Science of the Total Environment 932 (2024) 172943

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Meta-analysis of the Cetacea gut microbiome: Diversity, co-evolution, and interaction with the anthropogenic pathobiome

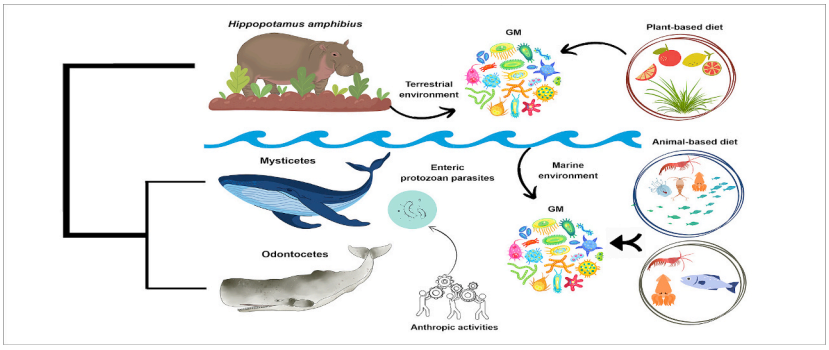
Elena Radaelli [a,b,1], Giorgia Palladino [a,b,1], Enrico Nanetti [a], Daniel Scicchitano [a,b], Simone Rampelli [a,b], Sabina Airoldi [c], Marco Candela [a,b,*], Marianna Marangi [d,**]

[a] Unit of Microbiome Science and Biotechnology, Department of Pharmacy and Biotechnology (FaBiT), Alma Mater Studiorum - University of Bologna, Via Belmeloro 6, 40126 Bologna, Italy
[b] Fano Marine Center, The Inter-Institute Center for Research on Marine Biodiversity, Resources and Biotechnologies, Viale Adriatico 1, 61032, Fano, Italy
[c] Tethys Research Institute, Viale G.B. Gadio 2, 20121 Milano, Italy
[d] Department of Clinical and Experimental Medicine, University of Foggia, Viale Luigi Pinto, 71122 Foggia, Italy

HIGHLIGHTS

- Cetaceans are pivotal apex predators in all marine ecosystems.
- In cetaceans, the gut microbiome shows signs of phylosymbiosis.
- Diet may also shape microbiome diversity in toothed and baleen whales.
- Anthropic impact in the sea can be inferred through zoonotic parasites detection.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Yolanda Picó

Keywords:
Cetacea
Gut microbiome
16S gene sequencing
Phylosymbiosis
Protozoan parasites

ABSTRACT

Despite their critical roles in marine ecosystems, only few studies have addressed the gut microbiome (GM) of cetaceans in a comprehensive way. Being long-living apex predators with a carnivorous diet but evolved from herbivorous ancestors, cetaceans are an ideal model for studying GM-host evolutionary drivers of symbiosis and represent a valuable proxy of overall marine ecosystem health. Here, we investigated the GM of eight different cetacean species, including both Odontocetes (toothed whales) and Mysticetes (baleen whales), by means of 16S rRNA-targeted amplicon sequencing. We collected faecal samples from free-ranging cetaceans circulating within the Pelagos Sanctuary (North-western Mediterranean Sea) and we also included publicly available cetacean gut microbiome sequences. Overall, we show a clear GM trajectory related to host phylogeny and taxonomy (i.e., phylosymbiosis), with remarkable GM variations which may reflect adaptations to different diets between baleen and toothed whales. While most samples were found to be infected by protozoan parasites of potential anthropic

* Correspondence to: Unit of Microbiome Science and Biotechnology, Department of Pharmacy and Biotechnology (FaBiT), Alma Mater Studiorum - University of Bologna, Via Belmeloro 6, 40126 Bologna, Italy.
** Corresponding author.
E-mail addresses: marco.candela@unibo.it (M. Candela), marianna.marangi@unifg.it (M. Marangi).
[1] These authors have contributed equally to this work and share first authorship.

https://doi.org/10.1016/j.scitotenv.2024.172943
Received 27 February 2024; Received in revised form 19 April 2024; Accepted 30 April 2024
Available online 6 May 2024

origin, we report that this phenomenon did not lead to severe GM dysbiosis. This study underlines the importance of both host phylogeny and diet in shaping the GM of cetaceans, highlighting the role of neutral processes as well as environmental factors in the establishment of this GM-host symbiosis. Furthermore, the presence of potentially human-derived protozoan parasites in faeces of free-ranging cetaceans emphasizes the importance of these animals as bioindicators of anthropic impact on marine ecosystems.

1. Introduction

The gut microbiome (GM) has been extensively studied in many species of terrestrial mammals (Ley et al., 2008; NIH HMP Working Group, 2009; McFall-Ngai et al., 2013; Xue et al., 2015; Moeller et al., 2016; Shiffman et al., 2017; Brice et al., 2019; Perofsky et al., 2019; Song et al., 2020; de Jonge et al., 2022), but few studies have been conducted on marine mammals, especially in the free-ranging state (Sanders et al., 2015; Soverini et al., 2016; Erwin et al., 2017; Godoy-Vitorino et al., 2021; Suzuki et al., 2019; Abdelrhman et al., 2020; Glaeser et al., 2022; Tian et al., 2020; Bai et al., 2021; Bai et al., 2022; Brown et al., 2022; Li et al., 2022a; Wang et al., 2022). Among marine mammals, the Cetacea infraorder (which comprises two parvorders, i.e., Odontocetes and Mysticetes) includes only exclusively aquatic species fully adapted to the aquatic environment, representing an ideal model for studying the evolutionary drivers of symbiosis between the GM and its host (Gatesy et al., 2013).

The emergence of the first cetaceans dates back to the Paleogene and, initially, they were classified within their dedicated order (Graur and Higgins, 1994; Gatesy et al., 1996; Gatesy et al., 2013). Following new molecular evidence – which suggested the origin of cetaceans from the artiodactyls – cetaceans were included within the Artiodactyla order, with the *Hippopotamidae* family encompassing their phylogenetically closest living relatives (Nikaido et al., 1999; Gatesy et al., 2013). The adaptation of cetaceans to the marine environment has naturally resulted in important ecological and morphological changes, following evolutionary convergence with different marine organisms (Gatesy et al., 2013; Foote et al., 2015). Particularly, for cetaceans, life in the marine environment required the acquisition of peculiar dietary habits, which represent an exception within their order. While terrestrial artiodactyls have an exclusively or predominantly herbivorous diet, cetaceans mainly feed on crustaceans, molluscs, and fish, with mysticetes feeding on more chitin-rich organisms (e.g., crustaceans and molluscs) compared to odontocetes (Langer., 2001; Sanders et al., 2015). In this regard, recent research has shown that cetacean guts exhibit functional and morphological similarities both to terrestrial carnivores, e.g., in terms of GM genes for protein digestion and biosynthesis, and to their herbivorous terrestrial relatives. They indeed possess a multichambered foregut and the microbiome functions for carbohydrate metabolism, combining carnivores and herbivores features as a possible adaptive response to their peculiar dietary niche in the marine environment (Herwig et al., 1984; Langer., 2001; Sanders et al., 2015). In this context, research has been performed aiming at comparing GM changes between mysticetes (baleen whales) and odontocetes (toothed whales), whose important differences in feeding apparatus led to a marked trophic separation (Ley et al., 2008; Thewissen et al., 2011; Gatesy et al., 2013). Studies suggested a general diversification of the cetacean GM according to parvorders (Muegge et al., 2011; Sanders et al., 2015; Bik et al., 2016; Erwin et al., 2017; McKenzie et al., 2017). Indeed, even if they rely on a carnivorous diet, mysticetes need to gain energy from chitin-rich foods (i.e., crustaceans and molluscs), requiring further adaptations, also in terms of the GM, for the extraction of energy from complex polysaccharides (Herwig et al., 1984; Langer., 2001; Sanders et al., 2015). On the other hand, the GM of odontocetes shows more carnivore-like features. For instance, the bottlenose dolphin (*Tursiops truncatus*), i.e., a fish and cephalopods-feeding species (Blanco et al., 2001), shows higher GM similarities with carnivorous marine fish compared to phylogenetically closer baleen whales (Soverini et al., 2016; Song et al., 2020; Wang et al., 2022).

Even if findings are promising and suggest the possible importance of the cetacean GM in the adaptation to marine life and in the specialization of the two cetacean parvorders (Wang et al., 2022), most studies are observational and do not provide an evolutionary perspective on the microbiome-host symbiosis. Therefore, further studies specifically focused on the discovery of evolutionary trajectories in GM-host symbiosis in cetaceans are needed.

Increasing attention has been recently drawn to the health status of cetaceans as a relevant proxy in monitoring the health of entire marine ecosystems (Marangi et al., 2021). Cetaceans play a central ecological role in marine habitats as bioindicator sentinels for ocean health, being long-living apex predators as well as primary and secondary consumers within the marine food web (Roman and McCarthy, 2010; Bik et al., 2016). In this scenario, investigating the distribution of pathogenic microorganisms, such as enteric parasites, in cetaceans, as well as their influence on gut health and GM structure, could act as a key strategy to monitor the well-being of these species and of the habitats they thrive in. In fact, the increase of anthropic pressure within coastal areas (e.g., from agricultural, animal, and human waste), together with global climate change, have heightened the rate of infectious diseases in marine environments, many of which can also be zoonotic (Marangi et al., 2022).

In our study, we investigate the GM of both toothed and baleen whales with a double goal: firstly, we want to provide some glimpses on the possible co-evolutionary dynamics between cetaceans and their hosted microbial communities. Secondly, we aim to assess the presence of enteric protozoan parasites potentially linked to anthropogenic activities. For this purpose, we explored and compared the GM compositional structure of apparently healthy, free-ranging individuals belonging to eight different species of cetaceans, including both mysticetes and odontocetes, sampled for the purpose of this study (Pelagos Sanctuary), as well as retrieved from the literature. Further, for the samples collected in the present study, we carried out a parasitological investigation in the animals' faecal samples considering three enteric zoonotic protozoan parasites, namely *Blastocystis* sp., *Giardia duodenalis* and *Cryptosporidium* spp. These are emerging pathogens of water and foodborne origin that can infect humans and a wide range of animals (Xiao and Feng, 2017), including marine wildlife, and have recently been reported for the first time in the Mediterranean Sea (Würsig, 2020; Marangi et al., 2022).

2. Materials and methods

2.1. Study area

The Pelagos Sanctuary is a marine protected area extending >87,500 km^2 in the North-Western Mediterranean Sea between the Italian peninsula, France and the Island of Sardinia, encompassing Corsica and the Tuscan Archipelago (Fig. 1). The Sanctuary waters include the Ligurian Sea and parts of the Corsican and Tyrrhenian Seas, and contain the internal maritime (15 %) and territorial waters (32 %) of France, Monaco and Italy, as well as the adjacent high seas (53 %). Within the Sanctuary area, the continental shelf is wide only in correspondence of limited coastal plains, i.e. eastern Corsica and Tuscany. Elsewhere, it is mostly narrow and disseminated with steep, deeply-cut submarine canyons (Notarbartolo-di-Sciara et al., 2008). Notably, this marine area is affected by anthropogenic stressors of high ecological impact. The Pelagos Sanctuary is close to the highly urbanized and

industrialized coasts of the Ligurian Sea, with a consequent high amount of waste in the waters (Vietti et al., 2010), and to the coasts of Tuscany, where numerous permanent and small-scale farming systems are present as a possible sources of release of enteric microorganisms by oral-faecal route through wastewater (Filippini et al., 2018; Guardone et al., 2020); moreover, the Tyrrhenian Sea is intensely used for aquaculture purposes (Glaeser et al., 2022).

2.2. Sampling

In the framework of a research project on the ecology of cetaceans, faecal samples of 18 individuals of sperm whale (*Physeter macrocephalus*, *Physeteridae*, Odontocetes), 5 individuals of long-finned pilot whale (*Globicephala melas*, *Delphinidae*, Odontocetes) and 11 individuals of fin whale (*Balaenoptera physalus*, *Balaenopteridae*, Mysticetes) were collected across the Ligurian waters of the Pelagos Sanctuary (Fig. 1). During several summer boat surveys, photo identification and floating faeces were collected using a fine nylon mesh net, avoiding direct contact and any disturbance of the animals (Marangi et al., 2021). Faecal samples were collected as quickly as possible upon defecation, immediately placed inside labelled sterile Falcon tubes and stored at −80 °C upon arrival to the laboratory until genomic DNA extraction. Faeces from one individual of Cuvier's beaked whale (*Ziphius cavirostris*, *Ziphidae*, Odontocetes) were also retrieved from a freshly stranded individual right after it was found dead along the Ligurian coast. For clarity purposes, we will be referring to these 35 samples as 'present study' samples. Sampling details of the present study are summarized in Supplementary Table S1.

2.3. 16S rRNA gene amplification and sequencing

Genomic DNA was extracted from the present study faecal samples (approximately 200 mg) by using the QiaAMP DNA Stool Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and then eluted in 200 μL of TE buffer. Extracted DNA was processed for 16S rRNA gene amplification and sequencing for prokaryotic community characterization. Library preparation was performed following the Illumina 16S Metagenomic Sequencing Library Preparation protocol (Illumina, San Diego, CA, USA). The V3–V4 hypervariable regions of the 16S rRNA gene were PCR amplified in a 50 μL final volume containing 25 ng of microbial DNA, 2× KAPA HiFi HotStart ReadyMix (Roche, Basel, Switzerland) and 200 nmol/L of 341F (5'-CCTACGGGNGGCWGCAG-3′) and 785R (5′-GACTACHVGGGTATCTAATCC-3′) primers with Illumina adapter overhang sequences added (Klindworth et al., 2013). The PCR thermocycle consisted of 3 min at 95 °C, 25 cycles of 30 s at 95 °C, 30 s at 55 °C and 30 s at 72 °C, and a final 5-min extension step at 72 °C (Turroni et al., 2016). PCR products were then purified with Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA). Indexed libraries were prepared by limited-cycle PCR, using Nextera technology (Illumina), and cleaned-up as described above. Libraries were quantified using the Qubit 3.0 fluorimeter (Invitrogen, Waltham, MA, USA), normalised to 4 nM and pooled. The sample pool was denatured with 0.2 N NaOH and diluted to a final concentration of 4.5 pM with a 20 % PhiX control. Sequencing was performed on an Illumina MiSeq platform using a 2 × 250 bp paired-end protocol, according to the manufacturer's instructions (Illumina).

2.4. Data acquisition

In order to integrate our dataset with already available data, we downloaded 16S rRNA gene data of cetacean GM from free-ranging and apparently healthy individuals. In particular, we took into account the previous works from Sanders et al. (2015) (MG-RAST database, accession code mgp3854), Glaeser et al. (2022) (NCBI Sequence Read Archive SRA, BioProject PRJNA627228) and Sherrill-Mix et al. (2018) (NCBI Sequence Read Archive SRA, BioProject PRJNA397174). From Sanders et al. (2015), we downloaded the 16S rRNA gene sequences of gut associated microbiomes of 7 North Atlantic right whales (*Eubalaena glacialis*), 1 sei whale (*Balaenoptera borealis*) and 4 humpback whales (*Megaptera novaeangliae*). From Glaeser et al. (2022), we downloaded the sequences of 1 individual of sei whale (*Balaenoptera borealis*), 2 blue

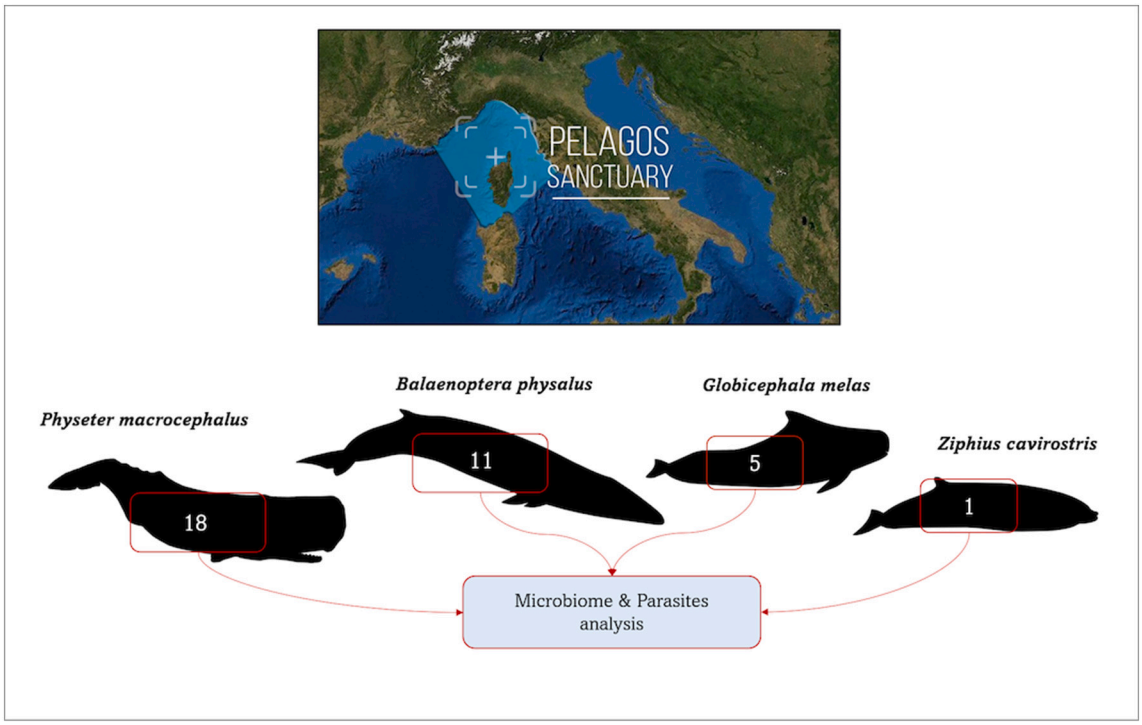


Fig. 1. – Sampling area and overall sampling summary of the present study. Location of the Pelagos Sanctuary in the Mediterranean Sea (https://www.wwfmmi.org/what_we_do/wildlife/the_pelagos_sanctuary/). The total number of samples collected for the present study from each cetacean species is reported.

whales (*Balaenoptera musculus*), 10 fin whales (*Balaenoptera physalus*) and 4 sperm whales (*Physeter macrocephalus*). From Sherrill-Mix et al. (2018), we downloaded the sequences of 1 individual of fin whale (*Balaenoptera physalus*). Besides cetacean sequences, we downloaded from Sherrill-Mix et al. (2018) and from Sanders et al. (2015), respectively, the 16S rRNA gene sequences of gut associated microbiomes of 4 different species of sharks (i.e., 5 bull sharks *Carcharhinus leucas*; 5 sandbar sharks *Carcharhinus plumbeus*; 2 tiger sharks *Galeocerdo cuvier*; 1 lemon shark *Negaprion brevirostris*) and 4 individuals of hippopotamus (*Hippopotamus amphibius*). All of these sequences were retrieved and integrated with the present study dataset. A schematic summary of the entire sample set (including both the present study samples and the publicly available sequences) is provided in Table 1.

2.5. Parasitological molecular analyses

Extracted genomic DNA was also used for parasitological molecular analyses focusing on 3 protozoan parasites, i.e., *Blastocystis* sp., *Giardia duodenalis* and *Cryptosporidium* spp. For the molecular and genetic characterization of *Blastocystis* sp., part of the 1800 bp *SSU-rDNA* (600 bps) was amplified following the PCR protocol as described in Scicluna et al., 2006 and Marangi et al., 2021. For the molecular and genetic characterization of *Giardia duodenalis* and *Cryptosporidium* spp., part of the TPI gene (~ 530 bps) and of the GP60 gene (~ 358 bps) was amplified following the nested-PCR protocol as described in Giangaspero et al., 2014 and Marangi et al., 2022.

The PCR fragments were run on 1.2 % agarose gel and positive samples were purified with Exonuclease I (EXO I) and Thermosensitive Alkaline Phosphatase (FAST AP) (Fermentas) enzymes according to the manufacturer's instructions. Purified amplicons were directly sequenced in both directions using the ABI PRIMS Big Dye Terminator v. 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, California, USA) with the same primers as the respective PCR reaction, according to the manufacturer's instructions. Obtained sequences were determined on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, USA) and the chromatograms were inspected by eye using the Finch TV software. Primer regions, as well as bad-quality regions, were removed.

Once the sequences had been cleaned up, each sequence was compared with the *Blastocystis*, *Giardia* and *Cryptosporidium* homologous nucleotide sequences available in GenBank databases using the Blastn program (https://blast.ncbi.nlm.nih.gov). Then, the obtained sequences corresponding to *Blastocystis SSU-rDNA* gene portion, *Giardia TPI* and *Cryptosporidium GP60* genes were gathered in a fasta file and aligned using the ClustalW implementation of the BioEdit software v. 7.0.5, and the alignment was adjusted manually, when necessary. Following this, we compared the sequences here obtained with those available on the GenBank database in order to determine the species. The presence or absence of any of these enteric protozoan parasites in the stools of the individuals of the present study is reported in Supplementary Table S2.

Table 1
Entire set of samples analyzed in this article.

Species	Study of origin	Samples analyzed (N)
Present study		
• *Physeter macrocephalus*	Present Study	18
• *Globicephala melas*	Present Study	5
• *Ziphius cavirostris*	Present Study	1
• *Balaenoptera physalus*	Present Study	11
		Tot 35
Publicly available data		
• *Eubalaena glacialis*	Sanders et al. (2015)	7
• *Balaenoptera borealis*	Sanders et al. (2015)	1
• *Megaptera novaeangliae*	Sanders et al. (2015)	4
• *Hippopotamus amphibius*	Sanders et al. (2015)	4
• *Balaenoptera borealis*	Glaeser et al. (2022)	1
• *Balaenoptera musculus*	Glaeser et al. (2022)	2
• *Balaenoptera physalus*	Glaeser et al. (2022)	10
• *Physeter macrocephalus*	Glaeser et al. (2022)	4
• *Balaenoptera physalus*	Sherrill-Mix et al. (2018)	1
• *Carcharhinus leucas*	Sherrill-Mix et al. (2018)	5
• *Carcharhinus plumbeus*	Sherrill-Mix et al. (2018)	5
• *Galeocerdo cuvier*	Sherrill-Mix et al. (2018)	2
• *Negaprion brevirostris*	Sherrill-Mix et al. (2018)	1

2.6. Bioinformatics and statistics

For the 16S rRNA gene analysis, raw sequences for a total of 35 cetacean samples from the present study and 47 samples from other studies were processed separately, according to the study of origin, using a pipeline combining PANDAseq (Masella et al., 2012) and QIIME 2 (Bolyen et al., 2019). High-quality reads were retained using the "fastq filter" function of the Usearch11 algorithm (Edgar, 2010) and clustered into amplicon sequence variants (ASVs) using DADA2 (Callahan et al., 2016). Normalisation was performed to the lowest number of reads per sample in each study and to 2485 reads per sample for the present study. Taxonomy was assigned using the VSEARCH classifier (Rognes et al., 2016) and the SILVA database (December 2017 release) as a reference (Quast et al., 2012). All sequences assigned to eukaryotes or unassigned were discarded. Overall, an average sequencing depth of 7.6 ± 6.6 thousand high-quality reads per sample was obtained, resulting in a total of 887 ASVs for the present study. For the samples retrieved from Sanders et al. (2015), 20.3 ± 7.9 thousand high-quality reads per sample were obtained, for a total of 3057 ASVs. From the study of Glaeser et al. (2022), 10.6 ± 4.5 thousand high-quality reads per sample were obtained, for a total of 1203 ASVs. Finally, for the samples retrieved from Sherrill-Mix et al. (2018), 34.6 ± 14.8 thousand high-quality reads per sample were obtained, resulting in a total of 1151 ASVs. Data from the different studies were merged at the different taxonomic levels (i.e., phyla, families, and genera) and subsequently normalised to the lowest relative abundance value of the present study, this being the one with the lowest normalisation value of reads per sample.

All statistical analyses were performed using R software (R Core Team; www.r-project.org), v. 4.1.2, with the packages "made4" (Culhane et al., 2005), "vegan" (Oksanen, 2013; https://CRAN.R-project.org/package=vegan), "vegan3d" (https://CRAN.R-project.org/package=vegan3d), "scatterplot3d" (Ligges and Mächler, 2002), "plot3D" (https://CRAN.R-project.org/package=plot3D), "rgl" (https://CRAN.R-project.org/package=rgl), "ggplot2" (Wickham et al., 2016), "Maaslin2" (Mallick et al., 2021), "reshape2" (Wickham, 2007), "RColorBrewer" (https://CRAN.R-project.org/package=RColorBrewer), "facto extra" (https://CRAN.R-project.org/package=factoextra) and "mixOmics" (Rohart et al., 2017). Alpha diversity was calculated as bacterial richness at the genus level using the function "estimateR" of the vegan package. The Wilcoxon rank-sum test was used to assess significant differences in alpha diversity distributions. Beta diversity was estimated by computing the Bray-Curtis distances and data separation in the Principal Coordinates Analysis (PCoA) was tested using a permutation test with pseudo-F ratio (function "adonis" in the vegan package). *P*-values, when necessary, were corrected for multiple testing using the Benjamini–Hochberg method, with a false discovery rate (FDR) ≤ 0.05 considered to be statistically significant. For parasitological analyses, weighted and unweighted UniFrac distances are shown in PCoA plots. In this case, alpha diversity is represented by means of 3 different metrics, namely Faith's Phylogenetic Diversity, number of observed ASVs and Shannon index.

The linear discriminant analysis (LDA) was performed through the tool LEfSe (Segata et al., 2011), which aimed to identify discriminant bacterial genera between the two cetacean parvorders (i.e., mysticetes and odontocetes) and the hippopotamus, retaining only microbial taxa with LDA score threshold of ±2 (on a log10 scale) and *p*-value ≤0.05. The PICRUSt2 tool (Douglas et al., 2020) was used to infer carbohydrate-active enzymes genes (CAZymes) within the ASVs assigned to the bacterial genera characteristic of the two cetacean parvorders and hippos, and previously identified by LEfSe, according to

the CAZy database (http://www.cazy.org/Glycoside-Hydrolases.html). The heatmap of presence/absence of these genes was created by means of the heatmap.2 function in R (ggplot package in R), reporting ASVs ascribed to each bacterial genus in rows and CAZymes EC numbers in columns.

The phylosymbiosis analysis was performed following the phylogenetic and microbial diversity trees construction. The gene-based phylogenetic tree of host samples was based on the mitochondrial cytochrome oxidase subunit 1 (COI) retrieved from NCBI; the alignment of sequences was performed by means of the MUSCLE tool (Edgar, 2004). Then, the gene-based tree was created through FastTree2 (Price and Dehal, 2010) and rooted on FigTree v. 1.4.4. (Rambaut, 2018). The microbiome diversity tree was an output of the QIIME2 pipeline for the microbial diversity metrics. The congruence between the two trees was then analyzed through a matrix correlation method (Mantel test with Pearson's correlations; p-value ≤ 0.05). Significant phylosymbiotic signal was calculated for each node by implementing probabilistic models in the Count program (Csűrös, 2010), that was originally designed to model the evolution of gene families through gene gain and loss.

3. Results

3.1. Microbiome community structure of odontocetes and mysticetes

Here, we performed a meta-analysis including a total of 65 GM samples belonging to 8 different cetacean species, of which 35 were newly sequenced in the present study and 30 were derived from public databases. The considered species were: *B. physalus*, *B. borealis*, *B. musculus*, *M. novaeangliae*, *E. glacialis*, *P. macrocephalus*, *G. melas* and *Z. cavirostris* (Supplementary Fig. 1). Alpha and beta diversity were computed comparing the two parvorders and the eight cetacean species. Notably, the cetacean GMs clearly segregate by both parvorder and species in the Bray-Curtis-based PCoAs (permutation test with pseudo-F ratio, $p \leq 0.001$) (Fig. 2). Additionally, alpha diversity shows significantly higher values for mysticetes compared to odontocetes (Wilcoxon rank-sum test) (Fig. 2). The study of origin does not seem to exert a strong influence on the overall GM profile of the cetaceans considered in this meta-analysis (permutation test with pseudo-F ratio), as demonstrated comparing the GMs of the two most widespread cetacean species (i.e., *P. macrocephalus* and *B. physalus*) across different studies (Supplementary Fig. 2).

Differences between odontocetes and mysticetes can be seen at all taxonomic ranks. More specifically, at the phylum level, mysticetes are characterized by two dominant phyla, namely Firmicutes and Bacteroidetes, while the sub-dominant phyla are Actinobacteria, Proteobacteria and Spirochaetae. Odontocetes included also Proteobacteria among the dominant components, with Proteobacteria, Firmicutes and Bacteroides being co-dominant in *P. macrocephalus* (mean relative abundance ± SD, 10.0 ± 21.2 %, 39.5 ± 22.5 % and 35.6 ± 22.7 %, respectively) and in *Z. cavirostris* (60.0 %, 11.1 % and 26.8 %, respectively), while Proteobacteria (39.4 ± 33.7 %), Firmicutes (27.7 ± 37.4 %) and Actinobacteria (17.4 ± 20.3 %) dominate the GM of *G. melas*.

At the family level, mysticetes are characterized by several dominant taxa. These include: *Clostridiaceae 1*, with the highest mean abundance value (43.6 ± 25.5 %) in *E. glacialis*, *Ruminococcaceae*, with the highest mean abundance value (36.6 ± 38.1 %) in *M. novaeangliae*, *Prevotellaceae*, with the highest mean abundance value (16.3 ± 18.2 %) in *B. borealis* and *Rikenellaceae*, with the highest mean abundance value (12.1 ± 17.0 %) in *B. musculus*. The sub-dominant families in mysticetes include *Bacteroidaceae*, *Pasterullaceae*, *Spirochetaceae* and *Erysipelotrichaceae*. The dominant gut microbial families in odontocetes are: *Rikenellaceae* (20.1 ± 19.9 %), *Ruminococcaceae* (15.8 ± 9.9 %) and *Bacteroidales BS11 gut group* (10.2 ± 9.8 %) in *P. macrocephalus*, *Planococcaceae* (20.4 ± 40.8 %) and *Burkholderiaceae* (14.4 ± 28.3 %) in *G. melas*, and *Enterobacteriaceae* (34.8 %), *Bacteroidaceae* (26.8 %) and *Pasterullaceae* (23.0 %) in *Z. cavirostris*. The sub-dominant families in odontocetes include *Lachnospiraceae*, *Christensenellaceae*, *Vibrionaceae* and *Micromonosporaceae*.

At lower taxonomic ranks, the cetacean GM showed recognizable species-specific differences. The dominant bacterial genera in mysticetes show a high degree of species-specificity. In the genus *Balaenoptera*, for instance, *Clostridium* sensu stricto *1* is the most dominant genus, with the highest mean abundance value (35.3 ± 49.9 %) in *B. musculus*. *E. glacialis* and *M. novaeangliae*, on the other hand, show the bacterium *Sarcina* as the most abundant genus (40.7 ± 23.3 % and 27.4 ± 33.0 %, respectively). Furthermore, in *M. novaeangliae*, the genus *Ruminococcaceae UCG-005* is among the most dominant taxa (19.5 ± 29.1 %). In mysticetes, sub-dominant genera also mirror species-specific patterns, and these include: *Bacteroides*, *Ruminococcaceae UCG-005* and *Ruminococcaceae UCG-014*. The bacterial genera found to be dominant in odontocetes are also species-specific. In *P. macrocephalus*, *Rikenellaceae RC9 gut group* is the dominant genus (17.9 ± 20.1 %). Dominant genera characterizing *G. melas* are *Lysinibacillus* (16.2 ± 32.4 %) and *Burkholderia-Paraburkholderia* (14.2 ± 28.5 %), while the ones characterizing *Z. cavirostris* are *Bacteroides* (26.8 %), *Morganella* (25.8 %) and *Actinobacillus* (13.1 %). The sub-dominant genera in odontocetes are *Christensenellaceae R-7 group*, *Photobacterium*, *Actinoplanes* and *Escherichia-Shigella*, again showing a certain degree of specificity in relation to the host species.

We applied linear discriminant analysis (LDA) effect size (LEfSe) (Segata et al., 2011), aiming to highlight the bacterial taxa that discriminated between mysticetes and odontocetes. According to LEfSe (Fig. 3), genera associated with odontocetes were *Rikenellaceae RC9 gut group*, *Christensenellaceae R-7 group*, *Dethiosulfovibrio*, *Staphylococcus*, *Tyzzerella 3*, *Enterococcus*, *Sphaerochaeta*, *Pyramidobacter*, *Escherichia-Shigella*, *Petrimonas*, *Ruminococcus 2*, *Gordonibacter*, *Anaerovorax*, *Lactobacillus*, *Anaerorhabdus furcosa group* and *Bilophila*.

On the other hand, mysticetes samples were enriched in *Ruminococcus torques group*, *Eubacterium fissicatena group*, *Anaerofustis*, *Peptococcus*, *Anaeroplasma*, *Hydrogenoanaerobacterium*, *Ruminococcaceae UCG-009*, *Prevotellaceae UCG-004*, *Prevotella 1*, *Lachnospiraceae AC2044 group*, *Eubacterium brachy group*, *Family XIII UCG-001*, *Atopobium*, *Faecalitalea*, *Granulicatella*, *Ruminococcaceae UCG-002*, *Oscillibacter*, *Sutterella*, *Ruminiclostridium 9*, *Treponema 2*, *Alloprevotella*, *Bacteroides*, *Ruminococcaceae UCG-005*, *Sarcina* and *Clostridium* sensu stricto *1*.

3.2. Inferred profile of carbohydrate-active enzymes for chitin and lignocellulose degradation in the GM taxa characterizing odontocetes and mysticetes

In order to provide some glimpses on the possible functional implications connected to the observed GM variations between odontocetes and mysticetes, the ASVs corresponding to the discriminant genera reported in Fig. 3 were obtained from the 35 cetaceans samples sequenced in the present study (see Supplementary Table S3 for a complete list of the bacterial ASVs). For each of the retrieved ASV, the matching profile of CAZymes for the degradation of lignocellulose and chitin was inferred. Specifically, for lignocellulose degradation enzymes, eight glycoside hydrolases (GHs), 17 carbohydrate esterases (CEs), 13 polysaccharide lyases (PLs) and 10 enzymes linked to auxiliary activities (AAs) were considered, while, for chitin degradation, eight enzymes belonging to the GH family and three belonging to the CE family were included (Supplementary Table S4). To provide an evolutionary perspective to our analysis, the ASVs corresponding to the taxa characterizing the *Hippopotamus amphibius* GM were also obtained from Sanders et al. (2015), and the matching profiles of lignocellulose and chitin degrading CAZymes were recovered as indicated above. In Fig. 4, we provide the heatmap of the obtained CAZymes profiles for the ASVs being discriminant of the GM of the hippopotamus, odontocetes and mysticetes. According to our findings, the ASVs characterizing the three different host species showed detectable differences in the profile of the CAZymes functions for the degradation of lignocellulose and chitin

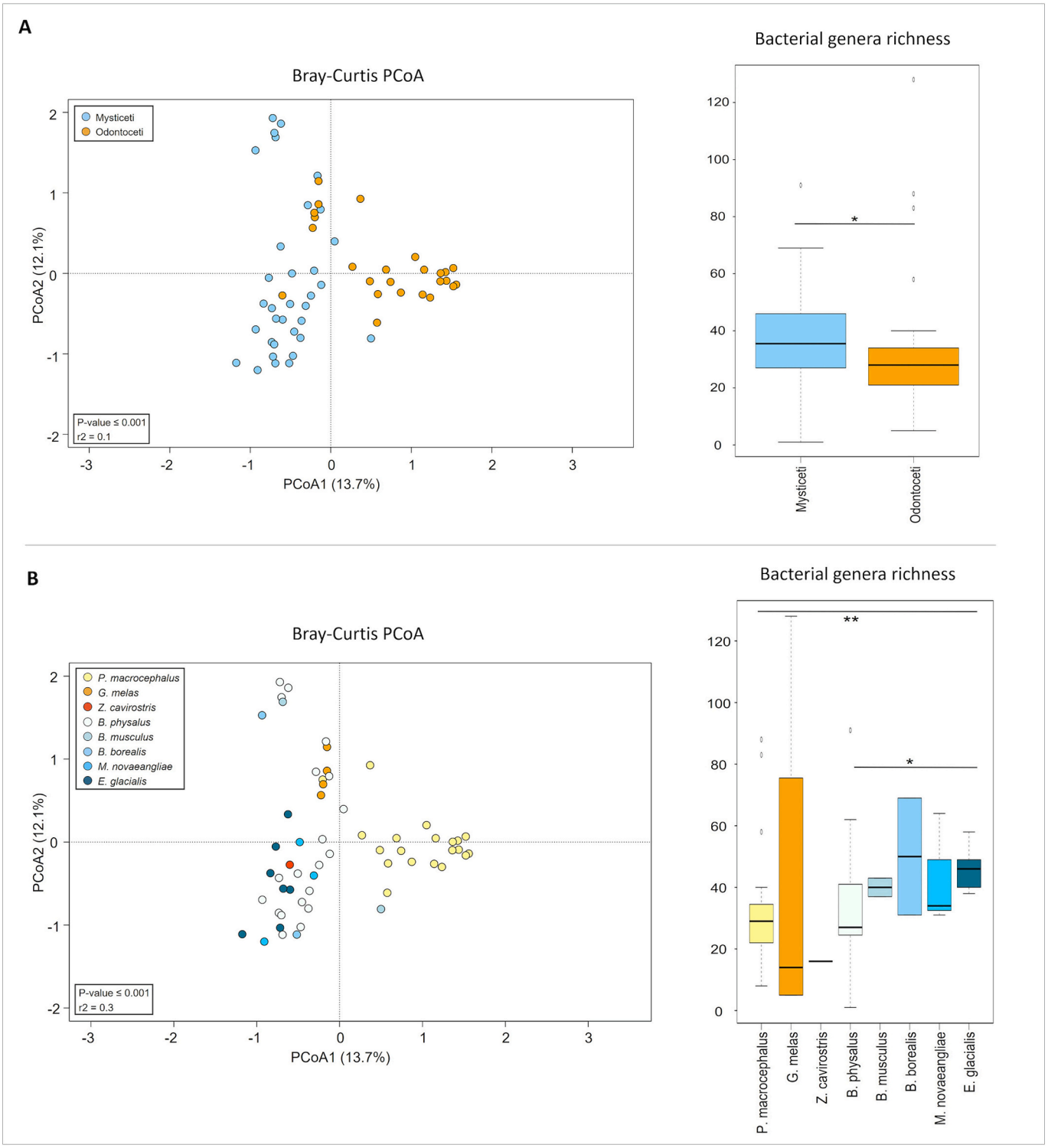


Fig. 2. – Microbiome community structure of odontocetes and mysticetes. (A) Representation of the entire set of samples (present study and publicly available) by Principal Coordinates Analysis (PCoA) based on Bray-Curtis distances of odontocetes (orange) and mysticetes (light blue) (permutation test with pseudo-F ratio, $p \leq 0.001$; left part of the image); the first and second principal components (PCoA1 and PCoA2) are plotted and the percentage of variance in the dataset explained by each axis is shown. On the right we show the corresponding boxplots of alpha diversity according to the bacterial genera richness in all samples (Wilcoxon rank-sum test, $*p \leq 0.05$). (B) Representation of the entire set of samples (present study and publicly available) by Principal Coordinates Analysis based on Bray-Curtis distances of the eight species considered in this article, namely *Physeter macrocephalus* (yellow), *Globicephala melas* (orange), *Ziphius cavirostris* (red), *Balaenoptera physalus* (white), *Balaenoptera musculus* (gray-blue), *Balaenoptera borealis* (light blue), *Megaptera novaeangliae* (blue) and *Eubalaena glacialis* (dark blue) (permutation test with pseudo-F ratio, $p \leq 0.001$; left part of the image); the first and second principal components (PCoA1 and PCoA2) are plotted and the percentage of variance in the dataset explained by each axis is shown. On the right we show the corresponding boxplots of alpha diversity according to the bacterial genera richness in all samples (Wilcoxon rank-sum test, $*p \leq 0.05$; $**p \leq 0.01$).

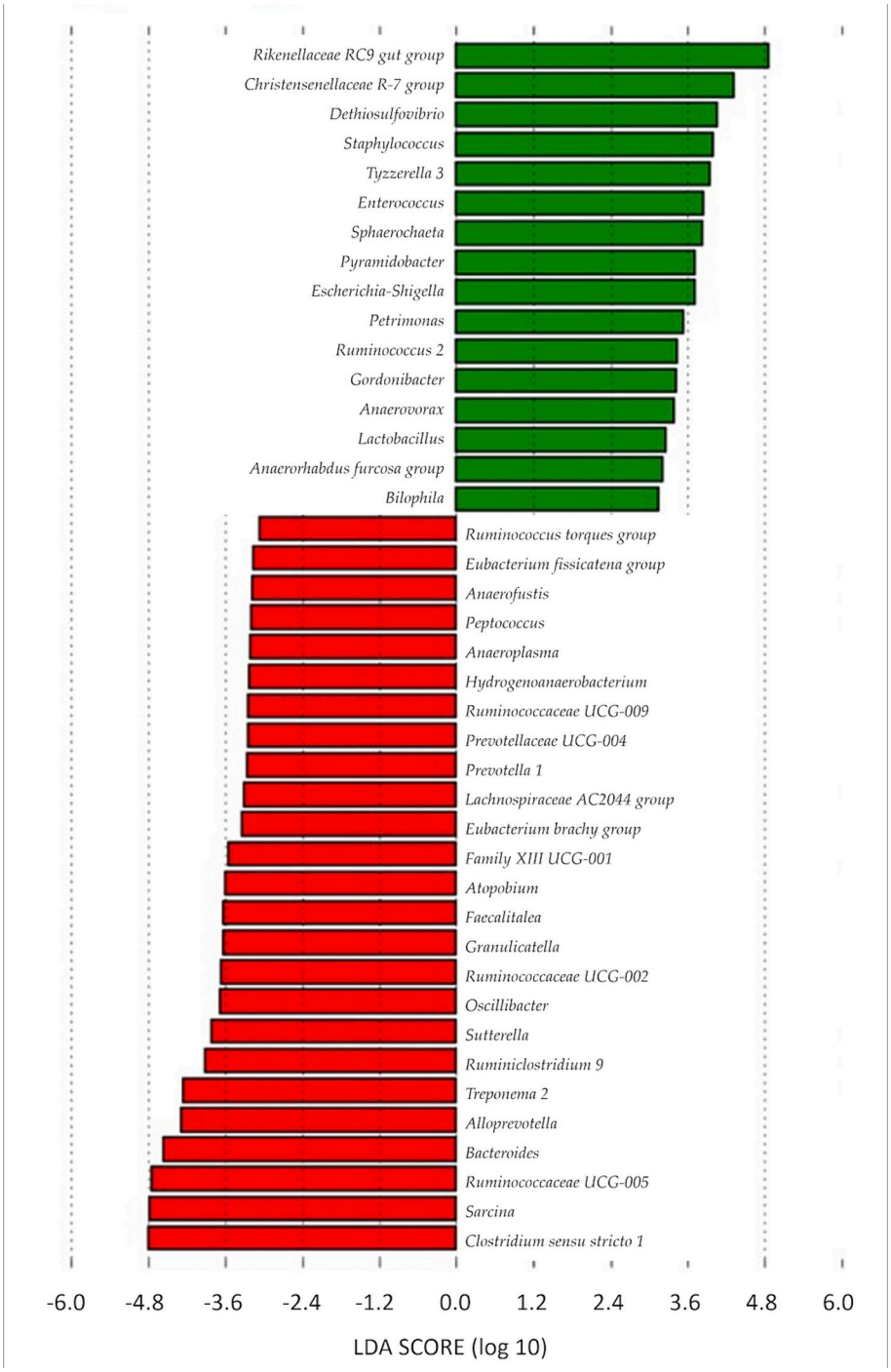


Fig. 3. – Discriminating genera between mysticetes and odontocetes. Linear discriminant analysis (LDA) scores of discriminating cetacean GM genera between the two cetacean parvorders, with mysticetes in red and odontocetes in green (the logarithmic threshold for discriminating features was set to 2.0 with $p \leq 0.05$). Plots were obtained by LDA effect size (LEfSe) analysis.

biomasses. As a parsimonious approach, we considered a set of host-specific ASVs as positive for a given CAZyme family only if least two ASVs were providing correspondent function. Consistent with this threshold, several CAZymes – such as pectin methylesterase, rhamnogalacturonan acetylesterase, UDP-3-O-acyl *N*-acetylglucosamine deacetylase, *N*-acetylglucosamine-6-phosphate deacetylase, β-glucosidase, mannase, cellulase, chitin disaccharide deacetylase (RE), and β-hexosaminidase – were represented in all the discriminant ASVs, irrespective of the host. On the contrary, a specific set of lignocellulose degrading enzymes, such as cephalosporin-C deacetylase, α-glucuronidase, *endo*-1,4-β-xylanase, and rhamnogalacturonan endolyase, were only shared by the ASVs characterizing hippos and mysticetes. The only enzyme exclusively represented in the ASVs characterizing mysticetes was chondroitin-sulfate ABC endolyase (AA), while the GH enzymes β-glucuronidase and mannosyl-glycoprotein endo-β-*N*-acetylglucosaminidase were specific of the ASVs discriminating odontocetes, being involved, respectively, in lignocellulose and chitin degradation. The only enzyme exclusively detected in ASVs characterizing both cetacean parvorders was eparin-sulfate lyase/heparin lyase III. Finally, the enzymes chitin deacetylase, involved in the degradation of chitin, and

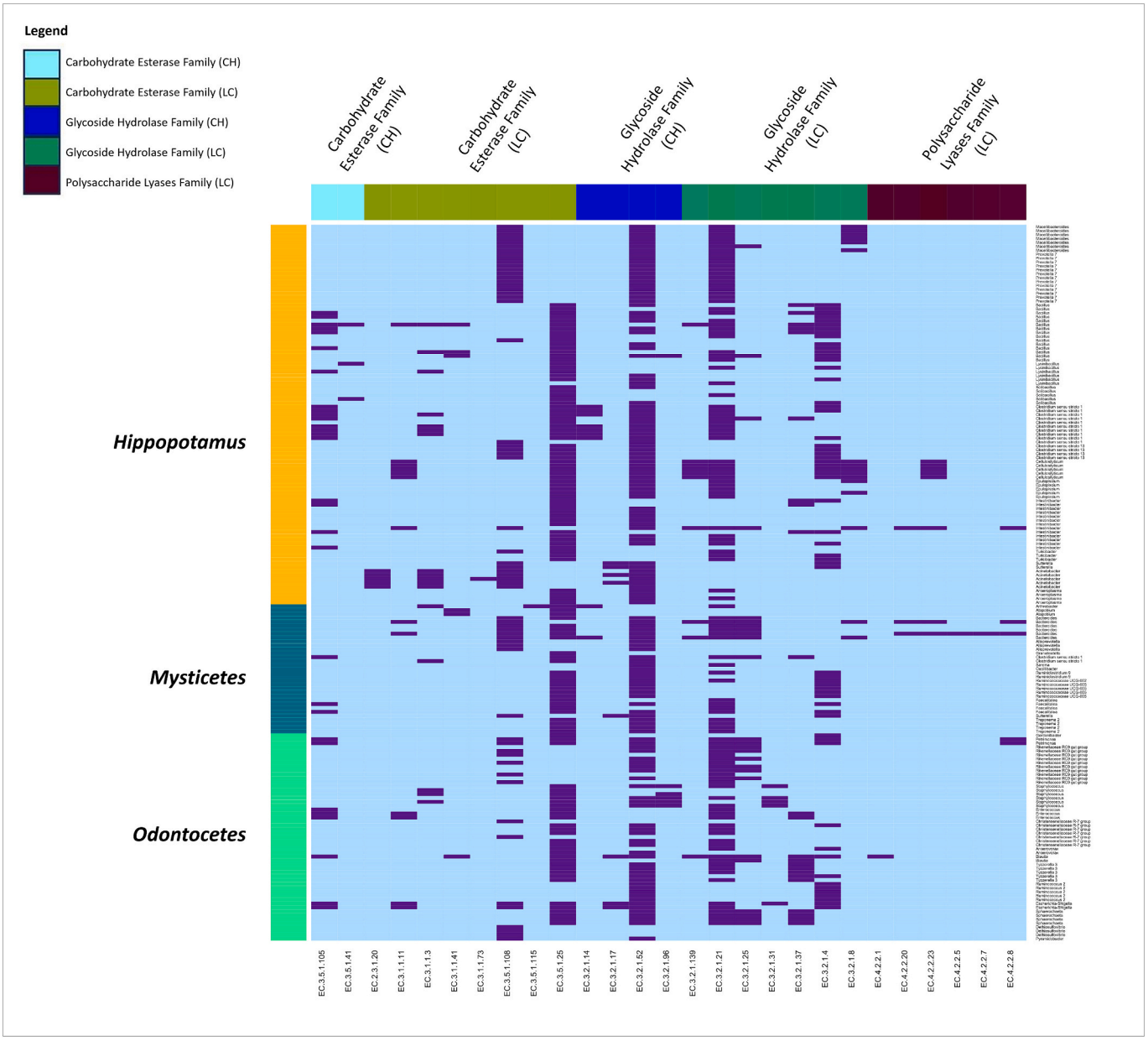


Fig. 4. – Different CAZymes layout between hippos, toothed, and baleen whales. Heatmap showing the presence/absence of different CAZymes families devoted to the deconstruction of complex biopolymers, namely lignocellulose (LC) and chitin (CH), in Mysticetes and Odontocetes from the present study (+ *Hippopotamus* from Sanders et al., 2015). Columns represent all CAZymes families grouped by the corresponding functional class (with EC number at the bottom and legend on the top-left), while discriminating ASVs taxonomy is listed in rows.

diacylglycerol O-acyltransferase, involved in the degradation of lignocellulose, were specific for the hippo-characterizing ASVs.

3.3. Phylosymbiotic trajectories in the cetacean GM

In order to emphasize possible congruence between microbiome variations and host phylogeny, we carried out a phylosymbiotic analysis. To this aim, a phylosymbiosis tree was constructed, specifically assessing the degree of congruence between the microbiome diversity and the hosts' phylogenetic relationships (Brooks et al., 2016; Lim and Bordenstein, 2020). For the phylosymbiosis assessment, four species of sharks were added as outgroup and the genus *Hippopotamus* was also considered for the construction of the tree, hippos being the closest living terrestrial relatives to cetaceans. The resulting phylosymbiotic tree (Fig. 5) showed different values of phylosymbiotic signals. The presence of a strong phylosymbiotic signal in almost all nodes of the tree, except for those marked with a black dot, is visible. Notably, the phylosymbiotic signal in our tree progressively increased for cetacean sub-clades, with most values of phylosymbiosis probability >90 % within the mysticete clade (Supplementary Table S5). As for cetaceans, the only weak signal was detected for the node that sees *Z. cavirostris* as sister taxon of all the other species, while the strongest phylosymbiotic signals in cetaceans can be found supporting all the sub-clades nested within the mysticetes. Furthermore, in Supplementary Table S5, bacterial taxa showing a high phylosymbiotic signal, with a maximum probability of phylosymbiosis higher than 50 % within the cetacean clade in our tree, are listed. Considering the common ancestor of all cetaceans analyzed in this article, the microbiome components showing the strongest phylosymbiotic signal were *Bacteroides* (56 %) and *Ruminiclostridium 9* (53 %).

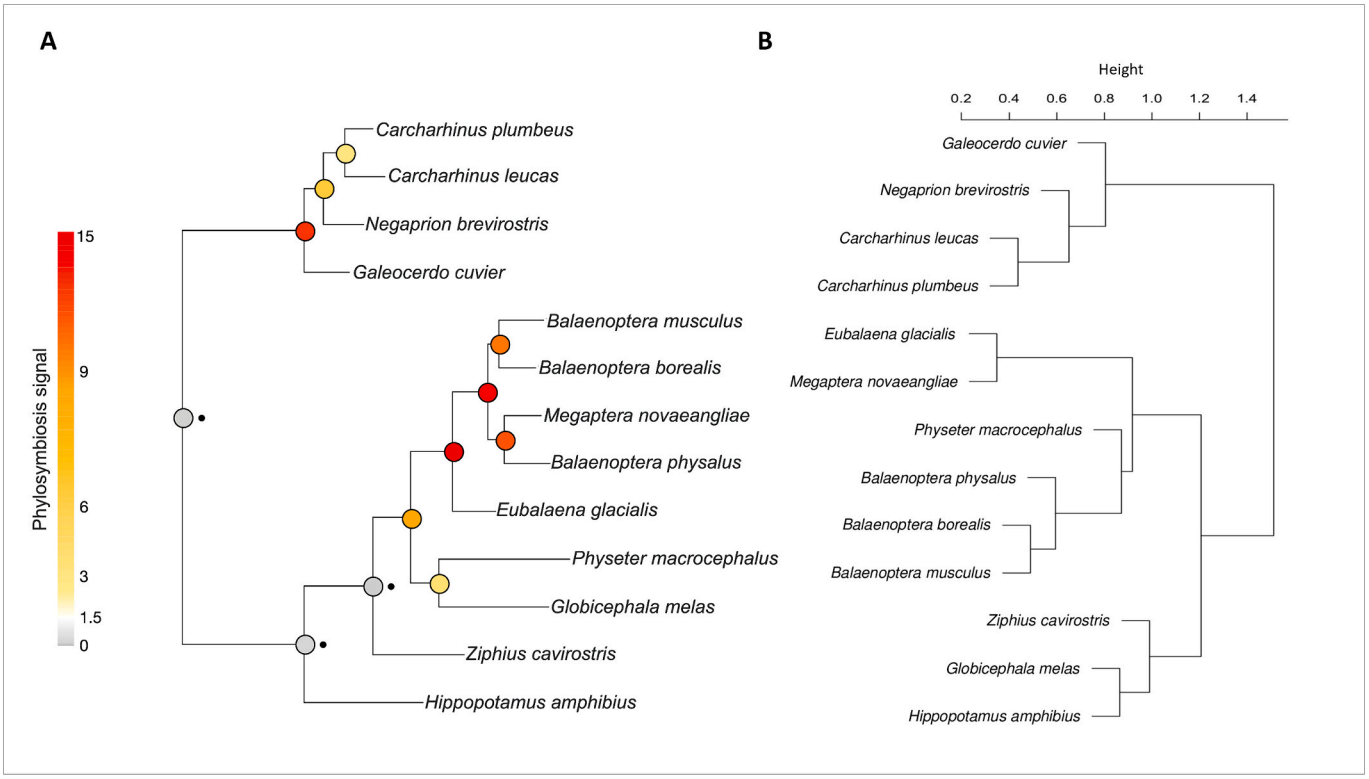


Fig. 5. – Phylosymbiosis tree showing evolutionary relationships among marine cetaceans and their gut microbial community. (A) The tree was created from the phylogenetic host tree, based on COI genes. (B) The microbial diversity tree was derived from the QIIME2 pipeline. Congruence between the trees was analyzed through the Mantel test with Pearson's correlation (*p*-value $\leq$0.05) and significant phylosymbiotic signal was calculated through the Count program. In panel A, the strength of the phylosymbiotic signal is highlighted by the color code on the left; black dots next to the tree nodes indicate the presence of weak phylosymbiosis.

3.4. Parasitological assessment

Next, faecal samples collected in the present study were subjected to molecular analyses for the detection of 3 enteric protozoan parasites, namely *Blastocystis* sp., *Giardia duodenalis* and *Cryptosporidium* spp. The cetacean species surveyed were *P. macrocephalus*, *G. melas*, and *Z. cavirostris* (Odontocetes) and *B. physalus* (Mysticetes).

Out of 30 individuals subjected to parasitological analysis, 20 (66.6 %) were found to be infected with one protozoan parasite (Table 2). Interestingly, none of the samples harbored more than one type of parasite. According to our data, the most widespread parasite was *Giardia duodenalis* (8 infected individuals), followed by *Blastocystis* sp. (7 infected individuals) and finally by *Cryptosporidium* spp. (5 infected individuals). Mysticetes, which in this case included only *B. physalus*, harbored one single parasite, i.e., *Giardia duodenalis* (5 infected individuals out of 9). On the contrary, Odontocetes seemed to be infected by all protozoans we searched for. Specifically, out of 16 individuals of *P. macrocephalus,* 11 were infected: 4 by *Blastocystis* sp., 4 by *Cryptosporidium* spp., and 3 by *Giardia duodenalis*; among *G. melas* samples, only one individual was parasite-free, with the others showing faecal presence of either *Blastocystis* sp. (2 individuals) or of *Cryptosporidium* spp. (1 individual); finally, *Z. cavirostris* was also infected by *Blastocystis*.

After molecular characterization of protozoan parasites sequences, *C. parvum* sequences were obtained for all the infected individuals.

As for microbial community structure in response to parasitosis, no significant differences were found ($p > 0.05$) when comparing infected and non-infected samples, both in terms of alpha diversity (Wilcoxon rank-sum test) and beta diversity (permutation test with pseudo-F ratio) (Fig. 6A; Fig. 6B and Supplementary Fig. 3). In order to deepen our understanding of potential protozoan parasite-driven GM dysbiosis, we also report bacterial families in cetacean GMs whose relative abundance significatively shifts when comparing infected versus non-infected cetaceans (Wilcoxon rank-sum test, $p \leq 0.05$), namely *Marinifilaceae*, *Clostridiaceae* and *Akkermansiaceae* (Fig. 6C). In particular, the *Marinifilaceae* family was the only bacterial taxon overrepresented in

Table 2
Summary of enteric parasites infection in samples of the present study.

		Mysticetes (n. 9)	Odontocetes (n. 21)		
		Balaenoptera physalus (n. 9)	*Physeter macrocephalus* (n. 16)	*Globicephala melas* (n. 4)	*Ziphius cavirostris* (n. 1)
Infected individuals based on parasite species	*Blastocystis* sp.	0	4	2	1
	Giardia duodenalis	5	3	0	0
	Cryptosporidium spp.	0	4	1	0
Overall infected samples	Infected samples according to the species	5 out of 9	11 out of 16	3 out of 4	1 out of 1
	Infected samples according to the parvorder	5 out of 9	15 out of 21		
	Infected samples in all cetaceans	20 out of 30			

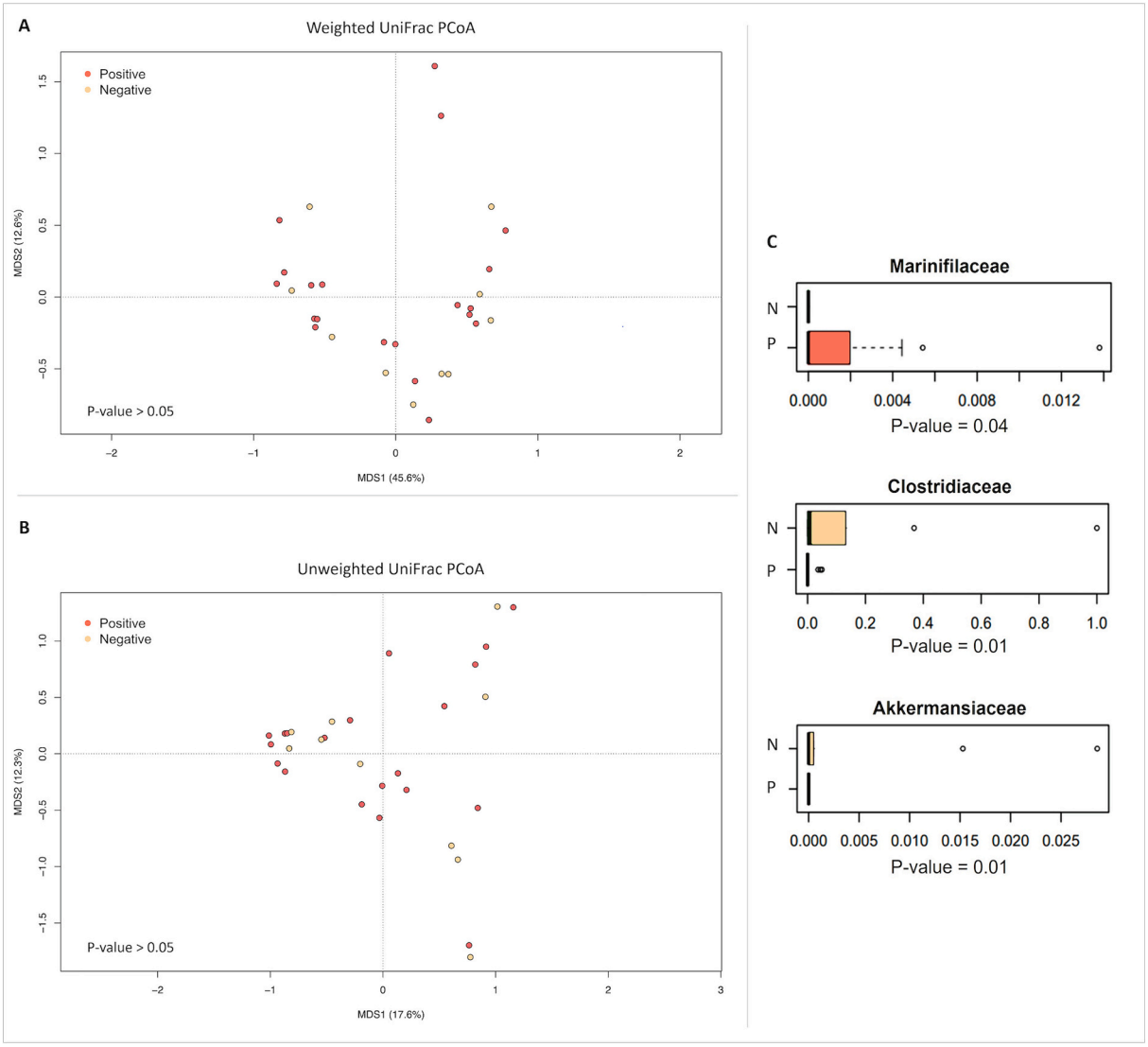


Fig. 6. – Cetacean GM structure and potential dysbiotic shifts following protozoan parasites infection. Principal Coordinates Analyses (PCoAs) based on weighted (A) and unweighted (B) UniFrac distances of the GM of infected (Positive; orange-red) vs non-infected (Negative; yellow) cetaceans from the present study (permutation test with pseudo-F ratio, $p > 0.05$). The first and second principal components (MDS1 and MDS2) are plotted and the percentage of variance in the dataset explained by each axis is shown. (C) Boxplots showing the distributions of bacterial gut families whose relative abundance varied significatively when comparing parasitized (P, positive) versus non-parasitized (N, negative) cetaceans. The central box represents the distance between the 25th and 75th percentiles. The median is marked with a black line. Whiskers identify the 10th and 90th percentiles. *P*-values are reported below each boxplot (Wilcoxon rank-sum test controlled for multiple testing using FDR).

infected individuals ($p = 0.04$), while *Clostridiaceae* and *Akkermansiaceae* were more abundant in uninfected samples ($p = 0.01$ for both families).

4. Discussion

This study provides some insights into the microbiome structure and biodiversity of eight cetacean species, with a focus on GM-host phylosymbiotic relationships, which are still poorly known, as well as on the GM CAZymes repertoire comparing both toothed and baleen whales. Furthermore, we report on the occurrence of 3 zoonotic enteric protozoan parasites in faecal samples from fin, sperm, pilot and Cuvier's beaked whales collected across the Pelagos Sanctuary (North-Western Mediterranean Sea).

We highlight a strong phylosymbiotic footprint for almost all nodes of the COI-based phylogeny of cetaceans, providing new clues on the co-evolutionary dynamics of the GM in both odontocetes and mysticetes. Since a sharp phylosymbiosis signal was clearly detected in cetaceans, especially within the mysticetes clade, we assume that host phylogeny and phylogenetic inertia have played a substantial role in the evolution of GM-host symbiosis across the biological history of Cetacea. These findings are in line with previous work that investigated how host phylogeny influences the skin microbiota of marine mammals, considering *Physeteridae*, *Delphinidae* (Odontocetes), *Balaenopteridae* (Mysticetes) and *Phocidae* (Pinnipedia) (Apprill et al., 2020). They describe that mysticetes show an overall stronger GM phylosymbiosis degree compared to odontocetes, suggesting the presence of more strict constraints in GM diversification within this parvorder. On the contrary, given difficulties in detecting phylosymbiosis at nodes characterized by greater phylogenetic distance (Lim and Bordenstein, 2020), we see weak signals in cetacean outgroups in our tree. However, the absence of a strong phylosymbiotic signal in the node representing the common ancestor of all cetacean species could be related to the use of only one sample of Cuvier's beaked whale in the present study.

Strengthening these observations, we found significant differences in cetacean GM structure when exploring the microbial communities in

terms of both alpha and beta diversity, as evidenced before for the cetacean skin, blow and faecal microbiomes (Apprill et al., 2020; Vendl et al., 2020; Marangi et al., 2021, respectively). According to our findings, these differences can be highlighted both for cetacean parvorders as well as for cetacean species. In particular, alpha diversity of bacterial genera shows significantly higher values for mysticetes compared to odontocetes. The bacterial genera characteristic of mysticetes, as elucidated by LEfSe, encompass taxa generally associated with high loads of complex carbohydrates in the diet. These include, for instance, *Ruminococcaceae*, *Lachnospiraceae*, *Eubacterium*, *Ruminiclostridium*, *Faecalibacterium* (Firmicutes), *Prevotella* and *Bacteroides* (Bacteroidetes), among others (Schwiertz et al., 2002; Lam et al., 2012; Precup and Vodnar, 2019; Houtman et al., 2022; Xie et al., 2022; Xu et al., 2022; Martín et al., 2023). The enrichment of these microorganisms in mysticetes could be related to the necessity for these animals to acquire a full set of GM functionalities for the digestion of the vast array of biopolymers which make up their broad diet, mainly consisting of chitin-rich Euphausiacea and Copepoda (Crustacea), various other planktonic organisms, but also shoaling fish which are engulfed while filter feeding (Kawamura, 1980). The increased relative abundance of Proteobacteria in odontocetes compared to mysticetes (especially in *G. melas* and *Z. cavirostris*), on the other hand, may reflect an evolutionary adaptation of odontocetes to active predatory behaviour and a diet based mostly on fish and cephalopods, such as squids (Soverini et al., 2016). Proteobacteria enriched in odontocetes GM include, for instance, the *Escherichia-Shigella* group. Besides, also within the phylum Firmicutes, odontocetes are characterized by some bacterial taxa previously reported in carnivorous animals (e.g., *Enterococcus* and *Lactobacillus*) (Soverini et al., 2016). To deepen our understanding of GM-host evolutionary dependencies and dynamics, we looked into the cetacean GM CAZymes asset, with a focus on both lignocellulose and chitin degradation, in both toothed and baleen whales. For sake of comparison, we included in this analysis also *Hippopotamus amphibius*. Indeed, lignocellulose and chitin represent the most abundant polymers on Earth (Raimundo et al., 2021) and are a part of the cetaceans' diet, as they are the basis of most planktonic organisms and as such widely distributed across all marine habitats (Agregán et al., 2022). Notably, we see that mysticetes, odontocetes and hippos share several GM lignocellulose and chitin degradation functions. As a matter of fact, despite some host-specific GM CAZymes which might be involved in fine-tuning digestive functions in relation to the diet of the animals, the chitin-rich nourishment of mysticetes and, to a minor extent, odontocetes, did not seem to strongly affect the GM CAZymes repertoire when compared to hippos. Moreover, the GM CAZymes similarities that we highlighted in mysticetes and odontocetes may lead to a reassessment of the role of chitin-containing organisms in odontocetes diet and of lignocellulose as a nourishment for both parvorders, suggesting a significant importance of these carbohydrates in odontocetes and mysticetes despite their different trophic niches. Shotgun metagenomics, in the future, will be pivotal to confirm these results more solidly.

Finally, the investigation on the occurrence of protozoan parasites *Blastocystis* sp., *Giardia duodenalis* and *C. parvum* in cetacean faeces from the Pelagos Sanctuary highlights a marked presence of these zoonotic protozoans across the sampled individuals (prevalence: 66.6 % of the total). Interestingly, all the infected individuals always show only one parasite in their stool, and this is in accordance with findings reported elsewhere within the *Delphinidae* family as well as in two different sea turtle species (Marangi et al., 2022). Moreover, we see that mysticetes, represented here by *B. physalus*, seem to be less susceptible to different protozoan infections than odontocetes, being positive only to *G. duodenalis*. Previously, other authors have investigated the presence of protozoan parasites, such as *G. duodenalis* and *C. parvum*, in cetacean specimens stranded and dead on the north-western coast of Spain (European Atlantic coast) (Reboredo-Fernández et al., 2015) and in common dolphins (*Delphinus delphis*) stranded along the Galician coast (Northwest Spain) (Reboredo-Fernández et al., 2014). Furthermore, the protozoan parasite *Blastocystis* sp. has been already detected in free-ranging fin and sperm whales (Marangi et al., 2021). Thus, this is the first study highlighting the occurrence of *Blastocystis* sp. in long-finned pilot whales and in the Cuvier's beaked whale. These results therefore extend the known host range of these zoonotic protozoan parasites.

Alpha and beta diversity analysis of the GM comparing infected and non-infected cetaceans underlines no strong signs of parasite-induced GM dysbiosis, except for three bacterial families whose members significantly differ in abundance between the two health states; these are *Marinifilaceae*, *Clostridiaceae* and *Akkermansiaceae*. The *Marinifilaceae*, found more abundant in infected individuals include anaerobic, sulfur-reducing taxa capable of both organic fermentation and anaerobic respiration (Li et al., 2022b). *Clostridiaceae* and *Akkermansiaceae*, on the contrary, are less represented in infected individuals. The *Clostridiaceae* family includes some well-known carbohydrates-fermenting and short chain fatty acids-producing microorganisms providing the host with pivotal nutritional and immunological substrates such as butyrate (Stoeva et al., 2021). The *Akkermansiaceae* are indeed mucin degraders contributing to the turnover of the mucus layers of the intestinal epithelium and are associated with metabolic health benefits for the host, including improved glucose metabolism and reduced inflammation (Walsh et al., 2014). Given the crucial role of these bacteria for the holobiont health, the effects of protozoan parasites on the GM of cetaceans deserve further investigation. Notably, the spread of these protozoans, possibly derived by human activities (Marangi et al., 2021; Maestrini et al., 2022; Moratal et al., 2022; Berrilli et al., 2023), across different cetacean species portrays a worrying picture of anthropic impact on marine mammals living in the Mediterranean Sea, even within the borders of a marine protected area. We thus state that monitoring the presence of these protozoans in cetaceans, as well as in other marine apex predators, may be a valuable proxy to assess the overall health of overexploited marine ecosystems, aiming to minimize, in a One Health perspective, the impact of anthropogenic activities and the spread of zoonotic microorganisms in the sea.

5. Conclusion

Our results clearly demonstrate that the cetacean GM, in terms of both taxonomic and functional characteristics, has been shaped over evolution by a vast array of drivers, including host phylogeny, environment as well as dietary habits, and this can be highlighted in the 2 parvorders that encompass all the present-day cetacean species, namely Odontocetes (toothed whales) and Mysticetes (baleen whales). As far as we can tell, this is the first work exploring the phylosymbiotic relationships of gut microbial communities within the Cetacea clade. To the best of our knowledge, this is also the first study analysing the cetacean GM structure in relation to the presence of zoonotic protozoan parasites and conducted on a large number of free-ranging individuals of 3 different cetacean species, hinting at cetaceans as proper bioindicators of overall ocean health in highly human-impacted marine habitats.

CRediT authorship contribution statement

Elena Radaelli: Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. **Giorgia Palladino:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. **Enrico Nanetti:** Writing – review & editing, Writing – original draft, Visualization, Data curation. **Daniel Scicchitano:** Writing – review & editing, Formal analysis, Data curation. **Simone Rampelli:** Writing – review & editing, Formal analysis, Data curation. **Sabina Airoldi:** Writing – review & editing, Visualization. **Marco Candela:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization. **Marianna Marangi:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

High-quality reads from all the present study samples were deposited in the European Nucleotide Archive (ENA) under the project accession number PRJEB73335.

Acknowledgment

We are grateful to all the researchers of Tethys Research Institute for their cooperation and collaboration. This work was supported by the "Controlling Microbiomes Circulations for Better Food Systems" (CIRCLES) project, which was funded by the European Union's Horizon 2020 research and innovation program under grant agreement no. 818290. The authors would like to thank Anna Selbmann, University of Iceland, for her important proofreading work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.172943.

References

Abdelrhman, K., Ciofini, A., Bacci, G., Mancusi, C., Mengoni, A., Ugolini, A., 2020. Exploring the resident gut microbiota of stranded odontocetes: high similarities between two dolphin species Tursiops truncatus and Stenella coeruleoalba. J. Mar. Biol. Assoc. U. K. 100 (7), 1181–1188. https://doi.org/10.1017/S0025315420000983.

Agregán, R., Lorenzo, J.M., Kumar, M., Shariati, M.A., Khan, M.U., Sarwar, A., Sultan, M., Rebezov, M., Usman, M., 2022. Anaerobic digestion of lignocellulose components: challenges and novel approaches. Energies 15 (22), 8413. https://doi.org/10.3390/en15228413.

Apprill, A., Miller, C.A., Van Cise, A.M., U'Ren, J.M., Leslie, M.S., Weber, L., Baird, R.W., Robbins, J., Landry, S., Bogomolni, A., Waring, G., 2020. Marine mammal skin microbiotas are influenced by host phylogeny. R. Soc. Open Sci. 7 (5), 192046 https://doi.org/10.1098/rsos.192046.

Bai, S., Zhang, P., Zhang, C., Du, J., Du, X., Zhu, C., Liu, J., Xie, P., Li, S., 2021. Comparative study of the gut microbiota among four different marine mammals in an aquarium. Front. Microbiol. 12, 769012 https://doi.org/10.3389/fmicb.2021.769012.

Bai, S., Zhang, P., Zhang, X., Yang, Z., Li, S., 2022. Gut microbial characterization of melon-headed whales (*Peponocephala electra*) stranded in China. Microorganisms 10 (3), 572. https://doi.org/10.3390/microorganisms10030572.

Berrilli, F., Papini, R.A., Barlaam, A., Normanno, G., Puccini, A., Procesi, I.G., Giangaspero, A., 2023. Final destination: The Mediterranean Sea, a vulnerable sea. The long journey of Giardia duodenalis cysts. Environmental Advances 13, 100414. https://doi.org/10.1016/j.envadv.2023.100414. ISSN 2666-7657.

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., Relman, D.A., 2016. Marine mammals harbor unique microbiotas shaped by and yet distinct from the sea. Nat. Commun. 7, 10516 https://doi.org/10.1038/ncomms10516.

Blanco, C., Salomón, O., Raga, J., 2001. Diet of the bottlenose dolphin (Tursiops truncatus) in the western Mediterranean Sea. J. Mar. Biol. Assoc. U. K. 81 (6), 1053–1058. https://doi.org/10.1017/S0025315401005057.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson 2nd, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37 (8), 852–857. https://doi.org/10.1038/s41587-019-0209-9. Erratum in: Nat Biotechnol. 2019 Sep;37(9):1091.

Brice, K.L., Trivedi, P., Jeffries, T.C., Blyton, M.D.J., Mitchell, C., Singh, B.K., Moore, B. D., 2019. The Koala (*Phascolarctos cinereus*) faecal microbiome differs with diet in a wild population. PeerJ 7, e6534. https://doi.org/10.7717/peerj.6534.

Brooks, A.W., Kohl, K.D., Brucker, R.M., van Opstal, E.J., Bordenstein, S.R., 2016. Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. PLoS Biol. Nov 18;14 (11), e2000225 https://doi.org/10.1371/journal.pbio.2000225.

Brown, K., Krumbeck, J., Herrick, K., Schmitt, T., 2022. Faecal microbiota analysis and transplantation in four oceanaria-based short-finned pilot whales Globicephala macrorhynchus. Journal of Zoo and Aquarium Research 10 (4), 194–201. Retrieved from. https://www.jzar.org/jzar/article/view/692.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., Holmes, S.P., 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nat. Methods 13 (7), 581–583. https://doi.org/10.1038/nmeth.3869.

Csűrös, M., 2010. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. Bioinformatics 26, 1910–1912. https://doi.org/10.1093/bioinformatics/btq315.

Culhane, A.C., Thioulouse, J., Perrière, G., Higgins, D.G., 2005. MADE4: an R package for multivariate analysis of gene expression data. Bioinformatics 21 (11), 2789–2790. https://doi.org/10.1093/bioinformatics/bti394.

Douglas, G.M., Maffei, V.J., Zaneveld, J.R., Yurgel, S.N., Brown, J.R., Taylor, C.M., Huttenhower, C., Langille, M.G.I., 2020. PICRUSt2 for prediction of metagenome functions. Nat. Biotechnol. 38 (6), 685–688. https://doi.org/10.1038/s41587-020-0548-6.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32 (5), 1792–1797. https://doi.org/10.1093/nar/gkh340.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26 (19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461.

Erwin, P.M., Rhodes, R.G., Kiser, K.B., Keenan-Bateman, T.F., McLellan, W.A., Pabst, D. A., 2017. High diversity and unique composition of gut microbiomes in pygmy (Kogia breviceps) and dwarf (K. Sima) sperm whales. Sci. Rep. Aug 3;7 (1), 7205. https://doi.org/10.1038/s41598-017-07425-z.

Filippini, R., Lardon, S., Bonari, E., Marraccini, E., 2018. Unraveling the contribution of periurban farming systems to urban food security in developed countries. Agron. Sustain. Dev. 21–38. https://doi.org/10.1007/s13593-018-0499-1.

Foote, A.D., Liu, Y., Thomas, G.W., Vinař, T., Alföldi, J., Deng, J., Dugan, S., van Elk, C. E., Hunter, M.E., Joshi, V., Khan, Z., Kovar, C., Lee, S.L., Lindblad-Toh, K., Mancia, A., Nielsen, R., Qin, X., Qu, J., Raney, B.J., Vijay, N., Wolf, J.B., Hahn, M.W., Muzny, D.M., Worley, K.C., Gilbert, M.T., Gibbs, R.A., 2015. Convergent evolution of the genomes of marine mammals. Nat. Genet. 47 (3), 272–275. https://doi.org/10.1038/ng.3198.

Gatesy, J., Hayashi, C., Cronin, M.A., Arctander, P., 1996. Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. Mol. Biol. Evol. 13 (7), 954–963. https://doi.org/10.1093/oxfordjournals.molbev.a025663.

Gatesy, J., Geisler, J.H., Chang, J., et al., 2013. A phylogenetic blueprint for a modern whale. Mol. Phylogenet. Evol. 66 (2), 479–506. https://doi.org/10.1016/j.ympev.2012.10.012.

Giangaspero, A., Papini, R., Marangi, M., Koehler, A.V., Gasser, R.B., 2014. *Cryptosporidium parvum* genotype IIa and *Giardia duodenalis* assemblage a in *Mytilus galloprovincialis* on sale at local food markets. Int. J. Food Microbiol. 3, 62–67. https://doi.org/10.1016/j.ijfoodmicro.2013.11.022.

Glaeser, S.P., Silva, L.M.R., Prieto, R., Silva, M.A., Franco, A., Kämpfer, P., Hermosilla, C., Taubert, A., Eisenberg, T., et al., 2022. A preliminary comparison on Faecal microbiomes of free-ranging large baleen (Balaenoptera musculus, B. Physalus, B. Borealis) and toothed (Physeter macrocephalus) whales. Microb. Ecol. 83 (1), 18–33. https://doi.org/10.1007/s00248-021-01729-4.

Godoy-Vitorino, F., Rodriguez-Hilario, A., Alves, A.L., Gonçalves, F., Cabrera-Colon, B., Mesquita, C.S., Soares-Castro, P., Ferreira, M., Marçalo, A., Vingada, J., Eira, C., Santos, P.M., 2017. The microbiome of a striped dolphin (Stenella coeruleoalba) stranded in Portugal [published correction appears in res Microbiol. 2021 Jan-Feb; 172(1):103564]. Res. Microbiol. 168 (1), 85–93. https://doi.org/10.1016/j.resmic.2016.08.004.

Graur, D., Higgins, D.G., 1994. Molecular evidence for the inclusion of cetaceans within the order Artiodactyla. Mol. Biol. Evol. 11 (3), 357–364. https://doi.org/10.1093/oxfordjournals.molbev.a040118.

Guardone, L., Susini, F., Castiglione, D., Ricci, E., Corradini, C., Guidi, A., Armani, A., 2020. Ascaridoid nematode larvae in wild gilthead seabream (Sparus aurata) and European seabass (Dicentrarchus labrax) caught in the Tyrrhenian Sea (Western Mediterranean Sea): a contribute towards the parasitological risk assessment on two commercially important fish species. Food Control 118 (107377). https://doi.org/10.1016/j.foodcont.2020.107377.

Herwig, R.P., Staley, J.T., Nerini, M.K., Braham, H.W., 1984. Baleen whales: preliminary evidence for forestomach microbial fermentation. Appl. Environ. Microbiol. Feb;47 (2), 421–423. https://doi.org/10.1128/aem.47.2.421-423.1984.

Houtman, T.A., Eckermann, H.A., Smidt, H., de Weerth, C., 2022. Gut microbiota and BMI throughout childhood: the role of firmicutes, bacteroidetes, and short-chain fatty acid producers. Sci. Rep. Feb 24;12 (1), 3140. https://doi.org/10.1038/s41598-022-07176-6.

de Jonge, N., Carlsen, B., Christensen, M.H., Pertoldi, C., Nielsen, J.L., 2022. The gut microbiome of 54 mammalian species. Front. Microbiol. Jun 16 (13), 886252 https://doi.org/10.3389/fmicb.2022.886252.

Kawamura, A.K.I.T.O., 1980. A review of food of balaenopterid whales. Sci. Rep. Whales Res. Inst 32, 155–197. Retrieved from https://api.semanticscholar.org/CorpusID:5027512.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., Glöckner, F.O., 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 41 (1), e1. https://doi.org/10.1093/nar/gks808.

Lam, Y.Y., Ha, C.W., Campbell, C.R., Mitchell, A.J., Dinudom, A., Oscarsson, J., Cook, D. I., Hunt, N.H., Caterson, I.D., Holmes, A.J., Storlien, L.H., 2012. Increased gut permeability and microbiota change associate with mesenteric fat inflammation and metabolic dysfunction in diet-induced obese mice. PLoS One 7 (3), e34233. https://doi.org/10.1371/journal.pone.0034233.

Langer., 2001. Evidence from the digestive tract on phylogenetic relationships in ungulates and whales. J. Zool. Syst. Evol. Res. 39, 77–90. https://doi.org/10.1046/j.1439-0469.2001.00158.x.

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., Gordon, J.I., 2008. Evolution of mammals and their gut microbes [published correction appears in science. 2008 Nov 21;322(5905):1188]. Science 320 (5883), 1647–1651. https://doi.org/10.1126/science.1155725.

Li, C., Xie, H., Sun, Y., et al., 2022a. Insights on gut and skin wound microbiome in stranded indo-Pacific finless porpoise (*Neophocaena phocaenoides*). Microorganisms 10 (7), 1295. Published 2022 Jun 27. https://doi.org/10.3390/microorganisms10071295.

Li, J., Dong, C., Lai, Q., Wang, G., Shao, Z., 2022b. Frequent occurrence and metabolic versatility of *Marinifilaceae* bacteria as key players in organic matter mineralization in global deep seas. mSystems Dec 20;7 (6), e0086422. https://doi.org/10.1128/msystems.00864-22.

Ligges, U., Mächler, M., 2002. Scatterplot3d-an r package for visualizing multivariate data. No. 2002, 22. https://doi.org/10.18637/jss.v008.i11. Technical report.

Lim, S.J., Bordenstein, S.R., 2020. An introduction to phylosymbiosis. Proc. R. Soc. B 287, 20192900. https://doi.org/10.1098/rspb.2019.2900.

Maestrini, M., Berrilli, F., Di Rosso, A., Coppola, F., Guadano Procesi, I., Mariacher, A., Felicioli, A., Perrucci, S., 2022. Zoonotic Giardia duodenalis genotypes and other gastrointestinal parasites in a badger population living in an Anthropized area of Central Italy. Pathogens Aug 11;11 (8), 906. https://doi.org/10.3390/pathogens11080906.

Mallick, H., Rahnavard, A., McIver, L.J., Ma, S., Zhang, Y., Nguyen, L.H., Tickle, T.L., Weingart, G., Ren, B., Schwager, E.H., Chatterjee, S., Thompson, K.N., Wilkinson, J. E., Subramanian, A., Lu, Y., Waldron, L., Paulson, J.N., Franzosa, E.A., Bravo, H.C., Huttenhower, C., 2021. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput. Biol. 17 (11), e1009442 https://doi.org/10.1371/journal.pcbi.1009442.

Marangi, M., Airoldi, S., Beneduce, L., Zaccone, C., 2021. Wild whale faecal samples as a proxy of anthropogenic impact. Sci. Report. 11, 5822. https://doi.org/10.1038/s41598-021-84966-4.

Marangi, M., Carlucci, R., Carlino, P., Fanizza, C., Cirelli, G., Maglietta, R., Beneduce, L., 2022. Dolphins and sea turtles may host zoonotic parasites and pathogenic bacteria as indicators of anthropic pressure in the Gulf of Taranto (northern Ionian Sea, central-eastern Mediterranean Sea). Vet. Res. Commun. 46, 1157–1166. https://doi.org/10.1007/s11259-022-10011-y.

Martín, R., Rios-Covian, D., Huillet, E., Auger, S., Khazaal, S., Bermúdez-Humarán, L.G., Sokol, H., Chatel, J.M., Langella, P., 2023. Faecalibacterium: a bacterial genus with promising human health applications. FEMS Microbiol. Rev. Jul 5;47 (4), fuad039 https://doi.org/10.1093/femsre/fuad039.

Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G., Neufeld, J.D., 2012. PANDAseq: paired-end assembler for illumina sequences. BMC bioinformatics 13, 1–7. https://doi.org/10.1186/1471-2105-13-31.

McFall-Ngai, M., Hadfield, M.G., Bosch, T.C., Carey, H.V., Domazet-Lošo, T., Douglas, A. E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S.F., Hentschel, U., King, N., Kjelleberg, S., Knoll, A.H., Kremer, N., Mazmanian, S.K., Metcalf, J.L., Nealson, K., Pierce, N.E., Rawls, J.F., Reid, A., Ruby, E.G., Rumpho, M., Sanders, J.G., Tautz, D., Wernegreen, J.J., 2013. Animals in a bacterial world, a new imperative for the life sciences. Proc. Natl. Acad. Sci. USA 110, 3229–3236. https://doi.org/10.1073/pnas.1218525110.

McKenzie, V.J., Song, S.J., Delsuc, F., Prest, T.L., Oliverio, A.M., Korpita, T.M., Alexiev, A., Amato, K.R., Metcalf, J.L., Kowalewski, M., Avenant, N.L., Link, A., Di Fiore, A., Seguin-Orlando, A., Feh, C., Orlando, L., Mendelson, J.R., Sanders, J., Knight, R., 2017. The effects of captivity on the mammalian gut microbiome. Integr. Comp. Biol. 57 (4), 690–704. https://doi.org/10.1093/icb/icx090.

Moeller, A.H., Caro-Quintero, A., Mjungu, D., Georgiev, A.V., Lonsdorf, E.V., Muller, M. N., Pusey, A.E., Peeters, M., Hahn, B.H., Ochman, H., 2016. Cospeciation of gut microbiota with hominids. Science 353, 380–382. https://doi.org/10.1126/science.aaf3951.

Moratal, S., Dea-Ayuela, M.A., Martí-Marco, A., Puigcercós, S., Marco-Hirs, N.M., Doménech, C., Corcuera, E., Cardells, J., Lizana, V., López-Ramon, J., 2022. Molecular characterization of cryptosporidium spp. in cultivated and wild marine fishes from Western Mediterranean with the first detection of zoonotic cryptosporidium ubiquitum. Animals (Basel) Apr 19;12 (9), 1052. https://doi.org/10.3390/ani12091052.

Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., González, A., Fontana, L., Henrissat, B., Knight, R., Gordon, J.I., 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science May 20;332 (6032), 970–974. https://doi.org/10.1126/science.1198719.

NIH HMP Working Group, Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J.A., Bonazzi, V., McEwen, J.E., Wetterstrand, K.A., Deal, C., Baker, C.C., Di Francesco, V., Howcroft, T.K., Karp, R.W., Lunsford, R.D., Wellington, C.R., Belachew, T., Wright, M., Giblin, C., David, H., Mills, M., Salomon, R., Mullins, C., Akolkar, B., Begg, L., Davis, C., Grandison, L., Humble, M., Khalsa, J., Little, A.R., Peavy, H., Pontzer, C., Portnoy, M., Sayre, M.H., Starke-Reed, P., Zakhari, S., Read, J., Watson, B., Guyer, M., 2009. The NIH Human Microbiome Project. Genome Res. Dec;19 (12), 2317–2323. https://doi.org/10.1101/gr.096651.109.

Nikaido, M., Rooney, A.P., Okada, N., 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interpersed elements: hippopotamuses are the closest extant relatives of whales. Proc. Natl. Acad. Sci. USA Aug 31;96 (18), 10261–10266. https://doi.org/10.1073/pnas.96.18.10261.

Notarbartolo-di-Sciara, G., Agardy, T., Hyrenbach, D., Scovazzi, T., Van Klaveren, P., 2008. The Pelagos sanctuary for Mediterranean marine mammals. Aquat. Conserv. Mar. Freshwat. Ecosyst. 18 (4), 367–391. https://doi.org/10.1002/aqc.855.

Oksanen, J., 2013. Vegan: ecological diversity. R project 368, 1–11. Retrieved from https://api.semanticscholar.org/CorpusID:8975795.

Perofsky, A.C., Lewis, R.J., Meyers, L.A., 2019. Terrestriality and bacterial transfer: a comparative study of gut microbiomes in sympatric Malagasy mammals. ISME J. 13 (1), 50–63. https://doi.org/10.1038/s41396-018-0251-5.

Precup, G., Vodnar, D.C., 2019. Gut Prevotella as a possible biomarker of diet and its eubiotic versus dysbiotic roles: a comprehensive literature review. Br. J. Nutr. Jul 28;122 (2), 131–140. https://doi.org/10.1017/S0007114519000680.

Price, M.N., Dehal, P.S., 2010. Arkin AP FastTree 2 –approximately maximum-likelihood trees for large alignments. PLoS One 5 (3), e9490. https://doi.org/10.1371/journal.pone.0009490.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41 (D1), D590–D596. https://doi.org/10.1093/nar/gks1219.

Raimundo, I., Silva, R., Meunier, L., Valente, S.M., Lago-Lestón, A., Keller-Costa, T., Costa, R., 2021. Functional metagenomics reveals differential chitin degradation and utilization features across free-living and host-associated marine microbiomes. Microbiome 9, 1–18. https://doi.org/10.1186/s40168-020-00970-2.

Rambaut, A., 2018. FigTree: tree figure drawing tool (Version 1.4.4). Retrieved from. http://tree.bio.ed.ac.uk/software/figtree.

Reboredo-Fernández, A., Gómez-Couso, H., Martínez-Cedeira, J.A., Cacciò, S.M., Ares-Mazás, E., 2014. Detection and molecular characterization of Giardia and Cryptosporidium in common dolphins (Delphinus delphis) stranded along the Galician coast (Northwest Spain). Vet. Parasitol. 202 (3–4), 132–137. https://doi.org/10.1016/j.vetpar.2014.03.018.

Reboredo-Fernández, A., Ares-Mazás, E., Martínez-Cedeira, J.A., Romero-Suances, R., Cacciò, S.M., Gómez-Couso, H., 2015. Giardia and Cryptosporidium in cetaceans on the European Atlantic coast. Parasitol. Res. 114 (2), 693–698. https://doi.org/10.1007/s00436-014-4235-8.

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584. https://doi.org/10.7717/peerj.2584.

Rohart, F., Gautier, B., Singh, A., Lê Cao, K.A., 2017. mixOmics: an R package for 'omics feature selection and multiple data integration. PLoS Comput. Biol. 13 (11), e1005752 https://doi.org/10.1371/journal.pcbi.1005752.

Roman, J., McCarthy, J.J., 2010. The whale pump: marine mammals enhance primary productivity in a coastal basin. PLoS One 5 (10), e13255. https://doi.org/10.1371/journal.pone.0013255.

Sanders, J.G., Beichman, A.C., Roman, J., Scott, J.J., Emerson, D., McCarthy, J.J., Girguis, P.R., 2015. Baleen whales host a unique gut microbiome with similarities to both carnivores and herbivores. Nat. Commun. 6 (1), 8285. https://doi.org/10.1038/ncomms9285.

Schwiertz, A., Lehmann, U., Jacobasch, G., Blaut, M., 2002. Influence of resistant starch on the SCFA production and cell counts of butyrate-producing Eubacterium spp. in the human intestine. J. Appl. Microbiol. 93 (1), 157–162. https://doi.org/10.1046/j.1365-2672.2002.01679.x.

Scicluna, S.M., Tawari, B., Clark, C.G., 2006. DNA barcoding of blastocystis. Protist 157, 77–85. https://doi.org/10.1016/j.protis.2005.12.001.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., Huttenhower, C., 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12, 1–18. https://doi.org/10.1186/gb-2011-12-6-r60.

Sherrill-Mix, S., McCormick, K., Lauder, A., Bailey, A., Zimmerman, L., Li, Y., Django, J. N., Bertolani, P., Colin, C., Hart, J.A., Hart, T.B., Georgiev, A.V., Sanz, C.M., Morgan, D.B., Atencia, R., Cox, D., Muller, M.N., Sommer, V., Piel, A.K., Stewart, F. A., Speede, S., Roman, J., Wu, G., Taylor, J., Bohm, R., Rose, H.M., Carlson, J., Mjungu, D., Schmidt, P., Gaughan, C., Bushman, J.I., Schmidt, E., Bittinger, K., Collman, R.G., Hahn, B.H., Bushman, F.D., 2018. Allometry and Ecology of the Bilaterian Gut Microbiome. mBio Mar 27;9 (2), e00319-18. https://doi.org/10.1128/mBio.00319-18.

Shiffman, M.E., Soo, R.M., Dennis, P.G., Morrison, M., Tyson, G.W., Hugenholtz, P., 2017. Gene and genome-centric analyses of koala and wombat fecal microbiomes point to metabolic specialization for *Eucalyptus* digestion. PeerJ 5, e4075. https://doi.org/10.7717/peerj.4075.

Song, S.J., Sanders, J.G., Delsuc, F., Metcalf, J., Amato, K., Taylor, M.W., Mazel, F., Lutz, H.L., Winker, K., Graves, G.R., Humphrey, G., Gilbert, J.A., Hackett, S.J., White, K.P., Skeen, H.R., Kurtis, S.M., Withrow, J., Braile, T., Miller, M., McCracken, K.G., Maley, J.M., Ezenwa, V.O., Williams, A., Blanton, J.M., McKenzie, V.J., Knight, R., 2020. Comparative analyses of vertebrate gut

microbiomes reveal convergence between birds and bats. mBio 11 (1), e02901–e02919. https://doi.org/10.1128/mBio.02901-19.
Soverini, M., Quercia, S., Biancani, B., Furlati, S., Turroni, S., Biagi, E., Consolandi, C., Peano, C., Severgnini, M., Rampelli, S., Brigidi, P., Candela, M., 2016. The bottlenose dolphin (Tursiops truncatus) faecal microbiota. FEMS Microbiol. Ecol. 92 (4), fiw055 https://doi.org/10.1093/femsec/fiw055.
Stoeva, M.K., Garcia-So, J., Justice, N., Myers, J., Tyagi, S., Nemchek, M., McMurdie, P. J., Kolterman, O., Eid, J., 2021. Butyrate-producing human gut symbiont, *Clostridium butyricum*, and its role in health and disease. Gut Microbes Jan-Dec;13 (1), 1–28. https://doi.org/10.1080/19490976.2021.1907272.
Suzuki, A., Segawa, T., Sawa, S., Nishitani, C., Ueda, K., Itou, T., Asahina, K., Suzuki, M., 2019. Comparison of the gut microbiota of captive common bottlenose dolphins Tursiops truncatus in three aquaria. J. Appl. Microbiol. 126 (1), 31–39. https://doi.org/10.1111/jam.14109.
Thewissen, J.G.M., Sensor, J.D., Clementz, M.T., Bajpai, S., 2011. Evolution of dental wear and diet during the origin of whales. Paleobiology 37, 655–669. https://doi.org/10.1666/10038.1.
Tian, J., Du, J., Lu, Z., Han, J., Wang, Z., Li, D., Guan, X., Wang, Z., 2020. Distribution of microbiota across different intestinal tract segments of a stranded dwarf minke whale, *Balaenoptera acutorostrata*. Microbiologyopen 9 (10), e1108. https://doi.org/10.1002/mbo3.1108.
Turroni, S., FIori, J., Rampelli, S., Schnorr, S.L., Consolandi, C., Barone, M., Biagi, E., Fanelli, F., Mezzullo, M., Crittenden, A.N., Henry, A.G., Brigidi, P., Candela, M., 2016. Fecal metabolome of the Hadza hunter-gatherers: a host-microbiome integrative view. Sci. Rep. 6 (1), 1–9. https://doi.org/10.1038/srep32826.
Vendl, C., Slavich, E., Nelson, T., Acevedo-Whitehouse, K., Montgomery, K., Ferrari, B., Thomas, T., Rogers, T., 2020. Does sociality drive diversity and composition of airway microbiota in cetaceans? Environ. Microbiol. Rep. 12 (3), 324–333. https://doi.org/10.1111/1758-2229.12835.
Vietti, R.C., Albertelli, G., Aliani, S., Bava, S., Bavestrello, G., Cecchi, L.B., Bianchi, C.N., Bozzo, E., Capello, M., Castellano, M., Cerrano, C., Chiantore, M., Corradi, N., Cocito, S., Cutroneo, L., Diviacco, G., Fabiano, M., Faimali, M., Ferrari, M., Gasparini, G.P., Locritani, M., Mangialajo, L., Marin, V., Moreno, M., Morri, C., Orsi, Relini L., Pane, L., Paoli, C., Petrillo, M., Povero, P., Pronzato, R., Relini, G., Santangelo, G., Tucci, S., Tunesi, L., Vacchi, M., Vassallo, P., Vezzulli, L., Wurtz, M., 2010. The Ligurian Sea: present status, problems and perspectives. Chem. Ecol. 26 (sup1), 319–340. https://doi.org/10.1080/02757541003689845.
Walsh, C.J., Guinane, C.M., O'Toole, P.W., Cotter, P.D., 2014. Beneficial modulation of the gut microbiota. FEBS Lett. Nov 17;588 (22), 4120–4130. https://doi.org/10.1016/j.febslet.2014.03.035.
Wang, X., Wu, X., Shang, Y., Mei, X., Zhou, S., Wei, Q., Sun, G., Dong, Y., Zhang, H., 2022. Convergent evolution of the gut microbiome in marine carnivores. Ecol. Evol. Oct 1;12 (10), e9373 https://doi.org/10.1002/ece3.9373.
Wickham, H., 2007. Meifly: Models explored interactively. In: Website ASA Sections on Statistical Computing and Graphics (Student Paper Award Winner 2007). Retrieved from https://api.semanticscholar.org/CorpusID:828411.
Wickham, H., Chang, W., Wickham, M.H., 2016. Package 'ggplot2'. Create Elegant Data Visualisations Using the Grammar of Graphics. Version 2 (1), 1–189. https://api.semanticscholar.org/CorpusID:225733766.
Würsig, B., 2020. Marine Mammals and Aquaculture, with Special Emphasis on Open Ocean Waters of New Zealand; Independent consultant. Draft report prepared for Fisheries New Zealand. https://www.mpi.govt.nz/dmsdocument/50254/direct.
Xiao, L., Feng, Y., 2017. Molecular epidemiologic tools for waterborne pathogens Cryptosporidium spp. and Giardia duodenalis. Food Waterborne Parasitol. 8 (9), 14–32. https://doi.org/10.1016/j.fawpar.2017.09.002.
Xie, J., Li, L.F., Dai, T.Y., Qi, X., Wang, Y., Zheng, T.Z., Gao, X.Y., Zhang, Y.J., Ai, Y., Ma, L., Chang, S.L., Luo, F.X., Tian, Y., Sheng, J., 2022. Short-chain fatty acids produced by Ruminococcaceae mediate α-linolenic acid promote intestinal stem cells proliferation. Mol. Nutr. Food Res. Jan;66 (1), e2100408 https://doi.org/10.1002/mnfr.202100408.
Xu, A.A., Kennedy, L.K., Hoffman, K., White, D.L., Kanwal, F., El-Serag, H.B., Petrosino, J.F., Jiao, L., 2022. Dietary fatty acid intake and the colonic gut microbiota in humans. Nutrients Jun 29;14 (13), 2722. https://doi.org/10.3390/nu14132722.
Xue, Z., Zhang, W., Wang, L., Hou, R., Zhang, M., Fei, L., Zhang, X., Huang, H., Bridgewater, L.C., Jiang, Y., Jiang, C., Zhao, L., Pang, X., Zhang, Z., 2015. The bamboo-eating giant panda harbors a carnivore-like gut microbiota, with excessive seasonal variations. mBio 6 (3), e00022-e15. https://doi.org/10.1128/mBio.00022-15.