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., 2019). 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, 2011; Hazen et al., 2019). However, to effectively use wild populations as sentinels, it is first necessary to establish a baseline.

1 | ORIGINAL ARTICLE

Patterns of the fecal microbiota in the Juan Fernández fur seal (Arctocephalus philippii)

Constanza Toro-Valdivieso1 | Frederick Toro2,3,4,5 | SamuelStubbs6 | Eduardo Castro-Nallar7 | Barbara Blacklaws1

1Department of Veterinary Medicine, University of Cambridge, Cambridge, UK
2Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, Chile
3Escuela de Medicina Veterinaria, Facultad de Recursos Naturales y Medicina Veterinaria, Universidad Santo Tomás, Viña del Mar, Chile
4ONG Panthalassa, Red de Estudios de Vertebrados Marinos en Chile, Santiago, Chile
5Ph.D. Program in Conservation Medicine, Facultad de Ecología y Recursos Naturales, Universidad Andrés Bello, Santiago, Chile
6Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
7Center for Bioinformatics and Integrative Biology, Universidad Andres Bello, Santiago, Chile

Correspondence
Constanza Toro-Valdivieso, Department of Veterinary Medicine, University of Cambridge, Cambridge, UK.
Email: cvt24@cam.ac.uk

Abstract

As apex predators, pinnipeds are considered to be useful bioindicators of marine and coastal environments. Endemic to a small archipelago in the South Pacific, the Juan Fernandez fur seal (JFFS) is one of the less-studied members of the pinniped family Otariidae. This study aimed to characterize the fecal microbiome of the JFFS for the first time, to establish a baseline for future studies of host-microbial-environment interactions and monitoring programs. During two consecutive reproductive seasons, 57 fecal samples were collected from seven different JFFS colonies within the Juan Fernandez Archipelago, Chile. Bacterial composition and abundance were characterized by sequencing the V4 region of the 16S rRNA gene. The overall microbiome composition was dominated by five phyla: Firmicutes (40% ±24), Fusobacteria (30% ±17), Bacteroidetes (22% ±10), Proteobacteria (6% ±4), and Actinobacteria (2% ±3). Alpha diversity was higher in Tierras Blancas. However, location was not found to be a dominant driver of microbial composition. Interestingly, the strongest signal in the data was a negative association between the genera Peptoclostridium and Fusobacterium, which explained 29.7% of the total microbial composition variability between samples. The genus Peptoclostridium has not been reported in other pinniped studies, and its role here is unclear, with interpretation challenging due to a lack of information regarding microbiome functionality in marine mammals. As a first insight into the JFFS fecal microbiome, these results contribute towards our understanding of the natural microbial diversity and composition in free-ranging pinnipeds.

KEYWORDS
Arctophoca philippii, marine mammals, microbiome, pinnipeds, scatology
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 technology (Redford et al., 2012; Trevelline et al., 2019). The gastrointestinal tract, especially the colon, is recognized as one of the largest microbial reservoirs (O’Hara & Shanahan, 2006). This microbial community fulfills 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, 2008b; Nicholson et al., 2012). Similar factors shaping the gut microbiome in terrestrial mammals influence that of marine mammals (Nelson et al., 2013; Stoffel et al., 2016; Pompa et al., 2011). These fur seals are the only native mammals in the archipelago and, like other pinnipeds, occupy upper trophic levels in the marine food web (Ochoa Acuña, 2012). The fecal microbiome has been characterized for eight pinniped species inhabiting the southern hemisphere, including three out of the eleven members forming the subfamily Arctocephalinae (fur seals): Arctocephalus pusillus doriferus (Smith et al., 2013), Arctocephalus actualis, and Arctophoca tropicalis (Medeiros et al., 2016). Also part of the Arctocephalinae subfamily is the Juan Fernandez fur seal (Arctophoca philippi philippi) (JFFS) which is 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 (Figure 1).

The archipelago is a hotspot for biodiversity with a high number of endemic species, including the JFFS (Aguayo et al., 1971; Friedlander et al., 2016; Pompa et al., 2011). These fur seals are the only native mammals in the archipelago and, like other pinnipeds, occupy upper trophic levels in the marine food web (Ochoa Acuña & Francis, 1995: Trites, 2019). Their feeding behavior, 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 characterize the JFFS fecal 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.
Samples were thawed on ice and centrifuged at 10,000 g for 15 min to pellet the sample out of RNAlater®. Genomic DNA was extracted from each pelleted sample (approximately 180 micrograms) using the MO BIO PowerSoil DNA Isolation kit (QIAGEN) according to the manufacturer’s instructions. Isolated DNA was quantified on a Qubit fluorometer (Invitrogen).

The bacterial 16S rRNA gene was polymerase chain reaction (PCR) amplified targeting a 250 bp region covering the V4 variable region. PCR amplification, barcode tagging, and library preparation were 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 the v2 chemistry producing 2 × 250 bp paired-end reads in the 2017 samples while the 2018 sequences were 2× 150 bp paired-end reads.

2.3 | Sequence data analysis and taxonomic classification

Raw sequence quality was manually assessed with FastQC v. 0.11.5 (Simon 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). Demultiplexed raw sequences were imported into QIIME2-2019.10 (Bolyen et al., 2019) where quality control, de-replication, read truncation, and paired read merging were 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 modeling and correcting Illumina sequencing errors. This step was carried out separately according to the year of collection. However, to normalize between datasets, the 250 bp reads produced from 2017 samples were truncated 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, representative sequences generated from both years were aligned in Geneious Prime® 2020.0.5 (https://www.geneious.com) by Multiple Alignment using the Fast Fourier Transform (MAFFT) plug-in with default settings and then assessed by eye (Katoh & Standley, 2013).

Next, a mid-point rooted, approximately-maximum-likelihood phylogenetic tree for diversity analysis was generated using the qiime2 phylogeny plug-in which uses 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).

2.4 | Data processing and statistical analysis

Statistical analysis was performed in duplicate, once using all available data and again with data corresponding to the core microbiome
only. The core microbiome was 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). To prepare the data by identifying unassigned ASVs and removing contaminants and samples with an insufficient depth of sampling prior to analysis, multiple filtering steps were applied to the data using the phylseq package (McMurdie & Holmes, 2013). (1) Unassigned ASVs at the Kingdom level were manually inspected 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 (Figure A1), three samples were identified as having an insufficient depth of sampling and 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). (3) Possible contamination signals were also removed by running a correlation analysis and comparing clusters with a list of previously identified reagent contaminants (Salter et al., 2014). (4) Finally, the data were rarefied using the same threshold used for filtering samples (Table A1) (McKnight et al., 2019).

The overall microbiota composition was characterized by summing the non-normalized read counts and obtaining the relative abundance at different taxonomic levels.

## 2.4.1 Alpha diversity

Estimates of within-sample diversity (alpha diversity) were calculated using the phylseq package. 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). The latter two approaches consider richness and 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.

Non-rarefied data were used to explore the alpha diversity. To compare locations, a one-way analysis of variance test (ANOVA) or a non-parametric Kruskal-Wallis test was performed for each estimate. ANOVA assumptions were tested by visualization of the data and statistical testing. A Shapiro-Wilk test was used to confirm normality and Levene’s test for heteroscedasticity. When exploring Shannon-Weiner and Simpson indices, sample 18JFFS23 (SC) was identified as an outlier (standard residual >3) and was removed for these indices only. Finally, data visualization suggested samples collected from TB differed from the other locations; thus, a post-hoc analysis was performed 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).

## 2.4.2 Beta diversity

To investigate variation between samples (beta diversity) two different distances were calculated using the rarefied full as well as the core datasets. Bray-Curtis dissimilarity distance was used to look at the differences between samples based on the ASVs abundances. Weighted UniFrac distance was used to explore the phylogenetic divergence between ASVs by also taking into account the abundance of these (with an emphasis on dominant ASVs). Respective distance matrices were visualized using principal coordinate analysis plots (PCoA).

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

Spearman’s rank correlation coefficient (\(\rho\)) was used to explore any possible associations between the different taxa and also between the first two components of the Bray-Curtis ordination analysis. Correlations were visualized in a correlation matrix plot, and only those significantly and strongly correlated (Rho (\(\rho\)) ≥ 0.6) were explored further. For this method, only the core microbiome dataset was used at the genus level.

## 3 RESULTS

Following the removal of low-quality sequences and merging the 2017 and 2018 datasets, a total of 2,074,038 paired reads, grouped into 595 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 (Table A1). Three samples were removed from the analysis due to rarefaction analysis indicating the insufficient depth of sequencing. The rarefied dataset ended up with 518 ASVs and a total of 754,974 reads.

### 3.1 Composition of the Juan Fernandez fur seal fecal microbiome

A total of 10 bacterial phyla were detected in the feces 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) (Table A2). Eighty-two bacterial families could be assigned, of which
14 had a relative abundance ≥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 (Figure 3a,b) (Table A3). Forty-six ASVs were present in at least 50% of the samples (Table A4). 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) (Table 1).

3.2 | Alpha diversity

Three alpha diversity indices (Chao1, Shannon-Weiner, and Simpson) were used to compare within-sample diversity between locations (Table A5). 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 Chao1 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 Tierras Blancas (Figure 4, Figure A2).

3.3 | Beta diversity

Based on weighted Unifrac dissimilarity distance, 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 the 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) was separated from the main cluster (cluster 1) along PC1 (Figure 5, Figure A3).

Based on the relative average abundance of the dominant phyla, evident differences in the overall microbial composition were visualized between the two clusters (Figure 6). 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, R² = 16.3%) and core datasets (∑F 13.6, Pr (>|F|) = 0.001, R² = 20.88%). SIMPER analysis identified five genera that together contributed 71% to the observed compositional difference between the clusters. As expected, both Fusobacterium and Peptostreptococcus 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 Table 2.

3.4 | 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 Ruminococaceae NK4A214 were also found to be influential for PC1 in Bray-Curtis analysis (Figure 7, Table A6). PCoA analyses showed strong negative correlations between PC1 and Bacteroides (Bray-Curtis, p = −0.67, p ≤ 0.001) and between PC1 and Fusobacterium (Bray-Curtis, p = −0.92, p ≤ 0.001 and weighted Unifrac, p = −0.94, p ≤ 0.001). Peptoclostridium, on the other hand, was positively correlated with PC1 (Bray-Curtis, p = 0.81, p ≤ 0.001, and weighted Unifrac, p = −0.75, p ≤ 0.001).

4 | 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 fecal microbiome in free-ranging pinnipeds and the first of JFFS. Our approach focused on characterizing the core members of the JFFS fecal 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.

Consistent with previous reports in other pinniped species, five phyla dominated the JFFS fecal microbiome: Firmicutes, Fusobacteria, Bacteroidetes, Proteobacteria, and Actinobacteria (Bik et al., 2016; Kim et al., 2020; Nelson, Rogers, & Brown, 2013; Numberger et al., 2016; Pacheco-Sandoval et al., 2019; Stoffel et al., 2020). When comparing our results to other southern pinnipeds, different microbial patterns were found in feces from other fur seal species (Medeiros et al., 2016; Smith et al., 2013). The fecal microbiome described for both the South American (Arctophoca australis australis) and the sub-Antarctic fur seals (Arctophoca tropicalis) is almost entirely dominated by Firmicutes (88.56% and 85.02%). Fusobacteria, on the other hand, represents less than 1% of the bacterial community for both species (Medeiros et al., 2016). The study involving these two species collected samples from dead juvenile individuals. Thus, it is expected to find altered microbiomes. Smith et al. (2013) characterized the fecal microbiome of Australian fur seal (Arctocephalus pusillus doriferus) pups and adult females. The adult samples showed similar proportions of Firmicutes, Bacteroidetes, and Actinobacteria as those observed here for JFFS. Fusobacteria was not detected in any
FIGURE 3 Composition of the Juan Fernandez fur seal fecal 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 1  Amplicon sequence variants present in at least 90% of the samples. Only three were present in all the samples. Unrarefied data were used to build this table. Abundance (abun) was calculated based on the total ASVs counts.

| ASV               | Phylum      | Family               | Genus               | Abun (%) |
|-------------------|-------------|----------------------|---------------------|----------|
| Present in all samples |             |                      |                     |          |
| 57729b2b058d8d5253d3e56e4f6386ca | Fusobacteria | Fusobacteriaceae | Fusobacterium | 14.93    |
| e8b1922518029c50c69add839142db03 | Fusobacteria | Fusobacteriaceae | Fusobacterium | 6.52     |
| c0dc53aad260a1b951b7f9996251c7c | Fusobacteria | Fusobacteriaceae | Fusobacterium | 3.73     |
| Present in at least 90% of the samples | F. |                      |                     |          |
| f347c63fc5e4aeab97531e656e3765e2a | Firmicutes | Peptostreptococcaceae | Peptoclostridium | 8.29     |
| 57f9edc6542ce6b78f352942d6774c6 | Bacteroidetes | Bacteroidaceae | Bacteroides | 4.28     |
| 31984a302dfe46b5e852fa473e682a4 | Bacteroidetes | Bacteroidaceae | Bacteroides | 4.26     |
| 1153942c5cc40d6ba5609222ded586fe | Firmicutes | Lachnospiraceae | Coprococcus 3 | 2.98     |
| 65dd9f625700a97a1ce9f5ee4e6cb | Firmicutes | Lachnospiraceae | Blautia | 2.18     |
| 435975b6d032d4b05233d9b94193b2ad | Firmicutes | Lachnospiraceae | [Ruminococcus] gauvreauii group | 1.93     |
| 0374c0e1f0654719b21d2701e9fa30 | Proteobacteria | Burkholderiaceae | Sutterella | 1.30     |
| 8e10797dedc288dcb0be61fe4b5a5dfb | Actinobacteria | Coriobacteriaceae | Collinsella | 1.16     |

FIGURE 4  Comparison of three different alpha diversity indices between the 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 were used to calculate the diversity estimates.

FIGURE 5  PCoA using Bray-Curtis dissimilarity distance matrix using the filtered rarefied core dataset. Samples clustered in two groups. (circles = cluster 1, triangles = cluster 2). Location is not driving the clustering.
of the adults. However, the authors only relied on fluorescent in situ hybridization (FISH) to detect these bacteria for this age group.

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 Bacteroidetes (all three in the case of JFFS and harbor seals) have been found to consistently dominate the overall microbial composition of pinnipeds, followed by Proteobacteria and Actinobacteria (Nelson, Rogers, & Brown, 2013; Pacheco-Sandoval et al., 2019). The latter two are usually at the 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 ours, 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 humans 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 (Hildebrandt et al., 2009; Pu et al., 2016; Turnbaugh et al., 2006). We 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, 2008b). 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 (Hildebrandt et al., 2009; Pu et al., 2016; Turnbaugh et al., 2006). 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 & Rogers, 2019).

The phylum Fusobacteria consists of facultative or strict anaerobes that produce various organic acids from amino acids or carbohydrates fermentation (Olsen, 2014). This phylum is usually found at the high relative abundance in the gut microbiomes of strict carnivores adapted to diets rich in proteins, purines, and polyunsaturated fatty acids (Guo et al., 2020; Zhu et al., 2018). 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-centered 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 (Garrett & Onderdonk, 2014; Huh & Roh, 2020; Potrykus et al., 2008). 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 (Guo et al., 2020; Zhu et al., 2018). 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 Bacteroidetes and Fusobacteria. This phenomenon

![Figure 6](image-url)  
**Figure 6** Relative average abundance of the dominant phyla according to the clusters identified with Bray Curtis dissimilarity. Showing only phyla with an average relative abundance ≥1%. The differences in microbial patterns can be identified from high taxonomic levels.

| Genus               | Mean cluster 1 | Mean cluster 2 | Mean 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    |

**Table 2** SIMPER analysis comparing the fecal microbiota composition of Juan Fernandez fur seal at the genus level. The table showing up to a cumulative contribution of 70%. Cluster averages were calculated based on total counts. Kruskal-Wallis results are only shown when reaching a significant difference.
suggests differences in nutritional needs and will be discussed later in the text.

4.1 | Within sample diversity

Initially, we hypothesized 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 (Figure 1). We found instead that BP did not differ from other less-visited locations such as EA and SC. Therefore, this finding is different from a previous report showing an association between exposure to anthropogenic stressors and reduced alpha diversity in harbor seals (Pacheco-Sandoval et al., 2019). The colony at TB was the only location with higher alpha diversity, indicating that samples collected from TB had a richer and more evenly distributed microbial composition than other samples. 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 (Li et al., 2016; Partrick et al., 2018). Lower diversity of the skin microbiome in denser populations was also observed in Arctocephalus gazella, a closely related species (Grosser et al., 2019).

The population density of JFFS and its effects on the microbiome have 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 feces (Wasser et al., 2000).

Despite the trend showing how TB differed from the other locations, only one of the three alpha diversity estimates (Simpson) showed TB to be statistically significantly different from the other locations. The other two diversity estimates (Chao1 richness and Shannon-Weiner) did not reach our significance cut-off. Both these estimates are affected by the detection of rare taxa, and larger libraries and sample sizes are more likely to input rare taxa into the data set. ANOVA was also used to compare locations with these diversity estimates. ANOVA is sensitive to differences in sample size, and therefore small group sizes may have affected statistical power.

### 4.2 Variation between samples

The Bray-Curtis dissimilarity PCoA revealed two distinct clusters. Seventy-five percent 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% (Delport et al., 2016; Nelson, Rogers, & Brown, 2013; 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 characterized 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 genus *Peptoclostridium* (previously known as *Clostridium*). 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 *Clostridoides 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 feces originated from sample contamination by surrounding environmental bacteria. Such high prevalence may be a sign of a deeper relationship between these uncharacterized 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 (Sernova et al., 2012). Juan Fernandez fur seal milk composition contains a higher proportion of lipids in comparison to many pinnipeds (~ 41%) (Ochoa-Acuña et al., 1999). Thus, if the fecal 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 analyzed from Australian fur seals were dominated by the class *Clostridia* in six and nine months 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 (preweaning 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 feces are still distinguishable from older individuals in color 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 a pre-weaning diet is not the explanation for the abundance of *Peptoclostridium*.

Differences between sexes may also be an explanation of the difference in samples. Otarids and Phocids such as northern and southern elephant seals exhibit an important degree of sexual size dimorphism (Ralls & Mesnick, 2009). Sex differences in foraging behavior and prey selection have also been reported (Andersen et al., 2013; Lewis et al., 2006; Ochoa Acuña & Francis, 1995). Based on the differences in diets, it is not surprising to find studies in gut microbial composition also showing sex-based differences. Samples collected from adult Southern elephant seals evidenced significant differences between adult males and females (Kim et al., 2020; Nelson, Rogers, & Brown, 2013). 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 (Kim et al., 2020; Nelson, Rogers, & Brown, 2013). The proportional changes are very similar to the one observed between clusters 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 sex as reported in northern elephant seal pups under a naturally controlled diet (Stoffel et al., 2020). Sexual dimorphism is a common mating strategy in otariids. Thus, otariids such as JFFS may show similar differences as the ones observed in elephant seals. This hypothesis could be confirmed using molecular methods for sex identification.

A commonality between the sex 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., 2008a; Nishida & Ochman, 2018). While pup survival relies 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 (Baby et al., 2014; Lea et al., 2002). Pacheco-Sandoval et al. (2019) showed that harbor seal fecal samples containing more lipid-rich prey had a much higher abundance of Firmicutes and lower Fusobacteria and Bacteroidetes (Pacheco-Sandoval et al., 2019). Molecular identification of prey species in fecal samples may therefore help to determine whether the diet is the driving factor behind the microbial differences observed here.

5 | CONCLUSION

This study characterized the fecal microbiome of the Juan Fernandez fur seal for the first time, including colonies from two of the three islands of the Juan Fernandez archipelago to which the species is endemic. Our findings showed that the overall microbiome composition was similar to compositions described for other pinnipeds. However, some of the samples showed a very different microbial composition pattern. This difference was mostly explained by an inverse relationship between Peptoclostridium and Fusobacterium abundance. Sex and its relationship to foraging behavior seem to be the most likely explanation of this phenomenon. However, additional studies investigating the relationship between sex, age, and prey are required to test this hypothesis. Overall, the results of this study provide a good baseline from which future hypothesis-based studies can be carried out, and it contributes to the understanding of host-microbial interaction in free-ranging, wild populations of pinnipeds. We highlight the need to expand knowledge in this field, particularly on microbial functionality, to understand its different members’ roles and compare microbial patterns between and within species.

ACKNOWLEDGMENTS

We want to thank SERNAPESCA and the Chilean National Forestry Commission (CONAF) in particular 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. Many thanks to Aerocardal for sponsoring flights to the Juan Fernández Archipelago, and OIKONOS, especially Héctor Gutierrez and Pablo Marriquê Angulo, for always giving a hand when needed. Ph.D. scholarship was provided to Constanza Toro-Valdivieso by The National Research and Development Agency of Chile (ANID) in partnership with Cambridge Trust. Financial support was provided by Newhaven College and the Department of Veterinary Medicine, University of Cambridge. Further support was provided by more than 60 people who donated to our study via the Crowdfunding platform.

CONFLICT OF INTEREST

The authors confirm that they have no conflicts of interest related to the content of this article.

AUTHOR CONTRIBUTIONS

Constanza Toro-Valdivieso: Conceptualization (equal); Formal analysis (lead); Funding acquisition (equal); Investigation (lead); Methodology (equal); Visualization (lead); Writing-original draft (lead); Writing-review & editing (equal). Frederick Toro: Formal analysis (supporting); Writing-review & editing (supporting). Samuel Stubbs: Formal analysis (supporting); Writing-original draft (supporting); Writing-review & editing (equal). Eduardo Castro-Nallar: Formal analysis (supporting); Funding acquisition (equal); Writing-review & editing (equal). Barbara Blacklaws: Conceptualization (equal); Formal analysis (supporting); Funding acquisition (equal); Methodology (equal); Supervision (lead); Writing-original draft (supporting); Writing-review & editing (equal).

ETHICS STATEMENT

All fecal samples were collected from the environment in a non-invasive manner. Disturbance of the colonies was kept to a minimum and no animal was handled or harmed in the process. Permits for the collection of samples were given by CONAF (Certificate 009217) and SERNAPESCA (R.E.X.N 43). Permission for the importation of samples into the United Kingdom was also obtained (ITIMP16.1158).

DATA AVAILABILITY STATEMENT

Raw reads data are publicly available in the European Nucleotide Archive (ENA) under the study accession PRJEB36555: https://www.ebi.ac.uk/ena/browser/view/PRJEB36555. All scripts used in this study can be accessed in GitHub at https://github.com/Cotis sima/JFFS_microbiome_first_characterisation.

ORCID

Constanza Toro-Valdivieso https://orcid.org/0000-0002-4485-5196

REFERENCES

Aguayo, A., Maturana, R., & Torres, D. (1971). El lobo fino de Juan Fernández. Revista de Biología Marina Y Oceanografía, 14(3), 125–149.
How to cite this article: Toro-Valdivieso, C., Toro, F., Stubbs, S., Castro-Nallar, E., & Blacklaws, B. (2021). Patterns of the fecal microbiota in the Juan Fernández fur seal (Arctocephalus philippii). MicrobiologyOpen, 10, e1215. https://doi.org/10.1002/mbo3.1215

FIGURE A1 Rarefaction curve estimating the number of ASVs (y-axis) for a given read count (x-axis). The vertical line indicates the cutoff at which samples were retained and rarefied.
FIGURE A2  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. Core rarefied data were used to calculate the diversity estimates.

FIGURE A3  PCoA using Bray-Curtis dissimilarity distance matrix using the filtered rarefied full dataset. Samples clustered in two groups (circles = cluster 1, triangles = cluster 2). Location is not driving the clustering.

TABLE A1  Inputs and outputs of each preprocessing step

| Preprocessing 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                            | 76,134                        | 0              | 2,074,038 |
| Filter ASVs (non-bacterial and ambiguous) | 57                | 577            | 2042                            | 76,134                        | 2081           | 2,071,957 |
| Filter samples                       | 54                | 577            | 13,981                          | 76,134                        | 8916           | 2,063,041 |
| Filter Contaminants                  | 54                | 558            | 13,981                          | 76,134                        | 278            | 2,062,763 |
| Rarefaction                           | 54                | 518            | 13,981                          | 13,981                        | 1,307,789      | 754,974   |
### TABLE A2 Bacterial phyla detected in Juan Fernandez fur seal feces

| Family counts | Total | Total counts rel. ab (%) | Mean rel. ab (%) | Rel. ab SD | Mean rel.ab (%) | Total ASV |
|---------------|-------|--------------------------|------------------|------------|-----------------|-----------|
| Firmicutes    | 863,365 | 41.85 | 40 | 24 | 296 |
| Fusobacteria  | 582,406 | 28.23 | 30 | 17 | 46 |
| Bacteroidetes | 455,251 | 22.07 | 22 | 10 | 94 |
| Proteobacteria| 113,805 | 5.52 | 6 | 4 | 74 |
| Actinobacteria| 30,597  | 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 A3 Summary of the bacterial family detected in feces of Juan Fernandez fur seal. Data are arranged in decreasing order based on counts mean

| Family | Total counts | Counts rel. ab (%) | Counts mean | Counts SD | Mean rel.ab (%) | Re. ab SD | Total ASV |
|--------|--------------|--------------------|-------------|-----------|-----------------|-----------|-----------|
| Fusobacteriaceae | 582,404 | 28.23 | 10,785.26 | 6958.72 | 30 | 17 | 45 |
| Bacteroidaceae | 320,047 | 15.52 | 5926.8 | 5319.43 | 15 | 10 | 28 |
| Ruminococcaceae | 310,109 | 15.03 | 5742.76 | 5206.41 | 15 | 13 | 139 |
| Lachnospiraceae | 213,725 | 10.36 | 3957.87 | 4195.97 | 9 | 8 | 61 |
| Peptostreptococcaceae | 193,151 | 9.36 | 3576.87 | 6353.37 | 9 | 16 | 16 |
| Rikenellaceae | 65,548 | 3.18 | 1213.85 | 1543.63 | 3 | 4 | 20 |
| Clostridiaceae | 60,276 | 2.92 | 1116.22 | 1294.81 | 1 | 3 | 11 |
| Burkholderiaceae | 47,544 | 2.3 | 880.44 | 849.59 | 2 | 2 | 8 |
| unclassified Gammaproteobacteria | 27,169 | 1.32 | 503.13 | 1116.46 | 1 | 2 | 9 |
| Acidaminococcaceae | 27,237 | 1.32 | 504.39 | 734.56 | 1 | 2 | 3 |
| Marinilaceae | 25,673 | 1.24 | 475.43 | 1022.21 | 1 | 2 | 13 |
| Prevotellaceae | 24,111 | 1.17 | 444.5 | 1606.42 | 1 | 4 | 4 |
| Coriobacteriaceae | 23,956 | 1.16 | 443.63 | 688.99 | 1 | 2 | 1 |
| Family XIII | 22,734 | 1.1 | 421 | 1294.81 | 1 | 3 | 11 |
| Clostridiales vadinBB60 group | 16,935 | 0.82 | 313.61 | 586.95 | 1 | 2 | 2 |
| Tannerellaceae | 15,153 | 0.73 | 280.61 | 647.81 | 1 | 2 | 8 |
| Succinivibrionaceae | 14,801 | 0.72 | 274.09 | 931.79 | 1 | 2 | 7 |
| Desulfovibrionaceae | 12,759 | 0.62 | 236.28 | 296.88 | 1 | 1 | 10 |
| Erysipelotrichaceae | 6926 | 0.34 | 128.26 | 160.05 | 0 | 1 | 7 |
| Akkermansiaiaceae | 6644 | 0.32 | 123.04 | 69.94 | 0 | 2 | 2 |
| Eggerthellaceae | 5951 | 0.29 | 110.2 | 244.78 | 0 | 1 | 4 |
| Helicobacteraceae | 5185 | 0.25 | 96.02 | 371.85 | 0 | 1 | 7 |
| Streptococcaceae | 4000 | 0.19 | 74.07 | 192.92 | 0 | 0 | 6 |
| unidentified Rhodospirllales | 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 |

(Continues)
| Family                  | Total counts | Counts rel. ab (%) | Counts mean | Counts SD | Mean rel.ab (%) | Re. ab SD | Total ASV |
|------------------------|--------------|--------------------|-------------|-----------|-----------------|-----------|-----------|
| Spongibacteraceae      | 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         |
| Viciuallaceae          | 662          | 0.03               | 12.26       | 71.52     | 0               | 0         | 2         |
| Flavobacteraceae       | 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           | 0.00               | 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         |
| Saprosiraceae          | 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 A4  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         |
| f347c63fc5e4aeb9753e165e3765e2a   | Firmicutes   | Peptostreptococcaceae | Peptoclostridium | 8.29         |
| e8b1922518029c50c69add839124db03  | Fusobacteria | Fusobacteriaceae     | Fusobacterium | 6.52         |
| 57f9edc6542ceeb78f352942d774c6     | Bacteroidetes | Bacteroidaceae       | Bacteroides   | 4.28         |
| 31984a302ffde46be5e852fa473e682a4   | Bacteroidetes | Bacteroidaceae       | Bacteroides   | 4.26         |
| b8d6a5a80d0d2s58f12f6ac7b9e0a1af   | Bacteroidetes | Bacteroidaceae       | Bacteroides   | 3.80         |
| c0dc53aad260a1b951b79999d6251c7c   | Fusobacteria | Fusobacteriaceae     | Fusobacterium | 3.73         |
| 1153942c5cc40d6ba5609222dedf86fe   | Firmicutes   | Lachnospiraceae      | Coprococcus 3 | 2.98         |
| 65dd9f625700a971ac9f5ee4e6cb       | Firmicutes   | Lachnospiraceae      | Blautia       | 2.18         |
| e176c3e42f3e3ce5d529c21ff5534e     | Firmicutes   | Clostridiaceae 1     | Clostridium sensu stricto | 1.95 |
| 43957b6d0323d8bb94193b2ad         | Firmicutes   | Lachnospiraceae      | [Ruminococcus] gauvreaui group | 1.93 |
| 1a73c668a4bb92b74a18b79f9ae63640   | Firmicutes   | Ruminococcaceae      | Ruminococcaceae UCG-005 | 1.75 |
| 5b87f47a447ef9a908507d2a0ed5b638   | Bacteroidetes | Rikenellaceae        | Alistipes     | 1.68         |
| bf41f11a100b11b4cbe9a5bdc591ea52   | Fusobacteria | Fusobacteriaceae     | Fusobacterium | 1.38         |
| 03f740ceaa1f0564719b212d701e5fa30  | Proteobacteria | Burkholderiaceae     | Sutterella    | 1.30         |
| 1188ef0328977f66e5179642f287ead    | Proteobacteria | Ruminococcaceae      | Ruminococcaceae UCG-005 | 1.29 |
| 256998f1bdf34e0c981cta484e751      | Firmicutes   | Lachnospiraceae      | unidentifie_Lachnospiraceae | 1.27 |
| 8e10797ddec28d9b9c6e64e5ba56db     | Firmicutes   | Ruminococcaceae      | Collinsella   | 1.16         |
| 2553cb6acfce16a17390555484369      | Firmicutes   | Ruminococcaceae      | [Eubacterium] coprostanoligenes group | 1.15 |
| b15e41c7f20b8dc40b0ed9f6c526885d   | Bacteroidetes | Prevotellaceae       | Alloprevotella | 1.14         |
| ca28c93154sf3db32df1c3ae132317     | Firmicutes   | Ruminococcaceae      | Ruminococcaceae UCG-005 | 1.12 |
| 76ded9f3adbc4155fb49e9dca2012ce81  | Firmicutes   | Ruminococcaceae      | Ruminococcaceae UCG-002 | 1.07 |
| c45b2a8ebeca2fca5036c312e8611416   | Bacteroidetes | Marinilaceae         | Odoribacter    | 1.07         |
| ce3476a906008973a3ab36ed068175d56 | Proteobacteria | Burkholderiaceae     | Sutterella    | 0.87         |
| 97525883636e3a001cb4d91c6cf7de06   | Proteobacteria | Acidaminococcaceae   | Phascolarctobacterium | 0.72 |
| 1cde60d0a8b17d66ed11653581f050    | Proteobacteria | Succinivibrionaceae  | Succinivibrio   | 0.68         |
| 6c4ce9ade2f56316c0ffac9587c173ec  | Firmicutes   | Lachnospiraceae      | unidentifie_Lachnospiraceae | 0.58 |
| 5851f40e4f4eb42b136d691b9a3bd2c83 | Bacteroidetes | Bacteroidaceae       | Bacteroides    | 0.55         |
| a0e5ee6d353d43229b5c9663cf05597    | Bacteroidetes | Bacteroidaceae       | Bacteroides    | 0.54         |
| 0ac8214c377777609c500f8576086b2e   | Firmicutes   | Lachnospiraceae      | Tyzzerella     | 0.44         |
| 420f3ebdb00de18846a5941b55a655e   | Bacteroidetes | Rikenellaceae        | Alistipes      | 0.44         |
| 0e7fd2a233c333cb8363b63a41b6bfc32  | Firmicutes   | Ruminococcaceae      | Ruminococcaceae NK4A214 group | 0.43 |
| df4080562977f20c5e5cc68907e39cc8  | Firmicutes   | Lachnospiraceae      | Tyzzerella     | 0.41         |
| c00129ca8377cb317764d4e03a9091d    | Fusobacteria | Fusobacteriaceae     | Fusobacterium | 0.41         |
| 93623ff46e3615ce4aa4f0a9554fd4de   | Proteobacteria | Desulfovibrionaceae  | unidentifie_Desulfovibrionaceae | 0.35 |
| a672e8b3e7eb3a28e5beabe661606ad2   | Firmicutes   | Ruminococcaceae      | unidentifie_Ruminococcaceae | 0.33 |
| 0c3d20387f4019f7f0d3c6b404019     | Firmicutes   | Ruminococcaceae      | Ruminoclostridium 9 | 0.30 |
| b6578d861d1c0e9230878a5a81c8501    | Proteobacteria | unidentifie_Gammaproteobacteria | unidentifie_Gammaproteobacteria | 0.25 |
| 57f0c2ba2627cebfe197aa991777cb0    | Bacteroidetes | Tannerellaceae       | Parabacteroides | 0.24 |

(Continues)
TABLE A5  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-normalized data were 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 A6  The selected value of the Spearman’s rank correlation performed on the rarefied core data including PC1 and 2 for each dissimilarity distance. The table reporting only the correlation that showed to be strong (0.6 ≤ |ρ| ≤0.79) and very strong, (0.8 ≤ |ρ| ≤1)

| Correlation pair       | p      | Strength | p     |
|------------------------|--------|----------|-------|
| Bacteroides            | -0.67  | Strong   | <0.001|
| Fusobacterium          | -0.92  | Very strong | <0.001|
| Peptoclostridium       | 0.81   | Very strong | <0.001|
| Ruminoclostridium 9    | 0.63   | Strong   | <0.001|
| Ruminococcaceae NK4A214 group | 0.61 | Strong   | <0.001|
| Odoribacter            | 0.62   | Strong   | <0.001|
| Parabacteroides        | 0.71   | Strong   | <0.001|
| Fusobacterium          | -0.63  | Strong   | <0.001|
| Ruminoclostridium 9    | 0.61   | Strong   | <0.001|
| Fusobacterium          | -0.94  | Very strong | <0.001|
| Peptoclostridium       | 0.75   | Strong   | <0.001|