source Genomics, Cambridge, UK). Sequences were visualised in Geneious Prime 2020.0.5 (2020).

2.6.4 Analysis of the mtDNA control region

First, Sanger sequencing chromatograms were checked manually in Geneious Prime 2020.0.5 (2020). Samples showing ambiguous sequences were reprocessed from the PCR step and a few samples were re-extracted. Samples that kept yielding ambiguous sequences despite multiple trials were not included in this study. Good quality sequences were included in the phylogenetic analysis which was performed in R version 4.1.0 (R Core Team, 2019).

I ran Clustal Omega Aligments using the Multiple Sequence Alignment (msa) package version 3.13 (Sievers *et al.*, 2011; Bodenhofer *et al.*, 2015). This first alignment included the JFFS sequences generated in this study with additional sequences from a previous study (Goldsworthy *et al.*, 2000). Additional sequences were downloaded from GenBank (accession numbers are listed in Supp. Tab. A.4 (Wynen *et al.*, 2001). Genbank only had five published JFFS mtDNA haplotypes for the mtDNA control region. Thus, I reconstructed the remaining eight based on the information provided by Goldsworthy *et al.* (2000). Haplotype identification according to the authors: ETEli, ETGOGO, ETHoley, LH190 and ETGaz were identified as AF384403.1 (APH1), AF384404.1 (APHI2), AF384405.1 (APHI3), AF384406.1 (APHI4) and AF384407.1 (APHI5) in Genbank. The alignment was manually checked and edited in Geneious. Then, I used the Haplotype package version 1.1.2 to identified haplotypes and haplotype frequencies. To limit the bias introduced by gaps, I excluded columns within the alignments where gaps were present in more 30% of the initial haplotypes (Dwivedi and Gadagkar, 2009).

General population genetic parameters were estimated using pegas: Population and Evolutionary Genetics Analysis System package version 1.0-1 (Paradis, 2010). Furthermore, means, standard deviations and confidence intervals calculated for haplotype diversity (H) and nucleotide diversity (π) were estimated using 1000 bootstraps. The haplotype diversity is the probability that two randomly sampled sequences are different. On the other hand, Nucleotide diversity is the average number of nucleotide differences per site among randomly chosen sequence pairs. This analysis included the sequences generated in this study. However, TB was the only location analysed in both the previous and the current studies. Thus, I performed the same analysis on the sequences previously identified in TB by Goldsworthy *et al.* (2000). Because the latter study was conducted more than two decades ago, I kept this group separated from the new TB sequences. Finally, I estimated the Tajima's D statistic to test if the diversity pattern observed in JFFS was consistent with the neutral expectation (H0).

Next, I conducted an Analysis of Molecular Variance (AMOVA) to test if there was genetic differences between the locations based on the mtDNA control region. For this purpose, I excluded the locations V and PI due to lack of samples. This analysis was carried out using the pegas package with 1000 permutations. Finally, using only the unique sequences (haplotypes) and haplotypes previously identified for other fur seal species (Supp. Tab. A.4), I performed a phylogenetic analysis in IQ-TREE v1.6.12, which uses Maximum Likelihood (ML) to infer phylogenetic trees (Minh *et al.*, 2020). First, the software performs a fast substitution model finder (Kalyaanamoorthy *et al.*, 2017). In this case, the best-fit model according to both the Akaike's Information Criteria (AIC) and Bayesian Information Criteria was HKY (Unequal transition/transversion rates and unequal base freq) + F (Empirical base frequencies) + I (allowing for a proportion of invariable sites.) + G4 (discrete gamma model with four categories). Additionally, I used the software's ultrafast bootstrap approximation with 1000 replicates to assess branch support (Hoang *et al.*, 2018). Unfortunately, IQ-TREE does not have the option of pairwise deletion to deal with gaps. Instead, gaps are treated as missing information. However, after the gap treatment mentioned earlier in the section, only a small number of gaps remained in the alignment. Thus, the remaining gaps will unlikely introduce significant noise to the phylogenetic reconstruction (Dwivedi and Gadagkar, 2009). Initially, I ran the analysis using only sequences identified for JFFS generated in both this study and Goldsworthy *et al.* (2000). Then, I extracted the the pairwise distance matrix generated by IQ-TREE and imported it in to R to conduct an Analysis of Molecular Variance (AMOVA) to test for genetic differences between the locations based on the mtDNA control region using the pegas package using 1000 permutations. The AMOVA requires and euclidean distance matrix, thus, I used the package ad4 1.7-18 to transform the matrix into an Euclidian distance matrix. Also, using pegas, I reconstructed the haplotype network for the data set to generate

in this study and a second one including the data generated previously. Different to inferring a phylogenetic tree which provides information on ancestry, haplotype networks infer the history of mutations between the alleles present in the sampled population. Using IQ-TREE, I conducted a final analysis, to infer a phylogenetic tree including only unique JFFS haplotype sequences and haploplotypes identified for other fur seal species. Accession numbers can be found in Supp. Tab. A.4.

2.6.5 Molecular sexing

Finally, I was interested in determining the sex of the faecal samples by performing a molecular sexing protocol. Further details on this protocol will be covered in Chapter 5. Briefly, The protocol used in this study for sex determination used a hybrid seminested PCR approach. First, individual PCRs were performed targeting the ZFX/ZFY and SRY genes (external fragments). The two PCR products were then combined and purified, and used as a template to conduct a multiplex qPCR using an internal set of primers. A positive SRY amplification indicates male samples while targeting ZFX/ZFY reduces the probability of misidentification of females due to PCR failure.

Robertson *et al.* (2018) optimised a qPCR approach using these two targets for molecular sexing of partially degraded pinniped tissue samples. Furthermore, the PinZFY primers developed by the authors were used in this study. However, new SRY primer sets had to be designed because the ones they used, previously developed by Fain and LeMay (1995), did not work in the JFFS tissue samples. Therefore, the new SRY primers (internal and external) were designed by first aligning different pinniped SYR sequences using Clustal Omega. The sequences included those used by Robertson *et al.* (2018) (all phocids) and an additional two Otariid SRY sequences, one from AFS and one from California sea lion (CSL)(Supp. Tab. A.3). Once the original SRY primers were identified (Fain and LeMay, 1995) a new set was developed in primer3 using the AFS SRY sequence. It was ensured that the new primer set targeted a slightly larger fragment than the original one so it would not affect the separation of the melting curves needed to differentiate the presence of both products (SRY and ZFX). The AFS was also used to design the external reverse primer for the nested PCR. Finally, the CSL sequence used by Robertson *et al.* (2018) was used in a BLAST search to extract any fur seal ZFXY sequences for developing external primers for the semi-nested PCR. An AFS ZFX sequence was used for this purpose.

All primer sets were first tested on tissue samples by running a PCR followed by gel electrophoresis and Sanger sequencing to confirm the primers successfully targeted the area of interest.

External PCRs were then optimised as described earlier in this section. Internal multiplexed qPCRs were performed in 15 μl reactions using the Kapa SYBR FAST universal qPCR kit v2.17 from Kapa Biosystems. Optimisation of the qPCR reaction focus on determining sample dilutions and primer concentrations.

Due to the effects of COVID-19 in education, undergraduate student from the school of Life Sciences were unable to get the hands on lab experience. For this reason, together with my supervisor, we organised a summer project around the molecular sexing of the JFFS faecal samples. Therefore, the lab work related to the molecular sex determination was performed by undergraduate students under my close supervision.

Chapter 3

First characterisation of the Juan Fernandez fur seal faecal microbiome – Who is there?

This chapter is based on the peer reviewed article "Patterns in the Juan Fernandez fur seal faecal microbiome". I am the first author of this paper and did all the data collection and analysis, and wrote approximately 95 % of the manuscript. The chapter is almost the same as the manuscript. However, minor modifications and additional information can be found.

3.1 Introduction

Marine environments are complex and interconnected systems subject to various environmental impacts. Pollution, climate change, disruption of the food network and pathogen dissemination are a few examples of problems currently affecting ocean integrity and function (Halpern *et al.*, 2019a). Integrated approaches at the macro- and micro-ecological levels are needed to properly understand and manage environmental threats in these kinds of complex systems. Identification and investigation of potential environmental sentinel species such as marine mammals can provide a better understanding of the deterioration or improvement of ocean health (Bossart, 2011b; Hazen *et al.*, 2019). However, to effectively use wild populations as sentinels, it is first necessary to establish baseline data to which to compare with future studies.

In the last couple of decades, the study of the microbiome in wild populations has increased, due to the profound impact of host-microbial interactions on host physiology and the growing affordability of sequencing technologies (Williams *et al.*, 2018; Trevelline *et al.*, 2019). The gastrointestinal tract, especially the colon, is recognised as one of the largest microbial reservoirs (O'Hara and Shanahan, 2006). This microbial community fulfils essential functions in digestion, metabolic activity and immunity, and differences in species composition and abundance can therefore provide much information about the host organism. For example, following its initial acquisition during birth and lactation, the microbiome is constantly modified by factors such as age, sex, and diet (Ley *et al.*, 2008a,a; Nicholson *et al.*, 2012). Similar factors shaping

the gut microbiome in terrestrial mammals influence that of marine mammals (Nelson *et al.*, 2013b; Smith *et al.*, 2013; Pacheco-Sandoval *et al.*, 2019; Stoffel *et al.*, 2020). However, studies have also shown substantial differences between marine and terrestrial mammal gut microbiomes, even when these two groups share a similar diet (e.g., herbivore, carnivore) (Nelson *et al.*, 2013a; Bik *et al.*, 2016). Thus, even though research into the microbiome of terrestrial mammals is at a relatively advanced stage, this information cannot be easily extrapolated to marine mammals whose microbiomes remain poorly understood particularly, those in non-captive, natural populations. Consistent characterisation of the core microbiome of these populations is therefore required as a fundamental baseline before we can attempt to understand it's functions, roles, interactions, and possible uses.

The Juan Fernandez fur seal (Arctophoca philippii philippii) (JFFS) is a marine mammal endemic to the Juan Fernandez Archipelago, a group of islands located in the middle of the Pacific Ocean 600 km away from the Chilean continental coast (Fig 1). The archipelago is a hotspot for biodiversity with a high number of endemic marine species, including the JFFS (Aguayo *et al.*, 1970; Pompa *et al.*, 2011; Friedlander *et al.*, 2016). These fur seals are the only native mammals to the archipelago and like other pinnipeds occupy upper trophic levels in the marine food web (Ochoa Acuna and Francis, 1995; Trites, 2019). Their feeding behaviour, lifespan, fat storage, and their amphibian lifestyle, which links marine and coastal environments, are some of the characteristics that make this species a great candidate to act as a marine bioindicator. However, despite showing a significant population recovery since the late 1960s and becoming an icon for local tourism, little is known about this species.

This study aimed to characterise the JFFS faecal microbiome for the first time, as a baseline for understanding the host-microbial interactions in this species. To investigate, we performed sequencing of the 16S rRNA gene, a highly conserved region of the bacterial genome, which provides a reliable overview of bacterial community composition.

3.2 Results

Fur seal faecal samples (n=57) were collected across two consecutive breeding seasons from seven different breeding colonies on the JFA. The samples were stored and shipped as in the Materials and Methods section 2.3. DNA from the samples was extracted, used as template for a 16S rRNA V4 region PCR (Kozich *et al.*, 2013) and then sequenced on the MiSeq platform as in section 2.4.

Following removal of low-quality sequences and merging the 2017 and 2018 datasets, a total of 2,074,038 paired reads, grouped into 595 amplicon sequence varianta (ASVs) were imported into R studio for statistical analysis. A total of 54 samples, with 2,062,763 sequences

clustered into 558 ASVs remained after the filtering steps (Supp. Tab. A.2). Three samples were removed from the analysis due to rarefaction analysis indicating insufficient depth of sequencing (Fig. 3.1). The rarefied dataset ended up with 518 ASVs and a total of 754,974

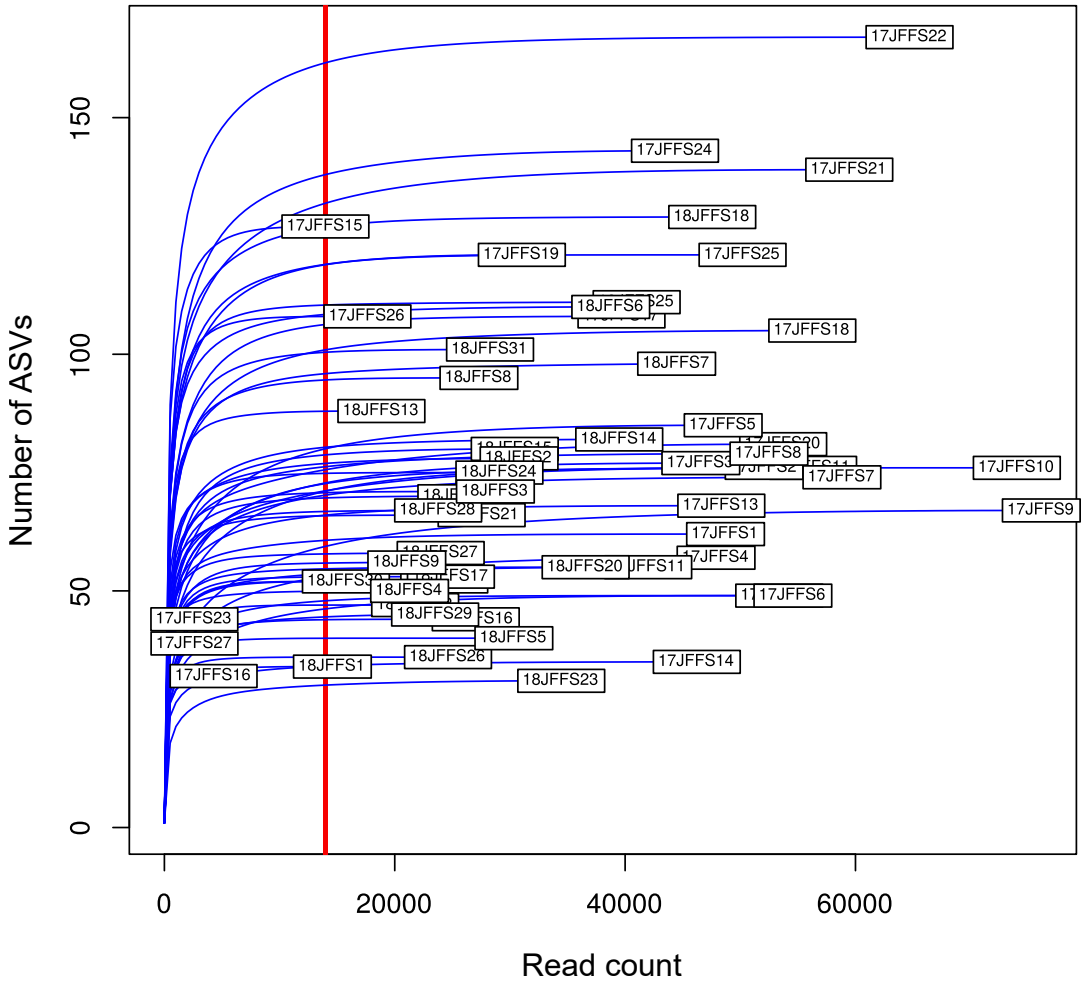

Figure 3.1. Rarefaction curve estimating the number of ASVs (y-axis) for a given read count (x-axis).
The vertical line indicates the cut-off at which samples were retained and rarefied.

reads.

3.2.1 Composition of the Juan Fernandez fur seal faecal microbiome

A total of 10 bacterial phyla were detected in the faeces of the JFFSs. From the total ASV counts *Firmicutes* (41.9 %), *Fusobacteria* (28.2 %), *Bacteroidetes* (22.1 %), *Proteobacteria* (5.5 %) and *Actinobacteria* (1.5 %) dominated the bacterial composition. The total ASV counts from individual samples were very similar to the average relative abundance: *Firmicutes* (40 %

±24), *Fusobacteria* (30 % ±17), *Bacteroidetes* (22 % ±10), *Proteobacteria* (6 % ±4) and *Actinobacteria* (2 % ±3) (Supp. Tab. A.5). Eighty-two bacterial families could be assigned, of which 14 had a relative abundance $\geq$ 1 % of the total ASV count. Five bacterial families accounted for 78.5 % of all read counts: *Fusobacteriaceae* (28.2 %) belonging to the phylum *Fusobacteria*, *Bacteroidaceae* (15.5 %) from the phylum *Bacteroidetes*, and *Ruminococcaceae* (15.0 %), *Lachnospiraceae* (10.4 %) and *Peptostreptococcaceae* (9.4 %) from the phylum *Firmicutes* (Fig. 3.2 A and B) (Supp. Tab. A.6). Forty-six ASVs were present in at least 50 % of the samples (Supp. Tab. A.7). While fourteen ASVs were present in > 90 % of samples, only three ASVs were present in all the samples, all of which were assigned to the genus *Fusobacterium* (14.9 %, 6.5 % and 3.7 % of the total reads respectively)(Tab. 3.1).

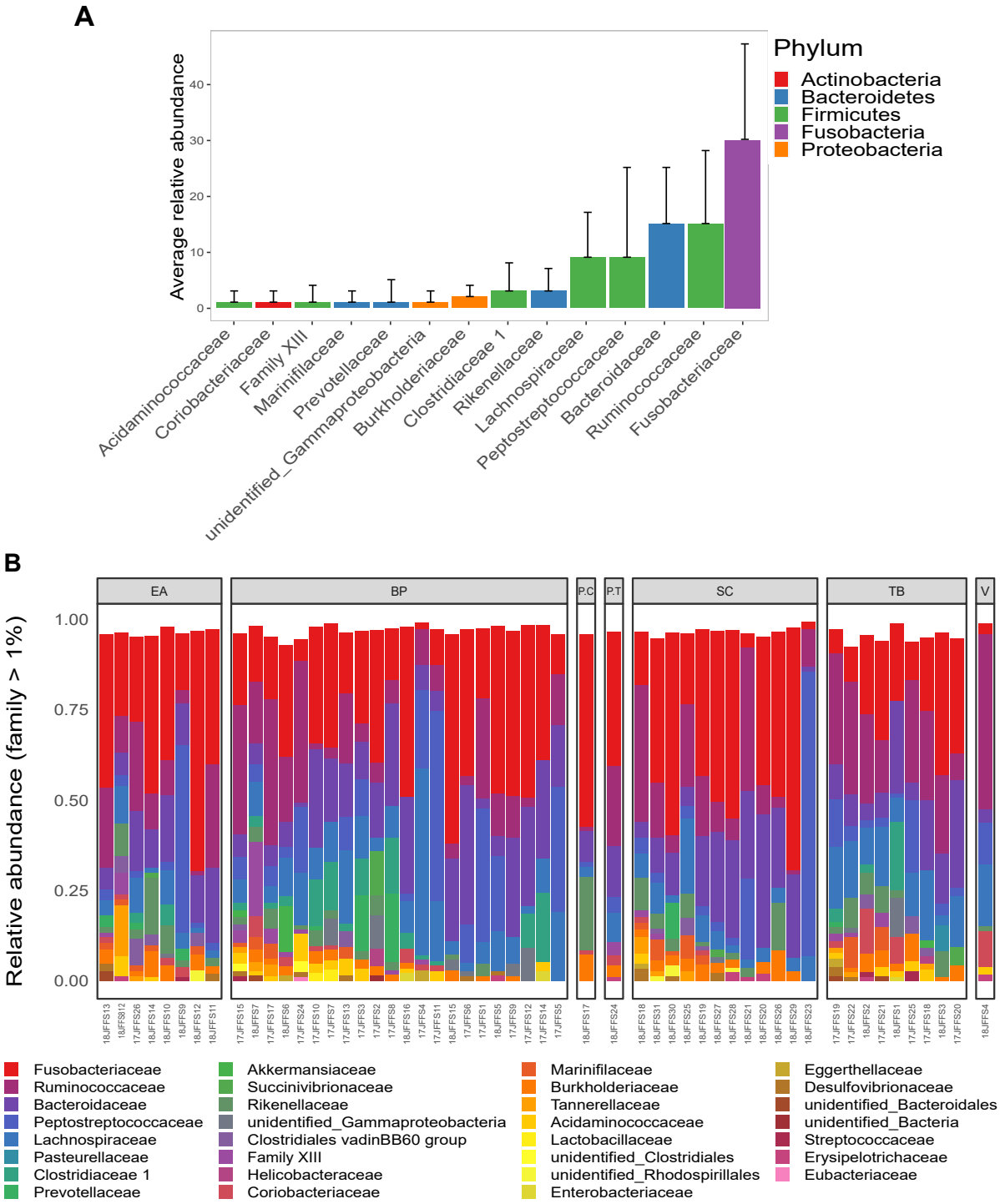

Figure 3.2. Composition of the Juan Fernandez fur seal faecal microbiome at the family level.

Only families with > 1% relative abundance are shown. A) Average relative abundance across all samples with standard deviations. B) Relative abundance per sample grouped by location: EA= El Arenal, BP= Bahia El Padre, PC = Piedra Carvajal, PT= Punta Truenos, SC= Santa Clara, TB= Tierras Blancas, V= Vaqueria.

Table 3.1. Amplicon sequence variants present in at least 90 % of the samples.
Unrarefied data was used to build this table. Abundances was calculated based on the total ASVs count.

Table 1

ASV	Phylum	Family	Genus
Present in all the samples			
57729b2b058d8d5253d3e56e4f6386ca	Fusobacteria	Fusobacteriaceae	Fusobacterium
e8b1922518029c50c69add839142db03	Fusobacteria	Fusobacteriaceae	Fusobacterium
c0dc53aad260a1b951b7f99966251c7c	Fusobacteria	Fusobacteriaceae	Fusobacterium
Present in at least 90% of the samples			
f347c63fc5e4aeb97531e656e3765e2a	Firmicutes	Peptostreptococcaceae	Peptoclostridium
57f9edc6542ce6b78ff352942d6774c6	Bacteroidetes	Bacteroidaceae	Bacteroides
31984a302fdfe46b5e852fa473e682a4	Bacteroidetes	Bacteroidaceae	Bacteroides
1153942c5cc40d6ba5609222ded586fe	Firmicutes	Lachnospiraceae	Coprococcus 3
65dd9f625700a97a1cce9f5eefe4e6cb	Firmicutes	Lachnospiraceae	Blautia
435975b6d032d4b05233d8b94193b2ad	Firmicutes	Lachnospiraceae	[Ruminococcus] gauvreauii group
03f74c0ea1f0654719b21d2701e9fa30	Proteobacteria	Burkholderiaceae	Sutterella
8e10797dedc288dbc0be61fe4b5a5dfb	Actinobacteria	Coriobacteriaceae	Collinsella

3.2.2 Alpha diversity

Three alpha diversity indices (Chao1, Shannon-Weiner and Simpson) were used to compare within-sample diversity between locations (Supp. Tab. A.8). Despite Tierras Blancas showing a trend towards higher diversity in all analyses, the one-way ANOVA results showed no significant differences between locations according to the Chao 1 index ($F(3/47) = 2.45$, $p = 0.07$, ges = 0.08) and Shannon-Weiner index ($F(3/46) = 2.65$, $p = 0.06$, ges = 0.09). The Simpson index (chi-squared = 8.26, $p < 0.05$, ges = not provided) on the other hand, showed a significant difference between locations. Post-hoc Dunnet's and Dunn's tests consistently showed that samples from TB had higher mean and mean rank values (respectively) than the other locations, especially when compared to BP (Fig. 3.3). When comparing the alpha

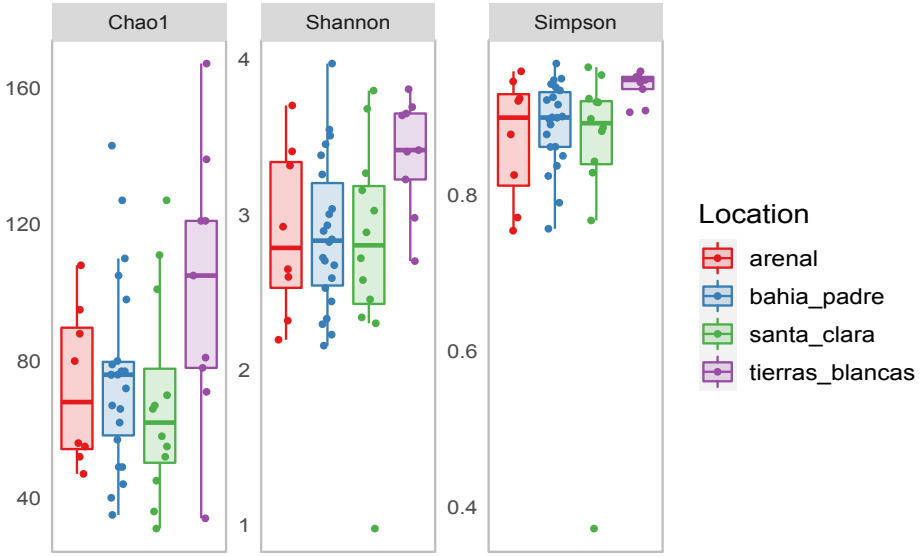

Figure 3.3. Comparison of three different alpha diversity indices between four reproductive colonies in the Juan Fernandez archipelago.
Samples collected from Tierras Blancas show a tendency to have higher levels of alpha diversity. Filtered rarefied data was used to calculate the diversity estimates.

diversity using the Jost Index (q=1), there was a significant difference in alpha diversity between locations (ANOVA, $F(3/46) = 3.66$, $p = 0.02$, ges = 0.14) after removing an identified outlier (17JFFS15). Furthermore, according to Dunnet's test, TB showed significantly higher alpha diversity when compared to BP ($p < 0.01$) and SC ($p = 0.03$) at 95 % confidence level).

3.2.3 Beta diversity

Based on weighted Unifrac dissimilarity distance as a measure of between sample diversity, 51.0 % (full dataset) and 53.8 % (core dataset) of the total variation between samples could be explained by the first principal component (PC1). No clustering of individual samples by location or year of collection was observed. Similarly, Bray-Curtis dissimilarity, which quantifies the differences in ASV abundance, found that the first principal components in both the full and core datasets explained 23.9 % and 29.8 % of the total variation respectively. In both data sets, a group of samples (cluster 2) were clearly separated from the main cluster (cluster 1) along PC1 (Fig. 3.4).

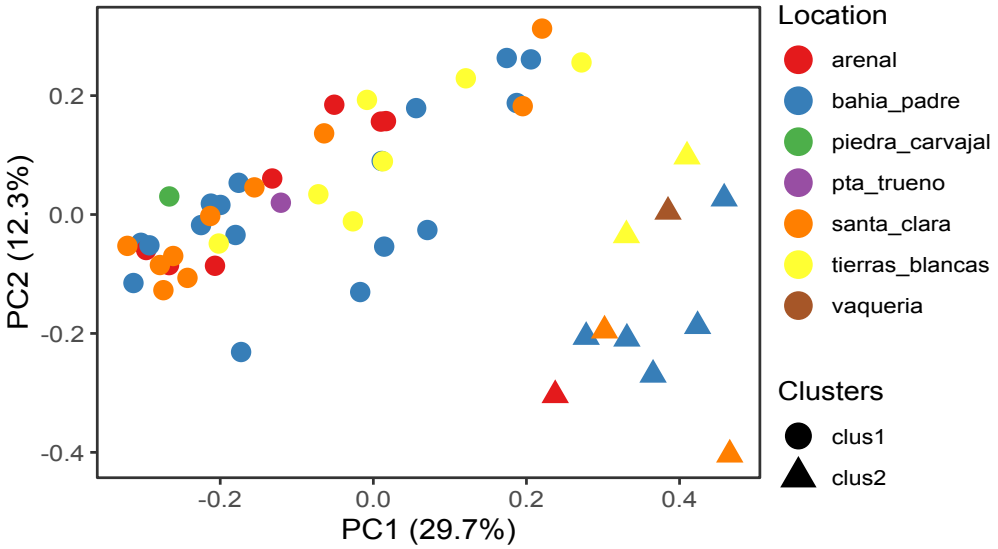

Figure 3.4. PCoA using Bray-Curtis dissimilarity distance matrix using the filtered rarefied core dataset.

Based on the relative average abundance of the dominant phyla, evident differences in the overall microbial composition were visualised between the two clusters (Fig. 3.5). PERMANOVA evidenced a significant difference in the microbial composition between the two clusters. This was consistent in both full (F= 10.1, Pr (>F) = 0.001, R^2 = 16.3 %) and core datasets (F= 13.6, Pr (>F) = 0.001, R^2 = 20.88 %). SIMPER analysis identified five genera that together contributed 71 % to the observed compositional difference between the clusters. As expected, both *Fusobacterium* and *Peptoclostridium* were the largest contributors (24 and 25 % respectively). Furthermore, the abundance of *Fusobacterium* and *Peptoclostridium* were significantly different between clusters. Full results of the SIMPER and Mann-Whitney U-tests are summarised in Tab. 3.2.

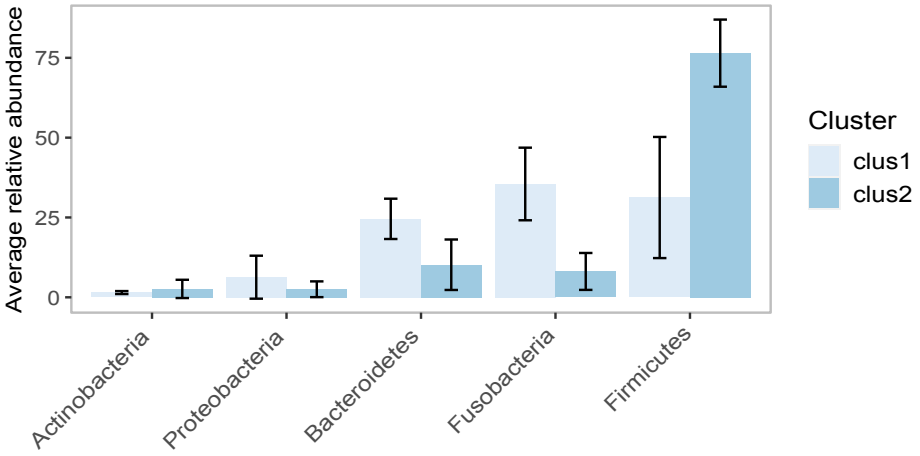

Figure 3.5. Relative average abundance of the dominant phyla according to the clusters identified with Bray Curtis dissimilarity.
Showing only phyla with an average relative abundance $\geq 1\%$. The differences in microbial patterns can be identified from high taxonomic levels.

Table 3.2. SIMPER analysis comparing the faecal microbiota composition of Juan Fernadez fur seal at the genus level. Cluster averages were calculated based on total counts. Kruskall-Wallis results are only shown when reaching significant difference.

Genus	av_cluster 1	av_cluster 2	Av.Diss	Contrib %	cum%	w	P-value
Peptoclostridium	3%	29%	17%	25	25	3	<0.001
Fusobacterium	34%	8%	17%	24	49	456	<0.001
Bacteroides	14%	6%	7%	10	59	365.5	0.006
Ruminococcaceae UCG-005	4%	7%	4%	6	65		No sig
[Ruminococcus] gauvreauii group	1%	6%	4%	5	70	124	0.06

As mentioned before, none of the PCoA evidenced any clustering associated to location. On the other hand, when perfoming a redundancy analysis (RA), samples clearly clustered according to the location they were collected from (Fig. 3.6). Following a similar pattern to

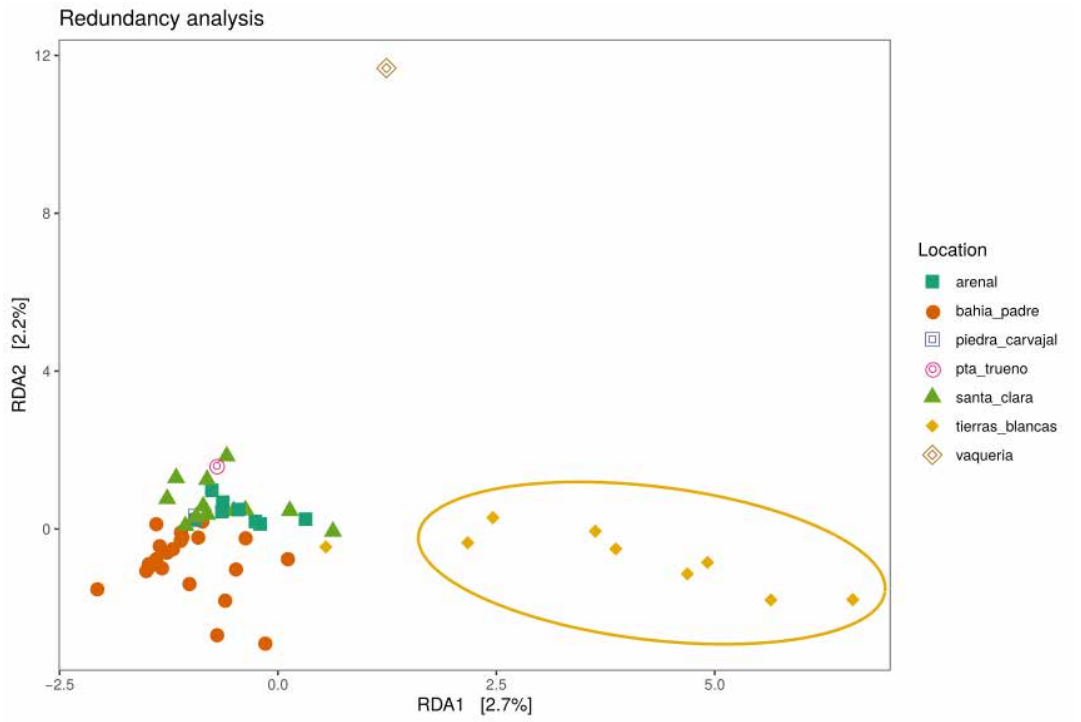

Figure 3.6. Redundancy analysis (RDA) comparing microbial composition of faecal samples collected in different locations.
Only phyla with an average relative abundance $\geq 1\%$ are shown.

the alpha diversity, Tierras Blancas was the most differentiated cluster. However, location only explained a small portion of the variance; 5.2 %, when combining the variation of the first two components.

3.2.4 Correlation analysis

Spearman correlation analysis revealed that the genera *Bacteroides*, *Fusobacterium* and *Peptoclostridium* were strong drivers of PC1 in both Bray-Curtis and Weighted Unifrac PCoA analyses. In addition, the genera *Ruminoclostridium* 9 and *Ruminococcaceae* NK4A214 were also found to be influential for PC1 in Bray-Curtis analysis (Fig. 3.7) (Supp. Tab. A.9). PCoA analyses showed strong negative correlations between PC1 and *Bacteroides* (Bray-Curtis, ρ = -0.67, p $\leq$ 0.001); and between PC1 and *Fusobacterium* (Bray-Curtis, ρ = -0.92, p $\leq$ 0.001 and weighted Unifrac, ρ = -0.94, p $\leq$ 0.001). *Peptoclostridium*, on the other hand, was positively

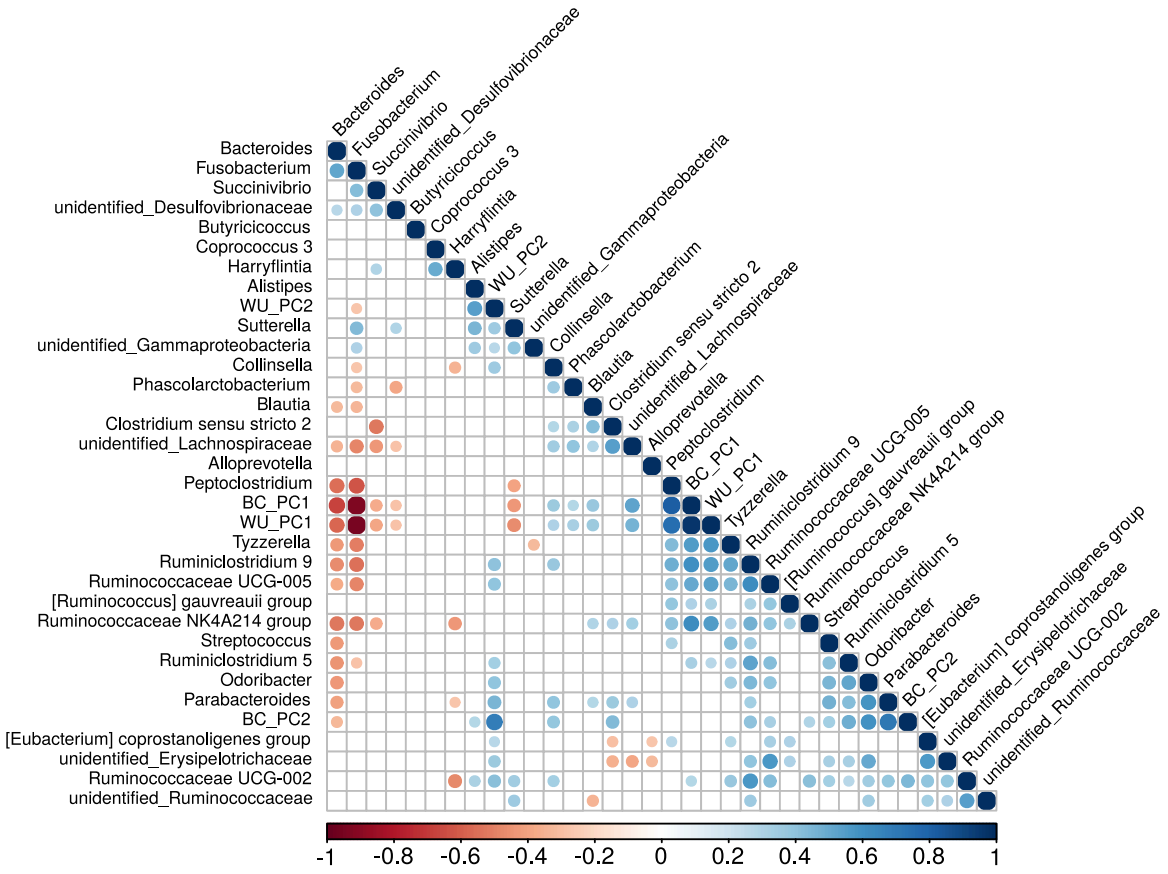

Figure 3.7. Spearman rank correlation correlogram between bacterial genera and the first two principal components generated from Unifraq and Bray-Curtis distances. The plot shows the direction (blue = positive, red = negative) and the strength (larger = stronger) of the correlation between each pair combination. Only significant correlations ($p \leq 0.05$) are represented with circles. WU_PC and BC_PC are the Weighted Unifraq and Bray-Curtis Principal Component distances respectively.

correlated with PC1 (Bray-Curtis, ρ = 0.81, p $\leq$ 0.001, and weighted Unifrac, ρ = -0.75, p $\leq$ 0.001).

3.2.5 Funtional prediction

Having identified the presence of different bacterial families in the JFFS faecal microbiome, I wanted to investigate what differences in function these bacteria may have in the two clusters. To look at the quality of the predictive functional analysis performed with PICRUSt2 , I first explored the Nearest Taxon Index (NSTI). Overall, the mean NSTI value per sample was 0.12 $\pm$ 0.03. Per location, the mean NSTI value calculated for cluster 1 was 0.12 $\pm$ 0.03 and for cluster 2 was 0.14 $\pm$ 0.03. As expected, these values were within the ranges observed in less characterised mammals (Langille *et al.*, 2013).

MetaCyc pathway prediction

Using PICRUST, at the top level, six MetaCyc pathways were predicted to be present in the JFFS faecal microbiome. These were dominated by biosynthesis (76 %), degradation / utilisation / assimilation (12 %), and generation of precursor metabolite and energy (10 %). From fifty-five pathways predicted at the second level, twenty-one had a relative abundance above 1 %. The most dominant of these pathways were nucleoside and nucleotide biosynthesis (mean 19.54 %), amino acid biosynthesis (mean 18.02 %), cofactor, prosthetic group, electron carrier and vitamin biosynthesis (mean 14.84 %) and, fatty acids and lipid biosynthesis (mean 7.33 %). Finally, at the lowest level, I performed a differential abundance analysis using ALDEx2. From 359 pathways inferred at the lowest level, only fifteen of them remained significantly different after the Benjamini-Hochberg correction of the p-value, where the effect size was larger than one, and the median difference was at least one fold change (Fig. 3.8 A). From these pathways, twelve were classified at the top level as Biosynthesis, two as Generation of Precursor Metabolite and Energy and one as Degradation/Utilisation/Assimilation. Only three pathways were predicted to be significantly more abundant in Cluster 2; two are involved in amino acid biosynthesis. The third one, Peptidoglycan Biosynthesis IV, is involved in cell structure Biosynthesis. Four out of the twelve significantly abundant pathways predicted in Cluster 1 were associated with Fatty Acid and Lipid Biosynthesis.

KEGG pathway prediction

As observed with MetaCyc database, six KEGG Pathways were predicted to be present in the JFFS faecal microbiome. Furthermore, 47 % of the predicted genes were classified within metabolism pathways, 22 % in genetic information processing, 13 % in environmental

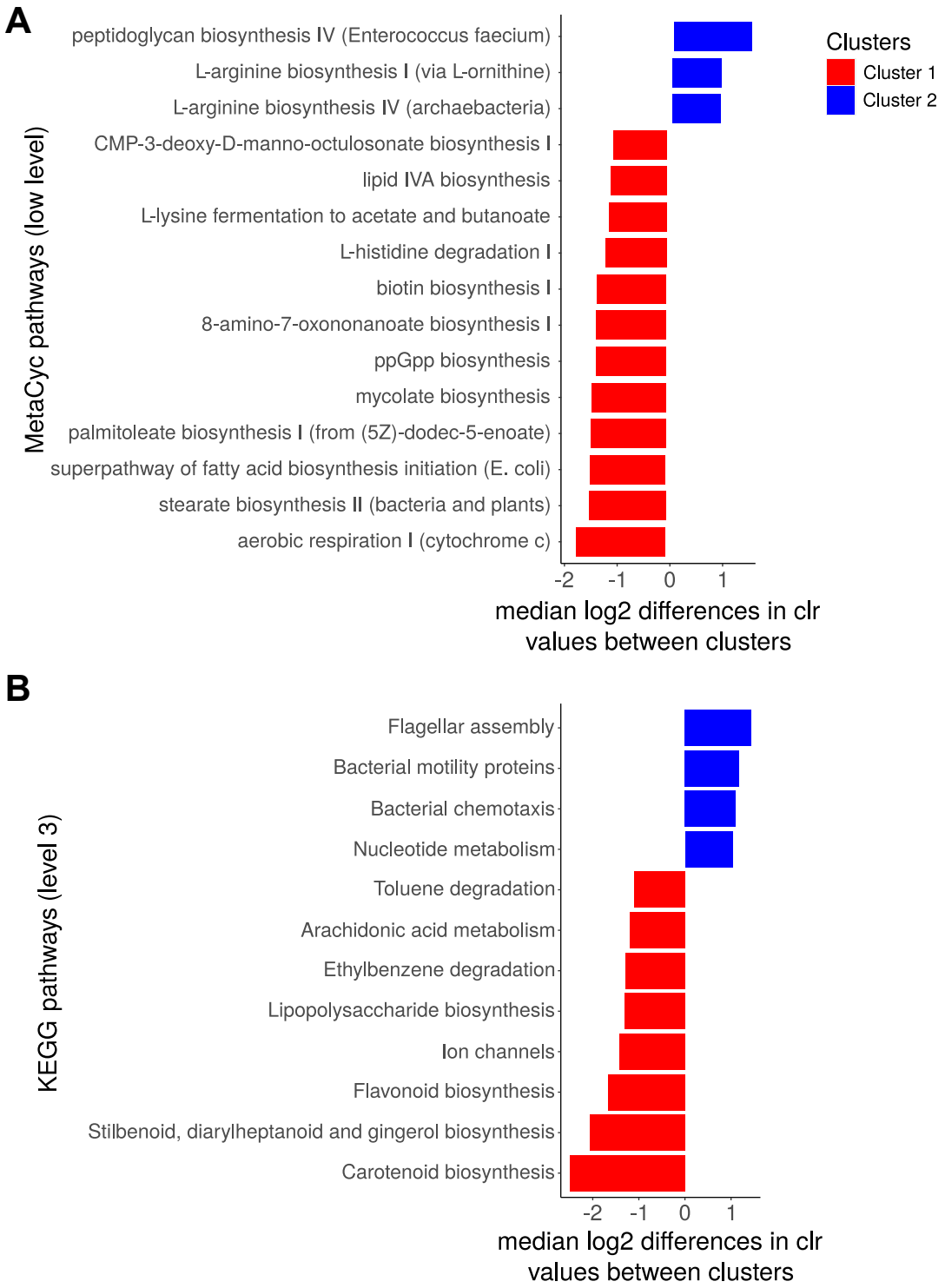

Figure 3.8. Divergent plot showing significant pathways identified in the JFFS faecal microbiome based on Metacyc pathways (A) and KEGGS pathways (B).
Only pathways with Benjamini-Hochberg corrected $p < 0.05$, effect size larger than one and with median difference larger than one fold change were considered significant.

information processing, 3 % in cellular processes, and 14 % were unclassified at Level 1. From thirty-nine pathways predicted at the second level, twenty-three had a relative abundance above 1 %. Six pathways at level 2 accounted for 50 % of the predictions; membrane transport (11.51 %), replication and repair (10.04 %), carbohydrate metabolism (9.98 %), amino acid metabolism (8.99), translation (6.48 %) and energy metabolism (5.71 %). ALDEx2 analysis predicted 259 KEGG functional outputs at level 1. Only twelve of these predictions were identified as significant features between clusters (Fig. 3.8 B). I used the same criteria applied for the MetaCyc differential abundance analysis to identify significant features. The most interesting results from this analysis were the significantly increased pathways classified as; xenobiotics biodegradation and metabolism, lipid metabolism, and metabolism of terpenoids and polyketides.

3.3 Discussion

Marine mammal microbiome studies of free-ranging wild populations are rare, with many of these studies being limited to a small number of individuals. Instead, most studies of marine mammals have relied on data from dead or captive animals. To our knowledge, this is one of the most extensive studies of the faecal microbiome in free-ranging pinnipeds and the first of JFFS. This approach focused on characterising the core members of the JFFS faecal microbiome, identified at the genus level, providing a baseline for understanding host-microbial interactions in this species. However, interpreting unexpected phenomena in a dataset such as ours is made difficult by a lack of literature with results generated using similar methodologies, as well as the various uncontrollable factors influencing wild populations.

3.3.1 Methodology

Because results often vary depending on the applied methodology, I thought it was relevant to discuss the reasoning behind my methodology and statistical analysis to give context to the results I obtained. In this study, I used 16S rRNA gene amplicon sequencing as a first approach to characterise the JFFS microbiome. This method aims to amplify conserved regions of the 16S rRNA gene present in most bacteria; in this case, I targeted the V4 region. The extensively curated databases, low bioinformatic requirements and affordability are some of the reasons why I believe this sequencing method is a good first approach to studying uncharacterised microbiomes of poorly studied hosts such as the JFFS. However, it is necessary to mention that amplicon sequencing has some important limitations (Janda and Abbott, 2007; Ranjan *et al.*, 2016; Durazzi *et al.*, 2021). For instance, only bacteria and archaea can be identified and at

a rather low resolution (genus level). Furthermore, the accuracy of the predicted taxonomic profiles depends on how well curated the reference database is, which algorithm was used to de-noise the sequences as well as the identity threshold that is used for clustering these sequences (e.g. 97 %, 99 %, 100 %) (Balvočiūtė and Huson, 2017; Edgar, 2018).

Here, I chose to use a 100 % identity clustering, also known as ASV, which is the output of the DADA2 de-noising algorithm. There are various benefits to using ASVs instead of operational taxonomic units (OTU), which are clusters based on similarity thresholds usually between 97 and 99 %. Using ASVs has been shown to increase the power, resolution and reproducibility of the analysis while also improving the identification of rare taxa (Callahan *et al.*, 2017; Nearing *et al.*, 2018). Furthermore, a recent study showed that methods using ASVs are better at distinguishing the true community from contamination (Caruso *et al.*, 2019). It is crucial, however, to mention that this method is not free from limitations. For instance, it may overestimate the abundance of those taxa carrying genomes containing multiple gene copy numbers. And, even though more accurate than OTUs, using ASVs does not guarantee ecological coherence because the taxonomic predictions are still done from a very short genetic fragment (Callahan *et al.*, 2017).

After processing the raw sequences and building the final feature table, the next challenge is analysing the data. Choosing a method to analyse microbiome data is not a straightforward process. There are many suggested methods and little agreement on what should be the best approach to interpret 16S rRNA gene sequencing outputs correctly (McMurdie and Holmes, 2014; Gloor *et al.*, 2017; Weiss *et al.*, 2017; Xia and Sun, 2017; Luz Calle, 2019; McKnight *et al.*, 2019). After carefully considering all the options, I decided to take a mixed approach adapted to the structure and characteristics of my data.

Filtering is a relevant first step that aims to control for spurious taxa from the data set. Usually, the filtering of ASVs is focused on discarding unclassified features and those with low abundance and prevalence. Even though this may limit the detrimental effects of contamination, chimaeras and sequencing errors, there is an unwanted effect of eliminating rare taxa that are genuinely part of the microbial community. On this aspect, DADA2 applies a powerful error model the can accurately identify and exclude features indicative of chimaeras and sequencing errors (Callahan *et al.*, 2016). Thus, the de-noising algorithm used by the software effectively reduces the probability of rare ASVs resulting from these issues (false positives). Because I relied on DADA2, I chose not to fully focus on low abundance or prevalence to direct the filtering step.

Regarding contamination, faecal samples are known to have large biomass, which naturally reduces the noise introduced by sequences originating from contamination (e.g. reagents). To deal with the contamination problem, I took an approach that uses Spearman correlation at

the genus level to identify patterns consistent with contamination. This identification method is also informed by genera previously described as contaminants (Salter *et al.*, 2014). With this method, I was able to identify and exclude a few contaminant genera which, as expected, were in very low prevalence/abundance. Finally, I had to deal with those ASVs which were not classified at a higher taxonomic rank (e.g. phylum level). To avoid excluding features that remained unclassified because they have not been described before (expected from poorly characterised microbiomes), I blasted all the unclassified features and discarded only those aligning to non-bacterial organisms (e.g. fur seal mitochondrial DNA) (Calvo *et al.*, 2018).

The final topic I would like to discuss regarding the methodology is my statistical approach. To initiate this discussion it is important to understand the nature of microbial data:

1) It is made from compositional discrete counts (Gloor *et al.*, 2017),

2) often there are large difference in library sizes even between samples processed at the same time, and

3) it contains many zeros, which complicates data transformations.

The challenges of working with such complex datasets have encouraged the search and development of statistical methods to explain better and more accurately the behaviour of the data (Gloor *et al.*, 2017; Xia and Sun, 2017; Luz Calle, 2019). However, as expected, choosing a method will depend on the data and the scientific questions.

Normalising the data is the first point of contention. There are many methods suggested to deal with differences in read depth (McKnight *et al.*, 2019). Here I will only focus on defending my decision of using rarefaction as my chosen normalisation method. Briefly, rarefying involves subsampling each sample down to a shared threshold, usually informed by a rarefaction curve, so that the number of reads is equal across all the samples. In recent years, this method sparked concerns from some authors. Some of the criticisms include discarding valid data, introducing uncertainty and the arbitrary selection of the library size (McMurdie and Holmes, 2014; Weiss *et al.*, 2017; Willis, 2019). However, many of the arguments against this method are made in the context of differential abundance analysis, an analysis I did not perform at the taxonomic level in this study.

Furthermore, there is evidence that rarefaction outperformed other normalisation strategies when performing diversity analysis (McKnight *et al.*, 2019; Cameron *et al.*, 2021). My study was mostly focused on exploring alpha and beta diversity to answer three questions: *Who is there?; Is there any evident patterns?; Which features are more likely to explain these patterns?* Thus, I concluded that rarefying was the best normalisation approach to explore the beta diversity. I do however, acknowledge the risks of unnecessarily discarding too much data

by selecting an inadequate library size in order to retain samples. I established the library size threshold based on the information provided by the rarefaction curve. Instead of setting a low threshold prioritising sample retention, I chose a point at which all the samples had reached a plateau and discarded three samples that were still in the exponential part of the curve.

3.3.2 Overall microbiome characterisation

Consistent with previous reports in other pinniped species, five phyla dominated the JFFS faecal microbiome: *Firmicutes*, *Fusobacteria*, *Bacteroidetes*, *Proteobacteria* and *Actinobacteria* (Nelson *et al.*, 2013b,a; Bik *et al.*, 2016; Numberger *et al.*, 2016; Pacheco-Sandoval *et al.*, 2019; Kim *et al.*, 2020; Stoffel *et al.*, 2020). When comparing this result to other southern pinnipeds, the most different microbial patterns were found in faeces from other fur seal species (Smith *et al.*, 2013; Medeiros *et al.*, 2016). The faecal microbiome described for both the South American (*Arctophoca australis australis*) and the subantarctic fur seals (*Arctophoca tropicalis*) was almost entirely dominated by *Firmicutes* (88.56 % and 85.02 %). *Fusobacteria*, on the other hand, represented less than 1 % of the bacterial community for both species (Medeiros *et al.*, 2016). The study involving these two species collected samples from juvenile individuals found dead. Thus, it is expected to find altered microbiomes. Smith *et al.* (2013) characterised the faecal microbiome of Australian fur seal (*Arctocephalus pusillus doriferus*) pups and female adults. The adult samples showed similar proportions of *Firmicutes*, *Bacteriodetes* and *Actinobacteria* as those observed for JFFS in my study. *Fusobacteria* was not detected in any of the adults. However, the authors only relied on fluorescent in situ hybridisation (FISH) for this age group to identify bacterial phyla.

Overall, pinniped gut microbiomes are very variable between and within species, possibly due to differences in their geographic range (e.g. polar versus subtropical), diet (benthic vs pelagic hunters, generalist versus specialist), or mating systems. One or more of *Fusobacteria*, *Firmicutes* and *Bacteroides* (all three in the case of JFFS and harbour seals), have been found to consistently dominate the overall microbial composition of pinnipeds, followed by *Proteobacteria* and *Actinobacteria* (Nelson *et al.*, 2013b; Pacheco-Sandoval *et al.*, 2019). The latter two are usually at lower abundance and *Actinobacteria*, in particular, has not been described in every pinniped species studied. Another interesting observation, common to all the studies reviewed, including this one, is that when *Firmicutes* dominates, the abundance of *Fusobacteria* and *Bacteroidetes* decreases, suggesting some degree of competition. The *Firmicutes*: *Bacteroidetes* ratio has been well documented in human and mice. In these land mammals, the ratio increases in response to diets high in lipids and decreases in response to large amounts of protein (Turnbaugh *et al.*, 2006; Hildebrandt *et al.*, 2009; Pu *et al.*, 2016). I

also observed changes in the relative abundance of *Fusobacteria* were similar to those observed in *Bacteroidetes*. This suggests some functionally redundant roles.

The phylum *Firmicutes* is common in mammalian gut microbiomes (Ley *et al.*, 2008a,b). Members of this taxonomic group are well known for their role in obesity in humans and mice, which is associated with an increase in *Firmicutes* and a decrease in *Bacteroidetes* (Turnbaugh *et al.*, 2006; Hildebrandt *et al.*, 2009; Pu *et al.*, 2016). The energy harvesting role of *Firmicutes* has also been identified in the zebrafish gut microbiome, where these bacteria are associated with an increase in lipid droplet numbers in epithelial cells (Semova *et al.*, 2012). Fat is fundamental for marine mammal survival, as it is needed for energy storage and thermoregulation and may explain why *Firmicutes* is consistently among the most dominant phyla across all pinniped species (Guerrero and Rogers, 2019).

The phylum *Fusobacteria* consists of facultative or strict anaerobes that produce various organic acids from amino acid or carbohydrate fermentation (Olsen, 2014). This phylum is usually found at high relative abundance in the gut microbiomes of strict carnivores adapted to diets rich in proteins, purines and poly-unsaturated fatty acids (Zhu *et al.*, 2018; Guo *et al.*, 2020). Similar to other marine carnivores, *Fusobacteria* was one of the most abundant phyla in JFFS (Pacheco-Sandoval *et al.*, 2019). Most of the knowledge generated around the specific role *Fusobacteria* may play in mammalian intestinal tracts is based on human-centred research. Even though some genus members seem to play a beneficial role in the human gut microbiome, the presence of relatively high levels of the genus *Fusobacterium* is more often associated with health issues (Potrykus *et al.*, 2008; Garrett and Onderdonk, 2014; Huh and Roh, 2020). Conversely, the high relative abundance of this bacterial genus in the gut of carnivores suggests a rather symbiotic relationship where *Fusobacterium* is likely to play a role in protein metabolism (Potrykus *et al.*, 2008).

Similar to *Fusobacteria*, the phylum *Bacteroidetes*, especially members of the genus *Bacteroides*, are associated with diets high in animal proteins (Zhu *et al.*, 2018; Guo *et al.*, 2020). This genus, known for its capacity to degrade animal-derived glycans (Eilam *et al.*, 2014), was the most abundant *Bacteroidetes*. Similar to previous reports, JFFS samples high in *Firmicutes* contained lower relative abundances of *Bacteriodetes* and *Fusobacteria*. This phenomenon suggests differences in dietary contituents and will be discussed later in the text.

3.3.3 Within sample diversity

Initially, I hypothesised that the alpha diversity of samples collected from BP, a key access point to Robinson Crusoe island, was going to be different from other colonies. BP is the most transited area in this study; it connects the airfield with the town and is a popular leisure location for the local community (Fig. 2.1). I found instead that BP did not differ from other

less-visited locations such as EA and SC. Therefore, this finding is different to a previous report showing an association between exposure to anthropogenic stressors and reduced alpha diversity in harbour seals (Pacheco-Sandoval *et al.*, 2019). The colony at TB was the only location with significantly higher alpha diversity, indicating that samples collected from TB had a richer and more evenly distributed microbial composition than other samples.

Despite the trend showing how TB differed from the other locations, two of the three alpha diversity estimates did not show statistical significance (Chao 1 richness and Shannon-Weiner diversity index). ANOVA, the statistical analysis used to compare locations based on these two estimates, is sensitive to differences in sample size resulting in reduced statistical power. The sample size effect could, therefore, partially explain the lack of statistical significance. It is also important to highlight the limitations of the diversity estimates used in this study. Chao1, Shannon-Weiner index and Simpson index are not be confused with diversity (Jost, 2006; Chao *et al.*, 2010). Instead, they provide a rough estimation of diversity and biases need to be considered. Richness, for instance, is particularly sensitive to differences in both library and sample size, meaning that larger library and sample sizes will result in more rare taxa. Although more robust, the Shannon-Weiner index is also largely influenced by rare taxa, which translates into similar sensitivities.

Additionally, I decided to explore alpha diversity based on true diversity. Transforming diversity indices into effective numbers (true diversity) is an uncommon practice in microbiome-related studies. However, there are many benefits to doing this transformation. For instance, the number of species becomes the standard metric unit across all indices, facilitating comparisons. It also deals with the non-linearity of indices such as Shannon-Weiner and Simpson (Jost, 2006; Chao *et al.*, 2010). Here I looked at effective numbers of order q = 1 given by the Shannon-Weiner index. The pattern I observed with this approach replicated what I showed earlier, where TB, once more, showed a significant difference when compared to the other locations, SC and BP specifically. Using the effective numbers also improved the statistical outcome, with the model now explaining up to 14 % of the variance.

Bacterial richness has been previously associated with population density due to the increase in microbial sharing (Li *et al.*, 2016). Alternative studies have suggested that overcrowding might also negatively affect microbial diversity due to higher levels of stress (Kozich *et al.*, 2013; Bharwani *et al.*, 2016). Lower diversity in denser fur seal populations was also observed in *Arctocephalus gazella*, a closely related species (Grosser *et al.*, 2019). I believe it is reasonable to say that the effects of density in microbial diversity are likely to depend on the social structure of the host species. It makes sense that species with complex social behaviour are likely to benefit from higher densities as family bonding will also contribute to decreasing stress levels and increasing overall well-being. On the other hand, species such as fur seals do

not seem to rely on intricate social connections between the colony members. Despite fur seals being classified as a gregarious species, it may be that colonies are formed based on domination rather than social bonds.

Population density of JFFS and its effects on the microbiome has not been studied. However, superficial observations from the field did not suggest differences in population density between the colonies. It may therefore be that other stressors were limiting alpha diversity in the other locations. For instance, the colony on TB was relatively sheltered compared to the other colonies, as it was situated on an open platform a few meters above sea level; in contrast, the other colonies were on narrow strips of land with greater exposure to sea storms, rockfalls and landslides. Additionally, the colony on TB is rarely visited by humans due to the complicated access. However, the effects of location on alpha diversity were marginal. Nevertheless, the stress hypothesis could be tested in future studies by measuring markers of stress (e.g. cortisol) in the faeces (Wasser *et al.*, 2000).

3.3.4 Variation between samples

Based on the unconstrained analyses I performed to explore the variation between samples, there was no evident clustering based on location. On the other hand, when I carried out the constrained analysis, RA, the samples grouped according to this variable, suggesting that colony location may be influencing the microbial composition in the JFFS faeces. However, this variable explained only a small portion of the total variability (5.2 %). Interestingly, the patterns were similar to those observed in the alpha diversity analysis, where samples collected from TB seemed to cluster further from the other locations on the y-axis, explaining most of the variability captured by RA (3 %). Samples collected from the other colonies grouped more closely but remained discernible; this was especially true for BP. Based on the consistent patterns identified by the RA and the alpha diversity analysis, the observed differences between locations are robust. However, this variable does not have much explanatory power on the overall microbial community composition. This lack of explanatory power makes sense considering that other factors such as diet have a larger influence on the microbial communities. Perhaps, these results may reflect the difference between the most and least anthropogenic location. In future work, it would be interesting to explore the possible factors influencing the location difference identified in this study. It might reveal signatures that can contribute to identifying effects links to anthropogenic factors, especially in BP.

The Bray-Curtis dissimilarity PCoA revealed two distinct clusters. Seventy-five per cent of the samples clustered together in what we named cluster 1. The remaining samples were grouped as cluster 2. This variation between clusters was mostly explained by the differences in the relative abundance of the genera *Fusobacterium* and *Peptoclostridium*. Samples in cluster

1 had a high relative abundance of *Fusobacterium* and very low *Peptoclostridium* relative abundance, whilst samples in cluster 2 showed the opposite pattern: increased *Peptoclostridium* and a significant drop in *Fusobacterium* relative abundance. To our knowledge, this is the first time the genus *Peptoclostridium* (phylum *Firmicutes*, class *Clostridia*) has been reported in a pinniped gut microbiome. The family *Peptostreptococcaceae*, to which *Peptoclostridium* belongs, has been reported in previous studies, but representing no more than 8 % of the total composition, and more often less than 4 % (Nelson *et al.*, 2013b; Delport *et al.*, 2016; Pacheco-Sandoval *et al.*, 2019). On average, *Peptoclostridium* represented 29 % of the microbial composition observed in Cluster 2 versus the average 3 % observed in Cluster 1.

The genus *Peptoclostridium* was initially proposed in 2013 and validated in 2016 (Galperin *et al.*, 2016). This poorly characterised taxonomic group is believed to metabolize amino acids and oligopeptides and has been isolated from both wastewater-mud and marine sediments (Galperin *et al.*, 2016). The SILVA 132 taxonomy reference database used in this study included 144 members in the *Peptoclostridium* clade from which only 11 were classified within the four known species of this genus (*P. litorale*, *P. acidaminophilum*, *P. paradoxum* and *P. thermoalcaliphilum*). The remaining clade members were classified as uncultured bacteria. It should be noted that depending on the taxonomic reference database used, the taxonomic classification regarding members of the genus *Peptoclostridium* may differ between studies. For instance, some studies may refer to species such as *Clostridiodes difficile* (previously known as *Clostridium*) as *Peptoclostridium difficile* (Pereira *et al.*, 2016). All four species included in the SILVA 132 database have been isolated from environments with little or no oxygen (Galperin *et al.*, 2016). Despite these species being linked to environmental samples, *Peptoclostridium* was found in at least 90 % of the samples. The particular condition required for this bacterial species to thrive makes it unlikely that the *Peptoclostridium* members found in JFFS faeces originated from sample contamination by surrounding environmental bacteria. Such high prevalence may be a sign of a deeper relationship between these uncharacterised bacteria and the host.

The microbiome is constantly reshaping through an individual's lifetime. Most of the changes occur within symbiotic margins responding to factors such as diet, reproductive state and age, but some changes may also result in dysbiosis and disease (Ley *et al.*, 2008b; Nicholson *et al.*, 2012). Despite the limited information available on free-range pinnipeds, a few hypotheses may be suggested to explain the significant changes observed between the two clusters reported in our study.

There is evidence that the mammalian gut microbiota changes over time. This difference is particularly evident between suckling and post-weaning stages, possibly due to dietary changes (milk vs solids). As discussed earlier, *Firmicutes* are known for their capacity to regulate

lipid absorption (Semova *et al.*, 2012). Juan Fernandez fur seal milk composition contains a higher proportion of lipids in comparison to many pinnipeds (∼41 %) (Ochoa-Acuna *et al.*, 1999). Thus, if the faecal samples from Cluster 2 were collected from pre-weaning pups (7-10 months old), it may be expected that a higher relative abundance of members of the phylum *Firmicutes* would be found. Similar to the microbial pattern observed in Cluster 2, samples analysed from Australian fur seals were dominated by the class *Clostridia* in six and nine month old pups (Smith *et al.*, 2013). In the same study, the families *Lachnospiraceae* and *Ruminococcaceae* were the most dominant family within this Class, while the overall relative abundance of *Peptostreptococcaceae* was less than 4 %. Despite age (pre-weaning diet) being a reasonable explanation for the difference observed in our dataset, this hypothesis arrives with a critical bias. Samples were collected between February and March, and at this point, pups would be no older than four months. At this stage, pup faeces are still distinguishable from older individuals by colour and consistency. Individuals from the previous reproductive season would be older than a year and milk would no longer form a part of their diet. This suggests that pre-weaning diet is not the explanation for the abundance of *Peptoclostridium*.

Differences between genders may also be an explanation of the difference in samples. In general, otariids and some phocids such as northern and southern elephant seals, exhibit an important degree of sexual size dimorphism (Ralls and Mesnick, 2009). Gender differences in foraging behaviour and prey selection have also been reported for various species including the JFFS (Ochoa Acuna and Francis, 1995; Lewis *et al.*, 2006; Andersen *et al.*, 2013). Based on the differences in diets, it is not surprising to find studies in gut microbial composition also showing gender-based differences. Samples collected from adult Southern elephant seals evidenced significant differences between adult males and females (Nelson *et al.*, 2013b; Kim *et al.*, 2020). The same studies did not find differences in leopard or Weddel seals, less sexually dimorphic phocids. Adult southern elephant seal females showed a significantly higher relative abundance of *Firmicutes* and less *Fusobacteria* and *Bacteroidetes* than males (Nelson *et al.*, 2013b; Kim *et al.*, 2020). The proportional changes are very similar to those observed between cluster 1 and 2 here. Cluster 2 shows patterns similar to those observed in females. It seems that the microbial community diverges early in life based on gender as reported in northern elephant seal pups under naturally controlled diet (Stoffel *et al.*, 2020). Sexual dimorphism is a common mating strategy in otariids. Thus, it is possible that otariids such as JFFS, show similar differences as the ones observed in elephant seals. This hypothesis could be confirmed by using molecular methods for gender identification.

A commonality between the gender and age hypotheses is their relationship to the diet. Differences in diet have been identified as one of the main drivers of gut microbiome diversity (Ley *et al.*, 2008b; Nelson *et al.*, 2013b; Nishida and Ochman, 2018). While pups rely

on lipid-rich milk, fish from the family *Myctophidae* are the most important prey of adult female JFFS (Francis *et al.*, 1998). Myctophids are known to be rich in fatty acids (Lea *et al.*, 2002; Baby *et al.*, 2014). Pacheco-Sandoval *et al.* (2019) showed that harbour seal faecal samples containing more lipid-rich preys had a much higher abundance of *Firmicutes* and lower *Fusobacteria* and *Bacteriodetes*. Molecular identification of prey species in faecal samples, may therefore help to determine whether diet is the driving factor behind the microbial differences observed here.

3.3.5 Functional prediction

Finally, the 16S gene amplicon approach is excellent to answer the question – Who is there? However, it does not provide any information about the functionality of that microbiome. To explore functionality, it is best to use a shotgun sequencing approach instead. However, this more advanced genomic method is often prohibitive for research groups with low resources, which is often the case for groups working on wild populations.

In recent years, new bioinformatics tools capable of predicting gene families from 16S rRNA gene amplicons have been developed. Here, I used the software Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt2) to predict the functional meta-genome of the JFFS faecal microbiome (Douglas *et al.*, 2020). I chose PICRUSt2 because it performs best at predicting functional profiles in non-primate mammal microbiomes compared to other tools (Douglas *et al.*, 2020). Briefly, Picrust2 functional predictions are generated from taxonomic profiles (ASVs) generated from 16S rRNA gene amplicon sequencing. The software then runs a phylogenetic analysis based on a reference tree containing 20 000 full prokaryotic 16S rRNA genes to predict the associated gene families (Langille *et al.*, 2013; Douglas *et al.*, 2020). Here, I decided to predict metabolic pathways using two databases; KEGGs and MetaCyc (Kanehisa and Goto, 2000; Krieger *et al.*, 2004). Although these databases share many reactions, KEGG contains interesting pathways that are not present in the latter. These pathways include xenobiotic degradation, glycan metabolism, and metabolism of terpenoids and polyketides (Altman *et al.*, 2013). On the other hand, MetaCyc is more up-to-date and contains more attributes, and fewer unbalanced reactions (Altman *et al.*, 2013).

With my approach, I wanted to identify those predicted pathways that were significantly different between the clusters detected in the taxonomic analysis. To deal with the background noise, I decided to take a very conservative statistical approach to recognise the most robust signals. The ALDEx2 analysis identified twelve significantly different KEGG orthology pathways. Carotenoid biosynthesis was particularly enriched in Cluster 1. This pathway is responsible for the production of beta-carotene, and vitamin A, an important antioxidant.

Evidence shows that carotenoids can regulate IgA production and play a role in immune system maturation (Lyu *et al.*, 2018; Rodriguez-Concepcion *et al.*, 2018).

It is known that carotenoids are synthesised by photosynthetic organisms, fungi, bacteria and a few invertebrates (Moran and Jarvik, 2010). Molluscs such as cephalopods; an important prey for JFFS, and crustaceans, an important prey for many cephalopods, are particularly rich in these lipophilic compounds (Fisher *et al.*, 2021). Furthermore, the presence of carotenoid-producing bacteria in the accessory nidamental gland of various female cephalopod species has been demonstrated (Grigioni *et al.*, 2000). A recent study in bivalves also found carotenoid-producing bacteria (Liu *et al.*, 2020). The same study showed these bacteria also had a high abundance of genes involved in the biosynthesis of fatty acids. Interestingly, according to both databases, the JFFS microbiome functional prediction also evidenced various fatty acids and lipid biosynthesis pathways enriched in Cluster 1.

Based on the taxonomic analysis, I was not surprised to see so many enriched pathways involved in fatty acid and lipid biosynthesis. Even though I did not focus on identifying prey content from the faecal samples, I speculate these functional signals to be directly linked to diet. Furthermore, I suspect that these results may reflect a cephalopod-rich diet as this prey is high in carotenoids and lipids.

Another interesting finding was the enrichment of xenobiotic degradation pathways in Cluster 1, specifically those involved in the degradation of aromatic compounds toluene and ethylbenzene. Hydrocarbon-degrading bacteria were identified in the 1980s (Atlas, 1981). Hydrocarbon contaminated environments are usually correlated with these bacteria, making them a good indicator of exposure (Chakraborty and Coates, 2004; van Dorst *et al.*, 2014). The over-representation of these pathways in cluster 1 suggests that the individuals in this cluster have been exposed to these aromatic hydrocarbons. At this point, there is little information on possible sources of contamination. However one is direct exposure from a contaminated environment, e.g. oils spills. Recent evidence showed that JFFS is particularly exposed to microplastics (Perez-Venegas *et al.*, 2018). However, plastic contamination does not necessarily reflect the environment surrounding the JFA. Based on satellite information, it is known that the JFFS feeding grounds are more than 500 km south of JFA (Francis *et al.*, 1998). This area, known as the South Pacific plastic patch, is largely affected by plastic debris (Eriksen *et al.*, 2013). Plastic debris is known to release toxic compounds such as hydrocarbons into marine environments.

Additionally, various studies have demonstrated the absorptive capacity of plastics, especially polystyrene and polypropylene, for mono-aromatic carbohydrates such as benzene and ethylbenzene (Lomonaco *et al.*, 2020). If these contaminated plastic particles enter the trophic network via accidental ingestion, these highly lipid-soluble components carried on the

plastic particles are likely to bioaccumulate and biomagnify (Scarlett et al., 2009). Lantern fish, another important prey for JFFS, is a mesopelagic lipid-rich species likely to ingest plastic debris (Ivar Do Sul and Costa, 2014). Here I hypothesise food sources and contamination of the hunting ground may be possible sources of hydrocarbon exposure.

Finally, the enriched pathways predicted for Cluster 2 were all linked to bacterial specific processes. These signatures may be suggesting some degree of dysbiosis. However, I was unable to find any robust evidence to support this claim. Base on the MetaCyc pathway official website (https://metacyc.org), the peptidoglycan biosynthesis IV pathway is involved in the synthesis of peptidoglycans found in gram-positive Enterococci. Pathways enriched according to the KEGG orthologs were mostly linked to bacterial motility, chemotaxis and nucleotide metabolism. The combination of both may be suggesting an increase in bacterial activity most likely driven by *Enterococcaceae*. Furthermore, the increased presence of L-arginine biosynthesis pathways identified by MetaCyc in Cluster 1 may also be supporting this hypothesis. L-arginine is a semi-essential amino acid with important roles in controlling intestinal homeostasis and contributing to limiting intestinal alteration (Fritz, 2013). These results, may suggest that the differences between the clusters may also result from pathological processes in cluster 2.

Even though PICRUSt2 gave interesting functional predictions, it is important to highlight that this method cannot directly identify genes. The predictions are only based on conserved genes within known bacterial genomes, which excludes horizontally transferred elements. This means that the reliability of the functional genes will depend on how represented the microbial communities are in the database. Thus, the results I provide here are aimed at obtaining some knowledge background on the functional metagenome from which I can generate hypotheses to test in the future using more reliable methods. Overall, the microbiome in Cluster 1 shows enrichment of pathways, such as carotenoids, linked to healthy functionalities. However, this cluster also shows an increase in xenobiotic pathways which may well indicate that JFFS are exposed to pollutants. This increase suggests the microbiome may be adapting to such conditions by increasing decontamination processes. If this is the case, it would be very interesting to explore this mammalian adaptation to the environment driven by their resident microorganism.

3.4 Data availability

Raw read data are publicly available in the European Nucleotide Archive (ENA) under the study accession PRJEB36555. All the scripts used in this study can be accessed in https://github.com/Cotissima/JFFS_microbiome_first_characterisation.

Chapter 4

Heavy metals: How bad is the pollution?

4.1 Introduction

Marine pollution is a significant threat to the ocean environment and its inhabitants. It has been a topic of great concern within the scientific community for many years. Monitoring pollutants and understanding its distribution, exposure and consequences in marine ecosystems as well as the effects on human health is complex. Unlike terrestrial environments, the marine world is fully interconnected through currents and movement patterns of organisms inhabiting the oceans.

Bioindicators have been widely used to monitor contaminant concentrations but there are problems with using this approach. For example, marine mammals are known to accumulate high levels of contaminants such as heavy metals but it is widely known that their exposure to these varies according to biological (sex or age) and ecological (especially those related to feeding habits) characteristics (Marcovecchio *et al.*, 1994; Bustamante *et al.*, 2004; Marangi *et al.*, 2021b). The ability to cope biologically with heavy metal exposure also seems to vary between species (Johnson *et al.*, 1978; Marcovecchio *et al.*, 1994; Varsha, 2013).

This chapter studies heavy metals exposure in the JFFS. Heavy metals are metal or metalloid elements of high relatively atomic weight. Some of these elements are classified as essential due to their known biological functions (e.g. zinc (Zn), and copper (Cu)). Other elements such as cadmium (Cd) and mercury (Hg) are considered non-essential and potentially toxic (Berry and Ralston, 2009; Tchounwou *et al.*, 2012; Nordberg *et al.*, 2015). These non-essential elements are of special concern as pollutants, due to their high toxicity and ability to damage multiple organs even when present at low concentrations.

Heavy metals have been associated with many detrimental effects in marine species including immunosuppression and poor reproductive success (Jaishankar *et al.*, 2014; Kakuschke *et al.*, 2011; Schaefer *et al.*, 2011; Das *et al.*, 2002). The source of heavy metals may be anthropogenic or natural (Christophoridis *et al.*, 2019; Bradney *et al.*, 2019; Garrett, 2000). Furthermore, research has shown that the release of Cd into marine environments has increased due to human activities such as mining and oil extraction (Christophoridis *et al.*, 2019) and the release of plastic polymers acting as vectors for heavy metals and other toxic compounds

(Bradney *et al.*, 2019; Brennecke *et al.*, 2016). Diet is the main route of exposure to heavy metals and other contaminants for marine mammals but different prey species may contribute differently to the amount of heavy metals incorporated from the diet.

Diet, as well as heavy metals, have a significant influence on the gut microbiome. Recently published data have suggested gut microbial responses to heavy metal exposure (Claus *et al.*, 2016; Breton *et al.*, 2013; Li *et al.*, 2019; Richardson *et al.*, 2018). Since the microbiome plays an important role in host fitness, survival and disease, characterisation of the functional microbiome is critical to better understand the role the gut microbiota plays in heavy metal exposure in marine mammals. However, to my knowledge only one study has looked at the gut microbiome in the context of heavy metal exposure in marine mammals.

The initial aim of this study was to explore the trace elements and heavy metal compositions of JFFS faeces in order to assess heavy metal exposure of the JFFS. This was then expanded to look at possible sources of these heavy metals in the diet and whether they were bioavailable. Finally, with the data from Chapter 3, I was also interested in exploring possible correlations between heavy metal exposure and the prevalence of specific microbial genera.

I performed a pilot quantitative heavy metal analysis in 14 JFFS faecal samples and compared them against the closely related AFS. The results showed that the JFFS were exposed to very high levels of toxic heavy metals, although the biological implications of these findings are not known.

4.2 Results

4.2.1 Trace Element Analysis of faecal samples

I started this analysis by exploring the trace elements in faecal samples collected from JFFS (n = 14) and AFS (n = 5). Trace element analysis was performed as described in section 2.5, by ICP-MS/MS, to provide the trace elements content (ng per g dry weight) of the faeces. Because of the collection method, I expected some degree of contamination of the faecal samples with the surrounding soil, even though I tried my best to limit soil contamination while sampling. This contamination is likely to affect the concentration of some elements in the faeces. I therefore also measured trace element levels in soil samples from the areas where JFFS faecal samples were collected, to allow sample contamination by the environment to be considered. Soil samples for the corresponding AFS faecal samples were not available for trace element and heavy metal analysis.

To limit the bias introduced by soil contamination, I ran both a principal component (PCA) and a cluster analysis on the total data set of all elements in the samples. In the first analyses, I

included faecal and soil samples to identify signals found predominantly in the faeces. The PCA reflected a clear clustering of faecal samples and soil samples, which explained 65.5 % of the total variation in the data. As shown in the biplot of Fig. 4.1 and the heatmap of Fig. 4.2, six trace elements, phosphorus (P), copper (Cu), zinc (Zn), selenium (Se), cadmium (Cd) and mercury (Hg) were primarily found in faeces. The heatmap of the two-dimensional Hierarchical Cluster Analysis also revealed three samples (18JFFS7, 18JFFS10, 18JFFS18) likely to be affected by soil contamination. However, I included these samples for further analysis because I only focused on the elements found low in soil but high in the faeces (See Table 4.1 for relevant trace element concentration medians and ranges). For the AFS faecal samples, because of the low concentrations of the selected elements present in these samples compared to JFFS, not being able to filter soil contaminants for this species does not affect the outcome of this study.

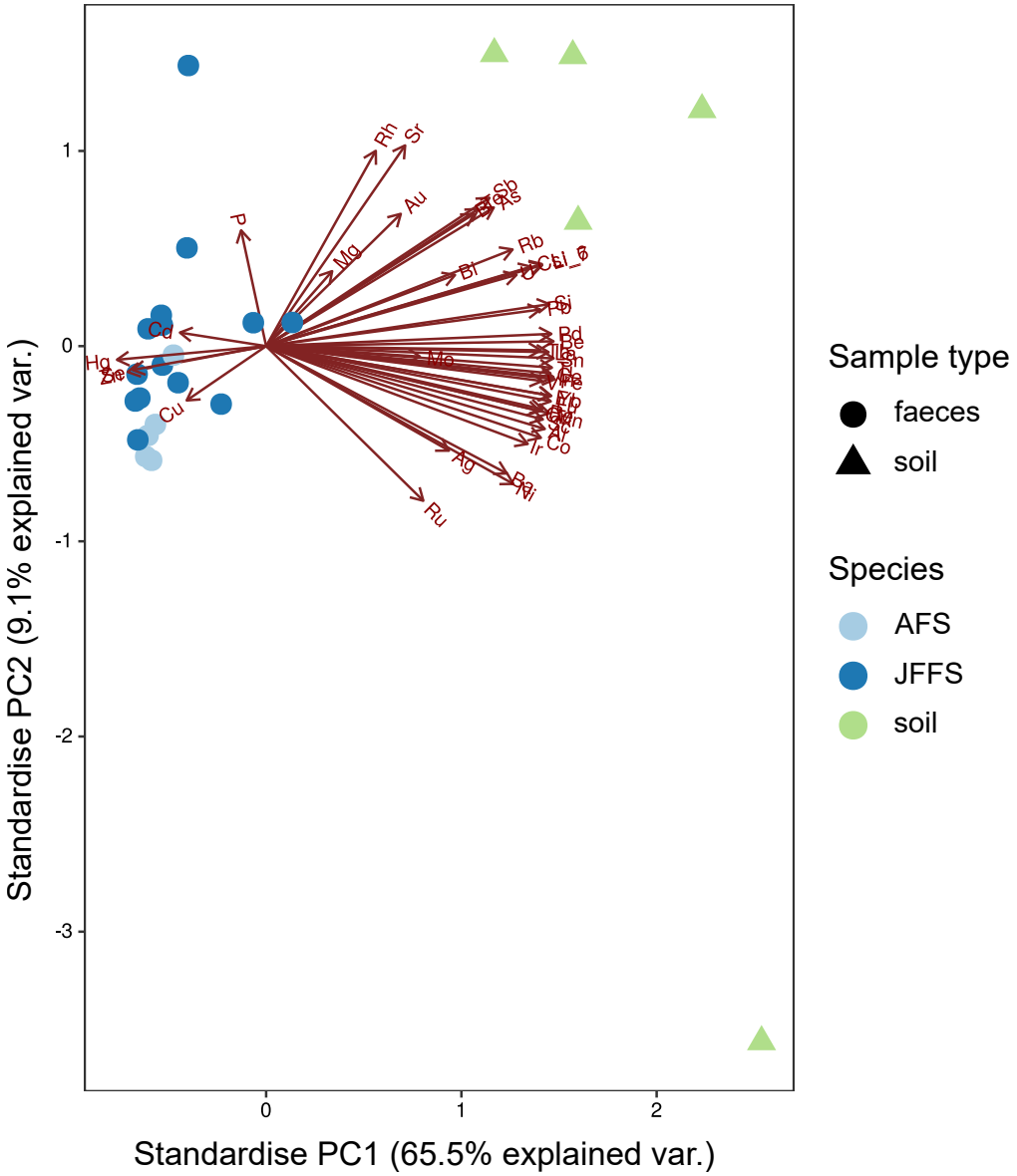

Figure 4.1. Principal component analysis (PCA) of the relationship between trace elements found in faecal and soil samples.

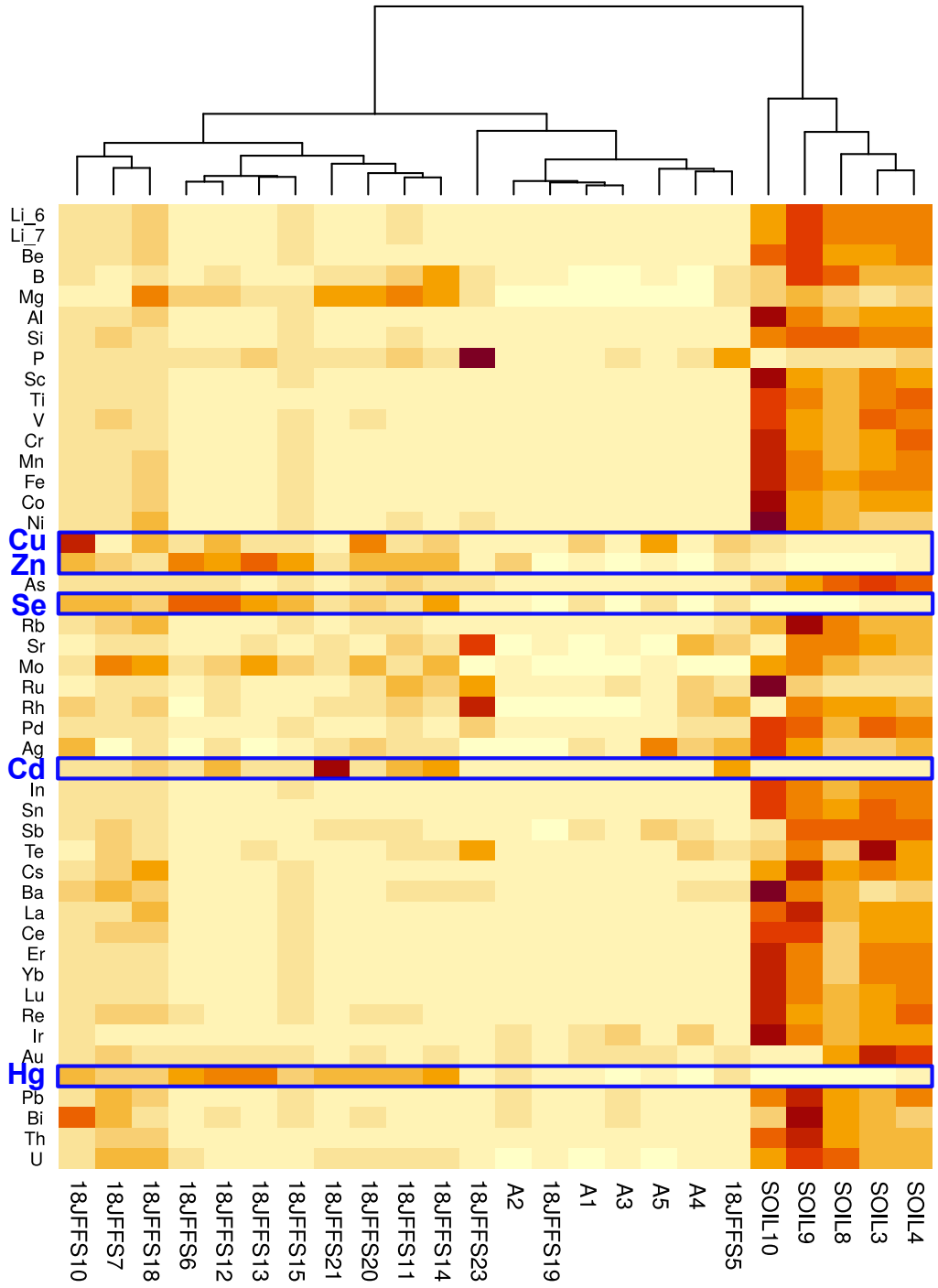

Figure 4.2. Heatmap showing the results of the two-dimensional Hierarchical Cluster analysis of faecal and soil samples.
Higher colour intensity indicates higher trace element concentrations. Trace elements are on the left, and sample names are on the bottom of the heatmap. Elements showing stronger signals in faecal samples (Cu, Zn, Se, Cd and Hg) are highlighted. The tree above the heatmap indicates the hierarchical-clustering of the samples.

Table 4.1. Trace element concentrations (means and ranges) measured in JFFS and AFS faecal samples in comparison with published levels in other marine species.

Only significantly different elements between species are shown. Additional, values taken from the literature were also included for perspective. All concentrations are shown based on dry weight.

Species	n		Cd (µg/g)	Hg (µg/g)	Zn (µg/g)	Se (µg/g)	P (mg/g)	Source
JFFS	14	median	33.34	0.81	857.47	19.07	10.15	This Study
A. philippii philippii		range	2.90 – 282.65	0.08 – 1.20	97.81 – 1600.17	1.75 – 38.88	1.27 – 111.59	
AFS	5	median	0.59	0.07	155.44	6.21	1.58	This Study
A. gazella		range	0.07 – 0.70	0.02 – 0.35	32.77 – 594.68	3.25 – 11.25	1.00 – 12.28	
Sperm whale	2	mean	0.45	1.49	96.92	10.63	-	Marangi et tal. 2021
Physeter macrocephalus		range	0.41 – 0.48	1.42 – 1.56	95.71 – 98.12	11.01 – 10.24		
Fin whale	2	mean	0.04	<LOQ	52.18	1.06	-	Marangi et tal. 2021
Balaenoptera physalus		range	0.03 – 0.04		43.73 – 60.62	0.84 – 1.27		
Little penguin	6	mean (sd)	0.73 (0.44)	0.30 (0.13)	-	5.10 (0.84)	-	Finger et al. 2107
Eudyptula minor		range	0.24 – 1.35	0.18 – 0.53	-	4.00 – 6.10		
Fish eating bat	35	mean	-	0.23	-	-	-	Drinkwater et al. 2021
Myotis vivesi		range		0.05 – 0.76				
Walrus	16	median	-	0.20	-	-	-	Rothenberg et al.2021
Odobenus rosmarus divergens		range		0.07 – 0.65				
Adult women (Amazon)	17	mean	-	0.05 (0.02)		-	-	Mendes et al. 2021
		range		0.02–0.11				
Children (Kasanda, Zambia)	88	median	0.16					Yabe et al. 2018
		range	0.07 – 0.43					

Next, I explored possible correlations between the six trace elements found at high levels in the faecal samples. Interestingly, Se, Zn and Hg showed significant and very strong positive correlations with each other (Fig. 4.3). As expected, when looking at the correlation matrix grouped by fur seal species (data not shown), AFS showed a similar pattern to JFFS for Se, Zn and Hg but these correlations were not significant (Fig. 4.3 B, C & D). This outcome is likely due to a lack of statistical power from the small sample size of AFS faeces.

Increasing scientific research has evidenced changes in the gut microbiome under heavy metal exposure (Richardson *et al.*, 2018; Breton *et al.*, 2013; Claus *et al.*, 2016). Furthermore, biotransformation of these elements by the gut microbiome has been shown in animal models (Yang *et al.*, 2021). For this reason, after looking at the correlation between the high faecal elements, I next explored the correlation between Hg and Cd, which are highly toxic heavy metals, and the prevalence of different genera in the microbiota, respectively. I used a rarefied feature table at the genus level obtained from the 16S rRNA gene amplicon sequencing (Chapter 3) for analysis. From a Spearman correlation analysis, four significant correlations were found. However, only two of these were strongly correlated ($r > 0.60$); Hg was positively correlated with the genus [*Ruminococcus*] *gauvreauii* group ($r = 0.65$, $p = 0.01$) and Cd showed a strong negative correlation with *Peptococcus* ($r = 0.61$, $p = 0.02$) (Fig. 4.4).

Se and Hg antagonism is well known as a mitigating mechanism for MeHg toxicity in animal tissues. Some authors have used Se/Hg molar ratio as an indicator of Se-driven Hg demethylation. This study has evidenced a strong correlation between Se and Hg. Thus, I performed a Spearman correlation analysis between the Hg:Se molar ratio and the genera found in the JFFS faecal samples to identify possible associations indicative of Se-driven demethylation. To obtain the molar ratio, I transformed each Hg and Se ug/g value by their respective atomic weights. This ratio was significantly and positively correlated with the genus *Bacteroides* ($r = 0.59$, $p = 0.03$) (Fig. 4.4 C).

Next, I was interested in exploring, between fur seal species, differences in terms of the concentration of the six elements found at higher levels in the faeces. Five of these elements were found to be significantly different between species based on Mann-Whitney non-parametric tests; Cd ($W = 0$, $p < 0.01$), Hg ($W = 4$, $p = < 0.01$), Se ($W = 8$, $p = 0.01$), Zn ($W = 8$, $p = 0.01$) and P ($W= 11$, $p = 0.03$) (Tables 4.1 and 4.2, Fig. 4.5). All these elements were found in significantly higher concentrations in JFFS samples. Levels of Cu were not found to be different between the species.

4.2.2 Analysis of prey samples

After looking at the trace elements found in faecal samples, I then focused on the trace elements levels in the JFFS diet. I mainly focused on Cd and Hg in fish and octopus samples. These

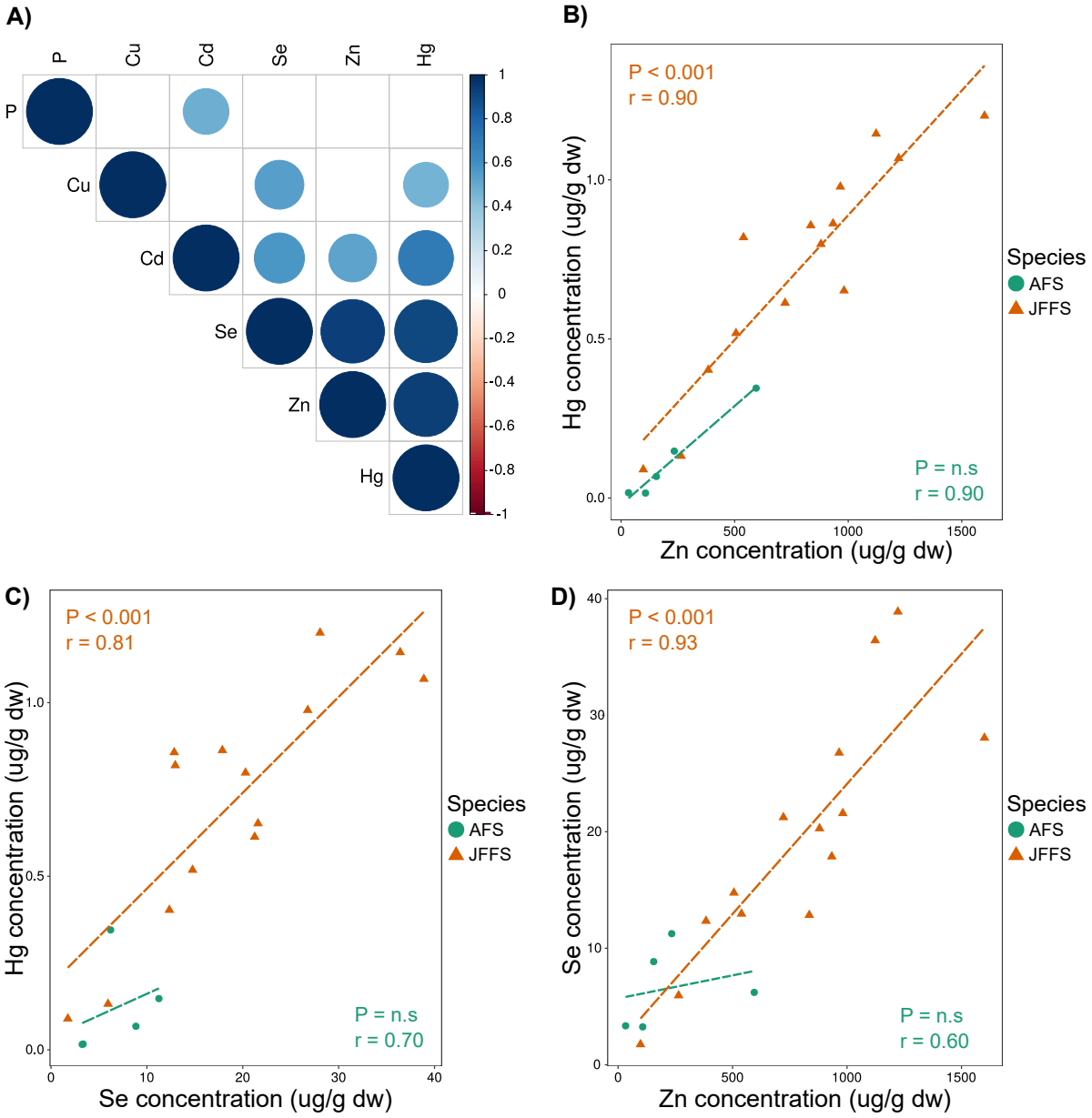

Figure 4.3. Spearman correlation analysis of elements in faecal samples.
A) Correlation matrix between the heavy metals found at high levels in all the faecal samples. The intensity of the colour and the size of the circles are proportional to the correlation coefficients with only correlations with p-value < 0.05 being shown. B) Correlation between mercury (Hg) and zinc (Zn). C) Correlation between Hg and selenium (Se). D) Correlation between Se and Zn. (Data points are coloured according to fur seal species (AFS and JFFS).

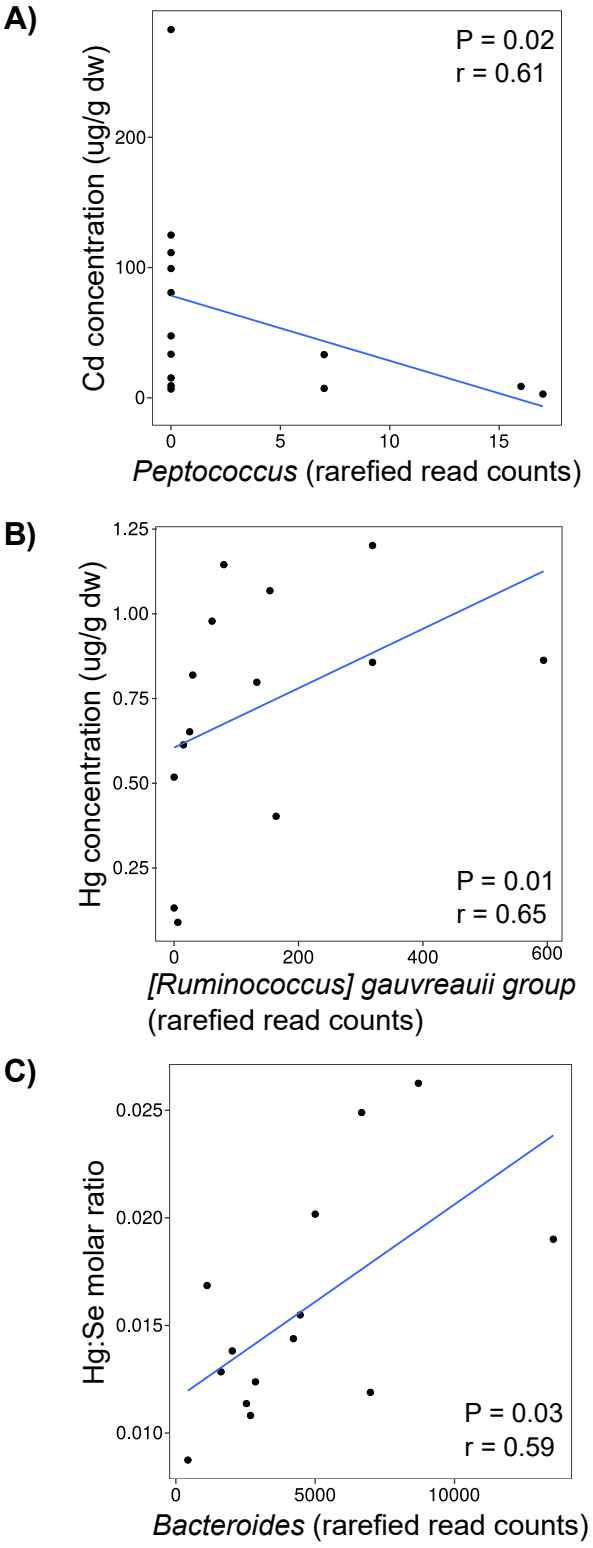

Figure 4.4. Spearman correlation analysis between heavy metals and genera found in JFFS faeces.
A) Correlation between Cd and *Peptococcus* B) Correlation between Hg and [*Ruminococcus*] *gauvreauii* group. C) Correlation between the Hg:Se molar ratio and *Bacteroides*.

Table 4.2. Summary of the Mann-Whitney tests for the siz elements found at high levels in fur seal faeces.
Each test was carried out with a total of 19 samples (JFFS = 14, AFS = 5). See medians in Table 4.1

Elements	W	Pval *	95% CI	
Cd	0	< 0.001	-6.2	-2.7
Hg	4	0.002	-920	-267.9
Zn	8	0.01	-607.7	-156.8
Se	8	0.01	-23.4	-3.5
P	11	0.03	-15.7	-1.7
Cu	25	n.s	-203	71.2

* Corrected p val

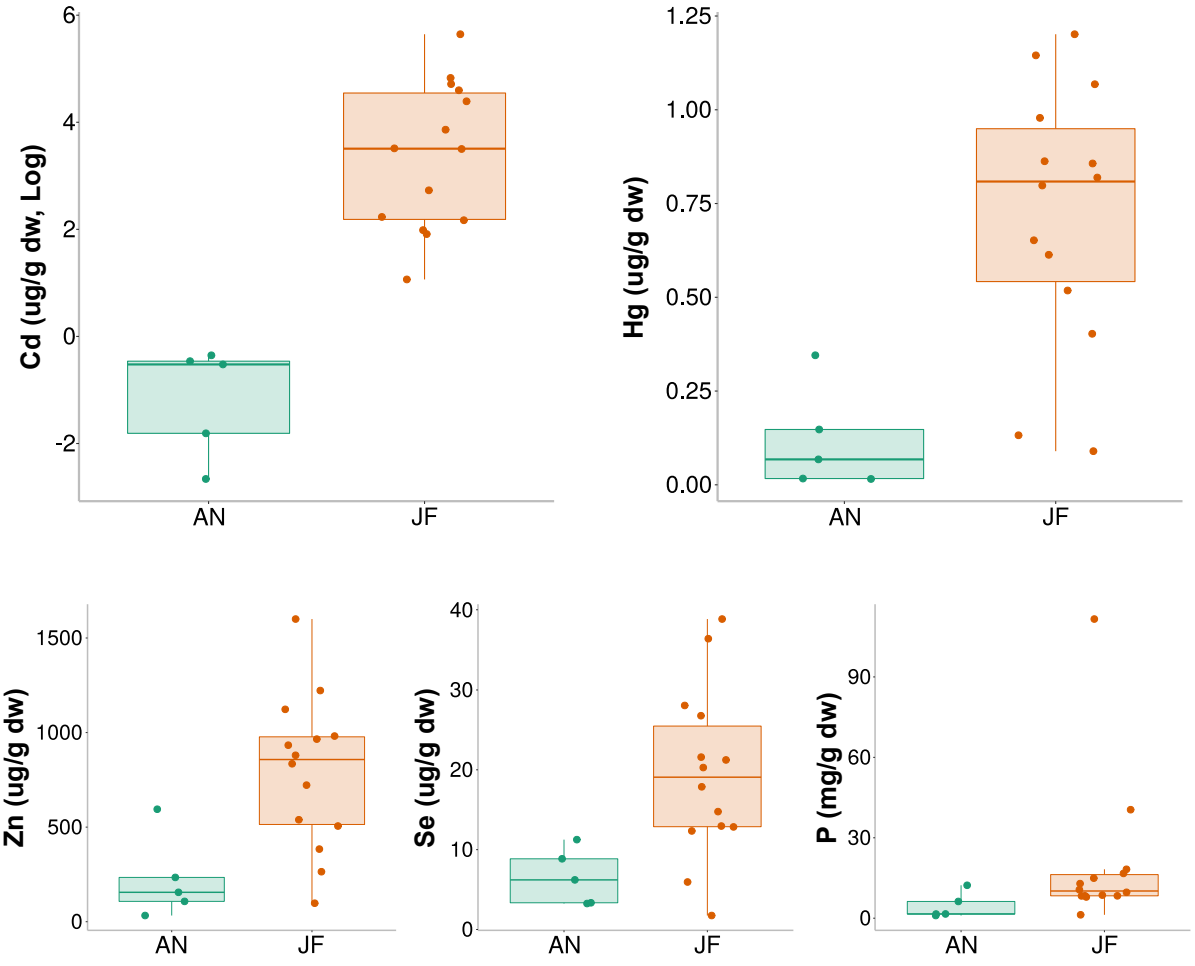

Figure 4.5. Boxplot showing the concentration of trace elements found in the faecal samples in the two fur seal species.
Only elements that were significantly different between species are shown (Cd, Hg, Zn, Se and P). JF = JFFS; AN = AFS.

elements are non-essential heavy metals with significant toxic potential. Myctophid fish and cephalopods such as octopus are described as important components of the JFFS diet (Ochoa Acuna and Francis, 1995). Moreover, Cd and Hg were found at high concentrations in the JFFS faecal samples. Table 4.3 summarises the findings in the prey samples. The octopus hepatopancreas had the highest levels of Hg and Cd of the analysed prey samples (medians = 101.3 ng/g ww and 76.6 μg/g ww, n = 3, respectively) suggesting this could be the source of Hg and Cd in the JFFS faeces.

Table 4.3. Trace element concentrations measured in JFFS diet.
Levels in fish and octopus are shown. Other published results of heavy metals in octopus, fish and krill are included for comparison. All concentrations are based on wet weight.

Prey	Origin	Sample type	N		Cd (µg/g)	Hg (ng/g)	Zn (µg/g)	Se (µg/g)	P (mg/g)	Ref.
Mytophids	SPG	Whole	11	median	0.26	6.6	10.3	1	6.1	
				range	0.12 – 0.64	2.0 – 66.5	8.4 – 24.4	0.6 – 2.2	4.5 – 13.4	
Octopus	JFA	Hepatopancreas	3	median	76.59	101.3	272.26	4.5	2.09	
Octopus vulgaris				range	55.81 – 133.28	41.5 – 330.7	248.17 – 399.11	1.8 – 5.6	1.90 – 2.12	
		Kindney	3	median	1.52	79	38.78	3.6	2.03	
				range	1.21 – 12.73	62.1 – 119.7	22.91 – 63.25	2 – 9.3	1.68 – 2.39	
		Gills	3	median	0.09	39.7	19.85	0.8	2.45	
				range	0.04 – 0.25	18.3 – 60	18.34 – 24.58	0.8 – 1.3	2.10 – 2.70	
Antarctic krill		Whole	4	mean	-	0.02	-	-	-	a
Euphasia superba				SD		0.01				
Mexican lamp fish*	Guaymas Basin	Whole	45	median	0.25	-	-	-	-	b
Triphoturus mexicanus				range	0.14 – 0.43					
Octopus	Bay of Biscay	Whole	13	mean	0.54	**285.52***	-	-	-	c, **d**
Octopus vulgaris				SD	0.24	**211 – 422.5**				
Octopus*	Bay of Biscay	Whole	13	mean	-	111.25	-	-	-	d
Octopus vulgaris				SD		74.25 – 165.5				
Octopus	Adriatic Sea	Hepatopancreas	10	mean	7.86	780	-	-	-	e
Octopus vulgaris				range	5.47 – 9.19	370 – 1380				

*Original values were transformed to wet weight asuming 75 % of water

Values in bold came from Ref. In bold

Ref: a) Cipro *et al.*, 2017, b) Figueiredo *et al.,* 2020, c) Bustamante *et al.*, 1998, d) Bustamante *et al.*, 2006, e) Storelli *et al.*, 2006

4.2.3 Analysis of bone samples

The trace element analysis I performed on the faeces samples revealed that JFFS were exposed to high levels of Hg and Cd. Based on the prey analysis, I was able to show that octopus and possibly other cephalopods are likely to be the primary source of these heavy metals. The next step was to look at the bioavailability of these elements, i.e. the possibility that they were absorbed from the fur seal gut. Bone is a natural store for heavy metals and as bone samples were readily available from dead JFF seals, I looked at Cd levels in twelve bone samples. These samples included five adult JFFS, five pup JFFS samples, one adult male grey seal (Orkney Islands) and one adult male AFS (South Georgia). The final two samples were included in this analysis only for a rough point of comparison.

This analysis revealed high concentrations of Cd in JFFS bones (Table 4.4). Due to the known interactions between Cd, Ca, P and Zn in bone (Buha *et al.*, 2019; Blumenthal *et al.*, 1995), I explored possible association between Cd and these essential bone elements. Contrary to my expectations, Cd concentrations did not correlate with any of these elements (Fig. 4.6). Finally, I looked at Si, an important sub-trace element in bone known for its role in bone health which is often affected by Cd even at minor concentrations. However, there was no correlation between Si and Cd (Fig. 4.6). Furthermore, none of the major elements in bone showed significant differences in concentration between age groups (Fig. 4.7). Overall, Cd concentrations were considerably higher in JFFS bones than in grey seal and AFS bone samples. Silicon concentration was apparently higher in JFFS bones than in AFS (see Table 4.4). However, more samples are needed from the AFS to confirm these differences statistically.

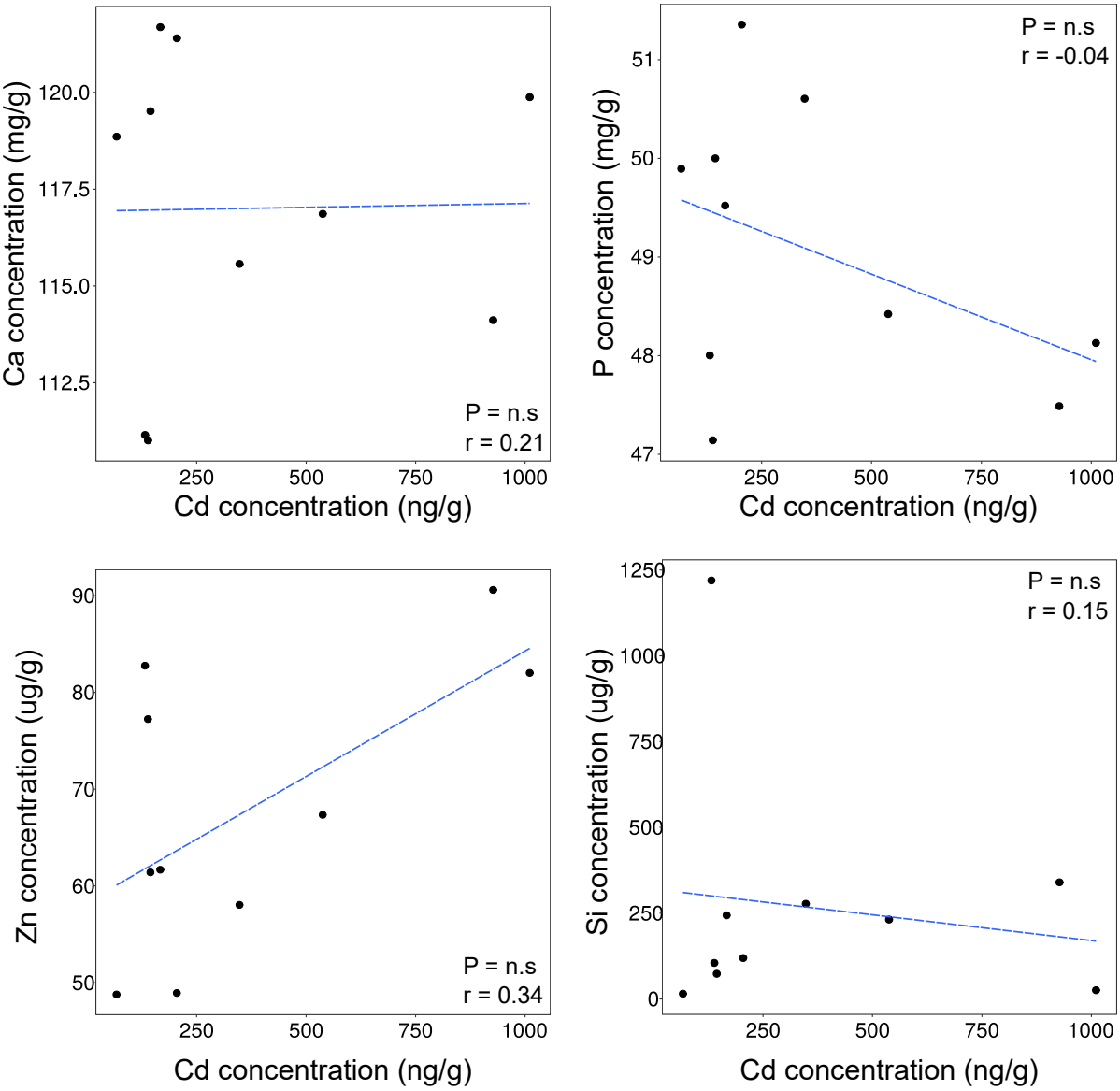

Figure 4.6. Behaviour of different elements (Ca, P, Zn and Si) found in bone samples in response to cadmium concentrations.
Scatter plots show there was only weak or no association between any of the elements and cadmium. The letters P and r indicate the rho and p-value resulting from the Spearman correlation analysis.

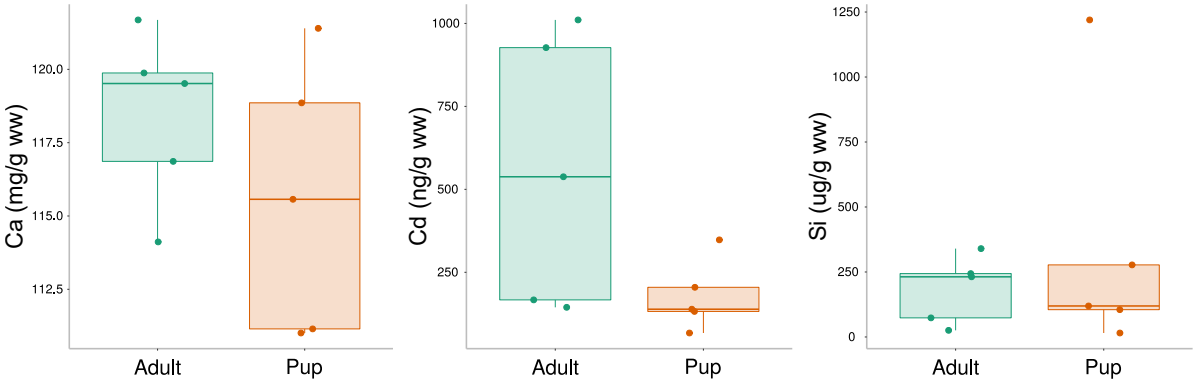

Figure 4.7. Boxplot showing the concentration of Ca, Cd and Si found in bone samples according to the estimated age group. Age was estimated by bone size.
None of the elements evidenced significant differences between groups.

Table 4.4. Trace element concentrations (means and ranges) measured in JFFS and AFS bone samples in comparison with published levels in other marine species.
Only significantly different elements between species are shown. All concentrations are shown based on dry weight.

Age groups	Age groups	n		Cd (ng/g)	Ca (mg/g)	P (mg/g)	Si (ng/g)	Ref.
JFFS	Adults	5	median	537.9	119.5	48.4	231.4	
			range	114.5 – 1010.6	114.1 – 121.7	47.5 – 50	25.5 – 339.9	
	Pups	5	median	138.8	115.6	49.9	119.3	
			range	67.1 – 347.8	111 – 121.4	47.1 – 51.4	15.3 – 1219.6	
	All	10	median	185.9	117.9	489.7	175.4	
			range	67.1 – 1010.6	111 – 121.7	471.4 – 513.6	15.3 – 1219.6	
Grey seal	Adult	1	avg replicates	14.6			246.5	
AFS	Adult	1	avg replicates	1.2			6.2	
striped dolphin*	Foetus	13	mean (Sd)	3 (2)				Honda *et al.*, 1986
Stenella coeruleoalba	Calf (male)	11		40 (30)				
	Immature (male)	6		120 (20)				
	Mature (male)	5		90 (20)				
	Mature (female)	5		180 (50)				
human*	Adult	37	median	5.75				Lanocha *et al.*, 2012
			range	0.25 – 67.25				
Dog*		24	median	14.5				Lanocha *et al.*, 2012
Canis familiaris			range	3.25 – 52.5				
Red fox*		12	median	33				Lanocha *et al.*, 2012
Vulpes vulpes			range	8.5 – 65				
bottlenose dolphins	Adult	15	mean	50				Lavery *et al.* 2009
Tursiops aduncus			range	5 – 330				

* Original values were transformed to wet weight assuming 75 % of water

4.3 Discussion

Even though JFFS inhabit a location free of polluting industries and with a low human population density, this non-migratory species is exposed to significant levels of contamination (i.e. heavy metals and microplastics) derived from natural causes as well as human activities. Furthermore, it is likely that in the case JFFS, foraging behaviour is one of the most critical risk factors of contamination exposure. For instance, the most important foraging grounds for JFFS are hundreds of kilometres away from their JFA territory (Francis *et al.*, 1998). These hunting areas are located around the most plastic-concentrated zone of the South Pacific Sub-tropical gyre (Francis *et al.*, 1998; Eriksen *et al.*, 2013; Markic *et al.*, 2018). Additionally, their most important prey, myctophids and cephalopods, have been associated with contaminant biomagnification and bioaccumulation leading to effects on top predators, such as marine mammal carnivores (Jakimska *et al.*, 2011). Furthermore, from a recent collaborative study, we could evidence that JFFS was the South American pinniped most exposed to microplastic fragments, most likely due to their foraging behaviour (Perez-Venegas *et al.*, 2018).

Here, I looked at faecal samples to explore the exposure of JFFS to heavy metals and compared them to an AFS control group. I also analysed example prey samples to investigate possible contaminant sources. Motivated by the results obtained from the faecal samples, I went on to explore Cd concentrations in bone samples. To my knowledge, this is the first study on heavy metal exposure in JFFS in more than two decades (Sepúlveda *et al.*, 1997). Additionally, this is the first report on heavy metals in a species of great relevance as both a local human food source and also to the local economy, such as octopus (*Octopus vulgaris*).

4.3.1 Findings in faecal samples

After controlling for potential soil contamination, only six elements were highly associated with the faecal samples. These included the heavy metals, Cd and Hg. Apart from Cu, all five remaining trace elements (Cd, Hg, Zn, Se and P) were significantly higher in JFFS compared to AFS faeces.

In marine mammals, diet is one of the most important sources of exposure for heavy metals and other pollutants (Marcovecchio *et al.*, 1994; De María *et al.*, 2021; Marangi *et al.*, 2021b). Thus, the trophic level at which marine species feed will largely influence the levels of exposure to contaminants. I believe that prey selection is likely to explain the differences observed between JFFS and AFS faecal samples. For instance, the levels of Cd, Hg, Zn and Se found in sperm whale (*Physeter macrocephalus*) faecal samples were higher than those found in its sympatric species, the Mediterranean fin whale (*Balaenoptera physalus*) (Marangi *et al.*, 2021b). Interestingly, AFS and fin whales feed predominantly on krill, whereas sperm whales

and JFFS prey mostly on cephalopods and fish (Ochoa Acuna and Francis, 1995; Francis *et al.*, 1998; Kirkman *et al.*, 2000; Roberts, 2003; Bentaleb *et al.*, 2011).

The levels of Hg in JFFS faecal samples were also higher than many of the previously reported concentrations in other mammalian species faecal samples (Finger *et al.*, 2017; Drinkwater *et al.*, 2021; Marangi *et al.*, 2021b; Mendes *et al.*, 2021; Rothenberg *et al.*, 2021). From the literature, only sperm whale faecal samples showed higher concentrations of Hg (Marangi *et al.*, 2021b). To give some perspective, Hg in JFFS was twelve times higher than AFS, four times higher than walruses and sixteen times higher than the levels found in adult human faeces from a small Amazonian community whose diets rely largely on mercury-contaminated fish (Mendes *et al.*, 2021; Rothenberg *et al.*, 2021).

Animals do have mechanisms to eliminate and detoxify toxic metals after exposures. A study looking at Hg elimination in fur and faeces in captive Baltic grey seals (*Halichoerus grypus*) found that faeces were an efficient elimination route for total Hg (Grajewska *et al.*, 2020). They estimated that faeces accounted for up to 48 % of Hg elimination; a similar rate was identified in bottlenose dolphins (Nigro *et al.*, 2002). Meanwhile, in this study, there was a strong correlation between Hg and Se concentrations in JFFS but not AFS faeces. Accumulating evidence strongly suggests Se may have an essential role in the detoxification process of methylmercury, the most toxic form of this heavy metal, as well as inorganic Hg (iHg) (Berry and Ralston, 2009; Khan and Wang, 2009). Furthermore, the formation of inert HgSe particles has been suggested as a critical Hg-detoxification strategy in marine animals, where either the particles themselves or a positive correlations between Hg and Se have been documented in various internal tissues, fur and feathers (Frodello *et al.*, 2000; Bustamante *et al.*, 2004; Ribeiro *et al.*, 2008; Sakamoto *et al.*, 2015; Squadrone *et al.*, 2015; Bolea-Fernandez *et al.*, 2019; Grajewska *et al.*, 2020). Thus, the correlation between Hg and Se found here may reflect the formation of HgSe particles as part of the detoxification mechanism.

The results of most significance within the faecal samples was the surprisingly high levels of Cd in JFFS. When comparing the median concentrations between species, this value was 57 times higher than that found in AFS samples. Even the minimum value in JFFS faecal samples was higher than any of the maximum values reported for other mammalian species (Table 4.1). Furthermore, JFFS faecal Cd levels were more than 200 times higher than those found in the faeces collected from children inhabiting a Cd-polluted mining town in Zambia (Yabe *et al.*, 2018). Despite the lower Cd levels detected in AFS faecal samples, compared to JFFS samples, these were not necessarily 'low' for mammals per se. The median value for this species was three times higher than the concentrations observed in the Zambian children previously mentioned and fourteen times higher than the levels found in fin whales (Yabe *et al.*, 2018; Marangi *et al.*, 2021b).

Heavy metals and the microbiome

In the last decade, there has been an increase in the number of studies looking at changes in the gut microbiota of animals exposed to various contaminants. Previous studies in mice have reported changes in the gut microbial composition after heavy metal (lead and Cd specifically) exposure (Breton *et al.*, 2013; Richardson *et al.*, 2018). For this reason, I was interested in finding out if there was any correlation between faecal Cd or Hg and members of the faecal microbiome known for those same individuals. From the Spearman correlation analysis, only two strong and significant correlations were identified; a positive one between Hg and the genus [*Ruminococcus*] *gauvreauii* group and a negative one between Cd and *Peptococcus*.

Initially, I thought the Cd-*Peptococcus* association could reflect either a genuine interaction or a confounding bias associated with prey selection. For instance, the same prey that would typically influence the abundance of *Peptococcus* may be also be contaminated with Cd. However, an experimental study in C57BL/6 mice also reported a decrease in this genus when mice were exposed to high levels of Cd (Li *et al.*, 2019). Additionally, Cd would have also impacted on the abundance of other genera if these were confounding variables. On the other hand, *Peptococcus* was a rare feature in the 16S rRNA gene sequencing output; only four out of ten samples had an abundance higher than zero. Thus, it is likely that sequencing depth failed to pick up the signal of rare genera such as *Peptococcus*; this means that the correlation output is based on only four samples. It is not possible, therefore, to determine whether the association between this genus and Cd is real, confounding or an artefact caused by the available information. It would be interesting to test the veracity of this association as it could inform about the role of the gut microbiome in the host response to heavy metal contamination in naturally exposed populations. Regarding the correlation observed between Hg and [*Ruminococcus*] *gauvreauii* group, I could not find any similar reports that could contribute to developing a reliable explanation for this finding.

However, I would like to focus on the positive correlation between the Hg/Se ratio and the genus *Bacteroides*. A study published in June 2021 identified 131 mercuric reductases (MerA) and three organomercury lyase (MerB) homologue genes within the phylum *Bacteroidetes* (Christakis *et al.*, 2021). These genes provide enzymatic mercury-detoxification mechanisms in prokaryotes in soil, plants and water (Schaefer *et al.*, 2004; Dang *et al.*, 2019; Christakis *et al.*, 2021). Furthermore, organisms carrying *Mer* operons are increasingly being used for bioremediation in Hg-contaminated environments (Kumari *et al.*, 2020). Interestingly, a study examining possible interactions between Hg and the gut microbiome in walruses found a negative correlation between methylmercury and the genus *Bacteroides* (Rothenberg *et al.*, 2021).

Another study published in 1978 also found evidence that *Bacteroides* cultures isolated from human intestines could metabolise methylmercury (Rowland *et al.*, 1978). Li *et al.* (2019) recently published a thorough review on the methylation and demethylation processes associated with different animals gut microbiomes. However, most research, especially regarding demethylation, has been conducted in either controlled animal experiments or *in vitro* (Li *et al.*, 2019; Yang *et al.*, 2021). Thus, the pathways in which gut microbiota demethylation processes may occur remain poorly understood. The positive association between the prevalence of the genus *Bacteroides* and methylmercury levels in walruses, combined with the negative association between this genus and the Hg/Se molar ratio, identified here in JFFS, may suggest this genus is contributing to mercury de-methylation processes in these pinniped species; possibly through *Mer* operons or other similar mechanisms. However, more research would be needed to test for this experimentally.

4.3.2 Prey samples

In marine mammals, heavy metal exposure will most likely depend on the trophic level at which species (or individuals within those species) feed (Marcovecchio *et al.*, 1994; Gray, 2002; Bustamante *et al.*, 2004; Ikemoto *et al.*, 2004; Ando *et al.*, 2005; Marangi *et al.*, 2021b). However, heavy metal exposure is also linked to specific environmental characteristics. For instance, a clear link has been shown between ocean depth and heavy metal contamination in different trophic webs (Chouvelon *et al.*, 2012; Moiseenko and Gashkina, 2020). Furthermore, Monteiro *et al.* (1998) showed that Hg accumulation in birds feeding on mesopelagic (below 200 meters sea depth) squid and fish were higher than when they fed on epipelagic fish and squid (above 200 meters sea depth) (Monteiro *et al.*, 1998).

The limited information on the JFFS diet suggests this species has a mesopelagic ichthyotentophagous diet, which means their diet is mainly based on mesopelagic fish, mainly myctophids, and squid (Ochoa Acuna and Francis, 1995). Additionally, the local community has reported that octopus inhabiting the coastal benthic floor of the archipelago forms part of the JFFS diet. Based on sample availability, I analysed heavy metal concentrations in two different types of prey known to be consumed by JFFS, *Symbolophorus sp*, a mesopelagic myctophid fish, and *Octopus vulgaris*, a benthic cephalopod.

As expected, the largest concentrations of Cd were found in octopus samples. As previously reported in the literature, the hepatopancreas, also known as the digestive gland, was the organ with the highest Cd levels (Bustamante *et al.*, 2002; Penicaud *et al.*, 2017). Furthermore, the Cd concentrations in the octopus hepatopancreas collected from the JFA was much higher compared to other studies, but Hg concentrations was relatively low (Bustamante *et al.*, 1998, 2006; Storelli *et al.*, 2006, 2010). Unfortunately, no other cephalopod prey species were

analysed in this study. However, due to their mesopelagic characteristic, I would hypothesise that these Cd levels would be even higher than in *O. vulgaris*. It is important to highlight that no robust conclusion can be made with regards to the prey data generated in this study. Not only was the sample size too small, but the information on the foraging behaviour of JFFS remains too limited, especially regarding octopuses. Therefore, it is not clear how often and how many octopuses an individual fur seal would hunt. Additionally, the lack of knowledge on heavy metal contamination in other cephalopod prey makes it difficult to identify how different prey species influence the Cd levels observed in JFFS faeces. To understand better how different preys contribute to heavy metal exposure, future work would need to include other cephalopod species such *Onychoteuthis banskii* (now reclassified as *Onychoteuthis aequimanus*), which were the most abundant squid beaks found in JFFS scats, especially in females (Ochoa Acuna and Francis, 1995). Future studies would benefit from stable isotope analysis such as δ13C and δ15N, to confirm the trophic level associated with each faecal sample.

When looking at Cd and Hg levels in myctophid samples, the levels looked unimportant compared to those found in octopus organs. However, this difference is likely to be inflated because whole fish were compared to individual organs in octopuses. Unfortunately, due to the size and fragility of the fish samples after defrosting, it was not possible to separate specific organs. A recent study on trace elements in myctophids from the Gulf of California showed similar levels to those presented here. These levels were among the highest compared to myctophids from other regions (Figueiredo *et al.*, 2020). Furthermore, Cd levels measured in whole individuals from various octopus species such as *O.vulgaris* (Bay of Biscay) and *Eledone cirrhosa* (Faroe Fsland), were only 1.8 and 10 times, higher than those observed in the myctophid samples from this study (Bustamante *et al.*, 1998). In comparison there was almost 300 times difference between the myctophids and octopus hepatopancreas analysed here. Thus, although cephalopods are likely to be the main Cd source, myctophids could still be contributing to Cd exposure in JFFS. On the other hand, as with octopus, Hg levels in myctophid samples analysed here compared to other studies were not particularly high (Cipro *et al.*, 2018). Thus, the source of Hg in JFFS remains unclear.

Regarding the AFS diet, it is known that this species feeds predominantly on krill. However, the composition of their diets may vary depending on environmental characteristics and differences in foraging behaviour between colonies. For instance, when water temperature increases, krill populations decline and AFS may then shift their prey selection toward cephalopods (Abreu *et al.*, 2019). Other researchers have also shown an association between prey preference and different AFS genetic lineages. For example, some colonies, such as the one included in this study, feed almost exclusively on krill, while others prey primarily on fish and practically no krill (Cleary *et al.*, 2019). Thus, the low Hg and Cd levels found in AFS faecal samples in this

study nicely reflect the colony preference for krill. This intra-species difference makes AFS an attractive target to study the effects of differences in heavy metal exposure on fur seals.

4.4 Bone samples

Based on the high levels of Cd in the faeces, it was logical to assume high levels of this heavy metal in bones, an important site for its accumulation. The ICP-MS data from the bone samples analysis confirmed the hypothesis that Cd was indeed bioavailable from the gut and accumulated in the bones. JFFS bone samples showed high levels of Cd compared to other marine and terrestrial mammalian species (Lavery *et al.*, 2009; Lanocha *et al.*, 2013; Honda and Littman, 2016). Cd is known for its detrimental effects on bone mineralisation (Blumenthal *et al.*, 1995; Buha *et al.*, 2019). For instance, high levels of this heavy metal negatively affect Zn and Ca concentrations in bones even at very low concentrations (Youness *et al.*, 2012; Nordberg *et al.*, 2015; Chen *et al.*, 2018; Buha *et al.*, 2019). However, despite the high bone Cd concentrations, there was no influence on the composition of these essential elements. These results suggest the bones in the JFFS are resilient to Cd toxicity.

Previous studies have shown that marine mammals exposed to high levels of heavy metals have developed mechanisms to tolerate such high levels of contamination. Most of these studies link heavy metal tolerance to increased expression of metallothioneins and show the important role of the liver and kidneys in the detoxification and excretion process (Wang *et al.*, 2014; Kehrig *et al.*, 2016; Polizzi *et al.*, 2017). To my knowledge, this is the first time bone resilience to high levels of Cd exposure have been reported in these species.

In the past 20 years, silicon (Si) has been reported to be associated with bone health (Jugdaohsingh, 2007). Although the exact role Si plays in bone is not understood, evidence suggests Si may be actively involved in the calcification process (Carlisle, 1970; Reffitt *et al.*, 2003; Kim *et al.*, 2013). Furthermore, the evidence clearly shows a positive association between increased silicon intake and higher bone mineral density (Reffitt *et al.*, 2003; Jugdaohsingh *et al.*, 2004; Hing *et al.*, 2006; Price *et al.*, 2013). Because of these reports in the literature, bone Si levels in JFFS was also analysed in this study. Unfortunately, there are no other reports on Si levels in marine mammals to compare with. For this reason, the cranial bones from one adult male AFS and one adult male grey seal were analysed for Si.

Interestingly, the AFS bone samples had minimal levels of Cd and Si compared to the JFFS and grey seal bone samples. The grey seal bone sample, on the other hand, had similar Si levels to the median values from JFFS bones but much lower Cd concentrations. Similar to JFFS, grey seals feed at a relatively high trophic level taking a wide variety of prey, including fish, cephalopods and crustaceans. Additionally, foraging ecology in grey seals differs significantly

between males and females (Tucker *et al.*, 2007). This difference has also been shown to influence heavy metal exposure in this species. Females have at least four times higher Cd levels compared to males in the kidney, liver, and muscle (Bustamante *et al.*, 2004).

When age was estimated by bone size, the Si concentration seemed to remain constant with age in JFFS in this study. However, I may not have estimated age properly using bone size and I could have placed juveniles in the adult group. Juvenile males and adult females do not differ much in size; thus, their bones are also similar in size. I did not differentiate the bones in terms of gender to overcome this problem.

From the above, I hypothesise that high levels of Si may be a general trait of species that feed at higher trophic levels regardless of the differences in feeding behaviour within the species. Of course, I need to highlight that this is a bold hypothesis based on a low number of samples. Future work should include more samples, look at species with different trophic ecologies, determine sex and somehow age the animals the samples came from. If this hypothesis was correct, I would expect to see similar Si levels in grey seal bones regardless of sex. Similarly, the AFS would consistently show low values of Si despite the different foraging ecology between species.

I did not explore the bioavailability of Hg in the JFFS by using relevant tissue samples e.g. liver, fat tissue, kidneys. Furthermore, prey samples included in this study did not define fully the source of this contamination. Nevertheless, I would expect high concentrations of this metal in key organs such as the liver, kidneys and muscles of JFFS. In addition, I would also expect to find high levels of metallothioneins, HgSe crystals and most of the methylmercury being stored in less sensitive cells such as muscle.

Finally, prey samples included in this study did provide clues as to the source of this heavy metal contamination. Identifying the source of Hg contamination could provide critical information to the local community. Diet is an important route of exposure for methylmercury. Although marine mammals show some resistance, humans, especially those eating large quantities of contaminated seafood, are particularly vulnerable to the negative effects. In the context of the JFA, the findings in this study might be of great relevance regarding food security. Carnivorous fish, octopus, and lobster are important parts of the local diet, but to my knowledge, there have been no studies on heavy metal contamination in the human food chain.

Chapter 5

Obtaining host DNA for JFFS faeces: trials and errors

5.1 Introduction

Genetic data can provide vital information such as genetic variation (e.g. population dispersal, historical events such as population bottlenecks and hybridization), an individual's fitness (e.g. inbreeding), population viability and even sex determination (DeYoung and Honeycutt, 2005). Despite the history of hunting resulting in a known population bottleneck and their endemism, little is known about the JFFS's population genetics. Furthermore, only one study has been carried out on this topic (Goldsworthy *et al.*, 2000). The lack of funding and complicated logistics may explain the stagnation of research development on this species. However, studying JFFS genetics is an urgent matter to inform policy makers on the legal protection of this endemic species as the current hunting ban is close to its legal end-date.

The local community is rather divided in terms of supporting, or not, the hunting of the JFFS. One of the main arguments held by those who support the allowance of the hunting is down to the JFFS individuals conflicting with the local fishermen and the sentiment that currently there are "too many" animals and that they need to be controlled. These ideas have been backed up by visual perception and by the increasing numbers reported in both sporadic unofficial and official censuses in the last 50 plus years (Torres, 1987; Osman and Moreno, 2017).

Even though the population growth has been evident, this does not necessarily reflect the effective population size nor does it tell us about the genetic fitness of the population. Understanding the JFFS genetic make up is vital now in light of the current climate change (Forcada and Hoffman, 2014b; Páez-Rosas *et al.*, 2021; Schumann *et al.*, 2013). Furthermore, the risk of relying only in the total numbers of individuals for assessing the state of the current JFFS population may result in allowing hunting on a species which may be under pressure, even though the pressure might not be reflected in total numbers yet.

Using non-invasive samples as a host DNA sources may facilitate the development of genetic studies in JFFS. For clarity, here I will consider non-invasive samples as those that do

not require direct interaction with the animals for collection (e.g. faeces, moulted hair). There are various benefits of using non-invasive samples for genetic studies. This contact-free strategy reduces animal stress and the physical risk to both animals and field researchers. Non-invasive samples are also more straightforward and quicker to collect and require considerably fewer logistics. Thus, more samples can be collected per sampling effort. Of the non-invasive samples, faeces can be particularly useful due to the diversity of information it contains. Information about diet, reproductive status, parasites, gut microbiome and, stress levels, are some of the topics that can be explored through analysing this kind of sample.

Using faeces for molecular studies on host genetics in wild animals is not new (Nsubuga *et al.*, 2004; Ramón-Laca *et al.*, 2015; Reddy *et al.*, 2012; Bourgeois *et al.*, 2019). However, faecal composition significantly varies between species and thus, sampling and extraction protocols that have been optimised in some species might under perform in another species, especially if the diet greatly differs (Eggert *et al.*, 2005).

Genetic studies in pinnipeds are relatively common, especially those targeting mtDNA. However, only few have used faeces as a DNA source. Reed *et al.* (1997a), Fietz *et al.* (2016), Zappes *et al.* (2017) and more recently Steinmetz *et al.* (2021) have successfully used faeces in free-range pinnipeds. All these studies included mtDNA as a target. Reed *et al.* (1997a), Fietz *et al.* (2016) and Zappes *et al.* (2017) also attempted to amplify nuclear DNA using a standard PCR protocol. Steinmetz *et al.* (2021), on the other hand, used real time PCR (qPCR) to amplify nuclear DNA for molecular sexing. qPCR is a more sensitive protocol, thus it requires less DNA. Finally, Reed *et al.* (1997a) and Steinmetz *et al.* (2021) successfully targeted microsatellite loci amplification. Microsatellites are by far the most challenging target due to the large DNA quantities needed for genotyping.

This study aimed to develop a protocol from sample collection to DNA amplification optimised for JFFS fur seal faecal samples. Here, I looked at the most effective sample collection and storage methods considering both mtDNA and nuclear DNA. This research is a significant contribution to the study of the JFFS. The optimised protocols will provided an effective and efficient strategy for carrying on the necessary genetic studies on the JFFS.

5.2 Results

5.2.1 Testing sample storage and DNA extraction kits

Host DNA can be extracted from faecal samples. However, quantity and quality tend to be poor. I therefore optimised different methodologies for collecting, storing and extracting samples for this work. Swab samples were collected from three dog faeces as described

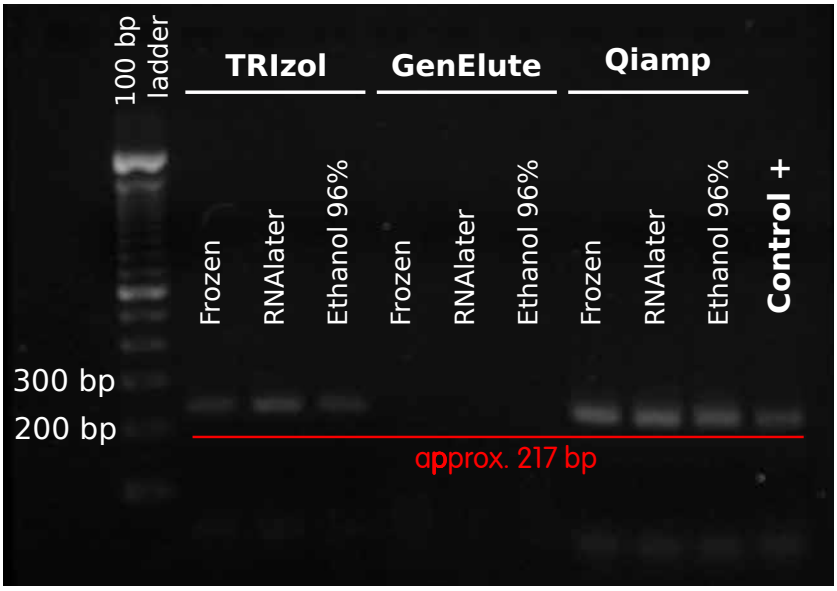

Figure 5.1. Testing storage and extraction methods of canine DNA obtained from faecal samples.
The storage of swab samples from the external surface of canine faeces by three different methods was compared by each of three DNA extraction methods. 2 % Agarose gel electrophoresis with PCR product from the rs22196038 locus located in the canine chromosome 12 (primers Chr16_SNP) in canine genomic DNA extracted from faecal samples is shown. The predicted size of the amplicon is 217 bp. The positive control was a DNA template extracted from a blood sample provided by Dr. Francisca Coddou.

in section 2.6 to test the performance of different storage methods (freezing, RNAlater and ethanol) and DNA extraction kits to conduct host genetic studies from faecal samples. Initially, four extraction kits were going to be tested. However, early technical problems with the Quick DNA Fecal/Soil Microbe MiniPrep Kit (Zymo Research) resulted in excluding this kit from the analysis. Furthermore, the filter in the Zymo-Spin III columns came off in all the tubes when loading the supernatant from the previous step. All dog faeces DNA extractions were PCR amplified using an in-house primer set targeting a nuclear region containing the rs22196038 locus located in canine chromosome 12. The expected PCR product size was 217 bp. Primers, Chr12_SNP, were designed and kindly provided by Dr. Francisca Coddou. She also provided a DNA template extracted from a random, unidentified dog blood sample to use as a positive control. Finally, PCR products were visualised on gel electrophoresis.

No bands were observed for any of the storage methods when using the GeneAll kit. On the other hand, samples extracted with Trizol or with the QIAmp DNA stool mini kit (from now on referred to as Qiamp kit) gave positive results for all three storage methods (Fig. 5.1). I decided to use RNAlater to collect the JFFS samples due to ease of use in the field. Additionally, as

mentioned in section 2.6.1, I finally extracted DNA using both Qiamp and Trizol (see section 5.2.3).

5.2.2 DNA Amplification

To explore the potential of JFFS faecal samples to conduct population genetic studies and sexing of samples reliably, one primer set targeting the left domain of the mtDNA control region, five primer sets targeting microsatellite loci, one primer set targeting the Zinc Finger X/Y (ZFX/Y) gene and a primer set targeting the SRY gene were used. From the five microsatellites loci targeted in this study, three (Agaz1, 2 and 3) were neutral while Agt10 and Agi5 where linked to the immune system (CXCL10 and CD44 respectively). In this section I will show the overall optimisation results. However, specific information regarding the amplification of the mtDNA control region and both ZFX/Y and SRY, can be found in sections 5.2.3 and 5.2.4 respectively.

Briefly, to increase the DNA amplification success in faecal samples, amplifications were carried out in a two step process. First, an external PCR was used to enrich the samples with the target DNA, followed by an internal PCR to increase the sensitivity of the assay. The internal PCR was either a nested (microsatellites) or a semi-nested (mtDNA, PinZFX/Y and SRY) protocol. Apart from the SRY primer set, all primer sets used for running internal PCRs were previously published. Supplementary table A.3 shows a description of the primers including the amplicon fragment size, the optimised annealing temperature and the reference.

To test and optimise the DNA amplification assays, I began by checking that the published primers produced PCR products on JFFS tissue sample DNA. For each PCR assay, a 2 % agar gel electrophoresis was used to check the PCR products sizes were in accordance with the respective predicted size. These products were purified and Sanger sequenced to confirm the primers amplified the correct locus. In the case of the microsatellites, sequences were clear into the repeat element, but the sequence quality decreased dramatically towards the end of the element, probably due to polymerase slippage (Fig. 5.2). The fragments therefore had to be sequenced in both directions. This gave specific JFFS sequence from the regions flanking the repeat element.

After target confirmation, tissue-amplified products were used as positive controls. Until this point, I set each annealing temperature as calculated on the online NEB Tm calculator (https://tmcalculator.neb.com). Next, a temperature optimisation was performed as described in section 2.6.3 (Fig. 5.3). Once the optimal annealing temperature for each published primer set was identified, I tried amplifying nuclear DNA from faecal samples (Fig. 5.4). First, I compared PCR products obtained from two different templates of the same sample (17JFFS6). The templates differed in that one was extracted from a faecal fragment while the other was extracted from a faecal swab. These DNA amplifications were performed in a single 40 cycle

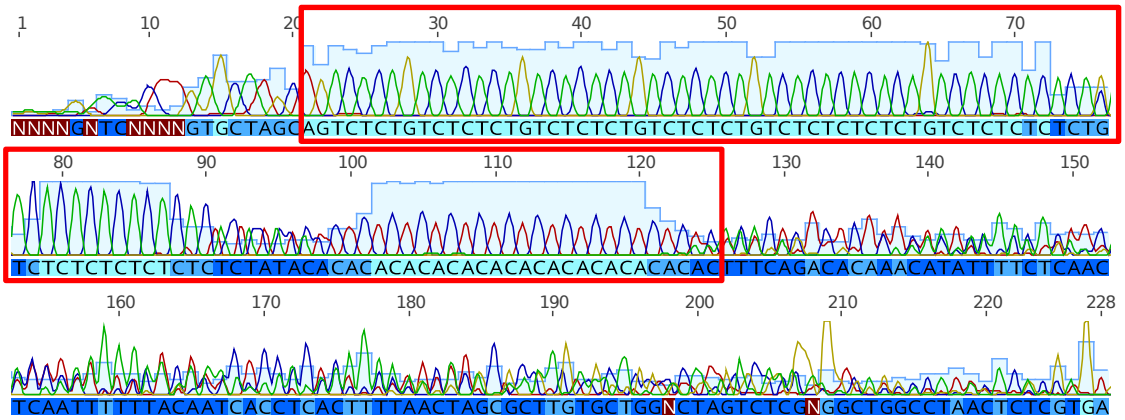

Figure 5.2. Chromatogram showing a segment of the internal Agaz-1 Sanger sequence amplified from a JFFS tissue sample.
The internal forward primer was used for sequencing. The repetitive motif is shown within the red rectangles. Expected motif: $(CT)_2GT(CT)_3GT(CT)_3GT(CT)_3GT(CT)_5GT(CT)_{12}TA(AC)_{15}$ (Hoffman, 2009).

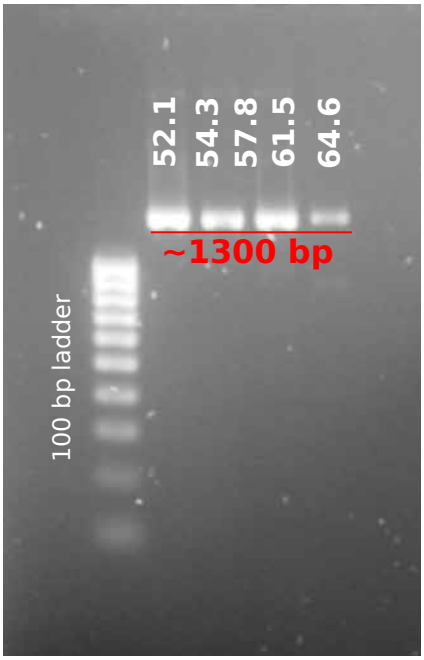

Figure 5.3. Temperature gradient to find the optimal annealing temperature for amplifying the mtDNA control region.
The 2 % agarose gel shows the PCR products generated for each tested temperature. PCRs were performed using the external primer set (T-Thr and T-Phe). The TrackIt™ 100 bp DNA Ladder, Invitrogen, was used here.

PCR run using the internal Agaz-1 primers. PCR products were checked using 2 % agarose gel electrophoresis. Overall, the DNA template extracted from a faecal fragment evidenced a larger number of non-specific PCR products (Fig. 5.4A). However, no bands matching the expected product size were observed when using either template (Fig. 5.4 A and B). Next, I wanted to see if performing a second 40 cycle PCR run, using the PCR products as a template, would increase the chances of success. Therefore, I input the PCR product generated from the swab sample into a new PCR reaction. Unfortunately, while loading the second round PCR product, I spilt some into the immediately adjacent wells contaminating the positive control and the first round PCR product. Despite the contamination, the 2 % agarose gel electrophoresis in Fig. 5.4C shows a band matching the expected product size when running the second PCR (80 cycles in total). However, non-specific products can also be observed. Furthermore, the specificity of the second round PCR reaction gel-purified product was confirmed with Sanger sequencing.

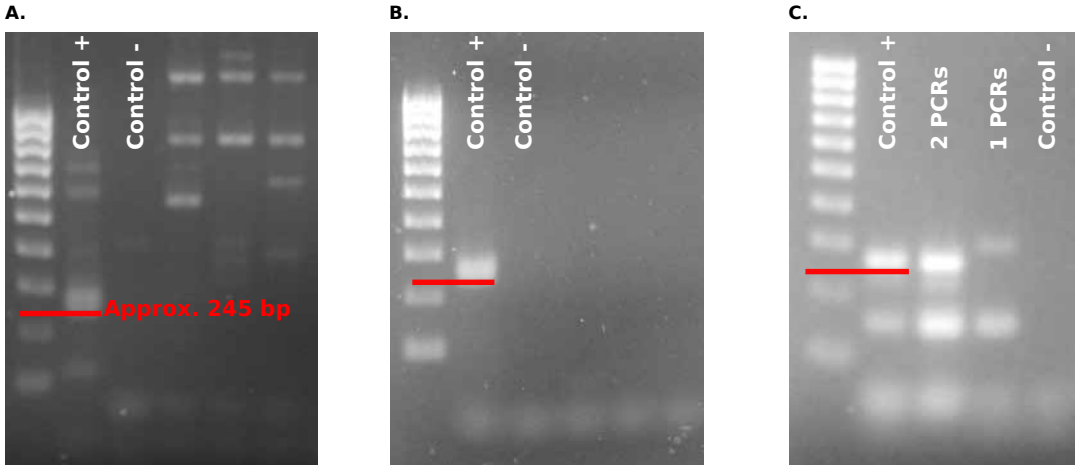

Figure 5.4. First trial for amplifying the Agaz-1 microsatellite from a faecal sample (17JFFS6) using published primers.
A) PCR amplification of DNA extracted from a faecal fragment. B) PCR amplification of DNA extracted from faecal swab. C) 2 % agar gel showing the effects of using PCR product as template to perform a second PCR round. The TrackIt 100 bp DNA Ladder, Invitrogen, was used here

Motivated by this result, I designed external PCR primers for each of the primer sets targeting microsatellite loci to try to increase the sensitivity, specificity and, therefore, the nuclear DNA amplification success by performing nested PCRs (see Supplementary Table A.3 and Appendix B for more details). In the case of the mtDNA control region, instead of developing new primers, I used the internal reverse primer (SCR) developed by Hoelzel *et al.* (1993) to carry out semi-nested PCRs (Fig.5.5). Similar to the optimisation protocol applied

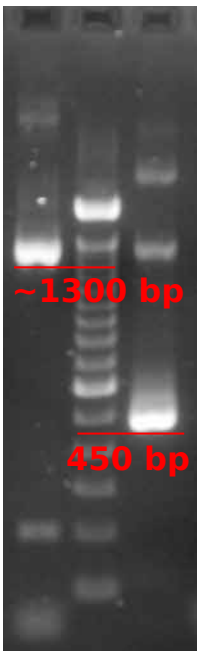

Figure 5.5. First semi-nested PCR amplification of the mtDNA control region from a tissue sample.
The 2 % agarose gel shows the difference between the external (left track) and internal (right track) product size. The primers T-Thr and T-Phe were used for the external reaction and T-Thr and SCR for the internal one (predicted product size 450 bp). The Ready-load 100 bp DNA Ladder, Invitogen, was used here (middle track).

to the internal primers, after confirming the specificity of the new primer pairs through gel electrophoresis and Sanger sequencing of tissue DNA PCR products, a temperature gradient assay was performed to find the optimal annealing temperature for each pair.

The next step was to perform nested PCRs (still on tissue samples) to check the performance of the nested PCR protocol on nuclear DNA amplification. I also tested different internal primer pair combinations: internal forward and internal reverse (nested), internal forward and external reverse, and external forward and internal reverse (semi-nested). However, regardless of the internal primer combination used, there were always multiple bands identified in the agarose gel electrophoresis of the internal PCR product when using the outer PCR product directly as a template for the internal reaction amplification. Furthermore, depending on the primer combination (nested or semi-nested), three or two bands, respectively, were observed (Fig. 5.6). This noise was most likely introduced by reagent leftovers present in the internal PCR template. For this reason, I compared the effects of purifying versus not purifying the external PCR product before using it as a template for the internal PCR (Fig. 5.6). This assay showed that purifying the external PCR product reduces the amount of noise observed in the internal PCR.

However, product purification was not the best alternative when using more sensitive methods such as qPCR (see section 5.2.4).

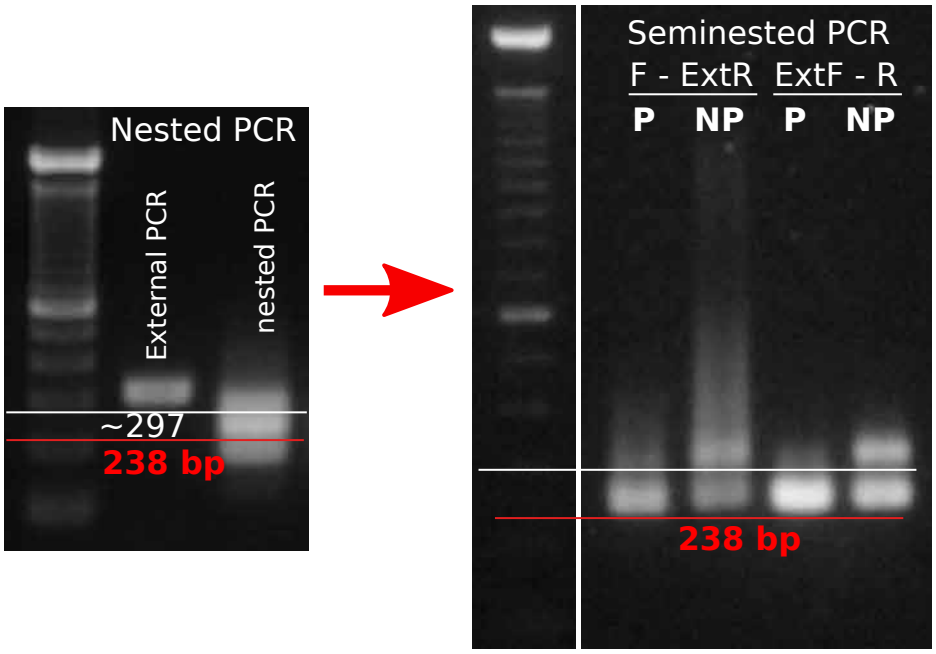

Figure 5.6. Agaz8 nested PCR Optimisation.
The effects of purification versus non-purification of the external PCR product before using it as a template for the internal PCR reaction were evaluated to control the noise introduced by reagent leftovers from the external PCR reaction. 2 % Agarose gel electrophoresis from PCR amplification of the Agaz8 microsatellite locus is shown. The letter P in the right figure indicates purified product and NP means non-purified product. The white line indicates the approximated size of the external PCR product. The Ready-load 100 bp DNA Ladder, Invitrogen, was used here.

Once the PCR optimisations were completed, I proceeded to perform DNA amplifications on the faecal samples. There are two important considerations when targeting host DNA from faecal samples. First, faecal samples contain high concentrations of PCR inhibitors such as bile salts and complex polysaccharides. Second, this sample type tends to be enriched with high concentrations of non-host DNA that are likely to interact with primers designed to target host DNA as I demonstrated earlier in this section. Despite using specialised DNA extraction kits such as the Qiamp kit used in this study, PCR inhibitors are likely to persist in the extract to some degree. Thus, diluting the DNA extract contributes to limiting the concentration and, therefore, the detrimental effects of PCR inhibitors. On the other hand, due to often low concentrations of host DNA, it is easy to over-dilute samples, especially affecting the concentration of nuclear DNA. Therefore, I first tried using a 1:4 dilution series of PCR template

starting at 200 ng (Fig. 5.7). However, I quickly realised that finding a dilution that would have worked for all the samples was unachievable due to the high variability between samples in terms of DNA content. Thus, I decided to run all the external PCRs using a 1:4 dilution of DNA template regardless of the stock DNA concentration. Furthermore, concentrations were adjusted only for those samples for which the internal PCR amplification failed.

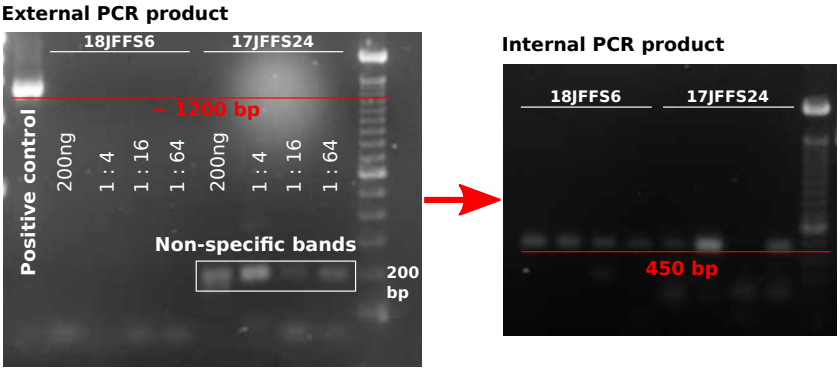

Figure 5.7. Agarose gel electrophoresis of external and internal PCR products of mtDNA amplification from JFFS faecal samples.
The external PCR was performed using a 1:4 dilution series of DNA templates. Primers T-Thr and T-Phe were used in the external reaction and T-The and SCR for the internal one. Ready-load 100 bp DNA Ladder, Invitrogen, was used.

Finally, I conducted PCR amplification in all 72 faecal samples by first targeting the mtDNA control region using the external primers (T-Thr and T-Phe). Due to the large copy numbers of mtDNA per cell, the amplification of this target was fairly successful with only one 40 cycle PCR run. Because the external products were sequenced using the internal SCR reverse primer, non-specific bands did not interfere with the sequencing as long as enough mtDNA was amplified to cover the Sanger sequencing DNA concentration requirements (see section 5.2.3. Thus, I only carried out a semi-nested PCR for this target if no product was observed in the gel electrophoresis or if the Sanger sequencing was ambiguous. Nuclear DNA amplifications were performed only on samples from which mtDNA sequences were retrieved (Fig. 5.8). Unfortunately, the microsatellite analysis has not been completed yet. The final analysis was going to be completed in Germany as part of a collaboration. However, due to COVID-19 and later Brexit I was unable to travel to Germany to do this part of the work. Therefore, samples are still in the Department of Veterinary Medicine waiting to be analysed.

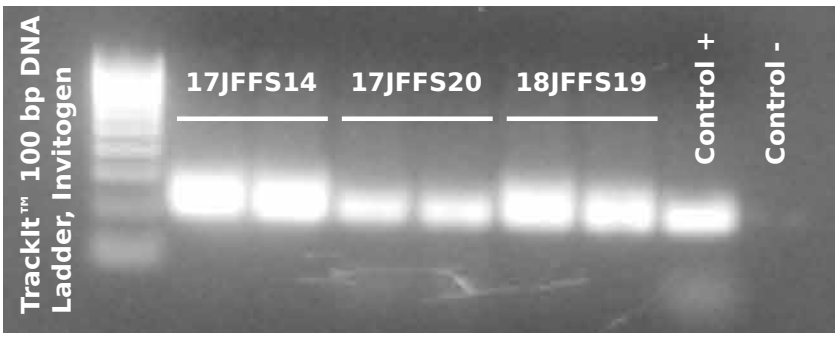

Figure 5.8. Agarose gel electrophoresis internal PCR products of Agaz1 microsatellite amplification from JFFS faecal samples.

5.2.3 MtDNA control region analysis

Extraction and amplification of JFFS mtDNA

This analysis targeted the left domain (5' end) of the mtDNA control region, specifically the area amplified and sequenced using the primer pair T-Thr and SCR (Fig. 5.9). A total of 72 DNA extracted samples were initially included in the mtDNA control region PCR amplification. When comparing the extraction kits used to obtain DNA from JFFS faecal samples, 83.7 % of the samples extracted with the Qiamp kit (N = 48) were used to amplify host DNA successfully.

Successfully amplified host DNA samples were defined based on the gel electrophoresis results (Fig. 5.5) and Sanger sequencing. Sequences matching fully or partially the reference sequence, in this case, one of the previously JFFS sequences published by Goldsworthy *et al.* (2000) (Supp. Tab. A.4), were considered successful. For instance, the bottom mtDNA control region sequence shown in Fig. 5.10 gave a slightly ambiguous sequence, although it was sufficiently readable to check that it matched the reference sequence to show the PCR was specific.

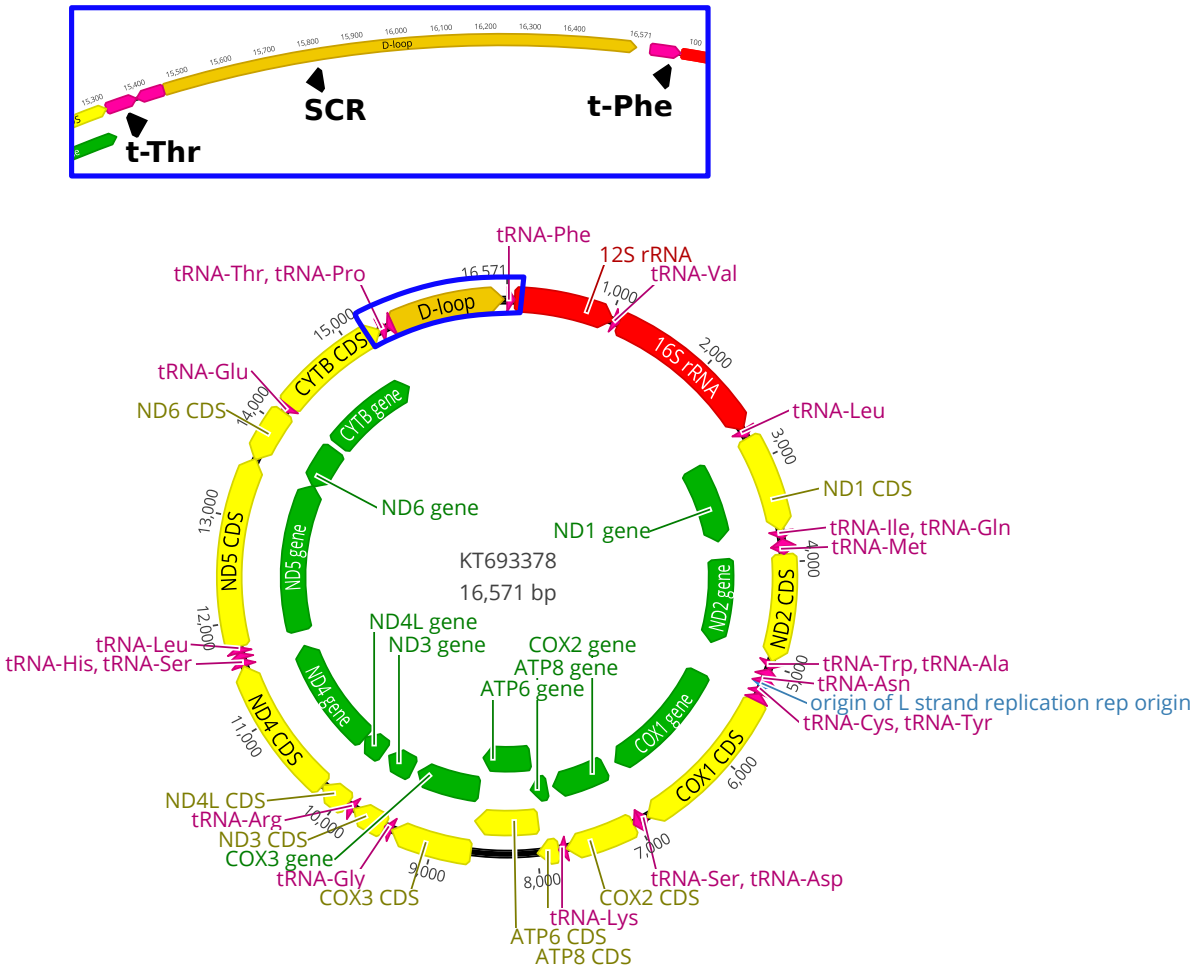

Figure 5.9. Schematic fur seal mtDNA representation.
The figure was generated using an annotated New Zealand fur seal (*Arctocephalus forsteri*) complete mitochondrion genome downloaded from Genbank (accession number KT693378). A blue rectangle delimits the segment targeted by the external primer pair. The position and direction of the primers are indicated with black arrowheads. The focus of this analysis was the area targeted by primer pair T-Thr and SCR.

Figure 5.10. Chromatogram showing a segment of mtDNA control region Sanger sequence.
Both sequences belong to the same sample (17JFFS18). The lower sequence is the first sequence obtained from the sample PCR product using primers T-Thr (forward) and T-Phe (reverse direction) and sequenced with the internal SCR-1 primer (reverse direction). The upper sequence was from the product of a semi-nested PCR using T-Thr and SCR-1 sequenced with the SCR-1 primer. The blue shading indicates the quality of the sequence read; the lighter the blue, the higher the quality. Sequences shown here were reverse complemented.

Due to the termination of the product line, those Qiamp extracts that failed to amplify host DNA, in addition to 16 extracts that did amplify host DNA but gave ambiguous sequences, were reextracted with TRIzol. In the end, 48 samples were extracted with TRIzol, including those 24 samples that needed to be re-extracted. Thus, 93.8 % of the samples extracted with Trizol showed positive mtDNA amplification. However, despite reprocessing, ten samples remained ambiguous, and eight failed to amplify any DNA. Therefore, only 54 (75 %) samples gave sufficiently good quality sequence to proceed with the sequence analysis.

Levels of polymorphism

In the end, 68 sequences, including the 54 faecal samples, the tissue control and 13 haplotype sequences previously published for this species were included in the sequence analysis (Goldsworthy *et al.*, 2000). The Clustal Omega alignment (all gaps present) was 295 bp long. However, the first 28 bp of the consensus sequences were part of the 3' half of the $tRNA^{pro}$. Sixty-five segregating sites were identified, of which 52 were nucleotide substitutions. Almost all substitutions were transitions (A/G or T/C). Furthermore, only two trans-versions were detected (A/T), one in consensus position 56 (sample 18JFFSB2) and one in position 188 (18JFFS2) (table 5.1). For this analysis, the alignment was modified by removing some of the gap sites as described in section 2.6.4. The modified alignment resulted in a 289 bp long consensus sequence with 59 polymorphic sites, of which 34 were Parsimony-Informative Sites (PIS) (11.8 % of all the sequence sites). PIS refers to positions with at least two different characters in which any of these characters are present in more than two sequences.

Initially, 27 haplotypes were identified. However, this number was reduced from 27 to 26 after the gap treatment. Furthermore, six consecutive polymorphic sites were excluded due to gap content from positions 99 to 104 of the original consensus sequence. I decided to retain nine sites with gaps because the gaps were present in only one haplotype per site. Thus, it was not worth excluding the entire column. As a result, the original haplotype 26 became part of haplotype 25 and haplotype 27 was renamed haplotype 26 (table 5.1). From the 26 identified haplotypes, fourteen were newly discovered. Only seven from the thirteen previously identified haplotypes were found again in this study (H4, H12, H14, H19, H23 and H24). Three haplotypes previously found only in Alejandro Selkirk island (AS), the only island not included in this study, were now found in RC island (H4, H24 and H25). All three were within the most prevalent haplotypes in the dataset.

Table 5.1. Juan Fernandez fur seal mtDNA control region haplotypes.
The variable positions in each haplotype (hap) are shown. Position numbers are based on the 295 consensus sequences. The first column indicates the haplotype name given in this study. The second column shows the haplotypes previously identified in Goldsworthy *et al.* (2000). Parentheses show the name on genbank (APHI1 to APHI5, accession numbers from AF384403.1 to AF384407.1). Columns in grey were deleted for further analysis to limit the bias introduced by the gaps. Row in grey became the same haplotype after gap treatment. Haplotype 26 was included in H25 and H27 was renamed H26.

hap name	Old hap name	13	37	57	90	94	95	96	97	98	99	100	101	102	103	104	105	106	107	110	114	120	121	132	136	138	140	164	165	166	167	169	170	171	173
H1		G	-	T	C	C	C	T	T	-	-	-	C	T	T	T	T	T	T	T	T	C	G	G	T	T	C	A	C	A	A	G	G	T	G
H2		.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	C	.	.	.	.	.	.	.	.	.	G	T	.	.	.	A	.	A
H3	LH190 (APH4)	.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	C	.	.	.	.	.	.	.	.	.	G	T	.	.	.	A	.	A
H4	ETGOGO (APH2)	.	A	A	T	-	-	-	-	.	.	.	-	-	-	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	G	.	A	.	.
H5	RC274	.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H6		.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	T	A	A	C	.	.	.	G	G	.	A	.	.
H7		.	A	A	.	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	A	.	.
H8	LH191	.	A	A	.	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	A	.	.	.
H9		.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	.	A	A	.	.	.	.	T	.	.	A	.	.
H10		.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	C	C	.	.	.	A	.	.	.	.	.	.	.	G	.	A	.	.
H11		.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	.	A	A	.	.	.	.	.	.	G	.	A	.	.
H12	ETAllred	.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	A	.	.
H13		.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	A	.	.
H14	RC280	.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	A	.	.
H15	RC276	.	A	A	T	-	.	.	.	T	.	.	-	-	-	.	.	.	C	.	.	T	A	.	.	-	-	.	T	.	.	.	.	-	.
H16		.	A	A	T	T	.	.	.	T	.	.	-	-	-	-	-	.	.	.	.	.	.	A	.	.	.	.	T	G	.	.	.	.	.
H17	LH197	.	A	A	T	T	.	.	.	T	T	T	T	.	.	.	.	C	C	C	C	T	A	A	.	.	.	G	T	.	.	.	.	.	.
H18		.	A	A	T	T	.	.	.	T	.	.	-	-	-	-	.	.	.	.	.	.	A	.	.	.	.	.	.	.	G	.	.	.	.
H19	LH194	.	A	A	T	T	.	.	.	T	.	.	-	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.
H20		.	A	A	T	T	.	.	.	T	.	.	-	-	-	-	.	.	C	.	.	T	A	A	.	.	.	.	.	.	.	.	.	.	.
H21		A	A	A	T	T	.	.	.	T	.	.	-	-	-	-	.	.	C	.	.	T	A	.	.	.	.	.	.	.	.	.	.	.	.
H22		A	A	A	T	T	.	.	.	T	.	.	-	-	-	-	.	.	C	.	.	T	A	A	.	.	.	.	.	.	.	.	.	.	.
H23	ETGaz (APH5)	A	A	A	T	T	.	.	.	T	.	.	-	-	-	-	.	.	C	.	.	T	A	A	.	.	.	.	.	.	.	.	.	.	.
H24	ETEli (APH1)	A	A	A	T	T	.	.	.	T	.	.	-	-	-	-	.	.	C	.	.	T	A	A	.	.	.	.	.	.	.	.	.	.	.
H25	ETPaul	.	A	A	T	T	.	.	.	T	.	.	-	-	-	-	.	.	C	.	.	T	A	.	.	.	.	.	T	.	.	.	.	.	.
H26 (25)		.	A	A	T	T	.	.	.	T	.	.	-	-	-	.	.	.	C	.	.	T	A	.	.	.	.	.	T	.	.	.	.	.	.
H27 (26)	ETHoley (APH3)	.	A	A	T	T	.	.	.	T	T	.	-	-	-	.	.	C	C	.	.	T	A	.	.	.	.	.	T	.	.	.	.	.	.

Table 5.2. Location and frequency of JFFS mitochondrial control region haplotypes. Each column within the delimited area indicates the frequency in which each haplotype was observed in each location. The final row summarises the number of locations in which each haplotype was identified.

Location (n)	H1	H2	H4	H6	H7	H9	H10	H11	H12	H13	H14	H16	H18	H19	H20	H21	H22	H23	H24	H25	haplotypes per location
BP (15)	0	0	1	0	1	0	0	2	1	2	1	1	1	1	0	1	0	0	0	3	11
EA (7)	0	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1	1	7
PI (4)	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	4
SC (12)	0	0	2	1	0	0	0	2	2	0	1	0	1	0	0	0	0	0	2	1	8
TB (16)	0	1	1	0	1	1	1	0	1	0	0	0	1	0	1	0	1	4	2	1	12
V (1)	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Total freq	1	1	4	1	3	3	1	4	5	4	2	1	3	1	1	1	1	6	5	7	
% freq	1.8	1.8	7.3	1.8	5.5	5.5	1.8	7.3	9.1	7.3	3.6	1.8	5.5	1.8	1.8	1.8	1.8	10.9	9.1	12.7	
n locations	1	1	3	1	3	3	1	2	4	3	2	1	3	1	1	1	1	3	3	5	

As expected, locations with more samples showed more haplotypes (Table 5.2), but when looking at the haplotype per number of samples ratio, this remained consistent across locations (around 0.7). The most widely distributed and prevalent haplotype was H25 (N = 7), present in five out of six locations and found in 13 % of the samples. Furthermore, the only location where this haplotype was not found was V, which only had one sample in the dataset. Interestingly, three out of seven samples with this haplotype were collected in BP, where this haplotype was also the most prevalent within this location. Another case of haplotype domination in the otherwise even-looking dataset was H23. Out of the six samples where this haplotype could be found, four (67.5 %) were collected in TB. Moreover, this haplotype was present in one-third of the samples collected in this location. Previously, H23 was the most prevalent haplotype in Alejandro Selkirk island (AS) and was only present in one out of the nine samples collected in TB by Goldsworthy *et al.* (2000).

Finally, the nucleotide diversity (π) remained consistent at 0.03 across locations and across studies. The overall haplotype diversity (H) was 0.927, slightly higher than the one previously reported. However, when comparing H from TB between studies, the values were almost the same. Tajima's D was seven times smaller than in the previous report (D = -0.317 and -0.048, respectively), suggesting population expansion. However, D did not reach significance (corrected P = 0.793).

MtDNA Control Region Phylogenetic analysis

The JFFS mtDNA control region unique sequences (haplotypes) derived from this study and those published by Goldsworthy *et al.* (2000) in addition to other Arctocephalus sequences

available on Genbank were used to perform a phylogenetic analysis (Supplementary Table A.4. Despite the poor branch support, the phylogenetic tree looked very similar to the one reconstructed 20 years ago (Goldsworthy *et al.*, 2000). JFFS mtDNA control region haplotypes clustered in two major haplogroups (Fig. 5.11). Furthermore, the majority of the most frequent haplotypes (H23, H24, H25) were all members of the same clade. The first clade from top to bottom in Fig. 5.11 only included four out of the thirteen previously identified haplotypes, and two of them were only found in samples from AS.

Next, I performed two haplotype network analyses, one using only the sequences generated in this study (Fig. 5.12) and a second one including the dataset generated by Goldsworthy *et al.* (2000) which included samples collected from AS island (Fig. 5.13). Both haplotype networks are in close agreement with the clustering observed in the phylogenetic tree. Overall, the network suggests H12 as a common ancestor to all the haplogroups. Interestingly, H12 was commonly found in TB 20 years ago but in this study it was found in only one sample out of eleven collected from the same location. However, this haplotype was the second most disperse haplotype (Table 5.2). In addition, haplotypes previously found in AS dominated the node cluster emerging to the left and above H12, while the cluster on the right mainly contained haplotypes only observed in this study. Finally, the analysis of molecular variance (AMOVA) did not evidence any significant genetic differences between islands or locations tested in this study.

5.2.4 Sexing of faecal samples

Due to profound differences between adult fur seal females and males, sex determination can provide critical information to study these populations. For instance, in the context of this study, patterns observed in the faecal microbiome or exposure risks to environmental pollutants may be influenced by the individual's sex. However, one of the limitations of working with faecal samples is the difficulty of linking individuals with samples which would allow the identification of visual queues for sex determination. To overcome this limitation, I decided to apply molecular techniques for sex determination directly from the faecal samples.

After reading the literature, I decided to base the PCR assays for sexing the JFFS samples on the paper by Robertson *et al.* (2018). This paper used a multiplex real-time PCR assay to amplify a region of the SRY gene to detect males and a region of the ZFX/Y region present in both males and females as an amplification control. The real-time PCR assay used the melt curve peak of the two products to differentiate between males and females: males have two peaks (Tm ca. 84 ° C and 88 ° C) and females one peak (Tm ca. 84 ° C).

Despite multiple trials, the SRY primer set (SRY53-3c and SRY53-3d) developed by Fain and LeMay (1995) and used by Robertson *et al.* (2018) failed to amplify DNA from tissue

samples. I therefore performed a search in GenBank for phocid and otariid SRY region sequences and aligned these and the primers above. The Clustal Omega alignment identified four nucleotides at the 5' end of the original reverse primer that did not match any of the sequences (Fig. 5.14). This mismatch was six nucleotides in the otariid sequences.

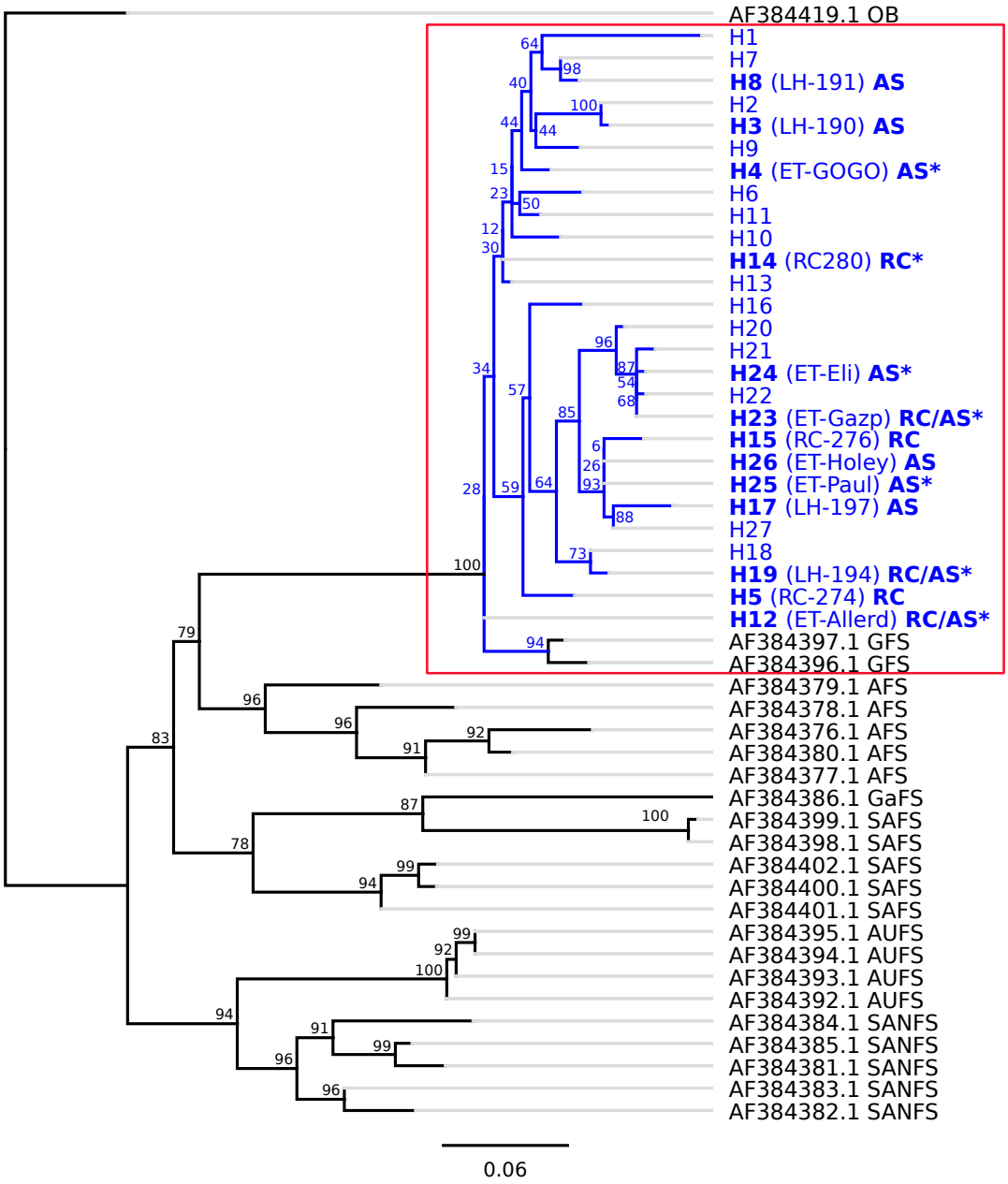

Figure 5.11. Phylogenetic reconstruction of the genus Arctocephalus with an emphasis on the Juan Fernandez fur seal (JFFS).
The area delimited in red highlights the clade shared by the JFFS and the Guadalupe fur seals (GFS). JFFS haplotypes are in blue. Haplotypes in bold indicate those previously identified by Goldsworthy *et al.* (2000), the original names are in parenthesis. Based on this previous report, samples found only in Alejandro Selkirk island can be identified by the AS letters, RC were found in Robinson Cruose island and RC/AS were found in both. The asterisk indicates those haplotypes found in the current study. Other species included in the phylogenetic reconstruction were: Antarctic fur seal (AFS), Galapagos fur seal (GaFS), South American fur seal (SAFS), Australian fur seal (AUFS) and Subantarctic fur seal (SANFS). The tree was rooted using a Southern seal lion haplotypes (OB).

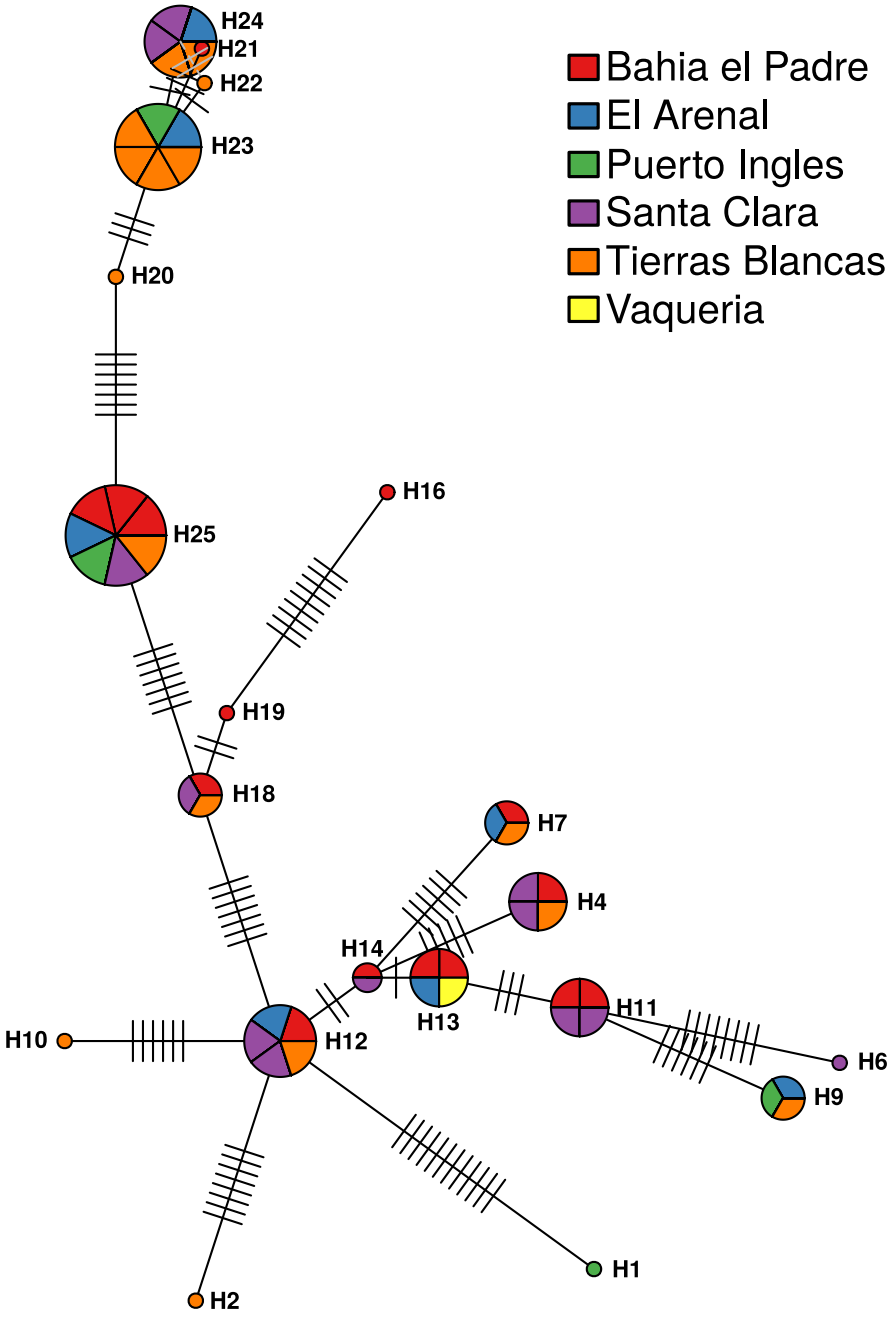

Figure 5.12. Haplotype network of the mtDNA control region using data generated in this study.
Each node represents a haplotype. The size of the node is relative to the number of times the haplotype was detected and each segment within the node represents one individual. The length of the lines connecting nodes is relative to the genetic distance between haplotypes and each cross through the line indicates a nucleotide difference. The grey crossed line connecting H22 and H24 indicates an alternative link.

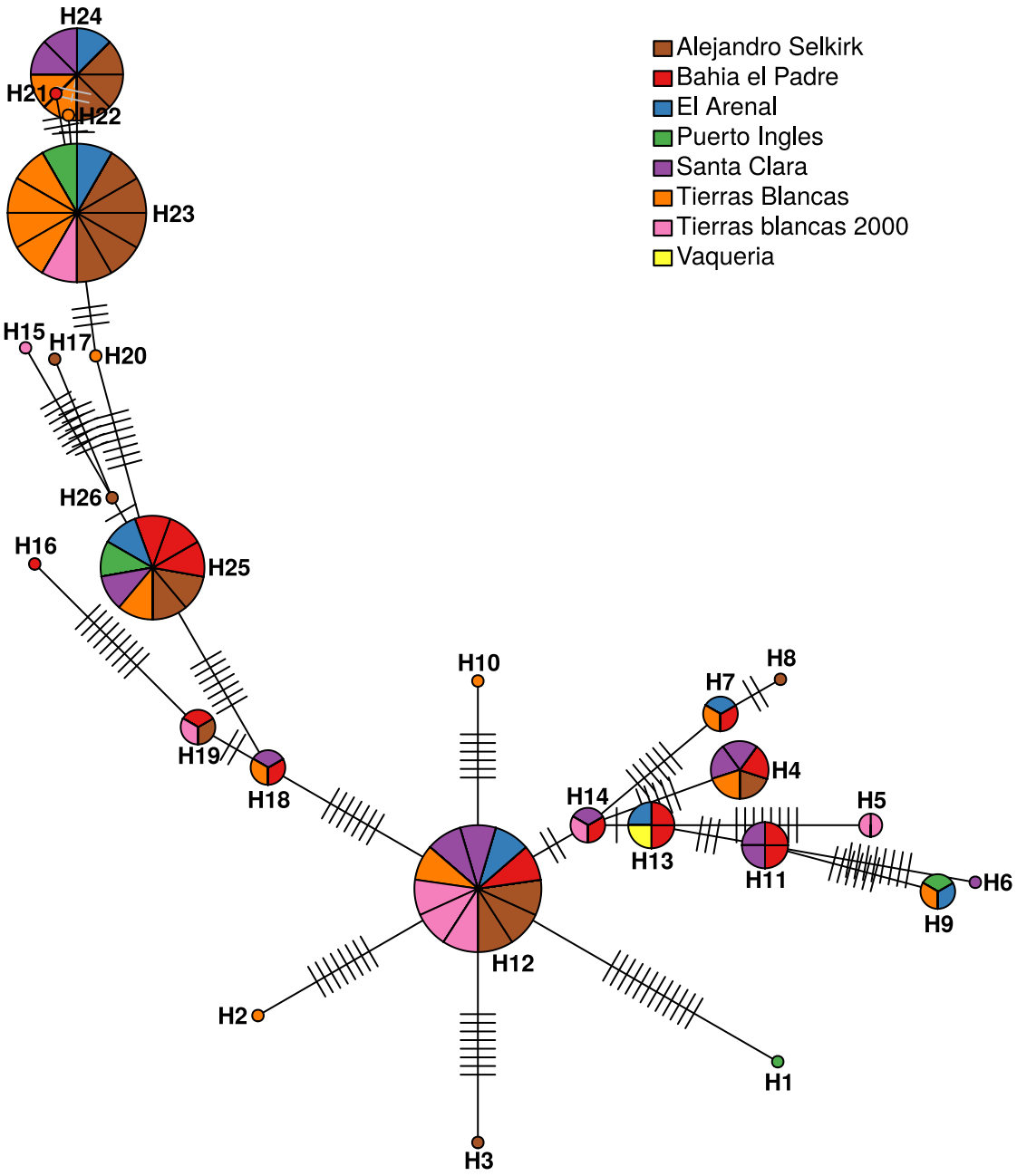

Figure 5.13. Haplotype network of the mtDNA control region combining data generated in this study and in Goldsworthy *et al.* (2000).
For more details refer to the caption in fig. 5.12

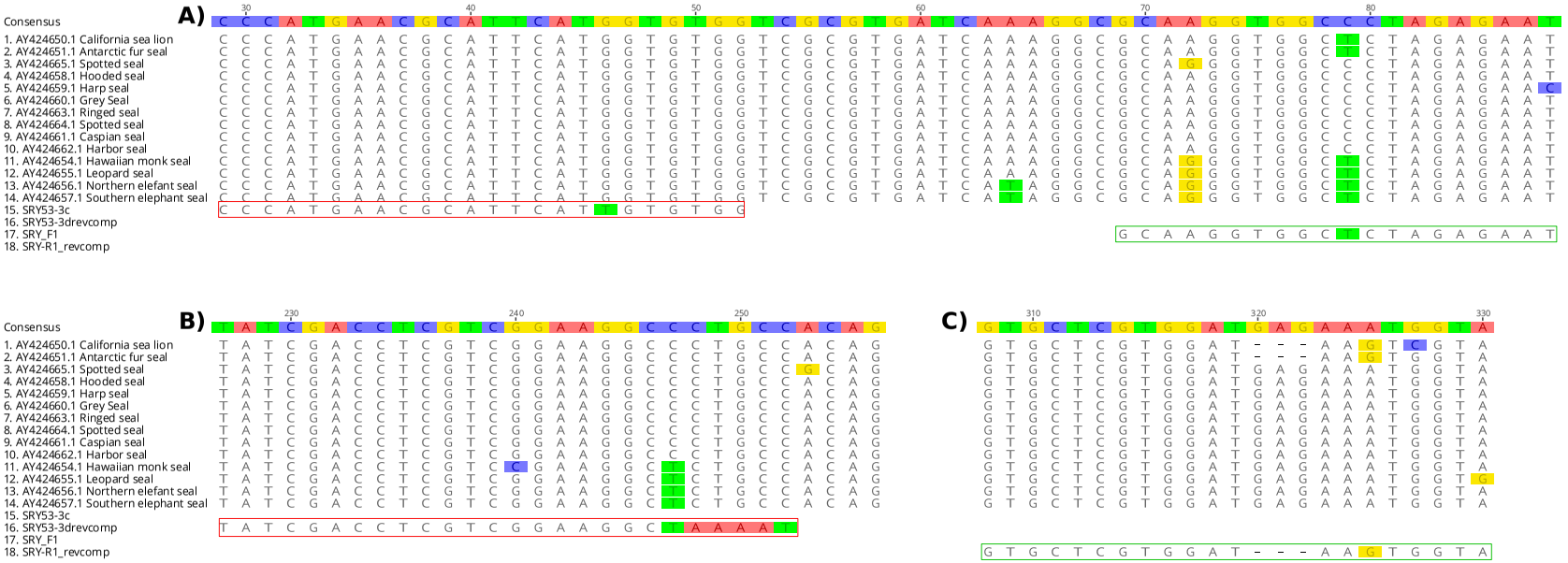

Figure 5.14. Pinniped SRY sequence alignment.
Three segments of the Clustal Omega alignment using otariid and phocid SRY sequences (Accession numbers in Supplementary Table A.4) were extracted for this figure. The first two sequences belong to Otariids. The remaining sequences belong to phocids. The positions of PCR primers are indicated. A) The upper alignment shows the area where the forward primers were located; B) and C) indicates the reverse primers. Rectangles in red highlight SRY53-3c and d. Primers developed in this study are highlighted in green (SRY-1). The numbers on top indicate the position in the consensus sequence.

This mismatch is likely to explain the amplification failure. Primer3 was used to design new primers using sequence AY424651.1 from AFS. This was done by limiting the primers to the coding region of the gene, and to give a primer Tm similar to those for the SRY PCR assay primers used in Robertson *et al.* (2018) (Tm optimum 57 ° C) and to try to maintain a qPCR product melt curve peak $\geq$ 88 ° C. This gave primers SRY-F1 and SRY-R1 (Supplementary Table A.3) and these were used successfully to amplify specific PCR product from JFFS tissue samples (Fig. 5.15A). These products were purified and sequenced using both PCR primers and the derived sequence is shown in Fig. 5.15B. The ZFX/Y PCR using primers PinZFY-forward and PinZFY-reverse successfully amplified JFFS tissue sample DNA to give a ca. 169 bp product (Fig. 5.15A) which was shown to be specific when sequenced (Fig. 5.15C).

As described in section 5.2.2, external primers were also developed for both SRY-1 and PinZFY primers. The same AFS sequence used to develop SRY-1 was also used to design a third SRY primer (SRY-2). SRY-2 is an external forward primer designed to be used in a semi-nested PCR with SRY-1 reverse. This PCR gave specific product of the correct size (378 bp) which when sequenced was shown to be specific for the correct region of DNA (data not shown). Finally, a blast search using known ZFX sequence identified the region 1901910 to 1902817 within the whole genome sequence of closely related AFS (accession number UIRR01000042). This was used to design a semi-nested PCR primer (Ext_pinZFY) for use with PinZFY-forward (302 bp ca. product size) (data not shown). Products from both external PCRs were purified together and used as a positive control in the following assays.

Next, a qPCR using the SRY-1 and the previously published PinZFY primers was performed to confirm the melting temperature of the two amplicons differed enough to identify the separation of the melting curves. The assays were carried out independently followed by a multiplex trial. The positive control was used as template. The melt curve for both amplicons peaked at lower temperature than those reported by Robertson *et al.* (2018), 80 ° C for PinZFY and 84 ° C for SRY-1. However, the separation between peaks was evident (Fig. 5.16).

The molecular sexing part of this analysis is still ongoing. The real-time PCR protocol required considerable optimisation. For instance, when used at the suggested primer concentrations for the PCR kit, the SRY primer PCR was particularly efficient at amplifying DNA and suppressing the signal generated by the PinZFY primer PCRs (Fig. 5.16A and C). Different concentrations of SRY and PinZFY primers were tested to overcome this problem. Trials showed the best combination was using 150 and 300 nM respectively per sample in a 15 ul reaction.

The next problem encountered was PCR contamination, especially for the SRY product, which kept being picked up by negative controls (Fig. 5.17). It was decided not to purified the external PCR products before the internal semi-nested PCR reaction to decrease the risk of

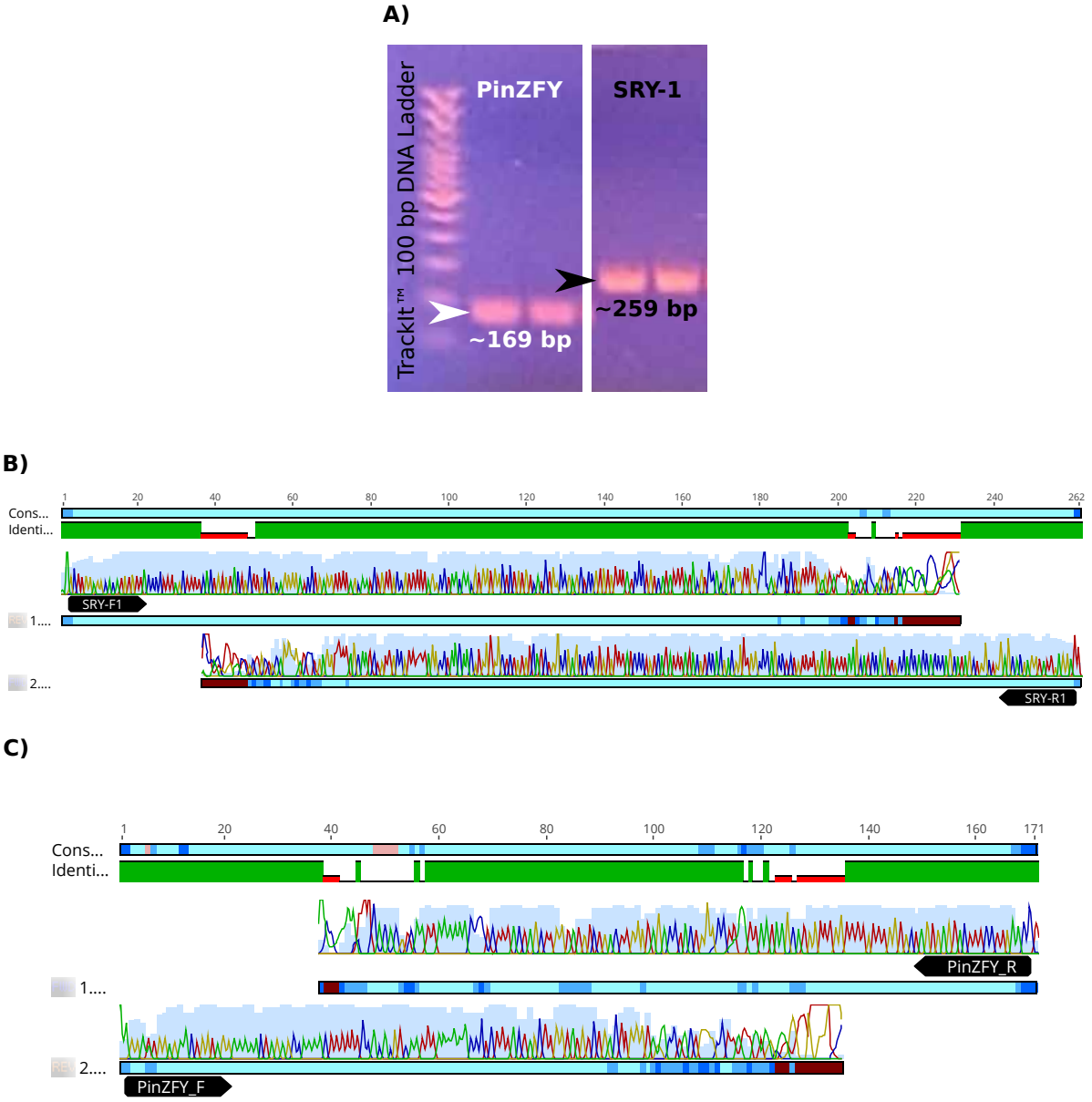

Figure 5.15. Testing the specificity of the primers PinZFY and SRY-1.
A) 2 % agar gel showing the product from PCR amplification of segments of the ZFX/Y (PinZFY primers) and SRY (SRY-1 primers) loci using JFFS tissue DNA. B) Chromatogram of unedited Sanger sequences showing the specificity of the SRY-1 primer set developed in this study. Both sequences belong to the same individual. The first one was sequenced with the reverse primer and was reverse complemented for the Clustal Omega alignment. The sequence at the bottom was sequenced with the forward primer. C) Similar to B) but using primers PinZFY. Primer location is indicated with a black arrow

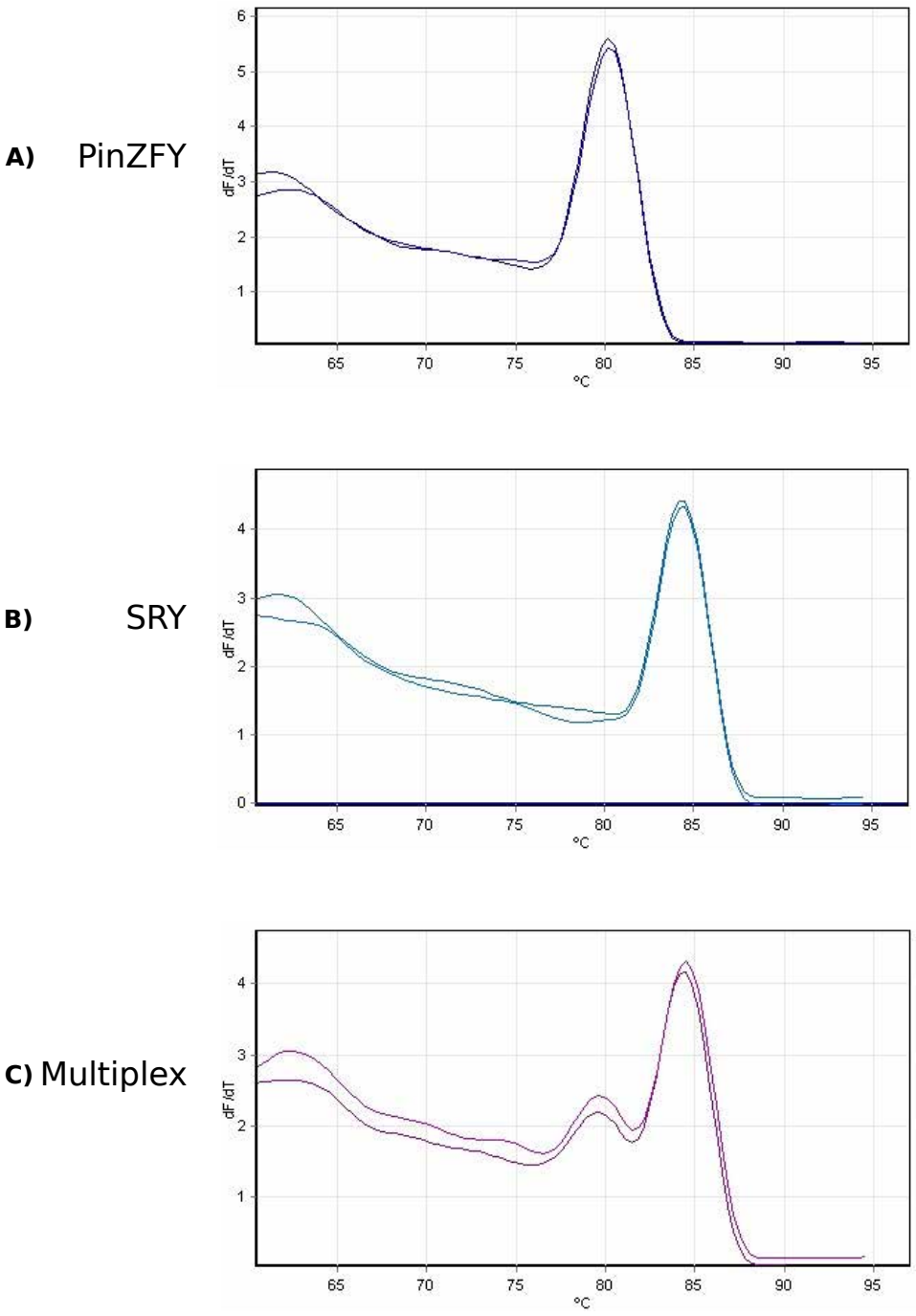

Figure 5.16. SYBR green Real-time PCR melting curves for ZFX/Y and SRY amplification.
Melting point temperatures (Tm) of the two amplicons, are graphically shown. A) DNA melt curve analysis of PinZFY amplicons (Tm 80 °C). B) DNA melt curve analysis of SRY amplicons (84 °C). C) DNA melt curve analysis resulting from the multiplex qPCR. Independent confirmed external PCR products for each target were combined in equal molarities as template in the reaction.

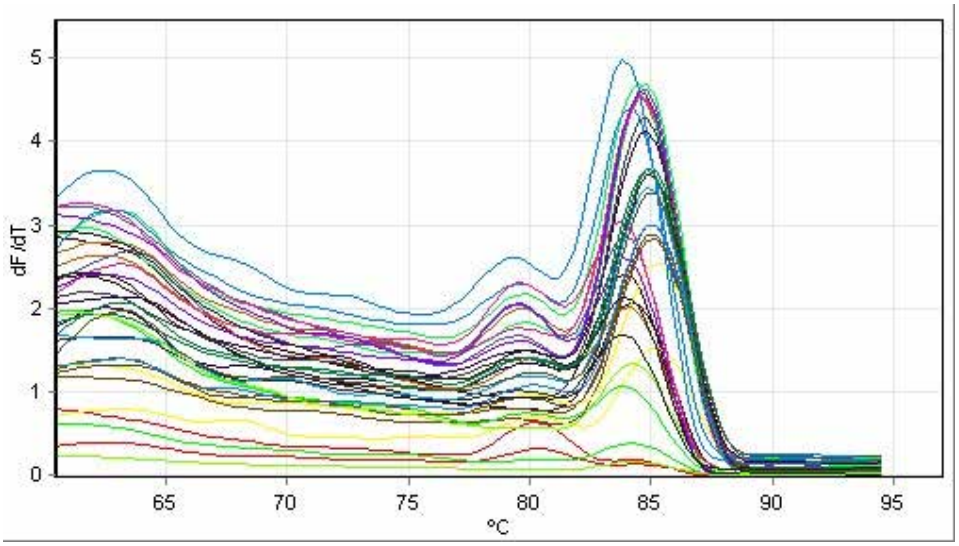

Figure 5.17. Multiplex SYBR green real-time PCR contamination with SRY amplicons. First large scale multiplex qPCR assay. Templates were generated by combined purification of the two external amplicons. All samples including the negative control showed two curves suggesting contamination of samples with SRY PCR product.

contamination. Instead, the products from the outer reactions were diluted 1 in 100 to limit any reagent carryovers from the previous amplification. This change in the protocol contributed to limiting the contamination picked up by the negative control. However, it did not disappear entirely.

Taking advantage of the sensitivity of real-time PCR, one final strategy was tested. To limit the concentration of PCR inhibitors present in direct faecal DNA extracts, these were diluted 1 in 10 before use in the real-time PCR with the inner primers. This change significantly improved the PCR outcome even though some of the samples do not show any amplification, possibly due to the low concentration of host DNA. A small melt curve peak at the SRY melt curve peak was seen in the identified female samples. This was also seen by Robertson *et al.* (2018).

Finally, the best approach to successfully conduct molecular sexing on the JFFS faecal samples is to use the diluted faecal DNA extract as the template in the inner PCRs and only use the semi-nested approach with diluted external PCR product when the first approach fails or gives ambiguous results. At the moment of writing this manuscript, only a handful of faecal samples had been tested.

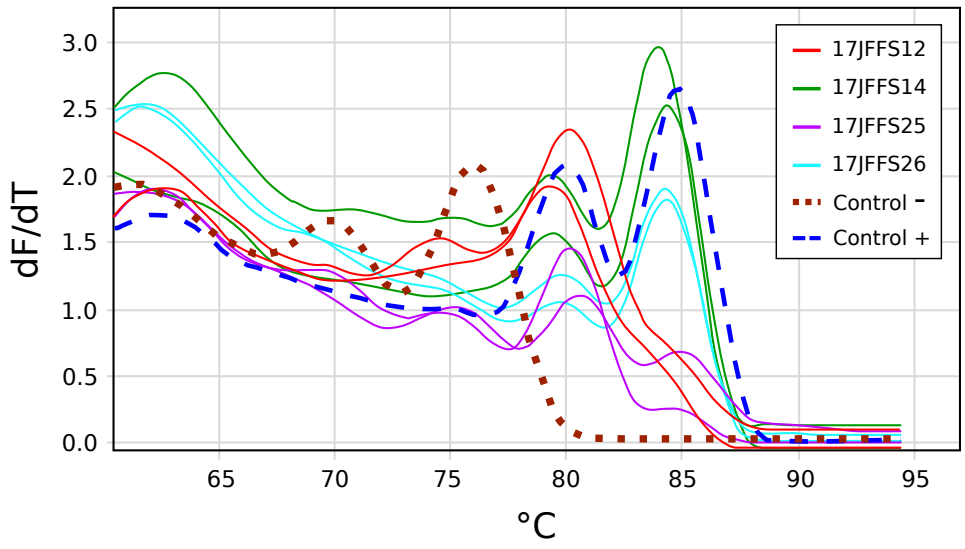

Figure 5.18. Molecular sex determination from JFFS faecal samples using multiplex SYBR green real-time PCR.
The melt curve analysis resulting from the optimised protocol allowed the identification of two males (17JFFS14 and 17JFFS26) and two females (17JFFS12 and 17JFFS 25). Negative control did not evidence contamination, and the positive control showed a more even signal between the two amplicons. Curve paths have been retraced for quality and clarity of the image.

5.3 Discussion

5.3.1 Sample storage and DNA extraction

Between the late 17th until the late 19th century, marine mammals were heavily hunted. In the case of the JFFS, it has been estimated that there were at least four million individuals when the hunting era started (Hubbs and Norris, 1971). By the end of the 1800's the population was reduced to a couple of hundreds and was eventually presumed extinct. However, the population has been recovering steadily since its rediscovery (Torres, 1987). Little is known about the genetic consequences of the population fluctuation and nothing is known about the current genetic make up of the JFFS (e.g. inbreeding, hybridisation). There is only one published study on JFFS genetics (Goldsworthy *et al.*, 2000).

Collecting samples from JFFS is logistically complicated and expensive, which may partly explain the lack of research development on the species. Working with faecal samples has many benefits when conducting genetic analysis in free-range wild animals. However, these types of samples contain host DNA content that is often low in quantity and quality. Here, I aimed to extract host DNA from the JFFS fur seal faecal sample to conduct the first genetic studies in this species in 20 years. In this study, I aimed to optimise a protocol to perform

genetic analysis on JFFS using faeces as the source of DNA. Here, I targeted the same mtDNA control region as Goldsworthy *et al.* (2000) and five microsatellite loci. I also investigated the use of two loci targeting X and Y chromosomes for molecular sexing of the samples.

Careful planning is needed from collection to DNA amplification to increase the success rate of isolating host DNA when using faecal samples. I collected samples from a very remote location. On some occasions, it took a few days from the moment of collection to putting the samples at -20 °C. Furthermore, the electricity supply on RC island is provided by an inefficient diesel engine that often fails to work. These working conditions meant that even when the samples could be placed in a regular domestic freezer, chances of power cut were high. The samples therefore had to 'survive' even if storage at -20 ° C could not be guaranteed. The simple test I conducted on dog faecal samples showed that host DNA could be amplified using any of the three storage methods I trialled (freezing, RNAlater and ethanol). However, based on the empirical results generated from the JFFS faecal samples, I can confidently say that storing the samples in RNAlater is a reliable and straightforward method, even when immediate freezing is not available.

Many manufacturers sell DNA extraction kits; however, faeces are difficult samples to work with due to the presence of PCR inhibitors such as fat and bile salts. However, faecal characteristics greatly vary between species and these kits are often optimised for extracting host DNA from human faeces or long microbial genomic DNA. Thus, some kits might not perform well with certain types of faecal samples. For this reason, I tested four different extraction kits on dog faeces. This experiment was performed using free kit samples from the manufacturers. Thus, this experiment was not focused on researching the most optimal method. Instead, I aimed at briefly screening the options available. Even at this early screening, the Zymo Research kit did not work well in any of the triplicates. Whether this was due to the specific nature of the faeces used is unknown.

JFFS faeces are very different to dog faeces; this was evident to me only during fieldwork. Nevertheless, I successfully amplified DNA extracted using TRIzol and the Qiamp kit. Both TRIzol and the QIAmp DNA stool mini kit could extract host DNA of reasonable quality from JFFS faeces. These faeces are exceptionally high in fat and tend to be very sticky, which was particularly problematic when handling faecal fragments (as opposed to swabs). Due to the fat content, extraction methods using TRIzol is likely to work better than commercial kits for which this type of sample has not been optimised. Therefore, I will likely keep using TRIzol as my chosen extraction method in the future.

I found that the most crucial factor likely to limit success relates to the samples themselves at the collection point e.g. sample age and swabbing strategy. For instance, the outside layer of faeces which probably has host DNA from sloughed cells is directly exposed to the environment,

and DNA is particularly sensitive to sun exposure. Furthermore, shed epithelial cells are not homogeneously distributed on the samples. Thus, it is essential to swab a large area of the samples, especially including the part that left the rectum last, which is likely to contain the most significant number of cells. Finally, swabbing as superficially as possible is likely limit the amount of "faecal" contamination that results in host DNA dilution and, a larger concentration of PCR inhibitors. Superficial scraping has also been used in previous studies on other pinniped species (Reed *et al.*, 1997a; Steinmetz *et al.*, 2021).

Working with faecal samples to conduct host genetic analysis is challenging and often involves many optimisation steps. However, once optimisation has been performed, taking more samples through the protocol should require little effort. For instance, from my experience mtDNA should always be the first target when studying host genetics from faecal samples, even if the eventual interest focuses on nuclear DNA. Due to the highly successful amplification rate observed in this and other studies (Reed *et al.*, 1997a; Fietz *et al.*, 2016; Zappes *et al.*, 2017; Steinmetz *et al.*, 2021), targeting mtDNA allows samples with no or marginal concentrations of host DNA to be identified and thus can be excluded early in the study. This approach was also suggested by Reed *et al.* (1997a). In my experience, finding the highest annealing temperature at which the primers can amplify DNA and then using touchdown PCR protocols definitely reduced some of the noise from non-host DNA by increasing the primer specificity.

Additionally, samples with low concentrations of DNA benefited from conducting nested PCR, as observed from the mtDNA amplification results. So far, early results from the molecular sexing analysis have confirmed that nuclear DNA can be amplified from the JFFS. However, sex was determined by the much more sensitive method (qPCR). Thus, the success of this assay, does not necessarily mean that targeting microsatellite will be equally successful due to the amount of DNA needed to carry out this analysis. Nevertheless, microsatellite analysis in other pinniped species have been successfully performed without the need of nesting (Reed *et al.*, 1997a; Fietz *et al.*, 2016). However, both authors reported samples were frozen shortly after collection which is not possible in JFFS. As I mentioned earlier, sample collection is likely to influence the starting DNA quantity and quality. However, based on agarose gel electrophoresis and Sanger sequencing, targeting microsatellite loci was more consistent when performing a nested PCR. This analysis has not yet been completed. This means that the PCR optimisation success remains unknown for the microsatellite analysis.

One of the easier sections of DNA to target for this type of work is mitochondrial DNA (mtDNA). This is present in multiple copies (100 to 100,000) within host cells and so has higher copy numbers in DNA samples than nuclear DNA (2 copies) in mammals. It is also very useful in allowing the study of population genetics as the mtDNA evolves more quickly than nuclear DNA (Stoneking, 2000). The control region, in particular, has been particularly

targeted, giving a good depth of known sequence in databases and allowing good definition of haplotypes to allow estimations of population structures (Goldsworthy *et al.*, 2000; Matthee *et al.*, 2006; Bernardi *et al.*, 1998; Bickham *et al.*, 1996; Rosel *et al.*, 2017). For this reason, this locus was chosen for this study.

Overall results generated in this study were very similar to those found by Goldsworthy *et al.* (2000). Some non-significant trends included a reduction in the number of haplotypes per sample and a sharp decrease of the Tajima'D value compared to the previous report. Negative D values are suggestive of population expansion. Even though this value was not statistically significant, it matches the evident population recovery observed and reported by the local community in the last 50 years. Furthermore, a brief communication on the JFFS population expansion was published on the National Forestry Commission official website in 2018 (shortened link to the article: shorturl.at/ekEN3). Thus, I expect that D will become significant if the population trend remains positive under ideal conditions.

Based on the mtDNA control region analysis performed in this study, no significant difference was observed between the colonies. A similar phenomenon was observed in the Cape fur seal (*Arctocephalus pusillus pusillus*). This pattern may suggest that JFFS fur seal females are not particularly philopatric (tend to return to the same location). One may argue that the samples colonies were too close to each other, making differences between colonies less evident. Nevertheless, it has been reported that AFS females return to the same place with an accuracy of a couple of meters. Thus, if JFFS would show such extreme philopatry, it would be expected to see some degree of differentiation between colonies even if this colony were separated by less than ten kilometres. Goldsworthy *et al.* (2000), on the other hand, did find significant differences between AS Island (not included in this study) and RC. It is more likely, then, that the lack of geographical differentiation observed in this study may, in part, be explained by the post-sealing recolonisation.

Due to the lack of genetic data from pre-sealing, it is not possible to compare the true genetic consequence of the population decline. However, based on the mtDNA data presented here and in Goldsworthy *et al.* (2000), genetic variability does not seem to be affected as severely as one would expect from a species that was once considered extinct, such as the case of the Guadalupe fur seal (*Arctocephalus philippii townsendii*) (Weber *et al.*, 2004). Furthermore, high mtDNA variability has also been reported in other otariids; victims of severe population declines due to the overhunting that lasted until the late 1800s but whose populations have also shown a post-sealing steady population recovery (Matthee *et al.*, 2006; Bickham *et al.*, 1996; Goldsworthy *et al.*, 2000). In the case of JFFS, as previously argued by Hubbs and Norris (1971) and Goldsworthy *et al.* (2000), it is likely that at the end of the sealing era, the population was considerably larger than previously thought. I would like to add that the

population may have also recovered faster, although unseen, due to accessibility difficulties. Nevertheless, a new study showed that severe population declines in many pinniped species do not necessarily imply reduced genetic diversity and population viability (Stoffel *et al.*, 2018).

Finally, a couple of field observations motivates further research, especially the analysis of nuclear DNA. One example is the presence of leucistic individuals showing lighter coloured fur, lack of pigmentation of the eye and blindness. These individuals were adult females, although I saw one male with similar fur colouration but because he was sleeping I could not confirmed eye colouration and blindness. I saw roughly 20 pups with these characteristics during the last two fieldwork seasons (2017-2018 and 2018-2019). However, I did not confirm their sex. Leucism may be suggestive of in-breeding and thus, it is worth investigating from the genetic and demographic (e.g. What is the proportion? Is this proportion increasing?) point of view.

Additionally, I observed juvenile AFS females and subantarctic fur seals (*Arctocephalus tropicalis*, SAFS) females and males among JFFS breeding colonies. Both species have been previously reported by Torres *et al.* (1984). The author hypothesised that AFS individuals came from AFS colonies from South Georgia or Cape Shirref aided by ocean currents. Unfortunately, and like the lighter coloured individuals, they are not being monitored. Thus, it is unknown how many of them can be usually found and whether there is a pattern in their visits (e.g. years with low prey abundance might push the AFS females to explore other locations). DNA studies may elucidate possible hybridisation cases as previously reported between AFS, SAFS and New Zealand fur seal (*Arctocephalus forsteri*) in Macquarie island (Lancaster *et al.*, 2006).

Although not complete, this study focused on host DNA amplification from JFFS faecal samples. This research is the first genetic insight of the JFFS in more than 20 years. Once completed, the information will greatly contribute to our understanding of the species' genetic history and its current status. Additionally, methods to explore mitochondrial and nuclear DNA, including conducting molecular sexing, have been optimised for this species. The optimisation means that future samples can now be analysed more easily. Finally, and most importantly, studying the modern JFFS population genetics will contribute to the identification of new conservation threats. For instance, population fluctuations due to El Niño hav been evidenced in Galapagos fur seals (Páez-Rosas *et al.*, 2021). Although not covered in that study, these population fluctuations may result in genetic consequences. Furthermore, Forcada and Hoffman (2014a) have evidenced a heterozygosity selection pressure in AFS populations due to climate. Currently, there is no consistent information on population fluctuation or genetic fitness in the species. All this information is urgently needed to inform new conservation measures as the hunting ban end date approaches.

Chapter 6

Final conclusions and future work

The JFFS is endemic to the JFA and the Desventuradas islands. However, the reproductive colonies can only be found in the former. In 1995, the Chilean government gave the species complete legal protection for 30 years, officially ending on November 9, 2025. This means that future conservation strategies need urgent discussion. Even though good research was carried out between the 1970s and early 2000s (Osman *et al.*, 2010; Francis *et al.*, 1998; Ochoa Acuna and Francis, 1995; Sepúlveda *et al.*, 1997; Goldsworthy *et al.*, 2000), no new information has emerged in the last 20 years. Furthermmore, the monitoring of this species remains, if anything, poor and mostly represented by intermittent animal censuses. The lack of funding and complicated logistics might explain, in part, the low research efforts.

The main aim of this project was to use non-invasive faecal samples to obtain information on the JFFS for the first time in 20 years. Here, I focused on three topics: faecal microbiome (chapter 3), heavy metal exposure (Chapter 4) and host genetics (chapter 5). But the potential of faecal samples is not limited to these three topics. Furthermore, JFFS faecal samples collected in this research were shared with collaborators to look at microplastic exposure in South American otariids. The resulting co-authored publication evidenced that the JFFS was the otariid most exposed to microplastics (Perez-Venegas *et al.*, 2020). Another collaborator is currently analysing samples for parasite identification. Thus, faeces are a cost-effective tool to study this species from various angles.

This chapter aims to (1) discuss the most relevant findings and their implications, (ii) describe potential areas for future works and (iii) discuss limitations.

6.1 Faecal Microbiome

I performed the first characterisation of the JFFS faecal microbiome. Overall, the core microbiome composition was similar to the compositions described for other pinnipeds. Colony location explained only a small percentage of the microbiome variability. This result was not surprising due to the proximity between the colonies. However, TB seemed to be notably different from the other locations. It would be interesting to include samples from the furthermost island, Alejandro Selkirk (AS, 180 km). Collecting samples from this location was part of the

initial planning of my research. Unfortunately, due to logistic difficulties, it was not possible to organise an expedition to AS until the reproductive season 2019 - 2020. Due to COVID 19, I decided not to participate in the expedition. Instead, CONAF park rangers and members of the local NGO Fundacion Endemica collected samples on my behalf. These samples still need to be analysed.

The largest proportion of the microbial composition variability was explained by an inverse relationship between *Peptoclostridium* and *Fusobacterium*• abundance, which translated into two clusters. Most of the samples were included in cluster 1, showing higher levels of *Fusobacterium* and lower of *Peptoclostridium*. I hypothesise diet to be the most likely explanation of this phenomenon. However, many factors influence foraging behaviour in pinnipeds, such as sex, age, and climate conditions. Furthermore, enriched pathways predicted with PICRUSt suggested better symbiosis in cluster 1 than in cluster 2. Therefore, additional studies investigating the relationship between sex, age and prey are required to test this hypothesis.

6.2 Heavy metals

In Chapter 2, I analysed heavy metal exposure in the JFFS. Hg and Cd were found in very high concentrations in JFFS faeces compared to AFS faeces and other previously reported species. Results obtained from both fur seals were in agreement with what is known about their foraging behaviour. Cd levels were exceptionally high in JFFS. Furthermore, the high concentrations of Cd in bone samples reflected its bioavailability. Surprisingly, no evidence of this heavy metal affecting other essential bone minerals suggests bone resilience to Cd toxicity in this species. Understanding the tolerance mechanism in JFFS bones could provide critical information for human medical research. JFFS bone samples also showed high levels of Si, an element previously associated with bone health. Perhaps this could be a starting point to explore tolerance mechanisms in marine mammal bones. Future work would require a larger number of samples from marine mammals feeding at different trophic levels.

Octopus samples analysed after discovery of the heavy metal levels in faecal samples are, to my knowledge, the first report of heavy metal contamination in this local species. Even though only a few samples were analysed, the results suggest contamination risk to the local community relying on these for food and economic resources.

In general, marine mammals have developed efficient strategies to counteract the high levels of heavy metals to which they are exposed. However, the thresholds at which these mechanisms become overloaded are not yet known. Furthermore, heavy metals are not the only contaminants present in the prey that could be detrimental to fur seal and human health (Perez-Venegas *et al.*, 2020). With the ever increasing pollutant concentrations in marine environments

it is important to keep monitoring and investigating how these transfer and change overtime. JFFS has now proven to be an important bioindicator for the region. Finally, this research can provide relevant information to policymakers if hunting restarts, especially regarding regulating possible JFFS-derived food products. The high levels of heavy metals found in JFFS should be considered a potential hazard if human consumption is reconsidered.

6.3 Host genetics

In the context of the hunting ban, population genetics is, in my opinion, one of the most important topics that need urgent consideration. The presence of leucistic individuals may, for instance, be indicative of poor genetic diversity. Leucism has been reported in other pinnipeds, especially in AFS and is currently considered a rare phenomena (Acevedo *et al.*, 2009; de Bruyn *et al.*, 2007; Romero and Tirira, 2017; Acevedo and Aguayo, 2008; Jones *et al.*, 2019; Bester *et al.*, 2008; Osinga *et al.*, 2010; du Toit *et al.*, 2019). Reported leucistic individuals are usually pups, while adult sightings are rare (Acevedo and Aguayo, 2008; Jones *et al.*, 2019). All reports of leucism in pinnipeds report lighter to white fur but normal coloured eyes. There are various characteristics that make the leucistic cases identified in JFFS of particular interest and concern. Even though leucism occurs in relatively small proportions, on the JFA its occurrence is noticeable in both pups and female adults. Furthermore, CONAF rangers have been aware of them for many years and they have noticed that every individual that had lighter fur was also blind. Blindness is perhaps the most concerning characteristic of this phenotype and has never been described in leucistic individuals from other species. Nevertheless, all sighted individuals look healthy and well-integrated in the colonies.

The optimised DNA amplification protocols presented in this dissertation are a significant start to understanding the genetic background of the JFFS. Furthermore, molecular sex determination becomes critical in studies such as this where the individuals from which samples have been collected cannot be identified directly. This is particularly important in highly dimorphic species such as the JFFS. A study on summer prey evidenced that JFFS adult females target different species than males (Ochoa Acuna and Francis, 1995). Due to the high influence diet has on the gut microbiome and exposure to contaminants such as heavy metals and microplastics, knowing the sex of the samples is important for good interpretation of the results.

6.4 Ongoing studies: Shotgun sequencing

A final interesting approach I was able to explore was shotgun sequencing. Unfortunately, I could not include the results in this dissertation because it is not possible for me to complete

the analysis within the PhD time frame. Nevertheless, I performed quick blast searches with a focus on eukaryotic 18S rRNA and mitochondrial DNA. As a result, I was able to extract the complete mitochondrial genome and 18S rRNA gene of JFFS. Unfortunately, full annotation of these contigs remains incomplete. Nevertheless, this opens an interesting opportunity to explore other segments of the JFFS nuclear DNA. Prey elements were also identified and following previous reports (Ochoa Acuna and Francis, 1995), most of the contigs were identified as cephalopods. However, the resolution remains poor. Finally, one of the contigs showed a 99 % identity with 100 % query match against the trematode parasite *Ogmogaster antarctica*. These early results already demonstrate the high potential and cost-effectiveness of shotgun sequencing to investigate different aspects of the JFFS by performing only one genetic analysis.

6.5 Final remarks

Due to the lack of updated information on the JFFS, the aims of this study were broad. This means that the patterns observed in this study are not necessarily conclusive. Instead, it opens a series of hypotheses that can be explored further in the future. Results interpretation was particularly challenging due to poor knowledge about diet and the inability to assign a sex to each sample. Diet and sex are important drivers of the gut microbiome. In addition, although strongly linked to sex in sexually dimorphic animals such as the JFFS, diet plays an important role in the transfer of pollutants such as heavy metals and microplastics. Thus, future work should explore prey selection. Ideally, prey selection should be studied at different time points within a year and annually to understand better the fluctuations in prey selection according to environmental and climatic phenomena. The limitation of sex identification, on the other hand, was targeted by the optimisation of a qPCR protocol for its molecular assignment. Once complete, sex will be included in the metadata, and its contribution to the data variability will be assessed.

This project has made clear the need to understand the foraging behaviour of the JFFS fur seal to understand better phenomena such as those observed in the faecal microbiome and heavy metal exposure. However, many gaps need to be filled. Future work should include a consistent evaluation of population abundance. Drones could be considered as a cost-effective way to assess JFFS population abundance and demographics (Krause and Hinke, 2021; Wood *et al.*, 2021). This method does not only enable a more precise evaluation of population abundance, but it also provides significantly more detail about those populations by facilitating. For instance, the identification of other species or leucistic individuals. Population genetics, as mentioned before, are also crucial as population abundance does not necessarily reflect population fitness. Continuity of species monitoring is critical. The lack of personnel and

funding could be overcome with the development of community science and to help the local community understand their local environment.

References

Abreu, J., Staniland, I., Rodrigues, C. F., Queirós, J. P., Pereira, J. M., and Xavier, J. C. (2019). Squid in the diet of Antarctic fur seals: potential links to oceanographic conditions and Antarctic krill abundance. *Marine Ecology Progress Series*, 628:211–221. (Cited on page 77.)

Acevedo, J. and Aguayo, M. (2008). Leucistic south american sea lion in chile, with a review of anomalously color in otariids. *Revista de Biologia marina y Oceanografia*, 43(2):413–417. (Cited on page 113.)

Acevedo, J., Torres, D., and Aguayo-Lobo, A. (2009). Rare piebald and partially leucistic antarctic fur seals, arctocephalus gazella, at cape shirreff, livingston island, antarctica. *Polar Biology*, 32(1):41–45. (Cited on page 113.)

Aguayo, A. and Maturana, R. (1970). Primer censo de lobos finos en el archipiélago de juan fernádezndez. *Biologia Pesquera*, 4:3–15. (Cited on page 4.)

Aguayo, A., Maturana, R., and Torres, D. (1970). El lobo fino de juan fernández. *Revista Bioloía Marina*, 14:135–149. (Cited on pages 4 and 32.)

Alfano, N., Courtiol, A., Vielgrader, H., Timms, P., Roca, A. L., and Greenwood, A. D. (2015). Variation in koala microbiomes within and between individuals: effect of body region and captivity status. *Scientific reports*, 5(1):1–12. (Cited on page 10.)

Allen, S., Brown, E., Faulkner, K., Gende, S., and Womble, J. (2011). Conserving pinnipeds in pacific ocean parks in response to climate change. *Park Science*, 2:48–52. (Cited on page 9.)

Altman, T., Travers, M., Kothari, A., Caspi, R., and Karp, P. D. (2013). A systematic comparison of the MetaCyc and KEGG pathway databases. *BMC Bioinformatics*, 14(1):1–15. (Cited on page 54.)

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410. (Cited on page 19.)

Amato, K. R., Yeoman, C. J., Kent, A., Righini, N., Carbonero, F., Estrada, A., Gaskins, H. R., Stumpf, R. M., Yildirim, S., Torralba, M., *et al.* (2013). Habitat degradation impacts black howler monkey (alouatta pigra) gastrointestinal microbiomes. *The ISME journal*, 7(7):1344–1353. (Cited on page 10.)

Amos, B. and Hoelzel, A. R. (1992). Applications of molecular genetic techniques to the conservation of small populations. *Biological Conservation*, 61(2):133–144. (Cited on page 11.)

Andersen, J. M., Skern-Mauritzen, M., Boehme, L., Wiersma, Y. F., Rosing-Asvid, A., Hammill, M. O., and Stenson, G. B. (2013). Investigating annual diving behaviour by hooded seals (cystophora cristata) within the northwest atlantic ocean. *PLoS ONE*, 8(11):e80438. (Cited on page 53.)

Ando, N., Isono, T., and Sakurai, Y. (2005). Trace elements in the teeth of Steller sea lions (Eumetopias jubatus) from the North Pacific. *Ecological Research*, 20(4):415–423. (Cited on page 76.)

Andrews, S. (2010). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. (Cited on page 19.)

Atlas, R. M. (1981). Microbial degradation of petroleum hydrocarbons: An environmental perspective. (Cited on page 55.)

Baby, L., Sankar, T. V., and Anandan, R. (2014). Comparison of lipid profile in three species of myctophids from the south west coast of kerala, India. *National Academy Science Letters*, 37(1):33–37. (Cited on page 54.)

Bahamonde, N. (1966). El ma y sus recursos. *Geografia economica de Chile, Primer apendice. Santiago, Chile: CORFO*, pages 81–85. (Cited on page 4.)

Balali-Mood, M., Naseri, K., Tahergorabi, Z., Khazdair, M. R., and Sadeghi, M. (2021). Toxic mechanisms of five heavy metals: Mercury, lead, chromium, cadmium, and arsenic. *Frontiers in pharmacology*, 12. (Cited on page 12.)

Balvočiūtė, M. and Huson, D. H. (2017). SILVA, RDP, Greengenes, NCBI and OTT — how do these taxonomies compare? *BMC Genomics*, 18(S2):114. (Cited on page 46.)

Bentaleb, I., Martin, C., Vrac, M., Mate, B., Mayzaud, P., Siret, D., de Stephanis, R., and Guinet, C. (2011). Foraging ecology of Mediterranean fin whales in a changing environment elucidated by satellite tracking and baleen plate stable isotopes. *Marine Ecology Progress Series*, 438:285–302. (Cited on page 74.)

Bernardi, G., Fain, S., Gallo-Reynoso, J., Figueroa-Carranza, A., and Le Boeuf, B. (1998). Genetic variability in guadalupe fur seals. *Journal of Heredity*, 89(4):301–305. (Cited on page 109.)

Berry, M. J. and Ralston, N. V. C. (2009). Mercury Toxicity and the Mitigating Role of Selenium. *EcoHealth 2009 5:4*, 5(4):456–459. (Cited on pages 57 and 74.)

Berta, A. and Churchill, M. (2012). Pinniped taxonomy: Review of currently recognized species and subspecies, and evidence used for their description. *Mammal Review*, 42(3):207–234. (Cited on page 3.)

Berta, A., Churchill, M., and Boessenecker, R. W. (2018). The origin and evolutionary biology of pinnipeds: seals, sea lions, and walruses. *Annual Review of Earth and Planetary Sciences*, 46:203–228. (Cited on pages 3 and 11.)

Bester, M. (2014). Marine mammals–natural and anthropogenic influences. In Freedman, B., editor, *Global Environmental Change*, pages 167–174. Springer, Dordrecht, Oxford. (Cited on page 6.)

Bester, M. N., de Clercq, H., Hofmeyr, G. G., and de Bruyn, P. N. (2008). Leucistic southern elephant seal at marion island? *Polar Biology*, 31(2):255–257. (Cited on page 113.)

Bharwani, A., Mian, M. F., Foster, J. A., Surette, M. G., Bienenstock, J., and Forsythe, P. (2016). Structural and functional consequences of chronic psychosocial stress on the microbiome and host. *Psychoneuroendocrinology*, 63:217–227. (Cited on page 50.)

Bickham, J. W., Patton, J. C., and Loughlin, T. R. (1996). High variability for control-region sequences in a marine mammal: implications for conservation and biogeography of steller sea lions (eumetopias jubatus). *Journal of mammalogy*, 77(1):95–108. (Cited on page 109.)

Bik, E. M., Costello, E. K., Switzer, A. D., Callahan, B. J., Holmes, S. P., Wells, R. S., Carlin, K. P., Jensen, E. D., Venn-Watson, S., and Relman, D. A. (2016). Marine mammals harbor unique microbiotas shaped by and yet distinct from the sea. *Nature Communications*, 7:10516. (Cited on pages 10, 32, and 48.)

Blekhman, R., Tang, K., Archie, E. A., Barreiro, L. B., Johnson, Z. P., Wilson, M. E., Kohn, J., Yuan, M. L., Gesquiere, L., Grieneisen, L. E., and Tung, J. (2016). Common methods for fecal sample storage in field studies yield consistent signatures of individual identity in microbiome sequencing data. *Scientific Reports*, 6(1):1–5. (Cited on page 17.)

Blumenthal, N. C., Cosma, V., Skyler, D., Legeros, J., and Waiters, M. (1995). The Effect of Cadmium on the Formation and Properties of Hydroxyapatite In Vitro and its Relation to Cadmium Toxicity in the Skeletal System. *Calcif Tissue Int*, 56:316–322. (Cited on pages 69 and 78.)

Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. *Bioinformatics*, 31(24):3997–3999. (Cited on page 27.)

Bolea-Fernandez, E., Rua-Ibarz, A., Krupp, E. M., Feldmann, J., and Vanhaecke, F. (2019). High-precision isotopic analysis sheds new light on mercury metabolism in long-finned pilot whales (Globicephala melas). *Scientific Reports 2019 9:1*, 9(1):1–10. (Cited on page 74.)

Bolyen, E., Rideout, J. R., Dillon, M. R., and Bokulich, N. A. e. a. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. (Cited on page 19.)

Boness, D. J. and Francis, J. M. (1991). The effect of thermoregulatory behaviour on the mating system of the juan fernandez fur seal, arctocephalus philippii. *Behaviour*, 119(1-2):104–126. (Cited on page 5.)

Borja, A. (2014). Grand challenges in marine ecosystems ecology. *Frontiers in Marine Science*, 0(FEB):1. (Cited on page 1.)

Bossart, G. D. (2011a). Marine mammals as sentinel species for oceans and human health. *Veterinary pathology*, 48(3):676–690. (Cited on page 1.)

Bossart, G. D. (2011b). Marine mammals as sentinel species for oceans and human health. *Veterinary Pathology*, 48(3):676–690. (Cited on page 31.)

Bourgeois, S., Kaden, J., Senn, H., Bunnefeld, N., Jeffery, K. J., Akomo-Okoue, E. F., Ogden, R., and McEwing, R. (2019). Improving cost-efficiency of faecal genotyping: New tools for elephant species. *PloS one*, 14(1):e0210811. (Cited on page 82.)

Bradney, L., Wijesekara, H., Palansooriya, K. N., Obadamudalige, N., Bolan, N. S., Ok, Y. S., Rinklebe, J., Kim, K.-H., and Kirkham, M. (2019). Particulate plastics as a vector for toxic trace-element uptake by aquatic and terrestrial organisms and human health risk. *Environment international*, 131:104937. (Cited on pages 12, 57, and 58.)

Brennecke, D., Duarte, B., Paiva, F., Caçador, I., and Canning-Clode, J. (2016). Microplastics as vector for heavy metal contamination from the marine environment. *Estuarine, Coastal and Shelf Science*, 178:189–195. (Cited on pages 12 and 58.)

Breton, J. Ô., Daniel, C., Dewulf, J., Pothion, S., Froux, N., Sauty, M., Thomas, P., Pot, B., and Foligné, B. (2013). Gut microbiota limits heavy metals burden caused by chronic oral exposure. *Toxicology Letters*, 222(2):132–138. (Cited on pages 58, 63, and 75.)

Buha, A., Jugdaohsingh, R., Matovic, V., Bulat, Z., Antonijevic, B., Kerns, J. G., Goodship, A., Hart, A., and Powell, J. J. (2019). Bone mineral health is sensitively related to environmental cadmium exposure- experimental and human data. *Environmental Research*, 176:108539. (Cited on pages 12, 69, and 78.)

Bustamante, P., Caurant, F., Fowler, S. W., and Miramand, P. (1998). Cephalopods as a vector for the transfer of cadmium to top marine predators in the north-east Atlantic Ocean. *Science of the Total Environment*, 220(1):71–80. (Cited on pages 76 and 77.)

Bustamante, P., Cosson, R. P., Gallien, I., Caurant, F., and Miramand, P. (2002). Cadmium detoxification processes in the digestive gland of cephalopods in relation to accumulated cadmium concentrations. *Marine Environmental Research*, 53(3):227–241. (Cited on page 76.)

Bustamante, P., Lahaye, V., Durnez, C., Churlaud, C., and Caurant, F. (2006). Total and organic Hg concentrations in cephalopods from the North Eastern Atlantic waters: Influence of geographical origin and feeding ecology. *Science of the Total Environment*, 368(2-3):585–596. (Cited on page 76.)

Bustamante, P., Morales, C. F., Mikkelsen, B., Dam, M., and Caurant, F. (2004). Trace element bioaccumulation in grey seals Halichoerus grypus from the Faroe Islands. *Marine Ecology Progress Series*, 267:291–301. (Cited on pages 57, 74, 76, and 79.)

Caccavo, J. A., Christiansen, H., Constable, A. J., Ghigliotti, L., Trebilco, R., Brooks, C. M., Cotté, C., Desvignes, T., Dornan, T., Jones, C. D., *et al.* (2021). Productivity and change in fish and squid in the southern ocean. *Frontiers in Ecology and Evolution*, page 351. (Cited on page 9.)

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal 2017 11:12*, 11(12):2639–2643. (Cited on page 46.)

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7):581–583. (Cited on pages 19 and 46.)

Calvo, A. Y., Manrique, J. M., and Jones, L. R. (2018). Rare unclassified 16S rRNA operational taxonomic units from the uncharted Engaño Bay (Argentinean Patagonia). *Canadian Journal of Microbiology*, 64(1):91–96. (Cited on page 47.)

Cameron, E. S., Schmidt, P. J., Tremblay, B. J.-M., Emelko, M. B., and Müller, K. M. (2021). To rarefy or not to rarefy: Enhancing diversity analysis of microbial communities through next-generation sequencing and rarefying repeatedly. *bioRxiv*, page 2020.09.09.290049. (Cited on page 47.)

Carlisle, E. M. (1970). Silicon: A possible factor in bone calcification. *Science*, 167(3916):279–280. (Cited on page 78.)

Caruso, V., Song, X. B., Asquith, M., and Karstens, L. (2019). Performance of microbiome sequence inference methods in environments with varying biomass. *mSystems*, 4. (Cited on page 46.)

Cattan, P., Yánez, J., and Torres, D. (1980). Helmintos parásitos del lobo fino, arctocephalus philippii (peters, 1866) de juan fernández. *Boletín Chileno de Parasitología*, 35(3-4):73–75. (Cited on page 5.)

Chakraborty, R. and Coates, J. D. (2004). Anaerobic degradation of monoaromatic hydrocarbons. *Applied Microbiology and Biotechnology 2004 64:4*, 64(4):437–446. (Cited on page 55.)

Chao, A., Chiu, C. H., and Jost, L. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1558):3599–3609. (Cited on page 50.)

Chen, X., Wang, Z., Zhu, G., Nordberg, G. F., Jin, T., and Ding, X. (2018). The association between cumulative cadmium intake and osteoporosis and risk of fracture in a Chinese population. *Journal of Exposure Science & Environmental Epidemiology 2018 29:3*, 29(3):435–443. (Cited on page 78.)

Chouvelon, T., Spitz, J., Caurant, F., Mèndez-Fernandez, P., Autier, J., Lassus-Débat, A., Chappuis, A., and Bustamante, P. (2012). Enhanced bioaccumulation of mercury in deep-sea fauna from the Bay of Biscay (north-east Atlantic) in relation to trophic positions identified by analysis of carbon and nitrogen stable isotopes. *Deep-Sea Research Part I: Oceanographic Research Papers*, 65:113–124. (Cited on page 76.)

Christakis, C. A., Barkay, T., and Boyd, E. S. (2021). Expanded Diversity and Phylogeny of mer Genes Broadens Mercury Resistance Paradigms and Reveals an Origin for MerA Among Thermophilic Archaea. *Frontiers in Microbiology*, 0:1645. (Cited on page 75.)

Christophoridis, C., Bourliva, A., Evgenakis, E., Papadopoulou, L., and Fytianos, K. (2019). Effects of anthropogenic activities on the levels of heavy metals in marine surface sediments of the thessaloniki bay, northern greece: Spatial distribution, sources and contamination assessment. *Microchemical Journal*, 149:104001. (Cited on page 57.)

Churchill, M., Boessenecker, R. W., and Clementz, M. T. (2014). Colonization of the southern hemisphere by fur seals and sea lions (carnivora: Otariidae) revealed by combined evidence phylogenetic and bayesian biogeographical analysis. *Zoological Journal of the Linnean Society*, 172(1):200–225. (Cited on page 9.)

Cipro, C. V., Cherel, Y., Bocher, P., Caurant, F., Miramand, P., and Bustamante, P. (2018). Trace elements in invertebrates and fish from Kerguelen waters, southern Indian Ocean. *Polar Biology*, 41(1):175–191. (Cited on page 77.)

Claus, S. P., Guillou, H., and Ellero-Simatos, S. (2016). The gut microbiota: a major player in the toxicity of environmental pollutants? *Npj biofilms and microbiomes*, 2(1):1–11. (Cited on pages 58 and 63.)

Cleary, A. C., Bester, M., Forcada, J., Goebel, M., Goldsworthy, S. D., Guinet, C., Hoffman, J. I., Kovacs, K. M., Lydersen, C., and Lowther, A. D. (2019). Prey differences drive local genetic adaptation in Antarctic fur seals. *Marine Ecology Progress Series*, 628:195–209. (Cited on page 77.)

Dang, F., Li, Z., and Zhong, H. (2019). Methylmercury and selenium interactions: Mechanisms and implications for soil remediation. *https://doi.org/10.1080/10643389.2019.1583051*, 49(19):1737–1768. (Cited on page 75.)

Das, K., Debacker, V., Pillet, S., and Bouquegneau, J.-M. (2002). Heavy metals in marine mammals. In *Toxicology of marine mammals*, pages 147–179. CRC Press. (Cited on page 57.)

Daskin, J. H. and Alford, R. A. (2012). Context-dependent symbioses and their potential roles in wildlife diseases. *Proceedings of the Royal Society B: Biological Sciences*, 279(1733):1457–1465. (Cited on page 10.)

de Bruyn, P. N., Pistorius, P. A., Tosh, C. A., and Bester, M. N. (2007). Leucistic antarctic fur seal arctocephalus gazella at marion island. *Polar Biology*, 30(10):1355–1358. (Cited on page 113.)

De María, M., Szteren, D., García-Alonso, J., de Rezende, C. E., Araújo Gonçalves, R., Godoy, J. M., and Barboza, F. R. (2021). Historic variation of trace elements in pinnipeds with spatially segregated trophic habits reveals differences in exposure to pollution. *Science of the Total Environment*, 750. (Cited on page 73.)

Delport, T. C., Power, M. L., Harcourt, R. G., Webster, K. N., and Tetu, S. G. (2016). Colony location and captivity influence the gut microbial community composition of the Australian sea lion (Neophoca cinerea). *Applied and Environmental Microbiology*, 82(12):3440–3449. (Cited on pages 10 and 52.)

Desforges, J.-P. W., Sonne, C., Levin, M., Siebert, U., De Guise, S., and Dietz, R. (2016). Immunotoxic effects of environmental pollutants in marine mammals. *Environment international*, 86:126–139. (Cited on page 12.)

DeYoung, R. W. and Honeycutt, R. L. (2005). The molecular toolbox: genetic techniques in wildlife ecology and management. *The Journal of Wildlife Management*, 69(4):1362–1384. (Cited on pages 11 and 81.)

Dietrich, M., Kearney, T., Seamark, E. C., Paweska, J. T., and Markotter, W. (2018). Synchronized shift of oral, faecal and urinary microbiotas in bats and natural infection dynamics during seasonal reproduction. *Royal Society open science*, 5(5):180041. (Cited on page 10.)

Douglas, G., Maffei, V., Zaneveld, J., Yurgel, S., Brown, J., Taylor, C., Huttenhower, C., and Langille, M. (2019). PICRUSt2: An improved and customizable approach for metagenome inference. *bioRxiv*, page 672295. (Cited on page 22.)

Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., Huttenhower, C., and Langille, M. G. (2020). PICRUSt2 for prediction of metagenome functions. (Cited on pages 22 and 54.)

Drinkwater, R., Goodwin, A., Cush, J., Korstian, J. M., Chumchal, M. M., Herrera M., L. G., Valdez, M., Otálora-Ardila, A., Flores-Martinez, J. J., and Clare, E. L. (2021). Molecular diet analysis of the marine fish-eating bat (Myotis vivesi) and potential mercury exposure. *Canadian Journal of Zoology*, pages 752–759. (Cited on page 74.)

du Toit, K. H., Mole, M. A., Wege, M., Reisinger, R. R., Oosthuizen, C. W., Shihlomule, Y. D., Jordaan, R. K., van Tonder, A., and de Bruyn, P. N. (2019). Anomalous lanugo coat colourations in sub-antarctic fur seal (arctocephalus tropicalis) pups born on marion island. *Polar Biology*, 42(5):1053–1057. (Cited on page 113.)

Durazzi, F., Sala, C., Castellani, G., Manfreda, G., Remondini, D., and De Cesare, A. (2021). Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. *Scientific Reports*, 11(1):3030. (Cited on page 45.)

Dwivedi, B. and Gadagkar, S. R. (2009). Phylogenetic inference under varying proportions of indel-induced alignment gaps. *BMC Evolutionary Biology*, 9:211 – 211. (Cited on pages 27 and 28.)

Edgar, R. C. (2018). Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. *PeerJ*, 2018(4):e4652. (Cited on page 46.)

Edwards, M. and Richardson, A. J. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature*, 430(7002):881–884. (Cited on page 9.)

Eggert, L. S., Maldonado, J. E., and Fleischer, R. C. (2005). Nucleic acid isolation from ecological samples—animal scat and other associated materials. *Methods in enzymology*, 395:73–82. (Cited on page 82.)

Eilam, O., Zarecki, R., Oberhardt, M., Ursell, L. K., Kupiec, M., Knight, R., Gophna, U., and Ruppin, E. (2014). Glycan degradation (GlyDeR) analysis predicts mammalian gut microbiota abundance and host diet-specific adaptations. *mBio*, 5(4). (Cited on page 49.)

Eriksen, M., Maximenko, N., Thiel, M., Cummins, A., Lattin, G., Wilson, S., Hafner, J., Zellers, A., and Rifman, S. (2013). Plastic pollution in the South Pacific subtropical gyre. *Marine Pollution Bulletin*, 68(1-2):71–76. (Cited on pages 55 and 73.)

Fackelmann, G. and Sommer, S. (2019). Microplastics and the gut microbiome: how chronically exposed species may suffer from gut dysbiosis. *Marine pollution bulletin*, 143:193–203. (Cited on page 10.)

Fain, S. and LeMay, J. (1995). Gender identification of humans and mammalian wildlife species from pcr amplified sex linked genes. *Proc Am Acad Forensic Sci*, 1:34. (Cited on pages 29 and 96.)

Fietz, K., Galatius, A., Teilmann, J., Dietz, R., Frie, A. K., Klimova, A., Palsbøll, P. J., Jensen, L. F., Graves, J. A., Hoffman, J. I., *et al.* (2016). Shift of grey seal subspecies boundaries in response to climate, culling and conservation. *Molecular Ecology*, 25(17):4097–4112. (Cited on pages 82 and 108.)

Figueiredo, C., Baptista, M., Grilo, T., Caetano, M., Markaida, U., Raimundo, J., and Rosa, R. (2020). Bioaccumulation of Trace Elements in Myctophids in the Oxygen Minimum Zone Ecosystem of the Gulf of California. *Oceans*, 1(1):34–46. (Cited on page 77.)

Finger, A., Lavers, J. L., Dann, P., Kowalczyk, N. D., Scarpaci, C., Nugegoda, D., and Orbell, J. D. (2017). Metals and metalloids in Little Penguin (Eudyptula minor) prey, blood and faeces. *Environmental Pollution*, 223:567–574. (Cited on page 74.)

Fisher, L. R., Kon, S. K., and Thompson, S. Y. (2021). Vitamin A and carotenoids in certain invertebrates: V. Mollusca: Cephalopoda. *J. Mar. biol. Ass. U.K*, 35:63–80. (Cited on page 55.)

Fisher, R. A., Corbet, A. S., and Williams, C. B. (1943). The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. *The Journal of Animal Ecology*, 12(1):42. (Cited on page 20.)

Forcada, J. and Hoffman, J. I. (2014a). Climate change selects for heterozygosity in a declining fur seal population. *Nature*, 511(7510):462–465. (Cited on pages 9 and 110.)

Forcada, J. and Hoffman, J. I. (2014b). Climate change selects for heterozygosity in a declining fur seal population. *Nature*, 511(7510):462–465. (Cited on page 81.)

Francis, J., Boness, D., and Ochoa-Acuña, H. (1998). A protracted foraging and attendance cycle in female Juan Fernandez fur seals. *Marine Mammal Science*, 14(3):552–574. (Cited on pages 5, 54, 55, 73, 74, and 111.)

Freund, M. B., Henley, B. J., Karoly, D. J., McGregor, H. V., Abram, N. J., and Dommenget, D. (2019). Higher frequency of central pacific el niño events in recent decades relative to past centuries. *Nature Geoscience*, 12(6):450–455. (Cited on page 9.)

Friedlander, A. M., Ballesteros, E., Caselle, J. E., Gaymer, C. F., Palma, A. T., Petit, I., Varas, E., Muñoz Wilson, A., and Sala, E. (2016). Marine biodiversity in juan fernández and desventuradas islands, chile: global endemism hotspots. *PLoS One*, 11(1):e0145059. (Cited on pages 2 and 32.)

Fritz, J. (2013). Arginine cools the inflamed gut. *Infection and Immunity*, 81:3500–3502. (Cited on page 56.)

Frodello, J. P., Roméo, M., and Viale, D. (2000). Distribution of mercury in the organs and tissues of five toothed-whale species of the Mediterranean. *Environmental Pollution*, 108(3):447–452. (Cited on page 74.)

Galperin, M. Y., Brover, V., Tolstoy, I., and Yutin, N. (2016). Phylogenomic analysis of the family peptostreptococcaceae (Clostridium cluster xi) and proposal for reclassification of Clostridium litorale (Fendrich et al. 1991) and Eubacterium acidaminophilum (Zindel et al. 1989) as peptoclostridium litorale gen. nov. *International Journal of Systematic and Evolutionary Microbiology*, 66(12):5506–5513. (Cited on page 52.)

Garrett, R. G. (2000). Natural sources of metals to the environment. *Human and Ecological Risk Assessment*, 6(6):945–963. (Cited on page 57.)

Garrett, W. S. and Onderdonk, A. B. (2014). Bacteroides, Prevotella, Porphyromonas, and Fusobacterium Species (and Other Medically Important Anaerobic Gram-Negative Bacilli). In *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases*, volume 2, pages 2773–2780. Elsevier Inc. (Cited on page 49.)

Geneious Prime 2020.0.5 (2020). https://www.geneious.com [Accessed: 18-09-2021]. (Cited on pages 19 and 27.)

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., and Egozcue, J. J. (2017). Microbiome Datasets Are Compositional: And This Is Not Optional. *Frontiers in Microbiology*, 8(NOV):2224. (Cited on pages 46 and 47.)

Goldsworthy, S., Francis, J., Boness, D., and Fleischer, R. (2000). Variation in the mitochondrial control region in the Juan Fernandez fur seal (Arctocephalus philippii). *Journal of Heredity*, 91(5):371–377. (Cited on pages xvi, 6, 26, 27, 28, 81, 90, 93, 94, 95, 96, 98, 100, 106, 107, 109, and 111.)

Grajewska, A., Falkowska, L., Saniewska, D., and Pawliczka, I. (2020). Fur and faeces – Routes of mercury elimination in the Baltic grey seal (Halichoerus grypus grypus). *Science of the Total Environment*, 717:137050. (Cited on page 74.)

Gray, J. S. (2002). Biomagnification in marine systems: the perspective of an ecologist. *Marine Pollution Bulletin*, 45(1-12):46–52. (Cited on pages 9, 12, and 76.)

Grigioni, S., Boucher-Rodoni, R., Demarta, A., Tonolla, M., and Peduzzi, R. (2000). Phylogenetic characterisation of bacterial symbionts in the accessory nidamental glands of the Sepioid sepia officinalis (Cephalopoda: Decapoda). *Marine Biology*, 136(2):217–222. (Cited on page 55.)

Grosser, S., Sauer, J., Paijmans, A. J., Caspers, B. A., Forcada, J., Wolf, J. B., and Hoffman, J. I. (2019). Fur seal microbiota are shaped by the social and physical environment, show mother–offspring similarities and are associated with host genetic quality. *Molecular Ecology*, 28(9):2406–2422. (Cited on page 50.)

Groussin, M., Mazel, F., Sanders, J. G., Smillie, C. S., Lavergne, S., Thuiller, W., and Alm, E. J. (2017). Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nature Communications*, 8(1):1–12. (Cited on page 10.)

Guerrero, A. I. and Rogers, T. L. (2019). From low to high latitudes: Changes in fatty acid desaturation in mammalian fat tissue suggest a thermoregulatory role. *BMC Evolutionary Biology*, 19(1):155. (Cited on page 49.)

Guo, X., Lei, H., Zhang, K., Ke, F., and Song, C. (2020). Diversification of animal gut microbes and NRPS gene clusters in some carnivores, herbivores and omnivores. *Biotechnology and Biotechnological Equipment*, 34(1):1280–1287. (Cited on page 49.)

Halpern, B. S., Frazier, M., Afflerbach, J., Lowndes, J. S., Micheli, F., O'Hara, C., Scarborough, C., and Selkoe, K. A. (2019a). Recent pace of change in human impact on the world's ocean. *Scientific Reports*, 9(1):1–8. (Cited on page 31.)

Halpern, B. S., Frazier, M., Afflerbach, J., Lowndes, J. S., Micheli, F., O'Hara, C., Scarborough, C., and Selkoe, K. A. (2019b). Recent pace of change in human impact on the world's ocean. *Scientific reports*, 9(1):1–8. (Cited on page 1.)

Hazen, E. L., Abrahms, B., Brodie, S., Carroll, G., Jacox, M. G., Savoca, M. S., Scales, K. L., Sydeman, W. J., and Bograd, S. J. (2019). Marine top predators as climate and ecosystem sentinels. (Cited on pages 9 and 31.)

Heckel, G. and Schramm, Y. (2021). *Ecology and Conservation of Pinnipeds in Latin America*. Springer. (Cited on page 5.)

Hildebrandt, M. A., Hoffmann, C., Sherrill-Mix, S. A., Keilbaugh, S. A., Hamady, M., Chen, Y. Y., Knight, R., Ahima, R. S., Bushman, F., and Wu, G. D. (2009). High-Fat Diet Determines the Composition of the Murine Gut Microbiome Independently of Obesity. *Gastroenterology*, 137(5). (Cited on pages 48 and 49.)

Hing, K. A., Revell, P. A., Smith, N., and Buckland, T. (2006). Effect of silicon level on rate, quality and progression of bone healing within silicate-substituted porous hydroxyapatite scaffolds. *Biomaterials*, 27(29):5014–5026. (Cited on page 78.)

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). Ufboot2: improving the ultrafast bootstrap approximation. *Molecular biology and evolution*, 35(2):518–522. (Cited on page 28.)

Hoelzel, A. and Green, A. (1992). Analysis of population-level variation by sequencing pcr-amplified dna. In Hoelzel, A., editor, *Practical Approach Series: Molecular Genetic Analysis of Populations*, pages 159–187. Oxford University Press, New York, NY. (Cited on page 26.)

Hoelzel, A. R., Halley, J., O'brien, S. J., Campagna, C., Arnbom, T., Le Boeuf, B., Rails, K., and Dover, G. A. (1993). Elephant Seal Genetic Variation and the Use of Simulation Models to Investigate Historical Population Bottlenecks. Technical report, British Antartic Survey. (Cited on pages 26 and 86.)

Hoffman, J. I. (2009). A panel of new microsatellite loci for genetic studies of antarctic fur seals and other otariids. *Conservation Genetics*, 10(4):989–992. (Cited on pages 11, 26, 85, and 157.)

Hoffman, J. I. (2011). Gene discovery in the Antarctic fur seal (Arctocephalus gazella) skin transcriptome. *Molecular Ecology Resources*, 11(4):703–710. (Cited on page 26.)

Hoffman, J. I. and Nichols, H. J. (2011). A novel approach for mining polymorphic microsatellite markers in silico. *PloS one*, 6(8):e23283. (Cited on page 154.)

Honda, K., Fujise, Y., Tatsukawa, R., Itano, K., and Miyazaki, N. (1986). Age-related accumulation of heavy metals in bone of the striped dolphin, Stenella coeruleoalba. *Marine Environmental Research*, 20(3):143–160.

Honda, K. and Littman, D. R. (2016). The microbiota in adaptive immune homeostasis and disease. (Cited on page 78.)

Hubbs, C. L. and Norris, K. S. (1971). Original teeming abundance, supposed extinction, and survival of the juan fernandez fur seal. *Antarctic Pinnipedia*, 18:35–52. (Cited on pages 4, 106, and 109.)

Huh, J. W. and Roh, T. Y. (2020). Opportunistic detection of Fusobacterium nucleatum as a marker for the early gut microbial dysbiosis. *BMC Microbiology*, 20(1). (Cited on page 49.)

Humble, E., Thorne, M. A., Forcada, J., and Hoffman, J. I. (2016). Transcriptomic snp discovery for custom genotyping arrays: impacts of sequence data, snp calling method and genotyping technology on the probability of validation success. *BMC Research Notes*, 9(1):1–12. (Cited on page 155.)

Ikemoto, T., Kunito, T., Watanabe, I., Yasunaga, G., Baba, N., Miyazaki, N., Petrov, E. A., and Tanabe, S. (2004). Comparison of trace element accumulation in Baikal seals (Pusa sibirica), Caspian seals (Pusa caspica) and northern fur seals (Callorhinus ursinus). *Environmental Pollution*, 127(1):83–97. (Cited on page 76.)

Ivar Do Sul, J. A. and Costa, M. F. (2014). The present and future of microplastic pollution in the marine environment. (Cited on page 56.)

Jaishankar, M., Tseten, T., Anbalagan, N., Mathew, B. B., and Beeregowda, K. N. (2014). Toxicity, mechanism and health effects of some heavy metals. *Interdisciplinary toxicology*, 7(2):60. (Cited on pages 12 and 57.)

Jakimska, A., Konieczka, P., Skóra, K., and Namieśnik, J. (2011). Bioaccumulation of metals in tissues of marine animals, part II: Metal concentrations in animal tissues. *Polish Journal of Environmental Studies*, 20(5):1127–1146. (Cited on page 73.)

Janda, J. M. and Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. (Cited on page 45.)

Johnson, M. S., Roberts, R. D., Hutton, M., and Inskip, M. J. (1978). Distribution of Lead, Zinc and Cadmium in Small Mammals from Polluted Environments. *Oikos*, 30(1):153. (Cited on page 57.)

Jones, C. W., Risi, M. M., Osborne, A., and Bester, M. N. (2019). First record of a leucistic sub-antarctic fur seal. *Polar Biology*, 42(6):1217–1220. (Cited on page 113.)

Jost, L. (2006). Entropy and diversity. *Opinion*, 2:363–375. (Cited on pages 20 and 50.)

Jugdaohsingh, R. (2007). Silicon and bone health. *Journal of Nutrition, Health and Aging*, 11(2):99–110. (Cited on page 78.)

Jugdaohsingh, R., Tucker, K. L., Qiao, N., Cupples, L. A., Kiel, D. P., and Powell, J. J. (2004). Dietary Silicon Intake Is Positively Associated With Bone Mineral Density in Men and Premenopausal Women of the Framingham Offspring Cohort. *Journal of Bone and Mineral Research*, 19(2):297–307. (Cited on page 78.)

Kakuschke, A., Valentine-Thon, E., Griesel, S., Fonfara, S., Siebert, U., and Prange, A. (2011). Are metal-induced hypersensitivities in harbor seals associated with liver function? *Marine pollution bulletin*, 62(8):1891–1894. (Cited on pages 12 and 57.)

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., and Jermiin, L. S. (2017). Modelfinder: fast model selection for accurate phylogenetic estimates. *Nature methods*, 14(6):587–589. (Cited on page 28.)

Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. (Cited on pages 22 and 54.)

Katoh, K. and Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30(4):772–780. (Cited on page 19.)

Kehrig, H. A., Hauser-Davis, R. A., Seixas, T. G., Pinheiro, A. B., and Di Beneditto, A. P. M. (2016). Mercury species, selenium, metallothioneins and glutathione in two dolphins from the southeastern Brazilian coast: Mercury detoxification and physiological differences in diving capacity. *Environmental Pollution*, 213:785–792. (Cited on page 78.)

Keller, C., Roos, C., Groeneveld, L., Fischer, J., and Zinner, D. (2010). Introgressive hybridization in southern african baboons shapes patterns of mtdna variation. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, 142(1):125–136. (Cited on page 11.)

Khan, M. A. and Wang, F. (2009). Mercury-selenium compounds and their toxicological significance: Toward a molecular understanding of the mercury-selenium antagonism. *Environmental Toxicology and Chemistry*, 28(8):1567–1577. (Cited on page 74.)

Kim, E. J., Bu, S. Y., Sung, M. K., and Choi, M. K. (2013). Effects of silicon on osteoblast activity and bone mineralization of MC3T3-E1 cells. *Biological Trace Element Research*, 152(1):105–112. (Cited on page 78.)

Kim, M., Cho, H., and Lee, W. Y. (2020). Distinct gut microbiotas between southern elephant seals and Weddell seals of Antarctica. *Journal of Microbiology*, 58(12):1018–1026. (Cited on pages 48 and 53.)

Kirkman, S. P., Wilson, W., Klages, N. T., Bester, M. N., and Isaksen, K. (2000). Diet and estimated food consumption of Antarctic fur seals at Bouvetoya during summer. *Polar Biology*, 23(11):745–752. (Cited on page 74.)

Korbie, D. J. and Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols 2008 3:9*, 3(9):1452–1456. (Cited on page 27.)

Kovacs, K. M., Aguilar, A., Aurioles, D., Burkanov, V., Campagna, C., Gales, N., Gelatt, T., Goldsworthy, S. D., Goodman, S. J., Hofmeyr, G. J., *et al.* (2012). Global threats to pinnipeds. *Marine Mammal Science*, 28(2):414–436. (Cited on pages 6 and 9.)

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and environmental microbiology*, 79(17):5112–20. (Cited on pages 18, 32, and 50.)

Krause, D. J. and Hinke, J. T. (2021). Finally within reach: A drone census of an important, but practically inaccessible, antarctic fur seal colony. *Aquatic Mammals*, 47(4):349–354. (Cited on page 114.)

Krieger, C. J., Zhang, P., Mueller, L. A., Wang, A., Paley, S., Arnaud, M., Pick, J., Rhee, S. Y., and Karp, P. D. (2004). MetaCyc: A multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Research*, 32(DATABASE ISS.):D438. (Cited on pages 22 and 54.)

Kumari, S., Amit, Jamwal, R., Mishra, N., and Singh, D. K. (2020). Recent developments in environmental mercury bioremediation and its toxicity: A review. *Environmental Nanotechnology, Monitoring & Management*, 13:100283. (Cited on page 75.)

Lancaster, M. L., Gemmell, N. J., Negro, S., Goldsworthy, S., and Sunnucks, P. (2006). Ménage à trois on macquarie island: hybridization among three species of fur seal (arctocephalus spp.) following historical population extinction. *Molecular Ecology*, 15(12):3681–3692. (Cited on page 110.)

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., and Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology 2013 31:9*, 31(9):814–821. (Cited on pages 22, 43, and 54.)

Lanocha, N., Kalisinska, E., Kosik-Bogacka, D. I., Budis, H., Sokolowski, S., and Bohatyrewicz, A. (2013). Comparison of metal concentrations in bones of long-living mammals. *Biological Trace Element Research*, 152(2):195–203. (Cited on page 78.)

Lavery, T. J., Kemper, C. M., Sanderson, K., Schultz, C. G., Coyle, P., Mitchell, J. G., and Seuront, L. (2009). Heavy metal toxicity of kidney and bone tissues in South Australian adult bottlenose dolphins (Tursiops aduncus). *Marine Environmental Research*, 67(1):1–7. (Cited on page 78.)

Lea, M. A., Nichols, P. D., and Wilson, G. (2002). Fatty acid composition of lipid-rich myctophids and mackerel icefish (Champsocephalus gunnari) - Southern Ocean food-web implications. *Polar Biology*, 25(11):843–854. (Cited on page 54.)

Legendre, P. and Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129(2):271–280. (Cited on page 21.)

Lewis, R., O'Connell, T. C., Lewis, M., Campagna, C., and Hoelzel, A. R. (2006). Sex-specific foraging strategies and resource partitioning in the southern elephant seal (Mirounga leonina). *Proceedings of the Royal Society B: Biological Sciences*, 273(1603):2901–2907. (Cited on page 53.)

Ley, R., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., Schlegel, M. L., Tucker, T. A., Schrenzel, M. D., Knight, R., and Gordon, J. I. (2008a). Evolution of mammals and their gut microbes. *Science*, 320(5883):1647–1651. (Cited on pages 10, 31, and 49.)

Ley, R., Lozupone, C., Hamady, M., Knight, R., and Gordon, J. (2008b). Worlds within worlds: Evolution of the vertebrate gut microbiota. *Nature Reviews Microbiology*, 6(10):776–788. (Cited on pages 49, 52, and 53.)

Li, H., Qu, J., Li, T., Li, J., Lin, Q., and Li, X. (2016). Pika Population Density Is Associated with the Composition and Diversity of Gut Microbiota. *Frontiers in Microbiology*, 7(MAY):758. (Cited on page 50.)

Li, X., Brejnrod, A. D., Ernst, M., Rykær, M., Herschend, J., Olsen, N. M. C., Dorrestein, P. C., Rensing, C., and Sørensen, S. J. (2019). Heavy metal exposure causes changes in the metabolic health-associated gut microbiome and metabolites. *Environment International*, 126:454–467. (Cited on pages 58, 75, and 76.)

Litzke, V., Ottensmann, M., Forcada, J., Heitzmann, L., and Ivan Hoffman, J. (2019). Heterozygosity at neutral and immune loci is not associated with neonatal mortality due to microbial infection in antarctic fur seals. *Ecology and evolution*, 9(14):7985–7996. (Cited on page 155.)

Liu, H., Liu, K., Fu, H., Ji, R., and Qu, X. (2020). Sunlight mediated cadmium release from colored microplastics containing cadmium pigment in aqueous phase. *Environmental Pollution*, 263:114484. (Cited on page 55.)

Lomonaco, T., Manco, E., Corti, A., La Nasa, J., Ghimenti, S., Biagini, D., Di Francesco, F., Modugno, F., Ceccarini, A., Fuoco, R., and Castelvetro, V. (2020). Release of harmful volatile organic compounds (VOCs) from photo-degraded plastic debris: A neglected source of environmental pollution. *Journal of Hazardous Materials*, 394:122596. (Cited on page 55.)

Luz Calle, M. (2019). Statistical analysis of metagenomics data. (Cited on pages 46 and 47.)

Lyu, Y., Wu, L., Wang, F., Shen, X., and Lin, D. (2018). Carotenoid supplementation and retinoic acid in immunoglobulin A regulation of the gut microbiota dysbiosis:. *https://doi.org/10.1177/1535370218763760*, 243(7):613–620. (Cited on page 55.)

Marangi, M., Airoldi, S., Beneduce, L., and Zaccone, C. (2021a). Wild whale faecal samples as a proxy of anthropogenic impact. *Scientific Reports*, 11(1):1–11. (Cited on page 10.)

Marangi, M., Airoldi, S., Beneduce, L., and Zaccone, C. (2021b). Wild whale faecal samples as a proxy of anthropogenic impact. *Scientific Reports 2021 11:1*, 11(1):1–11. (Cited on pages 57, 73, 74, and 76.)

Marcovecchio, J. E., Gerpe, M. S., Bastida, R. O., Rodríguez, D. H., and Morón, S. G. (1994). Environmental contamination and marine mammals in coastal waters from Argentina: an overview. *Science of the Total Environment*, 154(2-3):141–151. (Cited on pages 57, 73, and 76.)

Markic, A., Niemand, C., Bridson, J. H., Mazouni-Gaertner, N., Gaertner, J. C., Eriksen, M., and Bowen, M. (2018). Double trouble in the South Pacific subtropical gyre: Increased plastic ingestion by fish in the oceanic accumulation zone. *Marine Pollution Bulletin*, 136:547–564. (Cited on pages 1 and 73.)

Matthee, C., Fourie, F., Oosthuizen, W., Meyer, M., and Tolley, K. (2006). Mitochondrial dna sequence data of the cape fur seal (arctocephalus pusillus pusillus) suggest that population numbers may be affected by climatic shifts. *Marine Biology*, 148(4):899–905. (Cited on page 109.)

McKnight, D. T., Huerlimann, R., Bower, D. S., Schwarzkopf, L., Alford, R. A., and Zenger, K. R. (2019). Methods for normalizing microbiome data: An ecological perspective. *Methods in Ecology and Evolution*, 10(3):389–400. (Cited on pages 20, 46, and 47.)

McMurdie, P. J. and Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4):e61217. (Cited on page 19.)

McMurdie, P. J. and Holmes, S. (2014). Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Computational Biology*, 10(4):e1003531. (Cited on pages 46 and 47.)

Medeiros, A. W., Giongo, A., Valdez, F. P., Blaese de Amorin, D., Tavares, M., D'Azevedo, P. A., Franco, A. C., Frazzon, J., and Frazzon, A. P. G. (2016). Characterization of the faecal bacterial community of wild young South American (Arctocephalus australis) and Subantarctic fur seals (Arctocephalus tropicalis). *FEMS Microbiology Ecology*, 92(3):fiw029. (Cited on page 48.)

Mendes, V. A., de Carvalho, D. P., de Almeida, R., Maria Cristina, M. C. N., Pedrosa, O. P., de Sousa-Filho, I. F., Dórea, J. G., and Bastos, W. R. (2021). Mercury in blood, hair, and feces from subsistence fish-eating riverines of the Madeira River Basin (Western Amazon). *Journal of Trace Elements in Medicine and Biology*, 67:126773. (Cited on page 74.)

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, 37(5):1530–1534. (Cited on page 28.)

Moiseenko, T. I. and Gashkina, N. A. (2020). Distribution and bioaccumulation of heavy metals (Hg, Cd and Pb) in fish: influence of the aquatic environment and climate. *Environmental Research Letters*, 15(11):115013. (Cited on page 76.)

Monteiro, L. R., Granadeiro, J. P., and Furness, R. W. (1998). Relationship between mercury levels and diet in Azores seabirds. *Marine Ecology Progress Series*, 166:259–265. (Cited on page 76.)

Moore, S. E. (2008). Marine mammals as ecosystem sentinels. *Journal of Mammalogy*, 89(3):534–540. (Cited on page 1.)

Moran, N. A. and Jarvik, T. (2010). Lateral transfer of genes from fungi underlies carotenoid production in aphids. *Science*, 328(5978):624–627. (Cited on page 55.)

Morris, E. K., Caruso, T., Buscot, F., Fischer, M., Hancock, C., Maier, T. S., Meiners, T., Müller, C., Obermaier, E., Prati, D., Socher, S. A., Sonnemann, I., Wäschke, N., Wubet, T., Wurst, S., and Rillig, M. C. (2014). Choosing and using diversity indices: insights for ecological applications from the German Biodiversity Exploratories. *Ecology and evolution*, 4(18):3514–24. (Cited on page 20.)

Nearing, J. T., Douglas, G. M., Comeau, A. M., and Langille, M. G. (2018). Denoising the Denoisers: An independent evaluation of microbiome sequence error- correction approaches. *PeerJ*, 2018(8). (Cited on page 46.)

Nelson, T. M., Rogers, T. L., and Brown, M. V. (2013a). The gut bacterial community of mammals from marine and terrestrial habitats. *PLoS ONE*, 8(12):83655. (Cited on pages 32 and 48.)

Nelson, T. M., Rogers, T. L., Carlini, A. R., and Brown, M. V. (2013b). Diet and phylogeny shape the gut microbiota of Antarctic seals: A comparison of wild and captive animals. *Environmental Microbiology*, 15(4):1132–1145. (Cited on pages 32, 48, 52, and 53.)

Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., and Pettersson, S. (2012). Host-gut microbiota metabolic interactions. (Cited on pages 31 and 52.)

Nigro, M., Campana, A., Lanzillotta, E., and Ferrara, R. (2002). Mercury exposure and elimination rates in captive bottlenose dolphins. *Marine Pollution Bulletin*, 44(10):1071–1075. (Cited on page 74.)

Nishida, A. H. and Ochman, H. (2018). Rates of gut microbiome divergence in mammals. *Molecular Ecology*, 27(8):1884–1897. (Cited on page 53.)

Nordberg, G. F., Nogawa, K., and Nordberg, M. (2015). Cadmium. In Nordberg, G. F., Fowler, B. A., and Nordberg, M., editors, *Handbook on the Toxicology of Metals: Fourth Edition*, volume 1, chapter 32, pages 667–716. Academic Press, 4 edition. (Cited on pages 57 and 78.)

Nsubuga, A. M., Robbins, M. M., Roeder, A. D., Morin, P. A., Boesch, C., and Vigilant, L. (2004). Factors affecting the amount of genomic dna extracted from ape faeces and the identification of an improved sample storage method. *Molecular ecology*, 13(7):2089–2094. (Cited on page 82.)

Numberger, D., Herlemann, D. P. R., Jürgens, K., Dehnhardt, G., and Schulz-Vogt, H. (2016). Comparative analysis of the fecal bacterial community of five harbor seals (Phoca vitulina). *MicrobiologyOpen*, 5(5):782–792. (Cited on page 48.)

Ochoa Acuna, H. and Francis, J. M. (1995). Spring and summer prey of the Juan Fernandez fur seal, Arctocephalus philippii. *Canadian Journal of Zoology*, 73(8):1444–1452. (Cited on pages 5, 32, 53, 67, 74, 76, 77, 111, 113, and 114.)

Ochoa-Acuña, H., Francis, J. M., and Boness, D. J. (1998). Interannual variation in birth mass and postnatal growth rate of juan fernandez fur seals. *Canadian Journal of Zoology*, 76(5):978–983. (Cited on page 5.)

Ochoa-Acuna, H., Francis, J. M., and Oftedal, O. T. (1999). Influence of Long Intersuckling Interval on Composition of Milk in the Juan Fernández Fur Seal, Arctocephalus philippii. *Journal of Mammalogy*, 80(3):758–767. (Cited on page 53.)

O'Hara, A. M. and Shanahan, F. (2006). The gut flora as a forgotten organ. *EMBO reports*, 7(7):688–693. (Cited on pages 10 and 31.)

Olsen, I. (2014). The Family Fusobacteriaceae. In Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E., and Thompson, F., editors, *The Prokaryotes*, pages 109–132. Springer, Berlin, Heidelberg. (Cited on page 49.)

Osinga, N., van Voorst Vader, P., *et al.* (2010). Albinistic common seals (phoca vitulina) and melanistic grey seals (halichoerus grypus) rehabilitated in the netherlands. *Animal Biology*, 60(3):273–281. (Cited on page 113.)

Osman, L. and Moreno, C. (2017). Population ecology, trends and distribution of the juan fernandez fur seal, arctocephalus philippii (peters 1866) in chile. In Alava, J. J., editor, *Tropical pinnipeds*, pages 190–207. CRC Press, 1st edition. (Cited on pages 4, 6, 9, and 81.)

Osman, L. P., Moreno, C. A., and Trites, A. W. (2010). Growth rates and differential investment in male and female juan fernández fur seal pups. *Journal of Mammalogy*, 91(5):1188–1196. (Cited on pages 5 and 111.)

Pacheco-Sandoval, A., Schramm, Y., Heckel, G., Brassea-Pérez, E., Martínez-Porchas, M., and Lago-Lestón, A. (2019). The Pacific harbor seal gut microbiota in Mexico: Its relationship with diet and functional inferences. *PLoS ONE*, 14(8). (Cited on pages 32, 48, 49, 50, 52, and 54.)

Páez-Rosas, D., Torres, J., Espinoza, E., Marchetti, A., Seim, H., and Riofrío-Lazo, M. (2021). Declines and recovery in endangered galapagos pinnipeds during the el niño event. *Scientific reports*, 11(1):1–15. (Cited on pages 9, 81, and 110.)

Paradis, E. (2010). pegas an R package for population genetics with an integrated-modular approach. *Bioinformatics*, 26(3):419–420. (Cited on page 28.)

Patrick, L. (2002). Patrick l. mercury toxicity and antioxidants: Part 1: role of glutathione and alpha-lipoic acid in the treatment... *Alternative Medicine Review*, 7(6):456–471. (Cited on page 12.)

Penicaud, V., Lacoue-Labarthe, T., and Bustamante, P. (2017). Metal bioaccumulation and detoxification processes in cephalopods: A review. (Cited on page 76.)

Pereira, F. L., Oliveira Júnior, C. A., Silva, R. O., Dorella, F. A., Carvalho, A. F., Almeida, G. M., Leal, C. A., Lobato, F. C., and Figueiredo, H. C. (2016). Complete genome sequence of Peptoclostridium difficile strain Z31. *Gut Pathogens*, 8(1):11. (Cited on page 52.)

Perez-Venegas, D., Toro-Valdivieso, C., Meneses, D., Ayala, F., Seguel, M., Pavés, H., Cárdenas, S., Sepúlveda, M., and Glabán-Malagón, C. (2018). Microfiber ingestion in otariids along Perú-Chile coast: A possible sentinel of biomagnification? (Cited on pages 6, 9, 55, and 73.)

Perez-Venegas, D. J., Toro-Valdivieso, C., Ayala, F., Brito, B., Iturra, L., Arriagada, M., Seguel, M., Barrios, C., Sepúlveda, M., Oliva, D., *et al.* (2020). Monitoring the occurrence of microplastic ingestion in otariids along the peruvian and chilean coasts. *Marine pollution bulletin*, 153:110966. (Cited on pages 111 and 112.)

Polizzi, P., Romero, M. B., Chiodi Boudet, L. N., Ponce de León, A., Medici, S., Costas, A., Rodríguez, D., and Gerpe, M. (2017). Blood cadmium and metallothionein concentrations in females of two sympatric pinnipeds species. *Marine Pollution Bulletin*, 122(1-2):446–449. (Cited on page 78.)

Pompa, S., Ehrlich, P. R., and Ceballos, G. (2011). Global distribution and conservation of marine mammals. *Proceedings of the National Academy of Sciences*, 108(33):13600–13605. (Cited on pages 2 and 32.)

Porobic, J., Fulton, E. A., Parada, C., Frusher, S., Ernst, B., and Manríquez, P. (2019). The impact of fishing on a highly vulnerable ecosystem, the case of juan fernández ridge ecosystem. *PloS one*, 14(2):e0212485. (Cited on page 2.)

Potrykus, J., White, R. L., and Bearne, S. L. (2008). Proteomic investigation of amino acid catabolism in the indigenous gut anaerobe Fusobacterium varium. *PROTEOMICS*, 8(13):2691–2703. (Cited on page 49.)

Price, C. T., Koval, K. J., and Langford, J. R. (2013). Silicon: A review of its potential role in the prevention and treatment of postmenopausal osteoporosis. (Cited on page 78.)

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, 5(3):e9490. (Cited on page 19.)

Pu, S., Khazanehei, H., Jones, P. J., and Khafipour, E. (2016). Interactions between Obesity Status and Dietary Intake of Monounsaturated and Polyunsaturated Oils on Human Gut Microbiome Profiles in the Canola Oil Multicenter Intervention Trial (COMIT). *Frontiers in Microbiology*, 7(OCT):1612. (Cited on pages 48 and 49.)

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Rg Peplies, J., and Glö Ckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic*, 41(D1):D590–D596. (Cited on page 19.)

R Core Team (2019). R: The R Project for Statistical Computing. (Cited on pages 19, 24, and 27.)

Ralls, K. and Mesnick, S. L. (2009). Sexual Dimorphism. In Perrin, W. F., Bernd, W., and Thewissen, J. G. M., editors, *Encyclopedia of Marine Mammals*, pages 1005–1011. Elsevier Science & Technology, London, 2nd edition. (Cited on page 53.)

Ramón-Laca, A., Soriano, L., Gleeson, D., and Godoy, J. A. (2015). A simple and effective method for obtaining mammal DNA from faeces. *Wildlife Biology*, 21(4):195–203. (Cited on page 16.)

Ramón-Laca, A., Soriano, L., Gleeson, D., and Godoy, J. A. (2015). A simple and effective method for obtaining mammal dna from faeces. *Wildlife Biology*, 21(4):195–203. (Cited on page 82.)

Ranjan, R., Rani, A., Metwally, A., McGee, H. S., and Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and Biophysical Research Communications*, 469(4):967–977. (Cited on page 45.)

Reddy, P. A., Bhavanishankar, M., Bhagavatula, J., Harika, K., Mahla, R. S., and Shivaji, S. (2012). Improved methods of carnivore faecal sample preservation, dna extraction and quantification for accurate genotyping of wild tigers. *PLoS ONE*, 7(10). (Cited on page 82.)

Reed, J., Tollit, D., Thompson, P., and Amos, W. (1997a). Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. *Molecular ecology*, 6(3):225–234. (Cited on pages 82 and 108.)

Reed, J. Z., Tollit, D. J., Thompson, P. M., and Amos, W. (1997b). Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. *Molecular ecology*, 6(3):225–234. (Cited on page 25.)

Reffitt, D. M., Ogston, N., Jugdaohsingh, R., Cheung, H. F., Evans, B. A., Thompson, R. P., Powell, J. J., and Hampson, G. N. (2003). Orthosilicic acid stimulates collagen type 1 synthesis and osteoblastic differentiation in human osteoblast-like cells in vitro. *Bone*, 32(2):127–135. (Cited on page 78.)

Ribeiro, A. R., Eira, C., Torres, J., Mendes, P., Miquel, J., Soares, A. M. V. M., and Vingada, J. (2008). Toxic Element Concentrations in the Razorbill Alca torda (Charadriiformes, Alcidae) in Portugal. *Archives of Environmental Contamination and Toxicology 2008 56:3*, 56(3):588–595. (Cited on page 74.)

Richardson, J. B., Dancy, B. C., Horton, C. L., Lee, Y. S., Madejczyk, M. S., Xu, Z. Z., Ackermann, G., Humphrey, G., Palacios, G., Knight, R., and Lewis, J. A. (2018). Exposure to toxic metals triggers unique responses from the rat gut microbiota. *Scientific Reports*, 8(1):6578. (Cited on pages 58, 63, and 75.)

Roberts, S. M. (2003). Examination of the stomach contents from a Mediterranean sperm whale found south of Crete, Greece. *Journal of the Marine Biological Association of the United Kingdom*, 83(3):667–670. (Cited on page 74.)

Robertson, K. M., Lauf, M. L., and Morin, P. A. (2018). Genetic sexing of pinnipeds: a real-time, single step qPCR technique. *Conservation Genetics Resources*, 10(2):213–218. (Cited on pages 26, 29, 96, 102, and 105.)

Robinson, H., Thayer, J., Sydeman, W. J., and Weise, M. (2018). Changes in california sea lion diet during a period of substantial climate variability. *Marine Biology*, 165(10):1–12. (Cited on page 9.)

Rodriguez-Concepcion, M., Avalos, J., Bonet, M. L., Boronat, A., Gomez-Gomez, L., Hornero-Mendez, D., Limon, M. C., Meléndez-Martínez, A. J., Olmedilla-Alonso, B., Palou, A., Ribot, J., Rodrigo, M. J., Zacarias, L., and Zhu, C. (2018). A global perspective on carotenoids: Metabolism, biotechnology, and benefits for nutrition and health. *Progress in Lipid Research*, 70:62–93. (Cited on page 55.)

Romero, V. and Tirira, D. G. (2017). Leucistic antarctic fur seal (arctocephalus gazella) at robert island, south shetland islands, antarctica, with a note on colour morph nomenclature. *Polar Biology*, 40(9):1893–1897. (Cited on page 113.)

Rosel, P. E., Taylor, B. L., Hancock-Hanser, B. L., Morin, P. A., Archer, F. I., Lang, A. R., Mesnick, S. L., Pease, V. L., Perrin, W. F., Robertson, K. M., Leslie, M. S., Berta, A., Cipriano, F., Parsons, K. M., Viricel, A., Vollmer, N. L., and Martien, K. K. (2017). A review of molecular genetic markers and analytical approaches that have been used for delimiting marine mammal subspecies and species. *Marine Mammal Science*, 33:56–75. (Cited on pages 11 and 109.)

Rothenberg, S. E., Sweitzer, D. N., Rackerby, B. R., Couch, C. E., Cohen, L. A., Broughton, H. M., Steingass, S. M., and Beechler, B. R. (2021). Fecal Methylmercury Correlates With Gut Microbiota Taxa in Pacific Walruses (Odobenus rosmarus divergens). *Frontiers in Microbiology*, 0:1320. (Cited on pages 74 and 75.)

Rowland, I. R., Davies, M. J., and Grasso, P. (1978). Metabolism of methylmercuric chloride by the gastro-intestinal flora of the rat. *Xenobiotica*, 8(1):37–43. (Cited on page 76.)

Rubinoff, D., Cameron, S., and Will, K. (2006). A genomic perspective on the shortcomings of mitochondrial dna for "barcoding" identification. *Journal of heredity*, 97(6):581–594. (Cited on page 11.)

Sakamoto, M., Itai, T., Yasutake, A., Iwasaki, T., Yasunaga, G., Fujise, Y., Nakamura, M., Murata, K., Man Chan, H., Domingo, J. L., and Marumoto, M. (2015). Mercury speciation and selenium in toothed-whale muscles. *Environmental Research*, 143:55–61. (Cited on page 74.)

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J., and Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, 12(1):87. (Cited on pages 20 and 47.)

Schaefer, A. M., Stavros, H.-C. W., Bossart, G. D., Fair, P. A., Goldstein, J. D., and Reif, J. S. (2011). Associations between mercury and hepatic, renal, endocrine, and hematological parameters in atlantic bottlenose dolphins (tursiops truncatus) along the eastern coast of florida and south carolina. *Archives of environmental contamination and toxicology*, 61(4):688–695. (Cited on pages 12 and 57.)

Schaefer, J. K., Yagi, J., Reinfelder, J. R., Cardona, T., Ellickson, K. M., Tel-Or, S., and Barkay, T. (2004). Role of the bacterial organomercury lyase (MerB) in controlling methylmercury accumulation in mercury-contaminated natural waters. *Environmental Science and Technology*, 38(16):4304–4311. (Cited on page 75.)

Scheffer, V. B. (1958). Seals, sea lions, and walruses: a review of the pinnipedia. (Cited on page 3.)

Schumann, N., Gales, N. J., Harcourt, R. G., and Arnould, J. P. (2013). Impacts of climate change on australian marine mammals. *Australian Journal of Zoology*, 61(2):146–159. (Cited on page 81.)

Semova, I., Carten, J. D., Stombaugh, J., MacKey, L. C., Knight, R., Farber, S. A., and Rawls, J. F. (2012). Microbiota regulate intestinal absorption and metabolism of fatty acids in the zebrafish. *Cell Host and Microbe*, 12(3):277–288. (Cited on pages 49 and 53.)

Sepúlveda, M., Quiñones, R., Esparza, C., Carrasco, P., and Winckler, P. (2020). Vulnerability of a top marine predator to coastal storms: a relationship between hydrodynamic drivers and stranding rates of newborn pinnipeds. *Scientific reports*, 10(1):1–12. (Cited on page 9.)

Sepúlveda, M. S. (1998). Hookworms (uncinaria sp.) in juan fernandez fur seal pups (arctocephalus philippii) from alejandro selkirk island, chile. *The Journal of parasitology*, pages 1305–1307. (Cited on page 5.)

Sepúlveda, M. S. and Alcaíno, H. (1993). Fauna helmintológica en el lobo fino de juan fernández, arctocephalus philippii (peters, 1866). *Parasitol. día*, 17:19–24. (Cited on page 5.)

Sepúlveda, M. S., Ochoa-Acuña, H., and Sundlof, S. F. (1997). Heavy metal concentrations in Juan Fernandez fur seals (Arctocephalus philippii). *Marine Pollution Bulletin*, 34(8):663–665. (Cited on pages 5, 9, 73, and 111.)

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423. (Cited on page 20.)

Shi, W., Zhao, X., Han, Y., Che, Z., Chai, X., and Liu, G. (2016). Ocean acidification increases cadmium accumulation in marine bivalves: a potential threat to seafood safety. *Scientific Reports*, 6(1):1–8. (Cited on page 12.)

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1):539. (Cited on page 27.)

Simmonds, M. P. and Isaac, S. J. (2007). The impacts of climate change on marine mammals: early signs of significant problems. *Oryx*, 41(1):19–26. (Cited on pages 6 and 9.)

Simpson, E. H. (1949). Measurement of diversity [16]. (Cited on page 20.)

Sivertsen, E. (1954). A survey of the eared seals (family otariidae) with remarks of the antarctic seals collected by m/k« norvegia» in 1928-1929. *Det Norske Videnskapo-Akademi i Oslo*, 36:5–76. (Cited on page 3.)

Slade, R. W., Moritz, C., and Heideman, A. (1994). Multiple nuclear-gene phylogenies: application to pinnipeds and comparison with a mitochondrial dna gene phylogeny. *Molecular Biology and Evolution*, 11(3):341–356. (Cited on page 11.)

Smith, S. C., Chalker, A., Dewar, M. L., and Arnould, J. P. (2013). Age-related differences revealed in australian fur seal arctocephalus pusillus doriferus gut microbiota. *FEMS Microbiology Ecology*, 86(2):246–255. (Cited on pages 32, 48, and 53.)

Squadrone, S., Chiaravalle, E., Gavinelli, S., Monaco, G., Rizzi, M., and Abete, M. C. (2015). Analysis of mercury and methylmercury concentrations, and selenium: Mercury molar ratios for a toxicological assessment of sperm whales (Physeter macrocephalus) in the most recent stranding event along the Adriatic coast (Southern Italy, Mediterranean Sea). *Chemosphere*, 138:633–641. (Cited on page 74.)

Stappenbeck, T. S. and Virgin, H. W. (2016). Accounting for reciprocal host–microbiome interactions in experimental science. *Nature*, 534(7606):191–199. (Cited on page 10.)

Steinmetz, K., Murphy, S., Cadhla, O. Ó., Barnett, J., and Mirimin, L. (2021). Enabling pinniped conservation by means of non-invasive genetic population analysis. *Conservation Genetics Resources*, 13(2):131–142. (Cited on pages 82 and 108.)

Stoffel, M., Humble, E., Paijmans, A., Acevedo-Whitehouse, K., Chilvers, B. L., Dickerson, B., Galimberti, F., Gemmell, N. J., Goldsworthy, S., Nichols, H., *et al.* (2018). Demographic histories and genetic diversity across pinnipeds are shaped by human exploitation, ecology and life-history. *Nature Communications*, 9(1):1–12. (Cited on page 110.)

Stoffel, M. A., Acevedo-Whitehouse, K., Morales-Durán, N., Grosser, S., Chakarov, N., Krüger, O., Nichols, H. J., Elorriaga-Verplancken, F. R., and Hoffman, J. I. (2020). Early sexual dimorphism in the developing gut microbiome of northern elephant seals. *Molecular Ecology*, 29(11):2109–2122. (Cited on pages 32, 48, and 53.)

Stoneking, M. (2000). Hypervariable sites in the mtDNA control region are mutational hotspots. *American Journal of Human Genetics*, 67(4):1029–1032. (Cited on pages 11 and 108.)

Storelli, M. M., Garofalo, R., Giungato, D., and Giacominelli-Stuffler, R. (2010). Intake of essential and non-essential elements from consumption of octopus, cuttlefish and squid. *Food Additives and Contaminants: Part B Surveillance*, 3(1):14–18. (Cited on page 76.)

Storelli, M. M., Giacominelli-Stuffler, R., Storelli, A., and Marcotrigiano, G. O. (2006). Cadmium and mercury in cephalopod molluscs: Estimated weekly intake. *Food Additives and Contaminants*, 23(1):25–30. (Cited on page 76.)

Suzuki, A., Akuzawa, K., Kogi, K., Ueda, K., and Suzuki, M. (2021). Captive environment influences the composition and diversity of fecal microbiota in indo-pacific bottlenose dolphins, tursiops aduncus. *Marine Mammal Science*, 37(1):207–219. (Cited on page 10.)

Tatarenkov, A. and Avise, J. C. (2007). Rapid concerted evolution in animal mitochondrial dna. *Proceedings of the Royal Society B: Biological Sciences*, 274(1619):1795–1798. (Cited on page 11.)

Tchounwou, P. B., Yedjou, C. G., Patlolla, A. K., and Sutton, D. J. (2012). Heavy metal toxicity and the environment. (Cited on page 57.)

Torres, D. (1987). Juan fernandez fur seal, arctocephalus philippii. In *International symposium and workshop on the status, biology, and ecology of fur seals*. (Cited on pages 4, 81, and 106.)

Torres, D., Guerra, C., and Cárdenas, J. (1984). Primeros registros de arctocephalus gazella y nuevos hallazgos de arctocephalus tropicalis y leptonychotes weddelli en el archipiélago de juan fernández. *Ser. Cient. INACH*, 31:115–158. (Cited on pages 6 and 110.)

Trevelline, B. K., Fontaine, S. S., Hartup, B. K., and Kohl, K. D. (2019). Conservation biology needs a microbial renaissance: a call for the consideration of host-associated microbiota in wildlife management practices. *Proceedings of the Royal Society B*, 286(1895):20182448. (Cited on page 31.)

Trites, A. W. (2019). Marine Mammal Trophic Levels and Trophic Interactions. In Cochran, J. K., Bokuniewicz. J, H., , and Yager, P. L., editors, *Encyclopedia of Ocean Sciences*, chapter Marine Mam, pages 589–594. Elsevier Ltd, 3rd edition. (Cited on page 32.)

Tucker, S., Don Bowen, W., and Iverson, S. J. (2007). Dimensions of diet segregation in grey seals Halichoerus grypus revealed through stable isotopes of carbon (δ13C) and nitrogen (δ15N). *Marine Ecology Progress Series*, 339:271–282. (Cited on page 79.)

Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., and Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*, 444(7122):1027–1031. (Cited on pages 48 and 49.)

Uhen, M. D. (2007). Evolution of marine mammals: back to the sea after 300 million years. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology: Advances in Integrative Anatomy and Evolutionary Biology*, 290(6):514–522. (Cited on page 2.)

van Dorst, J., Siciliano, S. D., Winsley, T., Snape, I., and Ferrari, B. C. (2014). Bacterial targets as potential indicators of diesel fuel toxicity in subantarctic soils. *Applied and Environmental Microbiology*, 80(13):4021–4033. (Cited on page 55.)

Varsha, G. (2013). Mammalian Feces as Bio-Indicator of Heavy Metal contamination in Bikaner Zoological Garden, Rajasthan, India. *Res. J. Animal, Veterinary and Fishery Sci. International Science Congress Association*, 1(5):10–15. (Cited on page 57.)

Veron, S., Mouchet, M., Govaerts, R., Haevermans, T., and Pellens, R. (2019). Vulnerability to climate change of islands worldwide and its impact on the tree of life. *Scientific reports*, 9(1):1–14. (Cited on page 1.)

Vieira, M. L. C., Santini, L., Diniz, A. L., and Munhoz, C. d. F. (2016). Microsatellite markers: what they mean and why they are so useful. *Genetics and molecular biology*, 39:312–328. (Cited on page 11.)

Vlčková, K., Mrázek, J., Kopečný, J., and Petrželková, K. J. (2012). Evaluation of different storage methods to characterize the fecal bacterial communities of captive western lowland gorillas (Gorilla gorilla gorilla). *Journal of Microbiological Methods*, 91(1):45–51. (Cited on page 17.)

Von Schuckmann, K., Holland, E., Haugan, P., and Thomson, P. (2020). Ocean science, data, and services for the UN 2030 Sustainable Development Goals Relating the Sustainable Development Goal (SDG) 14 for Ocean and Life Below Water to the 16 remaining SDGs in the. *Marine Policy*, 121:104154. (Cited on page 1.)

Wang, W. C., Mao, H., Ma, D. D., and Yang, W. X. (2014). Characteristics, functions, and applications of metallothionein in aquatic vertebrates. (Cited on pages 12 and 78.)

Wasser, S. K., Hunt, K. E., Brown, J. L., Cooper, K., Crockett, C. M., Bechert, U., Millspaugh, J. J., Larson, S., and Monfort, S. L. (2000). A generalized fecal glucocorticoid assay for use in a diverse array of nondomestic mammalian and avian species. *General and Comparative Endocrinology*, 120(3):260–275. (Cited on page 51.)

Weber, D., Stewart, B., and Lehman, N. (2004). Genetic consequences of a severe population bottleneck in the guadalupe fur seal (arctocephalus townsendi). *Journal of Heredity*, 95(2):144–153. (Cited on page 109.)

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E. R., and Knight, R. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome*, 5(1). (Cited on pages 46 and 47.)

Williams, C. L., Caraballo-Rodríguez, A. M., Allaband, C., Zarrinpar, A., Knight, R., and Gauglitz, J. M. (2018). Wildlife-microbiome interactions and disease: exploring opportunities for disease mitigation across ecological scales. *Drug Discovery Today: Disease Models*, 28:105–115. (Cited on page 31.)

Willis, A. D. (2019). Rarefaction, alpha diversity, and statistics. *Frontiers in Microbiology*, 10(OCT):2407. (Cited on page 47.)

Wood, S. A., Robinson, P. W., Costa, D. P., and Beltran, R. S. (2021). Accuracy and precision of citizen scientist animal counts from drone imagery. *PloS one*, 16(2):e0244040. (Cited on page 114.)

Woodhams, D. C., Bletz, M. C., Becker, C. G., Bender, H. A., Buitrago-Rosas, D., Diebboll, H., Huynh, R., Kearns, P. J., Kueneman, J., Kurosawa, E., *et al.* (2020). Host-associated microbiomes are predicted by immune system complexity and climate. *Genome biology*, 21(1):1–20. (Cited on page 10.)

Wynen, L. P., Goldsworthy, S. D., Insley, S. J., Adams, M., Bickham, J. W., Francis, J., Gallo, J. P., Hoelzel, A., Majluf, P., White, R. W., and Slade, R. (2001). Phylogenetic Relationships within the Eared Seals (Otariidae: Carnivora): Implications for the Historical Biogeography of the Family. *Molecular Phylogenetics and Evolution*, 21(2):270–284. (Cited on pages 3 and 27.)

Xia, Y. and Sun, J. (2017). Hypothesis testing and statistical analysis of microbiome. (Cited on pages 46 and 47.)

Yabe, J., Nakayama, S. M., Ikenaka, Y., Yohannes, Y. B., Bortey-Sam, N., Kabalo, A. N., Ntapisha, J., Mizukawa, H., Umemura, T., and Ishizuka, M. (2018). Lead and cadmium excretion in feces and urine of children from polluted townships near a lead-zinc mine in Kabwe, Zambia. *Chemosphere*, 202:48–55. (Cited on page 74.)

Yang, T. T., Liu, Y., Tan, S., Wang, W. X., and Wang, X. (2021). The role of intestinal microbiota of the marine fish (Acanthopagrus latus) in mercury biotransformation. *Environmental Pollution*, 277:116768. (Cited on pages 63 and 76.)

Youness, E. R., Mohammed, N. A., and Morsy, F. A. (2012). Cadmium impact and osteoporosis: mechanism of action. *http://dx.doi.org/10.3109/15376516.2012.702796*, 22(7):560–567. (Cited on page 78.)

Yuan, Y., Zhang, Y., Zhang, P., Liu, C., Wang, J., Gao, H., Hoelzel, A. R., Seim, I., Lv, M., Lin, M., *et al.* (2021). Comparative genomics provides insights into the aquatic adaptations of mammals. *Proceedings of the National Academy of Sciences*, 118(37). (Cited on page 2.)

Zappes, I. A., Fabiani, A., Sbordoni, V., Rakaj, A., Palozzi, R., and Allegrucci, G. (2017). New data on weddell seal (leptonychotes weddellii) colonies: A genetic analysis of a top predator from the ross sea, antarctica. *Plos one*, 12(8):e0182922. (Cited on pages 82 and 108.)

Zhu, L., Wu, Q., Deng, C., Zhang, M., Zhang, C., Chen, H., Lu, G., and Wei, F. (2018). Adaptive evolution to a high purine and fat diet of carnivorans revealed by gut microbiomes and host genomes. *Environmental Microbiology*, 20(5):1711–1722. (Cited on page 49.)

Appendix A

Supplementary material

Table A.1. Summary information on the JFFS faecal samples for the studies.
The last three columns indicate whether the sample was used in a study: host genetics (Host_Gen), heavy metals (Heavy_M) and 16S microbiome analysis (16S). (An empty entry means NO.)

Sample ID	Year	Location	Host_Gen	Heavy_M	16S
17JFFS2	2017	BahiaPadre	YES		YES
17JFFS3	2017	BahiaPadre	YES		
17JFFS4	2017	BahiaPadre			YES
17JFFS5	2017	BahiaPadre			YES
17JFFS6	2017	BahiaPadre			
17JFFS7	2017	BahiaPadre	YES		YES
17JFFS8	2017	BahiaPadre	YES		YES
17JFFS9	2017	BahiaPadre			YES
17JFFS10	2017	BahiaPadre			YES
17JFFS11	2017	BahiaPadre			YES
17JFFS12	2017	BahiaPadre	YES		YES
17JFFS13	2017	BahiaPadre	YES		YES
17JFFS14	2017	BahiaPadre	YES		YES
17JFFS15	2017	BahiaPadre	YES		YES
17JFFS16	2017	BahiaPadre			YES
17JFFS17	2017	BahiaPadre	YES		YES
17JFFS18	2017	BahiaPadre	YES		YES
17JFFS19	2017	BahiaPadre	YES		YES
17JFFS20	2017	BahiaPadre	YES		
17JFFS21	2017	BahiaPadre	YES		YES
17JFFS22	2017	TierrasBlancas			YES
17JFFS23	2017	TierrasBlancas	YES		YES
17JFFS24	2017	TierrasBlancas			YES
17JFFS25	2017	TierrasBlancas	YES		YES
17JFFS26	2017	TierrasBlancas	YES		YES
17JFFS27	2017	TierrasBlancas	YES		YES
17JFFS28	2017	TierrasBlancas	YES		YES
17JFFS29	2017	TierrasBlancas	YES		YES

Sample ID	Year	Location	Host_Gen	Heavy_M	16S
17JFFS30	2017	ierrasBlanca	YES		YES
17JFFS31	2017	Arenal	YES		YES
18JFFS1	2018	ierrasBlanca	YES		YES
18JFFS2	2018	ierrasBlanca	YES		YES
18JFFS3	2018	ierrasBlanca	YES		YES
18JFFS4	2018	Vaqueria	YES		YES
18JFFS5	2018	BahiaPadre		YES	YES
18JFFS6	2018	BahiaPadre	YES	YES	YES
18JFFS7	2018	BahiaPadre		YES	YES
18JFFS8	2018	Arenal	YES		YES
18JFFS9	2018	Arenal	YES		YES
18JFFS10	2018	Arenal	YES	YES	YES
18JFFS11	2018	Arenal		YES	YES
18JFFS12	2018	Arenal		YES	YES
18JFFS13	2018	Arenal	YES	YES	YES
18JFFS14	2018	Arenal		YES	YES
18JFFS15	2018	BahiaPadre	YES	YES	YES
18JFFS16	2018	BahiaPadre			YES
18JFFS17	2018	iedraCarvajal			YES
18JFFS18	2018	LaMatriz		YES	YES
18JFFS19	2018	LaMatriz	YES	YES	YES
18JFFS20	2018	SantaClara	YES	YES	YES
18JFFS21	2018	SantaClara	YES		YES
18JFFS23	2018	SantaClara	YES	YES	YES
18JFFS24	2018	PtaTrueno			YES
18JFFS25	2018	SantaClara	YES		YES
18JFFS26	2018	SantaClara	YES		YES
18JFFS27	2018	SantaClara	YES		YES
18JFFS28	2018	SantaClara			YES
18JFFS29	2018	SantaClara	YES		YES
18JFFS30	2018	SantaClara			YES
18JFFS31	2018	SantaClara			YES
19JFFS1	2019	ierrasBlanca	YES		
19JFFS2	2019	ierrasBlanca	YES		
19JFFS3	2019	ierrasBlanca	YES		
19JFFS4	2019	ierrasBlanca	YES		
19JFFS5	2019	ierrasBlanca	YES		
19JFFS6	2019	ierrasBlancas			
19JFFS7	2019	ierrasBlanca	YES		
19JFFS8	2019	Arenal	YES		
19JFFS9	2019	Arenal	YES		
19JFFS10	2019	SantaClara			
19JFFS11	2019	SantaClara	YES		
19JFFS12	2019	SantaClara			
19JFFS13	2019	SantaClara	YES		
19JFFS14	2019	SantaClara	YES		
19JFFSB1	2019	Pto.Ingles	YES		
19JFFSB2	2019	Pto.Ingles	YES		
19JFFSNB1	2019	Pto.Ingles	YES		
19JFFSNB2	2019	Pto.Ingles	YES		

Table A.2. Effects of consecutive filtering on 16S rRNA gene sequencing data.
The first row shows the characteristics of the initial denoised, sequencing data input followed by row showing the output of the respective filtering step (identified in the first column). Rows are organised in a sequencial order. Thus, each row describes the input data for the next filtering step (next row).

prepocessing steps	Number of samples	Number of ASVs	Min. number of reads per sample	Max. number of read per sample	Filtered reads	Total
Raw	57	595	2042	76134	0	2074038
Filter ASVs (non-bacterial and ambiguous)	57	577	2042	76134	2081	2071957
Filter samples	54	577	13981	76134	8916	2063041
Filter Contaminants	54	558	13981	76134	278	2062763
Rarefaction	54	518	13981	13981	1307789	754974

Table A.3. Primer used to amplify host DNA.
The prefix Ext indicates a primer set designed for running an external PCR. Numbers in bold indicate the product melting temperature instead of the annealing temperature.

Application	Primer	Forward 5'-3'	Reverse 5'-3'	amplicon	Annealing T	obsevation	Reference
Microsatellite	Agaz-1	ACTCATGCCCTGCTTGAAAT	TGCTGGCCTAAGTCTCCTG	245	58	Neutral	Hoffman, 2009
	Ext_Agaz-1	GTAGGCAATTTTAGGGCAAG	GATCTAGGTCTGTATTCCCC	336	66		This study
	Agaz-2	CCCAAGTTTGACCCTCGATA	GGAAGGTGGGCCTTAGGTAT	238	66	Neutral	Hoffman, 2009
	Ext_Agaz-2	CTGGGGAGTAAAGGAGGTAGA	TGTGATGACAGGCTTCCACC	293	62		This study
	Agaz-8	GGGGAGCCCTGATAGAAATC	AGATTGATGGCCTGGGAAC	238	58	Neutral	Hoffman, 2009
	Ext_Agaz-8	GATCTCAAATCCAAATGCTG	GACAACTAATAGTCTCTCTGAG	297	62		This study
	Agt10	AAGGGGCCCATATTCTTCC	CATTGCACAGTCATGTGTGG	213	66	CXCL 10	Hoffman and Nichols, 2011
	Ext_Agt10	GCTTCTACTGCTATCCTCCC	TGAAACACATCCAGCATATCT	288	62		This study
	Agi05	TCCTCTTCTTCCTCCTCTTCC	AGAAGTCCCATTGGTCCTGG	217	66	CD44	Litzke et al. 2019
	Ext_Agi05	AAGCCCAGCCCTGACTAATG	GGCAAGTCCTCGGTCTTTAG	299	62		This study
mtDNA control region	T-Thr	TTCCCCGGTCTTGTAAACC					
	T-Phe		ATTTTCAGTGTCTTGCTTT	~1300			Hoelzel et al. 1993
	SCR		CCTGAAGTAAGAACCAGATG	450	56		
molecular sexing	PinZFY	TGGTAAGGTGTTCAGGATGG	GGAACTTGGTTTGATCACTCAT	169	**80**	ZFX/ZFY	Robertson et al. 2018
	SRY-1	GCAAGGTGGCTCTAGAGAAT	TACCACTTATCCACGAGCAC	259	**84 (83.6 – 84.3)**		This study
	Ext_PinXFY		TGACAGGGTTGCTCTATCTC	302	65		This study
	SRY-2	GACGAGCAATCCTACCTCAA		378	62		This study

Table A.4. Accession numbers for sequences used in the mtDNA control region analysis and for the sequences used to develop primers for molecular sexing.

Old taxonomy	Species name	Accession	
Arctocephalus gazella		AF384376.1	mtDNA control region
Arctocephalus gazella		AF384377.1	
Arctocephalus gazella	Antarctic fur seal	AF384378.1	
Arctocephalus gazella		AF384379.1	
Arctocephalus gazella		AF384380.1	
Arctocephalus tropicalis		AF384381.1	
Arctocephalus tropicalis		AF384382.1	
Arctocephalus tropicalis	Subantarctic fur seal	AF384383.1	
Arctocephalus tropicalis		AF384384.1	
Arctocephalus tropicalis		AF384385.1	
Arctocephalus galapagoensis	Galapagos fur seal	AF384386.1	
Callorhinus ursinus		AF384387.1	
Callorhinus ursinus		AF384388.1	
Callorhinus ursinus	Northern fur seal	AF384389.1	
Callorhinus ursinus		AF384390.1	
Callorhinus ursinus		AF384391.1	
Arctocephalus pusillus doriferus		AF384392.1	
Arctocephalus pusillus doriferus	Australian Fur Seal	AF384393.1	
Arctocephalus pusillus doriferus		AF384394.1	
Arctocephalus pusillus doriferus		AF384395.1	
Arctocephalus townsendi	Guadalupe fur seal	AF384396.1	
Arctocephalus townsendi		AF384397.1	
Arctocephalus australis		AF384398.1	
Arctocephalus australis		AF384399.1	
Arctocephalus australis	South American fur seal	AF384400.1	
Arctocephalus australis		AF384401.1	
Arctocephalus australis		AF384402.1	
Arctocephalus philippii		**AF384403.1**	
Arctocephalus philippii		**AF384404.1**	
Arctocephalus philippii	**Juan Fernandez fur seal**	**AF384405.1**	
Arctocephalus philippii		**AF384406.1**	
Arctocephalus philippii		**AF384407.1**	
Otaria byronia	South American sea lion	AF384419.1	
Mirounga leonina	Southern elephant seal	AY424657.1	SRY
Mirounga angustirostris	Northern elephant seal	AY424656.1	
Hydrurga leptonyx	Leopard seal	AY424655.1	
Monachus schauinslandi	Hawaiian monk seal	AY424654.1	
Phoca vitulina	Harbor seal	AY424662.1	
Phoca caspica	Caspian seal	AY424661.1	
Phoca largha	Spotted seal	AY424664.1	
Phoca hispida	Ringed seal	AY424663.1	
Halichoerus grypus	Gray seal	AY424660.1	
Phoca groenlandica	Harp seal	AY424659.1	
Cystophora cristata	Hooded seal	AY424658.1	
Erignathus barbatus Phoca	Spotted seal	AY424665.1	
Arctocephalus gazella	Antarctic fur seal	AY424651.1	
Zalophus californianus	California sea lion	AY424650.1	
Arctocephalus gazella	Antarctic fur seal	UIRR01000042.1 (location: 1901910 – 1902817)	ZFX/ZFY
Zalophus californianus	California sea lion	DQ811094	

Table A.5. Bacterial phyla detected in Juan Fernandez fur seal faeces.

Family	Total Counts	Total counts rel. ab (%)	Mean rel.ab (%)	rel. ab SD	total ASV
Firmicutes	863365	41.85	40	24	296
Fusobacteria	582406	28.23	30	17	46
Bacteroidetes	455251	22.07	22	10	94
Proteobacteria	113805	5.52	6	4	74
Actinobacteria	30597	1.48	2	3	21
Verrucomicrobia	6653	0.32	0	2	3
Epsilonbacteraeota	6554	0.32	0	1	10
Unidentified	2204	0.11	0	0	2
Tenericutes	1005	0.05	0	0	8
Lentisphaerae	900	0.04	0	0	3
Spirochaetes	34	0.00	0	0	3

Table A.6. Summary of bacterial family detected in faeces of Juan Fernandez fur seal. Data is arranged in decreasing order based on Counts mean

Family	Total Counts	Rel. ab (%)	Counts mean	Coutns SD	Mean Rel.ab (%)	re. ab SD	total ASV
Fusobacteriaceae	582404	28.23	10785.26	6958.72	30	17	45
Bacteroidaceae	320047	15.52	5926.8	5319.43	15	10	28
Ruminococcaceae	310109	15.03	5742.76	5206.41	15	13	139
Lachnospiraceae	213725	10.36	3957.87	4195.97	9	8	61
Peptostreptococcaceae	193151	9.36	3576.87	6353.37	9	16	16
Rikenellaceae	65548	3.18	1213.85	1543.63	3	4	20
Clostridiaceae 1	60276	2.92	1116.22	2385.25	3	5	16
Burkholderiaceae	47544	2.3	880.44	849.59	2	2	8
unidentified_Gammaproteobacteri	27169	1.32	503.13	1116.46	1	2	9
Acidaminococcaceae	27237	1.32	504.39	734.56	1	2	3
Marinifilaceae	25673	1.24	475.43	1022.21	1	2	13
Prevotellaceae	24111	1.17	446.5	1666.42	1	4	4
Coriobacteriaceae	23956	1.16	443.63	688.99	1	2	1
Family XIII	22734	1.1	421	1294.81	1	3	11
Clostridiales vadinBB60 group	16935	0.82	313.61	586.95	1	2	2
Tannerellaceae	15153	0.73	280.61	647.81	1	2	8
Succinivibrionaceae	14801	0.72	274.09	931.79	1	2	7
Desulfovibrionaceae	12759	0.62	236.28	296.88	1	1	10
Erysipelotrichaceae	6926	0.34	128.26	161.05	0	1	7
Akkermansiaceae	6644	0.32	123.04	696.94	0	2	2
Eggerthellaceae	5951	0.29	110.2	241.78	0	1	4
Helicobacteraceae	5185	0.25	96.02	371.85	0	1	7
Streptococcaceae	4000	0.19	74.07	192.29	0	0	6
unidentified_Rhodospirillales	3691	0.18	68.35	153.19	0	1	4
Lactobacillaceae	3649	0.18	67.57	336.59	0	1	3
unidentified_Bacteroidales	3395	0.16	62.87	168.66	0	0	5
Enterobacteriaceae	3289	0.16	60.91	195.2	0	1	6
unidentified_Clostridiales	2650	0.13	49.07	186.76	0	1	10
unidentified_Bacteria	2204	0.11	40.81	129.68	0	0	2
Pasteurellaceae	2192	0.11	40.59	285.04	0	1	6
Campylobacteraceae	1369	0.07	25.35	92.27	0	0	3
Spongiibacteraceae	1064	0.05	19.7	79.73	0	0	1
Nitrosomonadaceae	888	0.04	16.44	53.79	0	0	1
Mycoplasmataceae	881	0.04	16.31	92.11	0	0	7
Eubacteriaceae	806	0.04	14.93	70.4	0	0	1
Victivallaceae	662	0.03	12.26	71.52	0	0	2
Flavobacteriaceae	568	0.03	10.52	38.49	0	0	4
Barnesiellaceae	639	0.03	11.83	41.46	0	0	2
Peptococcaceae	438	0.02	8.11	22.97	0	0	2
Enterococcaceae	325	0.02	6.02	26.47	0	0	4
Vibrionaceae	113	0.01	2.09	9.78	0	0	2
vadinBE97	238	0.01	4.41	28.16	0	0	1
unidentified_Mollicutes RF39	124	0.01	2.3	16.87	0	0	1
Shewanellaceae	108	0.01	2	14.7	0	0	1
Corynebacteriaceae	246	0.01	4.56	14.72	0	0	3
Coriobacteriales Incertae Sedis	200	0.01	3.7	16.27	0	0	1
Christensenellaceae	180	0.01	3.33	8.89	0	0	2
Actinomycetaceae	188	0.01	3.48	9.79	0	0	6
Veillonellaceae	96	0	1.78	10.53	0	0	2
unidentified_Verrucomicrobiae	9	0	0.17	0.86	0	0	1
unidentified_Firmicutes	8	0	0.15	1.09	0	0	1
unidentified_Bacteroidia	2	0	0.04	0.27	0	0	1
unidentified_Actinobacteria	13	0	0.24	1.18	0	0	1
Thioalkalispiraceae	2	0	0.04	0.27	0	0	1
Staphylococcaceae	35	0	0.65	2.84	0	0	1
SC-I-84	3	0	0.06	0.41	0	0	1

Family	Total Counts	Rel. ab (%)	Counts mean	Coutns SD	Mean Rel.ab (%)	re. ab SD	total ASV
Saprospiraceae	2	0	0.04	0.27	0	0	1
Rhodobacteraceae	8	0	0.15	0.79	0	0	2
Rhodanobacteraceae	9	0	0.17	1.22	0	0	1
Rhizobiales Incertae Sedis	4	0	0.07	0.54	0	0	1
Pseudomonadaceae	16	0	0.3	1.24	0	0	3
Porphyromonadaceae	2	0	0.04	0.27	0	0	1
OCS116 clade	2	0	0.04	0.27	0	0	1
Nocardioidaceae	5	0	0.09	0.68	0	0	1
Neisseriaceae	80	0	1.48	9.69	0	0	2
Muribaculaceae	2	0	0.04	0.27	0	0	1
Moraxellaceae	12	0	0.22	1.16	0	0	3
Micrococcaceae	32	0	0.59	2.26	0	0	1
Leptotrichiaceae	2	0	0.04	0.27	0	0	1
Halomonadaceae	36	0	0.67	3.62	0	0	2
Gracilibacteraceae	29	0	0.54	2.96	0	0	2
Family XI	6	0	0.11	0.57	0	0	2
Dietziaceae	4	0	0.07	0.38	0	0	2
Desulfobulbaceae	3	0	0.06	0.41	0	0	1
Crocinitomicaceae	6	0	0.11	0.82	0	0	1
Chitinophagaceae	101	0	1.87	8.26	0	0	4
Carnobacteriaceae	10	0	0.19	0.97	0	0	2
Cardiobacteriaceae	12	0	0.22	1.21	0	0	1
Brachyspiraceae	34	0	0.63	3.02	0	0	3
Bacillaceae	40	0	0.74	3.6	0	0	3

Table A.7. Amplicon sequence variants present in at least 27 of the samples (50 %).
Relative abundance was calculated from the unrarefied data.

ASV	Phylum	Family	Genus	Abundance (%)
57729b2b058d8d5253d3e56e4f6386ca	*Fusobacteria*	*Fusobacteriaceae*	*Fusobacterium*	14.93
f347c63fc5e4aeb97531e656e3765e2a	*Firmicutes*	*Peptostreptococcaceae*	*Peptoclostridium*	8.29
e8b1922518029c50c69add839142db03	*Fusobacteria*	*Fusobacteriaceae*	*Fusobacterium*	6.52
57f9edc6542ce6b78ff352942d6774c6	*Bacteroidetes*	*Bacteroidaceae*	*Bacteroides*	4.28
31984a302fdfe46b5e852fa473e682a4	*Bacteroidetes*	*Bacteroidaceae*	*Bacteroides*	4.26
b8d6a5a80d025861f2afccb79e0a1aaf	*Bacteroidetes*	*Bacteroidaceae*	*Bacteroides*	3.80
c0dc53aad260a1b951b7f99966251c7c	*Fusobacteria*	*Fusobacteriaceae*	*Fusobacterium*	3.73
1153942c5cc40d6ba5609222ded586fe	*Firmicutes*	*Lachnospiraceae*	*Coprococcus 3*	2.98
65dd9f625700a97a1cce9f5eefe4e6cb	*Firmicutes*	*Lachnospiraceae*	*Blautia*	2.18
e176cb3e4c2f33cee5d529c21ff5534e	*Firmicutes*	*Clostridiaceae 1*	*Clostridium sensu stricto 2*	1.95
435975b6d032d4b05233d8b94193b2ad	*Firmicutes*	*Lachnospiraceae*	*[Ruminococcus] gauvreauii group*	1.93
1a73c668a4bb92b74a18b79f9ae63460	*Firmicutes*	*Ruminococcaceae*	*Ruminococcaceae UCG-005*	1.75
5b87f47a447ef9a905807d2abed5b638	*Bacteroidetes*	*Rikenellaceae*	*Alistipes*	1.68
bf4112a100b11b4cbe9a25bdc591ea52	*Fusobacteria*	*Fusobacteriaceae*	*Fusobacterium*	1.38
03f74c0ea1f0654719b21d2701e9fa30	*Proteobacteria*	*Burkholderiaceae*	*Sutterella*	1.30
1188ef0238977f665e179642f287aead	*Firmicutes*	*Ruminococcaceae*	*Ruminococcaceae UCG-005*	1.29
25699f81befd34e0c9d81cfa84f4e751	*Firmicutes*	*Lachnospiraceae*	*unidentified_Lachnospiraceae*	1.27
8e10797dedc288dbc0be61fe4b5a5dfb	*Actinobacteria*	*Coriobacteriaceae*	*Collinsella*	1.16
2553bcb6afcdea16b173909555484369	*Firmicutes*	*Ruminococcaceae*	*[Eubacterium] coprostanoligenes group*	1.15
b15e41c7f20b8dcd4b0ed9f6c526885d	*Bacteroidetes*	*Prevotellaceae*	*Alloprevotella*	1.14
ca28c391514fb33d2d2df1c3c8e12317	*Firmicutes*	*Ruminococcaceae*	*Ruminococcaceae UCG-005*	1.12
76ded93fadbc4155db4e9dcba2012c81	*Firmicutes*	*Ruminococcaceae*	*Ruminococcaceae UCG-002*	1.07
c45b2a8ebeca2fca503c6312e8611416	*Bacteroidetes*	*Marinifilaceae*	*Odoribacter*	1.07
ce3476a906008973a3ab56de06817d56	*Proteobacteria*	*Burkholderiaceae*	*Sutterella*	0.87
975258836de3a001cb4d91cf6cf7de06	*Firmicutes*	*Acidaminococcaceae*	*Phascolarctobacterium*	0.72
1cde608d0a8b17d6fed116653581f050	*Proteobacteria*	*Succinivibrionaceae*	*Succinivibrio*	0.68
6c4c9e8ad2f56316cfffac9587c173ec	*Firmicutes*	*Ruminococcaceae*	*Ruminococcaceae UCG-005*	0.58
58514f20ebf4be2b13d619ba3bd2cf83	*Bacteroidetes*	*Bacteroidaceae*	*Bacteroides*	0.55
a0eee6d353d432299b53c9663cf05597	*Bacteroidetes*	*Bacteroidaceae*	*Bacteroides*	0.54
0ac8214c377877609cd0f88567086b2e	*Firmicutes*	*Lachnospiraceae*	*Tyzzerella*	0.44
420f3edebd00de18846a5941b55a6d5e	*Bacteroidetes*	*Rikenellaceae*	*Alistipes*	0.44
0e7fdaa233c333cb8363b63a41bbfc32	*Firmicutes*	*Ruminococcaceae*	*Ruminococcaceae NK4A214 group*	0.43

ASV	Phylum	Family	Genus	Abundance (%)
df408056297f20c5ce5cc68907e39cc8	*Firmicutes*	*Lachnospiraceae*	*Tyzzerella*	0.41
c00129ca877cb31776ad4e4e03a9091d	*Fusobacteria*	*Fusobacteriaceae*	*Fusobacterium*	0.41
93623ff4fe3615ce4aa4f0a9554fd4de	*Proteobacteria*	*Desulfovibrionaceae*	*unidentified_Desulfovibrionaceae*	0.35
a672e8b3efeb3a28e5beabe661606ad2	*Firmicutes*	*Ruminococcaceae*	*unidentified_Ruminococcaceae*	0.33
0c3d2038714f019f70fdc3b6f4b40419	*Firmicutes*	*Ruminococcaceae*	*Ruminiclostridium 9*	0.30
b6578d861d1c0e923087c8a5a81c8501	*Proteobacteria*	*unidentified_Gammaproteobacteria*	*unidentified_Gammaproteobacteria*	0.25
57f0c2ba2627cebfea197aa991777cb0	*Bacteroidetes*	*Tannerellaceae*	*Parabacteroides*	0.24
305caa259fb99e3e9aa1eb5dac615002	*Firmicutes*	*Erysipelotrichaceae*	*unidentified_Erysipelotrichaceae*	0.19
cbeb8d4b3d3f4b0bfa328178582220a5	*Firmicutes*	*Streptococcaceae*	*Streptococcus*	0.17
66c7c8504838071f63638f7e03975cf27	*Proteobacteria*	*unidentified_Gammaproteobacteria*	*unidentified_Gammaproteobacteria*	0.15
78990f6a6e53bd64b9371e316ad97362	*Firmicutes*	*Ruminococcaceae*	*Butyricicoccus*	0.14
e8f48023e5081f948df1291acd8d356a	*Firmicutes*	*Lachnospiraceae*	*unidentified_Lachnospiraceae*	0.09
1c4ff74a77a35261b972eb21737647e9	*Firmicutes*	*Ruminococcaceae*	*Ruminiclostridium 5*	0.07
99ea1bdfe0e8b83616c6178b8fdbf1e4	Firmicutes	*Ruminococcaceae*	*Harryflintia*	0.05

Table A.8. Table reporting the mean values of Chao-1, Shannon-Weiner and Simpson (D) indexes and their standard deviation for each location.
Tierras Blancas consistently show higher values than the other three locations. Simpson here is used as 1-D. Thus, the higher the number, the more diverse. Non-normalised data was used to build this table.

Location	Chao1	Shannon-Weiner	Simpson
Arenal	72.6 ± 23.0	2.9 ± 0.5	0.87 ± 0.08
Bahia El Padres	75.7 ± 27.4	2.9 ± 0.5	0.89 ± 0.05
Santa Clara	68.3 ± 30.0	2.8 ± 0.7	0.85 ± 0.16
Tierras Blancas	101.9 ± 40.1	3.4 ± 0.4	0.94 ± 0.02

Table A.9. The selected value of the Spearman rank correlation performed on the rarefied core data, including PC1 and 2 for each dissimilarity distance.
The table is reporting only the correlation that showed to be strong ($0.6 \leq |\rho| \leq 0.79$) and very strong, ($0.8 \leq |\rho| \leq 1$).

Correlation pair		ρ	Strength	p
Bacteroides	Bray-Curtis PC1	-0.67	Strong	< 0.001
Fusobacterium	Bray-Curtis PC1	-0.92	Very strong	< 0.001
Peptoclostridium	Bray-Curtis PC1	0.81	Very strong	< 0.001
Ruminiclostridium 9	Bray-Curtis PC1	0.63	Strong	< 0.001
Ruminococcaceae NK4A214 group	Bray-Curtis PC1	0.61	Strong	< 0.001
Odoribacter	Bray-Curtis PC2	0.62	Strong	< 0.001
Parabacteroides	Bray-Curtis PC2	0.71	Strong	< 0.001
Fusobacterium	*Peptoclostridium*	-0.63	Strong	< 0.001
Ruminiclostridium 9	*Ruminococcaceae UCG-005*	0.61	Strong	< 0.001
Fusobacterium	Weighted Unifrac PC1	-0.94	Very strong	< 0.001
Peptoclostridium	Weighted Unifrac PC1	0.75	Strong	< 0.001

Appendix B

Microsatellite primers used in JFFS DNA amplification

This document contains all the primers included in the PhD thesis "Potential of faeces for non-invasive assessment of Juan Fernandez fur seals and their environment. Due to the nature of the samples, the DNA amplification will be carried out as a nested PCR. All the internal primers are taken from the literature (developed from the close related *Arctocephalus gazella*) while the external primers were created using primer3 (http://primer3.ut.ee/).

B.1 Protocol used for designing external primers for each internal primer pair

After selecting possible microsatellites from the literature, use Primer3 to find suitable external primer.

- Primer length: 18-22 (opt 20) bp
- Melting Temperature: 52-58
- GC%: 40-60
- Repeats (max 4 di-nucleotide repeats)
- Runs (no more than 4bp)

Before selecting, check the following info in NEB Tm calculator https://tmcalculator.neb.com:

- Melting Temperature: 50-58 best results. *Above 65 -> tendency to secondary annealing
- Primer annealing temperature

B.2 Immune microsatellites

B.2.1 Agt family

Internal primers were developed by (Hoffman and Nichols, 2011)

Agt10 internal primers

Agt10_F AAGGGGCCCATATTCTTCC

Agt10_R CATTGCACAGTCATGTGTGG

Agt10_Rcomp CCCACACATGACTGTGCAATG

Expected fragment size: 213

Repetitive motif: $(AT)_{9\text{-}11}$

Description: Best Blast hit against Pinnipedia was Chemokine C-X-C motif ligand 10, (CXCL 10).

Agt10 internal primers

Agt10_F GCTTCTACTGCTATCCTCCC

Agt10_R GAAACACATCCAGCATATCT

Agt10_Rcomp AGATATGCTGGATGTGTTTCA

Expected fragment size: 288

>JF746973.1 Arctocephalus gazella clone Agt-10 microsatellite sequence

```
  1 GGAGACACTC AATTGCTTAG ACTTTCTGAG CCTACTGCAG AGGAACCTCC AGTCACAGCA
 61 TCATGAACCA AAGTGCTGTT CTTATTTTCT GCCTTATCTT TCTGACTCTG ACTGGAACTC
121 AAGGAATACC TCTCTCTAGA ACTACACGCT GTACCTGTAT CAAGATTAGT GATGGATCTG
181 TAAATCTAAA GTCCTTAGAA AAACTTGAAG TGATTCCTGC AAGTCAATCT TGTCCACGTG
241 TTGAGATCAT TGCCACACTG AAAAAGAATG GGGAGAAAAG ATGTCTGAAT CCAGAGTCTA
301 AGAAAATCAA GATTTTATTG AGAGCAATTA GCAAGGAAAG ATCTAAAACA TCTCCTTGAA
361 CAGAGAAGGA TAATCTCTGT AATACTGATA AAGATAGACC AGAAAGAGAC TACCTCTGCC
421 ATCATTTCCC TGCACACAAT ATATGCAAGC CATTATTGTC CCTGGATTGC AGTTCTCTTA
481 AAAGGTGACC AACCACTGTC ACCAAATTAG CTGCTACTAC TCCTTGAGGG AGGTGGACGG
541 TTCATCACCC TGTCCTGGCA CTATAAGCTA TGCTGAGGTA CTACATTTTC AGTGAATGTA
601 CCAGATCCTA GCCCTACTAC TGACACTTTC CTCACCTGTC CTATCTTCTG TTATTAAGGG
661 ATATTTCCAC CTCTGGACTT ATCAGAGTTC TCAGAACTTC ACAAAGTACC CATAATACAA
721 TCTCCTTTTT TAAGAAAGAA CTTTACTCCA CGAGCTTCTA CTGCTATCCT CCCAAGGGGC
```

```
 781 CCATATTCTT CCAGTGGTTT TATATATATA TATATATATG AAACTCCAAA TACATAGAAG
 841 AACCTAGAAA TACCTGAAAA TGTATGTGCA AATACTATTT TTAACGAAAA GCTACATAAA
 901 ATAGGATTCT TAGATATACA TATTTCTTGT TTTCAGTGTT TATGGAGTAA CTTCTGTAAC
 961 TAAATACCAC ACATGACTGT GCAATGAAAA ATTTTAAAAT CTAGATATAT ACTCTGCATG
1021 TTATGTAAGA CAGATATGCT GGATGTGTTT CAAAATAAAA ATACTGTGCT CTCTTGGAGA
1081 TGTTAAGATA GATTATATAA CTGTTAGCAG ATCAAAAAAG TAATAAAA
```

B.2.2 Agi family

Internal primers developed by (Litzke *et al.*, 2019) using data generated by (Humble *et al.*, 2016). The original fasta file containing the contig is publicly available and can be downloaded from https://raw.githubusercontent.com/elhumble/snp_filtering/master/data/raw/joined_transcriptome.fasta

Agi05 internal primers

Agi05_F TCCTCTTCTTCCTCCTCTTCC

Agi05_R AGAAGTCCCATTGGTCCTGG

Agi05_Rcomp CCAGGACCAATGGGACTTCT

Expected fragment size: 217

Repetitive motif: $(TC)_{13}$

Description: The best Blast hit against Pinnipedia was CD44.

Agi05 external primers

ExtAgi05_F AAGCCCAGCCCTGACTAATG

ExtAgi05_R GGCAAGTCCTCGGTCTTTAG

ExtAgi05_Rcomp CTAAAGACCGAGGACTTGCC

Expected fragment size: 299

>AgU000254_v1.1

```
  1 AACGCtAGTA GTtACTtGGt tAgTTTTTTT ttttTtttgT TTCCTTTAGA CTTTTAATAG
 61 TAGCCTTTTG GTGTTTTCCA ATAAGTGCTT TCAACTGGGC AATATACATA TCATGCTTTC
121 CTCATTACTA CTGGTCCATC AATAAATATA CAAAAATCAG AGGAAGGGTG TGCTGTGACA
181 AATCAAAGTA ACAATGGTTA CTGTACATGC ATGAACAATG AGCTATTCTT TGAGAACTCA
241 ACATAAGTGA TTTATAGAAA GGCATACCAA TACATAAAGG TAAATGGACC ACGTTATTCT
301 TATTCTTAAG TACACTTGAC ATTTCTAGAG AAGTCAGAGA TTCCTTTCTC CCACTCCTCT
```

```
 361 GCTGTACCAG TATTGAATGA ACATTGGCTA TGAAATCAAT TCTTGGCCTC ATGTTCTCAA
 421 TATAAAACTT TGAGAGGTTC TGCTTGAGAA TTCCATTCTG CCAGGCTAGA ATCTACAAAA
 481 TATTGATATG AAGGAGTCTC TTGTCCCATT TTATTTTCAC CCAGTGCTAA GTTAAAGACA
 541 GCCCTCAGGA GACACAAAAG ATTTAGGAAA TTTTAGGAAA AATTGCATCT TAAGCTCTCT
 601 ATGCCATTGT TTCAAATGGG AGAGAATACA TCCTATTCCA TGAACTCTGG GCAGTATTGG
 661 AGGGGTATTG CATCCAAGGG CTGGAAGCTC TGGACCTCTC CCATtCTTAT TTGAGAACAC
 721 AGAGGCATAT CCAAGTGAGG GACTACAAAT CCTTCCTAAT AAAAGGAAAG ATCGTACTAA
 781 ATTCTAAAAA TAGAAAACAA TCACCTCTCT CAGCCTATTT ATTGGCAATC TCTAGAAAAG
 841 GAAAACCATC CCTAGCTGTG AGACTTgTAA GACTATCCAT GACCCTGAAG AAAATCTCGG
 901 GACAGATAGA ACCTGGCATA GCCCTCCTGA ACCGGATCCT GGAGTGGCTA AAGAGCTGGA
 961 ATGAGCaTGT CTGCAGCAGA CCAGGAAAGC TCTTGGAAAG CAGGGAGCAC TTGGACATGG
1021 TGACTGTGAA ACGGAGGGGC TAGATGTCCA TGTCTTTTCT TTCCACTTGA AAATAAAaGG
1081 AAAaCCAGAC AGCTGGGTAT GGAACTGGCC AATGAAGTTC ACAGAAGAGA AGGAAGTTTT
1141 CAAAGATCAT GATTGGGAGG AGGTTATGAA GGTGGAGAGG GGAGAACAGA AGTTGGAAAT
1201 GAGTGCAGAA AGTGGGCCTC CGTTTtAAAA aTagACCCTT CATAATCGCC CCAGGTATAA
1261 GCCCAGCCCT GACTAATGTG GgACCTATTT TCTtCTCCTC CTCCTTCTTT TCCTCTTCTT
1321 CCTCCTCTTC CTTCTCTtcT CTCTCTCTCT CTCTCTCTCT CtcTGGTCTT AGCAACATCA
1381 CCAGTTAAAC CtATCAGTTT AGGCAACAGA TGACTTTTtA TGGGGGCATA AGGAGAGCAG
1441 TCTCATTTAA TAGGAGTCCT TGGCTTATAG ATTCAATTTC TTCTATTACA AAGCATCTTT
1501 GGAAGTTCCA GGACCAATGG GACTTCTGag CCAATGACTA AAGACCGAGG ACTTGCCAAG
1561 CAAGGGATAA AGCCATTATT CTCTTATTTG GTCATAAACC CGCCTATCTG GTGCTAAGAG
1621 TCCAATGGTG GGAAGCCCTG TAACCTTCTG AGGGACTTCC CAGGCACTTA ACGTACCCTG
1681 AGTAGCAACT ATGTCGGGGA CAGCTGGCTT GCTAAAGTAT AAACAAATAC AGTTATAAAA
1741 ATATTCAGAC AGGATCTACA CAGCTGCCTC CCCCTTCCCC AACCTGTGGC AATCAGTTGG
1801 CTACTGTCTA TGCTCTCTTT CATTTCCATT GGCTTCTCTT TTCATTACAA GAGGTCAAAA
1861 GGTACTTCCA TATTGTTTTt GGACATCACT TCACAGGGTC AGTGTCTATG AAATGGGGTG
1921 CCACAAATGG TAGGTCATGC AATGTAATGT GACTTTTCAC AGCCTTTCAG GGCAGAGAAC
1981 TGGTGGCCAT TCTCACATAA AAGCAGCTGA GGTTCCTCCC CTCACAACTC TGGTGCTAGA
2041 ACAGAAGAAA GCTCTGGAGC CCcAGAGATA ACACAGCACT ATCTTCCTTG ATTCTCTTTG
2101 AGATTTGGTT GAGTGAGCCT GAATGGGGGG CTTAGTGGCT CTGAGATGAA GACTTTCAAG
2161 GAGGGCTTtC TCACATCTTC CTGATGGGTT TTtGGGAACC AGACCTTGCT GGTGAAAAAa
2221 CTGAAGGATC ACAGCACAAA GGGTTAAAGT tGGAAGAAGT ACAGATGTTT ATTATGAATC
2281 AGCTTAAACC CTTTATGCCT CTGACAAAGT aCCAAAaAAA AAAAaTaCAT ACTGATCAAA
2341 GGACCTCCAG GGTTTAATAT TTCAAAAACA CAGATAAATA GCTTCCTACA GATAAATAGC
2401 TTCACCTTTT GGGTATTTCC CAGAAGCATG TGAAAAATTT CCAGGGGCTC TGTAGAAGGG
2461 GATGTGTAAC CAGTACATTC TGATCCCCTA CCCCACTGgA AAGAAATGCC CAGGGACTAG
```

```
2521 GCTGCAGGAG GAGGGGAGGG ATGTGGACCC ATTCAAATAT CTTCCCAGGG ACCAATCCCT
2581 CTTAGCCCTG GGGGATGTCC TTAGCTGGCT GCTGTTGCCT GGGAGGAGGA GTGGAGGGAA
2641 GGTCGGGGAT CAGGAAGATG TATTTGTGGC TTTTGTTAGA AGCCATCCAG AGCATGCCCT
2701 CAGGGGTGGA GGCTTCTAGG TGGGAACTGG AGCAGTCAAA TTCTTGCAGG TTATCATTTG
2761 GGCCCTCATT TCAGAAAGCA ATGCTGTTTT TATAAACTGG ACCTGACTTT AAAAATAAAA
2821 CGCAAAAACA aaaAGGAGTA GAAAATTTTA TACTAAAAaA aGATCCCCAA TGAAACAATC
2881 AGTAGCACAT TGCATCTGTG AAGTGTCCCA GCTCCTTATC ATGGTTATGT CTCCAACTGC
2941 CGTTTCTCTC CAAGGTCACC ACGGCGTGGG TGTTACACCC CAATCTTCAT GTCCACATTC
3001 TGCAGGTTCC GCGTCTCGTC GGCTGTCATA TACTGATCTG GAGTCTCTGA GGGCTCCTTG
3061 TTCACCAAGT GCACCATCTC CTGAGACTTG CTGGCCTCTC CGTTGATTCC GCTTGGCTTC
3121 CGGTCCCCcA CTGCTCCGTT GCCATTGTTG ATCACCAGCT TCTTCTTCTG CCCACACCTT
3181 CTTCGACTAT TGACAGCAAT GCAAACGGCA AGAATCAAAG CCAAGGCCAG GAGGGAGGCC
3241 AAGATGATGA GCCATTCTGG AATTTGAGGT TTCCGCATAG GAAATTCCGA GGTCGTGTTT
3301 GCCCCACCTT CTTGACTCCC ACTTGAGTGG CCAGCTGATT CTGATCCATG AGTGGTATGG
3361 GACGTTCCAC TGGGGTGAAC TGATGAGTCT TGGTC
```

B.3 Neutral microsatellites (Agaz family)

nternal primers developed by Hoffman (2009).

B.3.1 Agaz1

Agaz1 internal primers

Agaz1_F ACTCATGCCCTGCTTGAAAT

Agaz1_R CAGGAGACTTAGGCCAGCA

Agaz1_Rcomp TGCTGGCCTAAGTCTCCTG

Expected fragment size: 245

Repetitive motif: $(CT)_2GT(CT)_3GT(CT)_3GT(CT)_3GT(CT)_5GT(CT)_{12}TA(AC)_{15}$

Agaz1 external primers

ExtAgaz1_F GTAGGCAATTTTAGGGCAAG

ExtAgaz1_R GATCTAGGTCTGTATTCCCC

ExtAgaz1_Rcomp GGGGAATACAGACCTAGATC

Expected fragment size: 336

>HM142903.1 Arctocephalus gazella microsatellite Agaz-1 sequence

```
  1 GATCAAGCAT GGACCTGTGG CTAAAGAAAG ACCCCAGGCA GAGGGACACA AGTGTGAATA
 61 CAAAGAGGTG TTTGTAGGCA ATTTTAGGGC AAGCACTGAT AAGTGGGACA CTAACAATGT
121 CTGCTCCAGT CCACCCCTTG CGCCACCCAC CTCCACTCAT GCCCTGCTTG AAATTAATCT
181 GTTTTGTTAG TGATTCTCCA AAGTGCTAGC AAGTCTCTGT CTCTCTGTCT CTCTGTCTCT
241 CTGTCTCTCT CTCTGTCTCT CTCTCTCTCT CTCTCTCTCT TAACACACAC ACACACACAC
301 ACACACACAC ACAATTTCTG AAGCAAAATT ACCTCTTCTA CTCCCCATTT TAGAATCACC
361 TCACTTCCAG CTAGCGCTTG TGCTGGCCTA AGTCTCCTGC TCGGTGGGGG AATACAGACC
421 TAGATC
```

B.3.2 Agaz2

Agaz2 internal primers

Agaz2_F CCCAAGTTTGACCCTCGATA

Agaz2_R GGAAGGTGGGCCTTAGGTAT

Agaz2_Rcomp ATACCTAAGGCCCACCTTCC

Expected fragment size: 238

Repetitive motif: $(AC)_{24}$

Agaz2 external primers

ExtAgaz2_F CTGGGGAGTAAAGGAGGTAGA

ExtAgaz2_R TGTGATGACAGGCTTCCACC

ExtAgaz2_Rcomp GGTGGAAGCCTGTCATCACA

Expected fragment size: 293

>HM142904.1 Arctocephalus gazella microsatellite Agaz-2 sequence

```
  1 GATCCCTGTT GTGCTAATTG TCTTCACATT CATCTAATCC ACACCACCGC CTGCAAAGTC
 61 AGTATTTTTA TATCCACGTA CACCAAAAGA AAACAGAAGG TAGAACAGTA AAATAACCTG
121 TGCAAAGTCT CCTGGGGAGT AAAGGAGGTA GATTTAAGCC CAAGTTTGAC CCTCGATACC
181 ATGCTTTTTA TCTCTATGCT TAGAGACCCT TGTGTTCAAA TTCACTCTTG TTACGTATTT
241 TTAAAGTTCC TATTGTGAAA TGTCATCATG GTTGAACGTA ACTAATGCCT CATGAAACAC
301 ACACACACAC ACACACACAC ACACACACAC ACACACACAC ACACTGTACG TAAAGGTACA
361 GACAGATATA AAACAGATAC CTAAGGCCCA CCTTCCCAGG AAGAGGTGGA AGCCTGTCAT
421 CACACTACAC GTGTCCCTGG TGTCTTTTTT ATGATGACAC TCTGTCTACC TGGAG
```

B.3.3 Agaz8

Agaz8 internal primers

Agaz8_F GGGGAGCCCTGATAGAAATC

Agaz8_R AGATTGATGGCCTGGGAAC

Agaz8_Rcomp GTTCCCAGGCCATCAATCT

Expected fragment size: 238

Repetitive motif: $(AC)_{22}$

Agaz8 external primers

ExtAgaz8_F GATCTCAAATCCAAATGCTG

ExtAgaz8_R GACAACTAATAGTCTCTCTGAG

ExtAgaz8_Rcomp CTCAGAGAGACTATTAGTTGTC

Expected fragment size: 297

>HM142910.1 Arctocephalus gazella microsatellite Agaz-8 sequence

```
  1 GATCTCAAAT CCAAATGCTG TACAAGCAAG TGCCTGCATA AGGGGGATAA AAGGGTGGTA
 61 TAGGGGAGCC CTGATAGAAA TCAGGAAACA TATAAATGCC CTTTTAAGAC TTTTATTCCA
121 AATCGTTTAG ACATCACACA CACACACACA CACACACACA CACACACACA CACACACACG
181 CAAGCTCACT TCAGTTATTC AAATTTGTTC CCAGGCCATC AATCTGTAGT GCTTGATATT
241 TCAACATAAT TATTTTAAAA CATTCTCAGA GAGACTATTA GTTGTCCTAA TCACTCA
```