

OPEN

Microbiota fingerprints within the oral cavity of cetaceans as indicators for population biomonitoring

Pedro Soares-Castro¹, Helena Araújo-Rodrigues¹, Filipa Godoy-Vitorino², Marisa Ferreira^{1,3}, Pablo Covelo⁴, Alfredo López^{4,5}, José Vingada^{3,5}, Catarina Eira^{3,5} & Pedro Miguel Santos¹

The composition of mammalian microbiota has been related with the host health status. In this study, we assessed the oral microbiome of 3 cetacean species most commonly found stranded in Iberian Atlantic waters (*Delphinus delphis*, *Stenella coeruleoalba* and *Phocoena phocoena*), using 16S rDNA-amplicon metabarcoding. All oral microbiomes were dominated by *Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Fusobacteria* bacteria, which were also predominant in the oral cavity of *Tursiops truncatus*. A Constrained Canonical Analysis (CCA) showed that the major factors shaping the composition of 38 oral microbiomes (p-value < 0.05) were: (i) animal species and (ii) age class, segregating adults and juveniles. The correlation analysis also grouped the microbiomes by animal stranding location and health status. Similar discriminatory patterns were detected using the data from a previous study on *Tursiops truncatus*, indicating that this correlation approach may facilitate data comparisons between different studies on several cetacean species. This study identified a total of 15 bacterial genera and 27 OTUs discriminating between the observed CCA groups, which can be further explored as microbiota fingerprints to develop (i) specific diagnostic assays for cetacean population conservation and (ii) bio-monitoring approaches to assess the health of marine ecosystems from the Iberian Atlantic basin, using cetaceans as bioindicators.

The ecological sustainability of aquatic environments is being dramatically threatened by climate change, habitat deterioration and a vast array of human-driven activities. Human impacts affect directly ocean life, often leading marine species to an endangered status¹⁻³. Cetaceans, particularly the Odontoceti, occupy high trophic levels (predators, primary or secondary consumers) and, consequently, their condition and wellbeing reflect the health and status of lower trophic levels⁴⁻⁷.

The common dolphin (*Delphinus delphis*) is one of the most abundant cetacean species in Atlantic Iberian waters⁸, sharing the habitat with other members of the family Delphinidae, such as the bottlenose dolphin (*Tursiops truncatus*) and the striped dolphin (*Stenella coeruleoalba*). The more coastal delphinids also share their habitat with the harbour porpoise (*Phocoena phocoena*), a member of the family Phocoenidae. The highest abundances of striped dolphins in Portugal are markedly found offshore, whereas harbor porpoises are mostly concentrated near the coast. On the other hand, important common dolphin abundances can be found both near the coast and offshore, sometimes forming mixed groups with striped dolphins⁹.

Recent surveys showed that the harbour porpoise population is severely declining in Iberian Atlantic waters⁹. There is a growing concern about the impact of various hazards to marine wildlife, and recent EU legal frameworks (2008/56/EC, 2010/477/EU, 2017/848/EU) highlight the need for marine populations to achieve Good Environmental Status. The more near-coastal animals may be exposed to a wide variety of anthropogenic

¹Department of Biology and Centre for Molecular and Environmental Biology (CBMA), University of Minho, Campus de Gualtar, 4710-087, Braga, Portugal. ²University of Puerto Rico, School of Medicine, Department of Microbiology and Medical Zoology, Microbial Ecology and Genomics Lab, GPO Box 365067, San Juan, Puerto Rico, 00936-5067, USA. ³Portuguese Wildlife Society (SPVS), Quiaios, Field Station, Apartado 16 EC Quiaios, 3081-101, Figueira da Foz, Portugal. ⁴Coordinadora para o Estudo dos Mamíferos Mariños (CEMMA), P.O. Box 15, 36380, Pontevedra, Gondomar, Spain. ⁵Department of Biology and CESAM, University of Aveiro, 3810-193, Aveiro, Portugal. Correspondence and requests for materials should be addressed to P.M.S. (email: psantos@bio.uminho.pt)

contaminants and pathogens. Systematic studies on marine mammal health and disease are crucial to support conservation and management measures, particularly considering species such as the harbour porpoise.

In Continental Portugal, since the year 2000 when regional dedicated marine animals stranding networks started operating until 2016, a relatively high cetacean stranding rate was registered. The annual average rates amount to 236 stranded cetaceans, even though only about 500 km of the Portuguese coast has been monitored on a regular basis and the major cause for dead strandings (nearly 45%) was accidental capture⁹ and also confirmed disease (e.g. viral and bacterial infections^{10–12}). However, in many cases (about 36%), the observed cetacean mortality causes remained undetermined, due to the number of stranded animals found in advanced decomposition stages.

Mammalian microbial communities have critical roles in the nutrition of the host, resistance to colonization by pathogens and the maturation of the host immune system^{13,14}. Moreover, the microbiota composition of each animal species is shaped by the interactions with the immune system, according to the animal physiology, diet and social habits, environmental context, among others¹⁵, and has been related with either health or disease statuses of the host^{16,17}.

Understanding host-microbe interactions in cetaceans may contribute to the identification of compromised populations, microbial markers of disease and ultimately to scientifically based population management decisions. In this sense, high throughput sequencing technologies paved the way towards implementing innovative monitoring approaches such as the environmental surveillance of the ocean ecosystem^{18,19}.

There are currently twenty-one studies on the cetacean microbiota available, mostly focused on bottlenose dolphin (*Tursiops truncatus*) samples retrieved in different regions of the world, including captive and in the wild individuals^{20–29}. Godoy-Vitorino *et al.*¹² published the first report of a female striped dolphin (*Stenella coeruleoalba*) microbiome from different body sites. Other cetacean-related studies include the killer whale and humpback whale, which showed associations between skin microbiota and episodic migrations²⁷.

In the present study, we assessed the oral microbiome of the 3 small cetacean species most commonly found stranded in the Atlantic Iberian coast (*S. coeruleoalba*, *D. delphis*, and *P. phocoena*), the latter two have never had their microbiomes characterized before. The oral cavity is a body site of easy accessibility due to its non-invasive nature, allowing a fast sampling procedure, which can be systematically used in either dead or live animals, while contributing for the preservation of sample integrity. A correlation analysis was applied to determine whether each of the analyzed cetacean species holds unique features in their microbiome profile, allowing species segregation based on their microbiome. The microbiome profiles were also evaluated (regardless of cetacean species) to identify discriminatory fingerprints according to age class, ecology of the animal or cause of death. The correlation analysis was also tested using 16S rDNA-amplicon data from a previous study focused on bottlenose dolphins¹⁷ to evaluate if the proposed data analysis would detect similar discriminatory patterns, using data from different experimental setups. Ultimately, we hypothesize that the correlation analysis approach used in the present work may be applied to 16S rDNA-amplicon datasets obtained with different sequencing technologies using samples from different geographic locations, facilitating data comparisons between different research studies on the microbiome composition of several cetacean species.

Results

Structure of the oral community of 3 species of Odontoceti cetaceans. In the present work, the standardized pipeline established by Godoy-Vitorino *et al.*¹² was used to study the microbiome composition from the oral cavity of 3 species of Odontoceti cetaceans (Supplementary Table S1).

A total of 4430357 Miseq paired reads were filtered and merged into 2904009 high quality-filtered sequences comprising 251 bp of the V4 region of the 16S rRNA gene. After the removal of unclassified sequences, a total of 2539933 sequences were binned and classified into 1329 OTUs (Supplementary Table S2). In total, 29 bacterial phyla were identified, being distributed by 70 classes, 215 families and 483 assigned genus (Supplementary Table S2). All oral microbial communities were dominated by members belonging to the *Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Fusobacteria* phyla (Supplementary Fig. S1), which were also predominantly detected in the oral cavity of bottlenose dolphins (*Tursiops truncatus*), described by Bik *et al.*¹⁷.

The microbial community from the oral cavity differs among species and at different development stages of the animal. The sampled bacterial communities were subjected to a Canonical Correspondence Analysis (CCA) to discriminate the biological and ecological factors underlying their variability (CCA is sensitive to the less abundant and unique species in the samples).

The ordination according to the animal phylogeny (Fig. 1) grouped the samples of the same species together, clustered separately from the other two species (CCA p-value = 0.001 when considering bacterial species).

Although this clustering profile suggested the existence of interspecific differences in the microbial communities of the oral cavity, the significance was only observed for the comparison between the *P. phocoena* samples vs. the *D. delphis*-*S. coeruleoalba* group (CCA1 component p-value = 0.001; CCA2 component p-value = 0.224). Samples from both species *D. delphis* and *S. coeruleoalba* were in fact clustered more closely, which also reflected their closer phylogenetic relationship: *D. delphis* and *S. coeruleoalba* belong to the family *Delphinidae* while *P. phocoena* belong to the family *Phocoenidae* (the phylogenetic trees based on the nucleotide sequence of the mitochondrial genes coding for the cytochrome c oxidase polypeptide I and the cytochrome b are presented in the Supplementary Fig. S2).

Samples from *D. delphis* showed higher average number of OTUs and Shannon index values than the samples collected from *P. phocoena* and *S. coeruleoalba* (although not statistically significant), which could suggest a greater interspecific variation of the microbiota colonizing the oral cavity of these animals (Fig. 1b). The oral microbiome associated to *D. delphis*, *P. phocoena* and *S. coeruleoalba* were comprised by a total of 1134 OTUs, 903 OTUs and only 756 OTUs, respectively (Fig. 1c). From the 1329 OTUs detected in this study, the 3 species shared

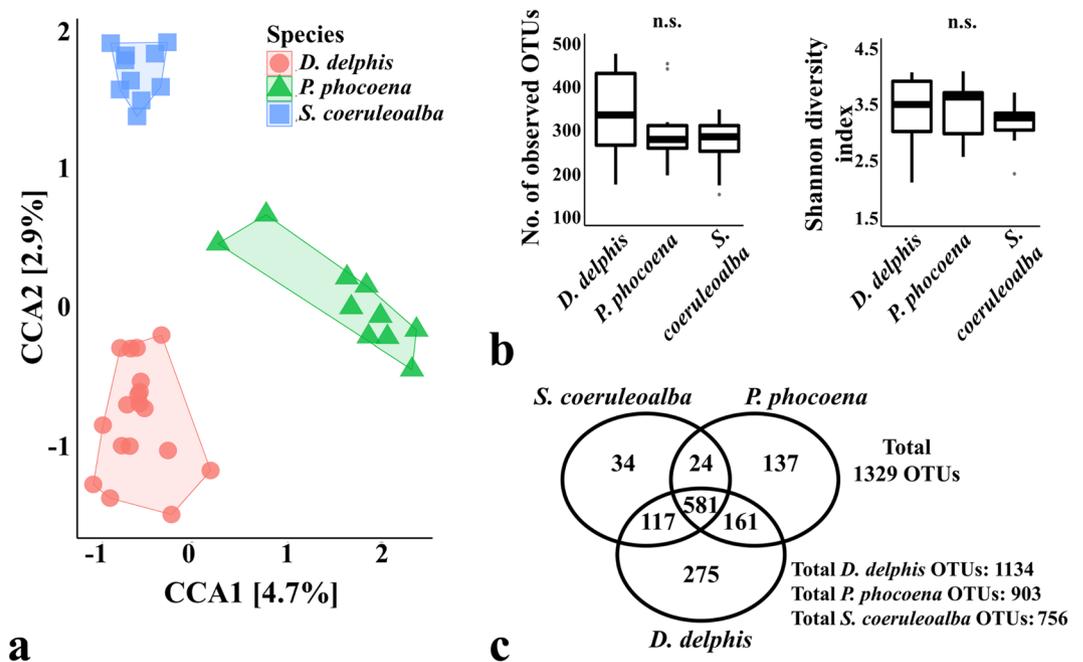


Figure 1. Comparison of the oral microbiomes of cetaceans according to the species of the sampled animals. Panel (a) shows the Canonical Correspondence Analysis (CCA) (p -value = 0.001) which was performed after subsampling the OTU table to even sequencing depth and at the bacterial species level, with Hellinger transformation of the abundances. Similar results were obtained when grouping the OTUs by bacterial genera. The colour frames in both CCA plots group the samples according to the constrained variable. The major contributions of the “host species” variable are shown as % in the first and second component of the CCA plot (CCA1 and CCA2, respectively). Panel (b) shows the richness and diversity measures calculated between species (as the average observed OTUs and Shannon diversity index of each sample), using the OTU table normalized by total sum scaling. Significance of these alpha-diversity metrics was tested with Kruskal-Wallis chi-squared test followed by pairwise Wilcoxon test between groups (n.s., statistically not significant). Panel (c) shows the number of total OTUs in each group and the number of OTUs shared between groups. *Delphinus delphis* ($n = 18$), *Phocoena phocoena* ($n = 10$) and *Stenella coeruleoalba* ($n = 10$).

581 OTUs (around 44%). Whereas the microbial community of *D. delphis* presented similar number of common OTUs with the other 2 species (117 with *S. coeruleoalba* and 161 with *P. phocoena*), the *S. coeruleoalba* and *P. phocoena* communities shared only 24 OTUs.

A significant clustering was also observed in the CCA plot according to the development stage of the sampled individuals (Fig. 2a; CCA p -value = 0.001), in which all samples from juvenile animals, regardless of their species, were grouped separately from the more closely located groups of adult and sub-adult animals (CCA1 component p -value = 0.012; CCA2 component p -value = 0.167).

The differential clustering shown in Fig. 2a was also observed between immature and mature specimens, when constraining the CCA according to the sexual maturity of the animals (CCA p -value = 0.033, data not shown), which supports the idea of an evolving microbiome throughout the maturation and aging.

According to Fig. 2b, the composition of the microbial community from adult and mature animals was composed by a higher average number of OTUs and Shannon diversity index, when compared to the oral community of juvenile and immature animals (p -value for pairwise Wilcoxon rank tests for both alpha-diversity metrics: adult vs. juvenile < 0.008; mature vs. immature < 0.002).

A total of 1154 OTUs were detected in adults, 919 OTUs in sub-adult animals and 874 OTUs in juveniles (Fig. 2c), from which 655 OTUs were common among the 3 groups. Juveniles and sub-adult animals shared only 27 OTUs (3%), whereas a larger fraction was shared between each one of these two groups and adult animals (>12%).

The composition of the oral microbial community of cetaceans is also affected by the biogeography and cause of death of the animal. The CCA in Fig. 3 shows differences between samples collected from the northern and western Atlantic Iberian coast (CCA p -value = 0.046). This clustering may suggest the existence of 2 communities of cetaceans that may have a different biogeography, thus reflecting different social habits, different use of space or the utilization of different food resources available throughout the Atlantic Iberian coast. Nevertheless, the alpha-diversity metrics did not show significant differences between both locations (Fig. 3b) and they share approximately 80% of the total number of OTUs (Fig. 3c).

The samples collected from animals due to accidental capture (bycatches) by fisheries and from stranded diseased animals were also clustered into two distinct groups (CCA p -value = 0.013), as observed in Fig. 4.

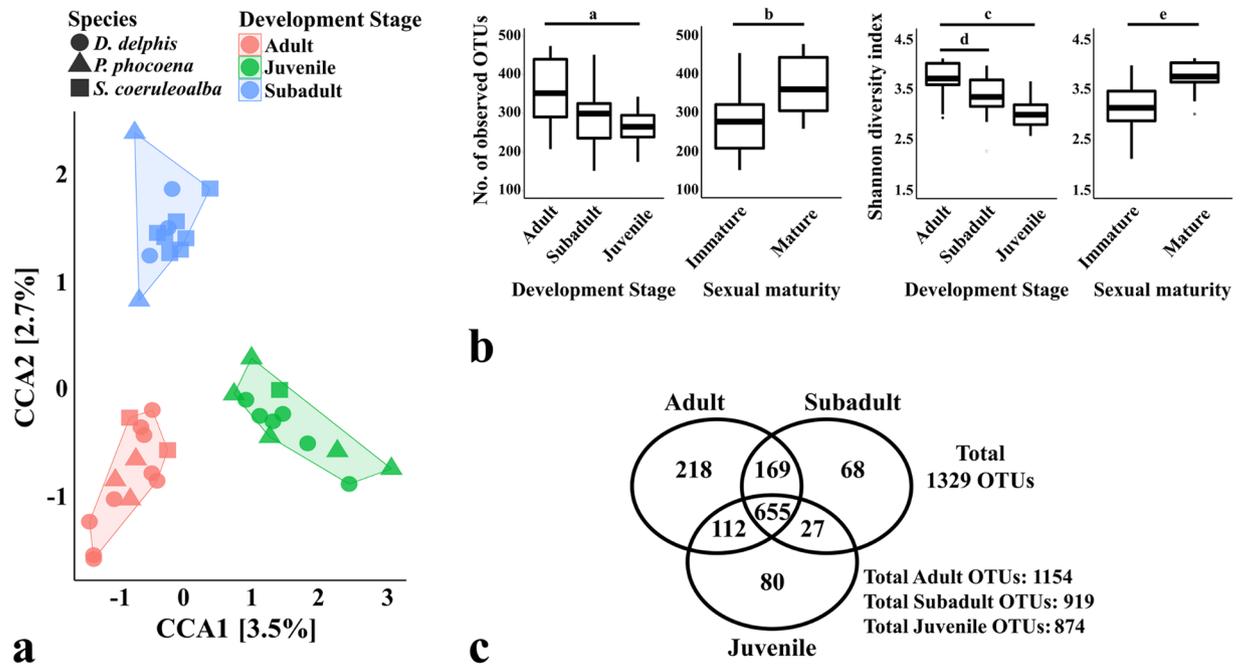


Figure 2. Comparison of the oral microbiomes of cetaceans according to the development stage/age class. The Canonical Correspondence Analysis (CCA) in panel (a) (p -value = 0.001) was performed after subsampling the OTU table to even sequencing depth and at the bacterial species level, with Hellinger transformation of the abundances. Similar results were obtained when grouping the OTUs by bacterial genera. The colour frames in both CCA plots group the samples according to the constrained variable and the symbols represent the species of the sampled animals: *D. delphis* (●), *P. phocoena* (▲) or *S. coeruleoalba* (■). The major contributions of the “development stage/age class” variable are shown as % in the first and second component of the CCA plot (CCA1 and CCA2, respectively). Panel (b) shows the richness and diversity measures calculated between species (as the average observed OTUs and Shannon diversity index of each sample), using the OTU table normalized by total sum scaling. Significance of these alpha-diversity metrics was tested with Kruskal-Wallis chi-squared test followed by pairwise Wilcoxon test between groups (a, p -value = 0.008; b, p -value = 0.002; c, p -value = 0.001; d, p -value = 0.002; e, p -value = 0.001). Panel c shows the number of total OTUs in each group and the number of OTUs shared between groups. Adult (n = 14), Subadult (n = 12) and Juvenile animals (n = 12).

Similarly to what is observed in other mammals, there may be a tight relation between the overall health status of the cetacean and the its oral microbiome composition^{30–35}. Even though all samples were derived from strandings (that resulted in death), the diseased animals seemed to comprise a less diverse oral microbiome, with a total of 873 OTUs, whereas the group of accidental captured (bycatches) totalled 1223 OTUs (Fig. 4). The health status of stranded cetaceans, which do not show signs of bycatch, is usually poor and associated with illness (pathologies and infections), or trauma. The weakened animal may end up in shallow waters, resulting in stranding and, usually, death.

The discriminatory power of the oral microbial community is also observed in datasets of healthy cetaceans.

In order to validate the discriminatory power of the variables under study in healthy cetaceans, the sequencing data from 25 animals, obtained by Bik *et al.*¹⁷, was also analysed in this work (Supplementary Table S3). Although composed solely by bottlenose dolphins (*T. truncatus*), the two sampled populations, from San Diego Bay (San Diego, USA) and Sarasota Bay (Florida, USA), also comprised female and male animals, at different ages (classified as stages of the development in the present study). Their oral microbial communities were also sampled with swabs and the extracted DNA was sequenced by pyrosequencing chemistry, using primers targeting the V3-V4-V5 region of the 16S rRNA gene. The OTU tables from both studies were merged to remove redundant taxa and subsampled to an even sequencing depth of 1019 sequences. Since only 54 OTUs shared the same representative sequence (Supplementary Table S4), the constrained ordinations were carried out at the bacterial genus level. From a total of 353 classified genera, 63 were detected in both studies.

Bik *et al.*¹⁷ did not find a clustering effect for animal age (2–51 years old) or gender, using the non-metric multidimensional scaling based on the Bray-Curtis distances of the *T. truncatus* microbial communities. Similarly, the ordination of the data obtained in the present study using Bray-Curtis distances also did not result in a clear clustering of samples according to the tested variables (data not shown).

The CCA of the oral microbiotas from both studies clustered the samples into different groups, according to the animal species (CCA1 component p -value = 0.001; CCA2 component p -value = 0.001) and their location (CCA1 component p -value = 0.001; CCA2 component p -value = 0.001), as shown in Supplementary Fig. S3. Strikingly, the CCA grouped the cetaceans into distinct groups, when constrained by the health status of the animal (Fig. 5), in which the dead animals from our study were clustered together, apart from the healthy animals

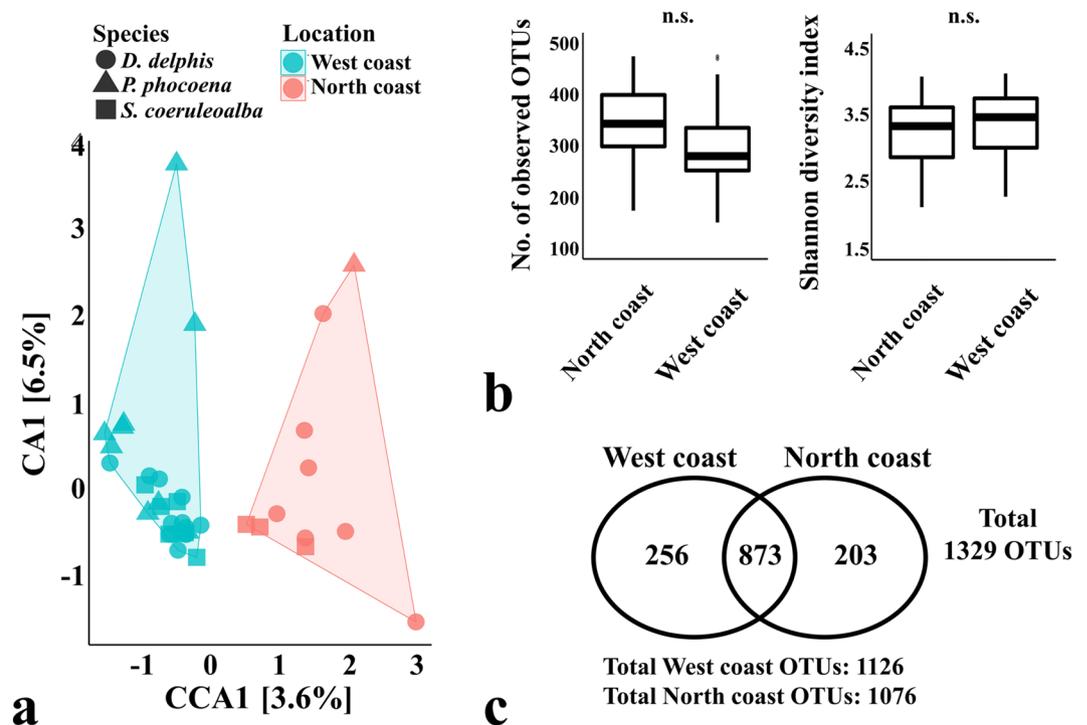


Figure 3. Comparison of the oral microbiomes of cetaceans according to the stranding location of the specimen. The Canonical Correspondence Analysis (CCA) in panel (a) (p -value = 0.046) was performed after subsampling the OTU table to even sequencing depth and at the bacterial species level, with Hellinger transformation of the abundances. Similar results were obtained when grouping the OTUs by bacterial genera. The colour frames in both CCA plots group the samples according to the constrained variable and the symbols represent the species of the sampled animals: *D. delphis* (●), *P. phocoena* (▲) or *S. coeruleoalba* (■). The major contributions of the “stranding location” variable are shown as % in the first and second component of the CCA plot (CCA1 and CCA2, respectively). Panel (b) shows the richness and diversity measures calculated between species (as the average observed OTUs and Shannon diversity index of each sample), using the OTU table normalized by total sum scaling. Significance of these alpha-diversity metrics was tested with Kruskal-Wallis chi-squared test followed by pairwise Wilcoxon test between groups (n.s., statistically not significant). Panel (c) shows the number of total OTUs in each group and the number of OTUs shared between groups. Animals sampled in the northern Atlantic Iberian coast ($n = 11$) and animals sampled in the western Atlantic Iberian coast ($n = 27$).

sampled by Bik *et al.*¹⁷ (CCA1 component p -value = 0.001; CCA2 component p -value = 0.015). We do not discard the role of host-specific factors especially our samples, as most of the diseased animals were *Stenella coeruleoalba*, except for one *D. delphis*.

Significant differences in CCA clustering of the merged datasets were also maintained between adults and juveniles (CCA1 component p -value = 0.012; CCA2 component p -value = 0.005; Supplementary Fig. S3) and between mature and immature animals (CCA p -value = 0.001; CCA plot not shown).

Commensal oral community among the 3 cetacean species sampled in this study. Given the different clusters obtained by CCA (Figs 1–4), we filtered the microbial profiles to identify the core community among the sampled cetaceans (Supplementary Table S5). Contrary to the total number of OTUs comprising the microbial communities of the 3 species, the number of OTUs present in all samples of each species (the species core microbiomes) was composed by a small fraction of the total OTUs (Fig. 6a), which evidenced the high heterogeneity of the samples. A similar trend was found grouping the samples by age class, sexual maturity, stranding location and cause of death (data not shown).

The core microbiome of *D. delphis*, *P. phocoena* and *S. coeruleoalba* were composed by 21, 18 and 39 OTUs, respectively (Supplementary Table S5), but their genus-level richness did not differ significantly (data not shown). Regardless of the animal species, the core microbiome of all 38 samples was composed by 12 OTUs, such as *Phocoenobacter* (OTU0), *Porphyromonas* (OTU4), *Oceanivirga* (OTU3), *Arcobacter* (OTU6) and *Fusobacterium* (OTU2, OTU266), as well as members of the genera *Campylobacter* (OTU32), *Ralstonia* (OTU83) and *Maritimimonas* (OTU16), 2 OTUs of the *Cardiobacteriaceae* family (OTU5, OTU157) and 1 OTU of the *Bacteroidales* S24-7 family (OTU244) (Fig. 6b and Supplementary Table S5). Some of these taxa have been detected in the oral cavities of other cetaceans and may represent the core symbiotic community colonizing the oral cavity of cetaceans^{17,29}.

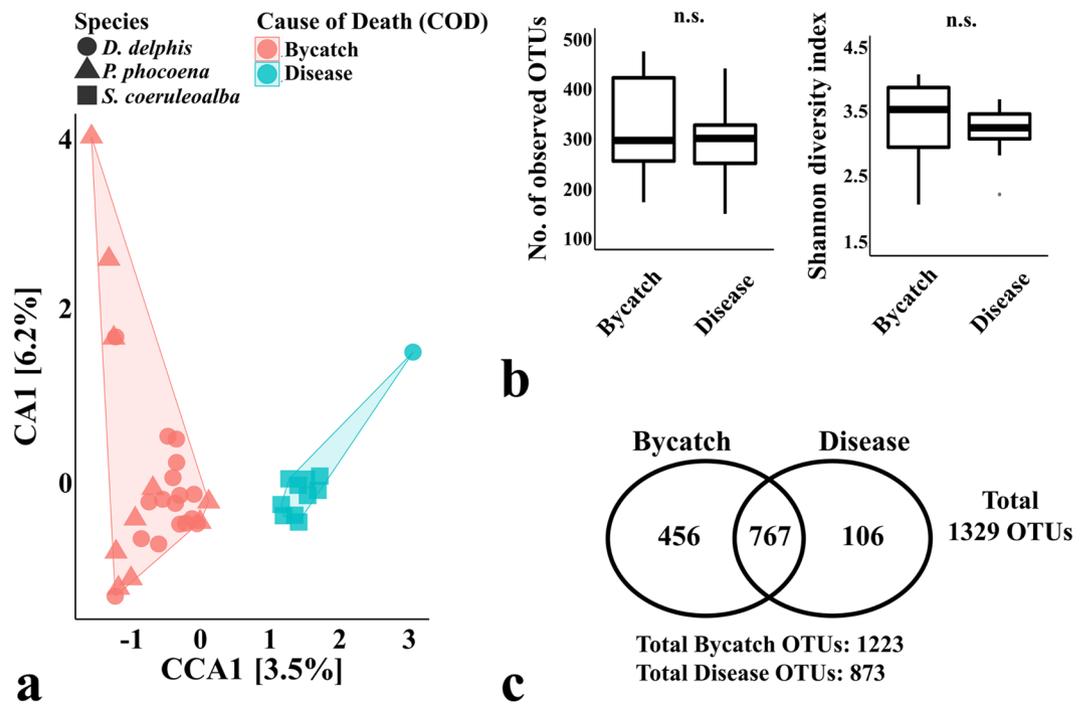


Figure 4. Comparison of the oral microbiomes of cetaceans according to their cause of death. The Canonical Correspondence Analysis (CCA) in panel (a) (p -value = 0.013) was performed after subsampling the OTU table to even sequencing depth and at the bacterial species level, with Hellinger transformation of the abundances. Similar results were obtained when grouping the OTUs by bacterial genera. The colour frames in both CCA plots group the samples according to the constrained variable and the symbols represent the species of the sampled animals: *D. delphis* (●), *P. phocoena* (▲) or *S. coeruleoalba* (■). The major contributions of the “cause of death” variable are shown as % in the first and second component of the CCA plot (CCA1 and CCA2, respectively). Panel (b) shows the richness and diversity measures calculated between species (as the average observed OTUs and Shannon diversity index of each sample), using the OTU table normalized by total sum scaling. Significance of these alpha-diversity metrics was tested with Kruskal-Wallis chi-squared test followed by pairwise Wilcoxon test between groups (n.s., statistically not significant). Panel c shows the number of total OTUs in each group and the number of OTUs shared between groups. Sampled animals resulting from bycatch ($n = 27$) and stranded diseased animals ($n = 11$).

Microbial fingerprints within the oral cavity of cetaceans. The taxa driving the observed CCA were identified by (i) a sPLS-DA to identify the major bacterial genera contributing to the separation of samples, followed by the Indicator Species Analysis of the selected genera and respective OTUs, and (ii) the linear discriminant analysis (LDA) effect size algorithm (LEfSe), as detailed in Supplementary Table S6.

Only the taxa showing statistical significance at the genus and OTU level (p -value < 0.05) in both approaches were considered potential microbial signatures of the associated groups (Fig. 7).

From a total of 21 bacterial genera and 40 OTUs that were considered potential bacterial signatures between the sampled cetacean species (Supplementary Table S6), 5 genera and 9 OTUs were detected as statistically significant by both approaches (Fig. 7). The bacterial genera *Dethiosulfovibrio* (OTU141, with a 94% nucleotide identity corresponding to the anaerobic *Dethiosulfovibrio marinus* strain WS100³⁶; and *Marinicella* (OTU255, with a 96% nucleotide identity corresponding to the marine *Marinicella pacifica* strain sw153³⁷) were associated to harbour porpoise samples (*P. phocoena*). Two *Bdellovibrio* OTUs (OTU56 and OTU338) with low nucleotide homology to known species (<90%) were highly represented in the clustering of common dolphins and striped dolphins (*D. delphis*-*S. coeruleoalba*). *Bdellovibrio* OTU338 was associated with the *D. delphis* group and the *Bdellovibrio* OTU56 with *S. coeruleoalba*. Moreover, *S. coeruleoalba* samples were also associated with the predominance of *Phocoenobacter* (4 OTUs, in particular with OTU495, with a 98% nucleotide homology to *Phocoenobacter uteri* strain M1063U/93) and *Porphyromonas* genera (OTU4, which was identified as a member of the core microbiome of the 3 cetaceans). However, the *Bdellovibrio* OTU56, the *Phocoenobacter* and *Porphyromonas* OTUs, in particular *Porphyromonas* OTU4, were also associated with diseased animals (as shown in Fig. 7), which suggested that the abundance and balance of these taxa in the community may be determinant to their pathogenic potential. The genus *Gemella* and the *Gemella* OTU18 were also associated with the group of diseased animals of the present study.

To identify taxa that could discriminate healthy from diseased animals, the Indicator Species Analysis and LEfSe were applied to the combined CCA including the animals from this study and those from Bik *et al.*¹⁷, as observed in Fig. 5. When clustering was performed according to species or location (Supplementary Fig. S3), the samples associated to the groups of *T. truncatus* in USA populations (Sarasota Bay and San Diego Bay) were the same samples leading to the healthy specimens’ group on the CCA (Fig. 5). After filtering the bacterial

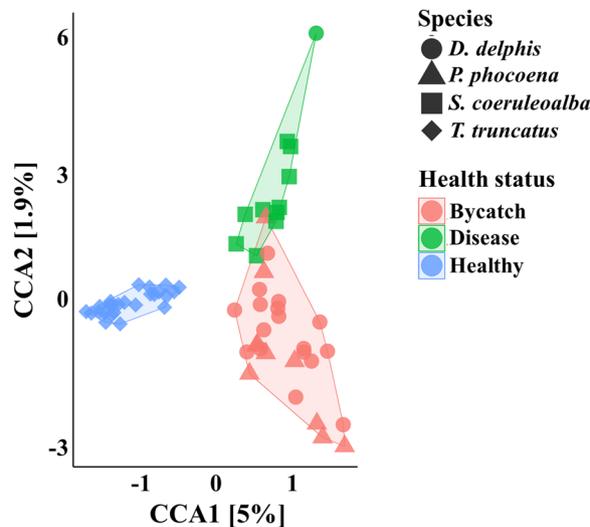


Figure 5. CCA of the oral microbiomes of the 4 cetacean species, according to their health status. Canonical Correspondance Analysis of the oral microbial communities of cetaceans sampled in this study and by Bik *et al.*¹⁷, constrained according to the health status of the sampled animals (p-value = 0.001). The color frames in the plot group the samples according to the constrained variable, comprising 3 groups: dead animals derived from bycatch (n = 27), stranded diseased animals (n = 11) and healthy animals sampled by capture and release procedure (n = 25) in Bik *et al.*¹⁷. Both ordinations were performed after subsampling the OTU tables to even sequencing depth of 1019 sequences and at the genus level, with Hellinger transformation of the abundances. In both panels, the species of the sampled animals are represented by symbols: *D. delphis* (●), *P. phocoena* (▲) or *S. coeruleoalba* (■) or *T. truncatus* (◆). The major contributions of the “health status” variable are shown as % in the first and second component of the CCA plot (CCA1 and CCA2, respectively).

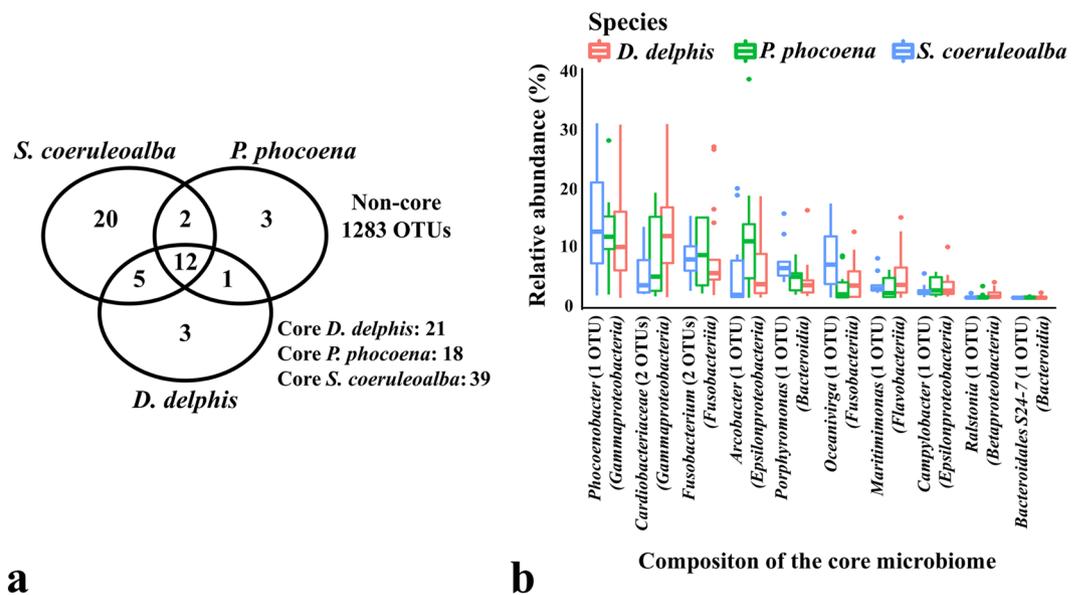


Figure 6. Core OTUs among the different cetacean species (a) and present in all 38 samples (b). In panel (a), the number of OTUs comprising the core microbiome of each group is represented by the OTUs present in all respective samples. In panel (b), the taxa abundance of the OTUs present in all samples is shown as the average relative frequency for each genera (or family when a genus was not assigned).

genera and OTUs contributing to the clustering according to these 3 variables (species, location and health status), the OTU79, OTU383, OTU1228, OTU1348 from *Porphyromonas* genus were identified as potential signature of the diseased animals (p-value < 0.05 by the Indicator Species Analysis and LEfSe). Moreover, the genera *Bradyrhizobium* (OTU419), *Escherichia* (OTU130) and *Pseudomonas* (OTU69, OTU155, OTU168) were also considered as potential indicators of disease when using the merged datasets. Furthermore, they were not detected in the oral microbiome of healthy *T. truncatus*.

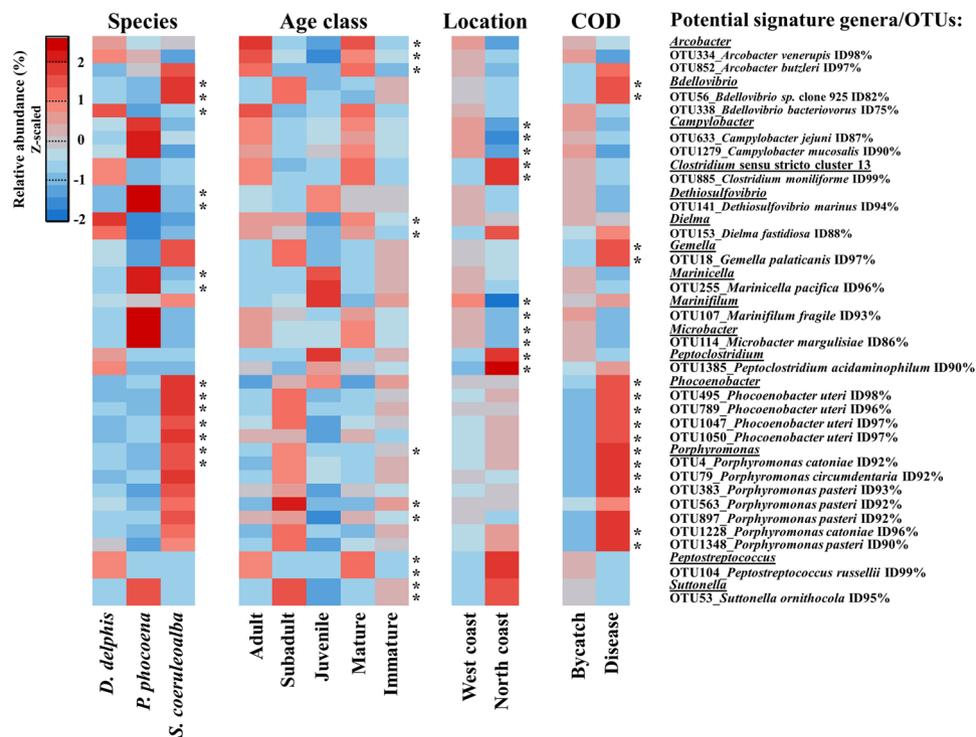


Figure 7. Bacterial genera and OTUs with different relative abundance between the groups of variables under study. The significance of potential microbial fingerprints (highlighted with “*”, p-value < 0.05) were identified by the Indicator Species Analysis and the linear discriminant analysis (LDA) effect size algorithm (LEfSe), detailed in Supplementary Table S6. Only the taxa showing significant differential abundance at the genus and OTU levels by both methods were considered. For representation and due to a broad range of values, the relative abundances were Z-scaled to highlight their comparisons between groups. Taxonomic validation of the representative sequences of the selected OTUs was performed with BLASTN analysis against the NCBI non-redundant nucleotide database.

The clustering of the healthy animals was specifically associated with the presence of 4 OTUs of *Desulfobacterium* spp., 2 OTUs of *Tannerella* spp. and 1 OTU of *Thalassobius* spp., none of them present in the oral communities of the diseased animals sampled in this study. The genera *Marinifilum* (17 OTUs), *Maritimimonas* (10 OTUs), *Microbacter* (4 OTUs), *Sphaerochaeta* (2 OTUs) and *Thiothrix* (4 OTUs) were also considered as potential signature taxa of the non-diseased status of the studied cetaceans, by being present in healthy or accidentally captured animals and contributing with several OTUs to the CCA clustering (Fig. 5 and Supplementary Table S7). Nevertheless, due to the low number of common OTUs and bacterial genera between both studies, the significance of potential signature taxa associated with a healthy vs. diseased status may have been biased, by the absence of several taxa in one of the tested groups.

A total of 20 genera and 38 OTUs were considered potential bacterial signatures for the different age classes (development stage and sexual maturity) of the sampled animals (Supplementary Table S6), from which 5 genera and 7 OTUs were common for both methods used to identify the taxa driving the observed CCA (Fig. 7). The bacterial genera *Arcobacter* (OTU334 with a 98% nucleotide homology to *Arcobacter venerupis* strain CECT7836T and OTU852 with a 97% homology to *Arcobacter butzleri* strain NCTC12481), *Dielma* (OTU153 with a 88% homology to *Dielma fastidiosa* strain JC13) and *Peptostreptococcus* (OTU104 with a 99% homology to *Peptostreptococcus russellii* strain ING2-D1D) were associated with adult/mature animals. Members of the genera *Porphyromonas* (OTU563 and OTU897, both with a 92% homology to *Porphyromonas pasteri* strain KUFDS01) and *Suttonella* (OTU53 with a 95% homology to *Suttonella ornithocola* strain B6/99/2) were associated with animals in the subadult stage.

The separation according to the location of the animals was associated to 22 genera and 27 OTUs (Supplementary Table S6), with a significant overrepresentation of the genera *Campylobacter* (OTU663 and OTU1279 with <90% homology to known species), *Marinifilum* (OTU107 with 93% homology to *Marinifilum fragile* strain JC2469), *Microbacter* (OTU114 with 86% homology to *Microbacter margulisiae* strain ADRI) in the animals of the western Atlantic Iberian coast, whereas members of *Clostridium sensu stricto* cluster 1 (OTU885 with 99% homology to *Clostridium moniliforme* strain HYN0057) and *Peptoclostridium* (OTU1385 with 90% homology to *Peptoclostridium acidaminophilum* strain a1-2) were associated to the animals of the northern Atlantic Iberian coast (Fig. 7).

Discussion

Metagenomic studies in mammals have been mainly focused on the gut-related microbiome of terrestrial animals. Variations in the microbial community structure have been associated with the (i) host phylogeny, in which different species co-inhabiting the same niche have shown species- or population-specific gut microbiomes^{38–40}, (ii) host age and development stage^{41–43}, host gender^{44,45}, host biogeography and social habits^{43,46–51}. Little is known about the inter- and/or intraspecific variability of the microbiomes of cetaceans. Due to the high heterogeneity of the samples from the 38 Odontoceti used in the present study, clustering patterns of their oral microbiomes were not clearly observed with ordinations based on Bray–Curtis, as also reported by Bik *et al.*¹⁷. The ecology of these animals has been described to drive genetic differentiation between populations of cetaceans inhabiting the Atlantic Ocean, the North Sea and the Mediterranean Sea^{52–55}, which could also contribute to different predispositions for the oral colonization of specific taxa. Apprill *et al.*²⁷ compared the skin microbiome of humpback whales from subtropical waters (near Hawaiian Islands) and Alaska waters, and described the geographic distribution of the animals as a shaping factor for the differences observed in the structure of these microbial communities, exposed to different conditions of the surrounding environment.

The heterogeneity of the oral samples in the present study was evidenced by the variation of the alpha-diversity values and by the low number of core OTUs predicted for each species (1.9%, 1.9% and 5.2% of the total OTUs of *D. delphis*, *P. phocoena* and *S. coeruleoalba*, respectively). This heterogeneity was also observed in Bik *et al.*¹⁷, since only 2 OTUs were present in all oral samples of *T. truncatus*. According to Apprill and co-workers²⁶, the drone-captured blowhole samples from humpback whales *Megaptera novaeangliae* revealed a microbiome composed by 25 core OTUs and only 2 bacterial genera comprised the core skin microbiome of this whale species²⁷. In contrast, Erwin and co-workers reported 147 and 155 OTUs composing the core gut microbiome of the pygmy (*Kogia breviceps*) and dwarf (*Kogia sima*) sperm whales²⁸. These reports evidence a greater variability associated with the microbiome of body sites exposed to the external environment and their concomitant sampling methods, in contrast with the microbiome of internal body sites (e.g. gut), which might be prone to a lower degree of variation regarding their microbial richness. In this respect, this initial analysis could suggest that evaluating the oral cavity microbiota composition may not constitute the best method to detect overall variations in the health status of cetacean populations, unless a large number of samples is available. However, these are pioneering results, which can be used in future non-invasive biomonitoring studies as a comparison baseline from which health changes or deviations will be evaluated.

The core microbiome of all 38 samples analyzed in this study is composed by 12 OTUs, most of them related to human pathogens. *Pasteurellaceae* members, such as the *Phocoenobacter* OTU0, can be commensals. This OTU showed a 99% nucleotide identity with *Phocoenobacter uteri*, a species that was related to septicemia in marine mammals^{56,57}. However this species was also found in apparently healthy individuals¹⁷. *Porphyromonas* OTU4 showed a 92% nucleotide homology to *Porphyromonas catoniae* strain ATCC 51270, whose abundance was associated with a caries-free oral status in humans⁵⁸, and was detected in the oral cavity of the healthy *T. truncatus* reported by Bik *et al.*¹⁷. Similarly, even though commensal *Fusobacterium* strains have been associated to several pathologies in humans (e.g. inflammatory bowel disease, oral carcinoma)^{59,60} and in a diseased striped dolphin (Godoy-Vitorino, 2017), 16 *Fusobacterium* OTUs were detected in the oral microbiomes of healthy bottlenose dolphins¹⁷. *Campylobacteraceae* genera, such as *Arcobacter* and *Campylobacter*, have been associated with human and animal illness^{61,62} but have also been found in marine sediments⁶³ as well as in the dental plaque of non-diseased captive delphinids²⁹.

Cardiobacteriaceae members have been associated with bacteremia and wound infections in humans⁶⁴. They were detected in the blowhole, tongue and oral cavity of the dolphin reported by Godoy-Vitorino *et al.*¹², as indicators of sepsis. Nevertheless, 51 OTUs belonging to the *Cardiobacteriaceae* family were detected in oral cavity of healthy bottlenose dolphins¹⁷. The *Bacteroidales* S24-7 family (*Candidatus Homeothermaceae*) has been predominantly found in the gut of homeothermic animals (e.g. human, mice and other rodents, koala, pig) and are described as mainly carbohydrate-fermentative and microaerophilic or anaerobic bacteria⁶⁵. The ubiquitous presence of these 12 OTUs in all 38 samples (Fig. 6 and Supplementary Table S5), regardless animal species, age class, stranding location or the animal health status, is evidence that these bacteria may show low virulence levels in healthy animals, whose pathogenicity may be only potentiated in immunocompromised individuals. The identification of a commensal microbiota catalogue will be the cornerstone to set a baseline microbial profile of the oral cavity and more accurately determine harmful variations in the microbiome composition.

In the present study, the analysis of the oral microbiomes of cetaceans with a Canonical Constrained Analysis showed that the phylogeny and ecology of these animals shaped the microbial community structure of the oral cavity. The major variables driving the separation of the communities were the animal species and the development stage, which explained approximately 8% and 6% of the CCA variation, as shown in the axis CCA1 and CCA2 components of the CCA plots (Figs 1 and 2).

The evidenced divergence between the microbial communities according to the animal species (Fig. 1) may reflect their different use of the marine environment (striped dolphins are mostly found offshore whereas porpoises near the coast). The association of specimens from *S. coeruleoalba* and *D. delphis* has been also reported in the Mediterranean sea and in other areas of the Atlantic ocean^{66,67}. The occurrence of mixed groups of *coeruleoalba* and *D. delphis* may contribute to a more similar microbiota between animals and attenuates interspecific differences, mainly because individuals share the same areas, similar food resources, they may physically interact, and they are exposed to the same environmental chemicals. To mitigate eventual bias regarding the inter-species analysis of bacterial richness and diversity derived from different species group size (common dolphin group includes 18 animals, in comparison to 10 harbour porpoises and 10 striped dolphins), the same clustering analysis was performed after normalizing the “Species” groups to 10 animals. Supplementary Fig. S5 shows that the number of sampled animals did not affect the CCA clustering pattern (p-value = 0.001 in all panels), neither resulted in statistically significant differences in the comparisons of the average number of OTUs and Shannon

index values between the difference cetaceans species. The variation in the degree of data dispersion was the only effect observed in Supplementary Fig. S5b,c, which derived from the different contribution in OTUs that each animal might bring to the combined analysis, as a result of the biological variation inherent to the type of samples. A similar trend was observed in the sample number normalizations carried out for the other groups of variables under study (data not shown).

Reports of delphinid schools predominantly composed by animals of similar ages are not uncommon and support the clustering observed^{68,69}. For instance, in the western Pacific, *S. coeruleoalba* schools have been described to be composed either by juveniles, adults or by both types of animals⁶⁸. Juvenile schools may inhabit coastal waters, which usually offer more protection from predators. Younger members (calves) can remain in adult schools until completing 2 years of age and then join juvenile schools, until reaching adulthood. After reaching sexual maturity, sub-adult animals may join breeding or non-breeding adult schools. Breeding schools are usually comprised by sexually mature adult females and may also include adult males. Sub-adult and adult males in breeding schools may leave, after most females are pregnant. Thus, the breeding school evolves into a non-breeding adult school, which will later include their calves⁶⁸. Age-related differences in the gut microbiome were also described for other mammals such as the ring-tailed lemurs, Zucker rat and the Australian fur seal^{41–43}. Although common and striped dolphins mixed groups can be found at times, the separation according to host species was expected. This oral microbiome community separation is, in fact, a good indication of the applicability of such data in the future to assess health population changes within species. The oral microbiome community separation according to development stage is even more important, since it indicates that deviations (health imbalances) may be connected to and affect by the dynamics of population structure.

The statistically significant CCA clustering according to stranding location might be biased by the animals' ability to travel large distances along the coast or away from the coast (the microbial profile of the water or of the prey that could associate the microbial profiles and the potential signature genera to the different areas is not known). Therefore, microbial profiles of the main prey for each cetacean species, as well as microbial profiles of seawater in known cetacean occurrence hotspot areas, are needed. A significant CCA clustering of the samples based on bacterial genera was not obtained when the gender of the animals was used as biological variable (female vs. male animals). The lack of clustering according to host gender was expected considering that the analysed cetaceans species are composed of highly social individuals (less so for harbour porpoises) often forming mixed-sex groups. Significant differences were only detected when the microbial communities were compared at the family level (CCA p-value = 0.031). This is one of the more important results of this study, considering the severe conservation status of the porpoise population in Portugal. Microbiome fingerprinting for Phocoenids (or *Phocoena phocoena*, the only Phocoenidae representative in the North Atlantic) will prove very valuable in the future since the population is likely to undergo local extinction within decades⁹. If the population continues to decline, monitoring the evolution of the Phocoenidae health status will be critical to identifying impacts, sources and solutions that may contribute to halting the population decline.

The CCA approach presented in this study also organized the pyrosequencing datasets from Bik *et al.*¹⁷ into different groups, when constrained according to the location (p-value = 0.001, using bacterial genus or OTUs), development stage and sexual maturity (p-value using bacterial genus = 0.205; p-value using OTUs = 0.038) of the animals (Supplementary Fig. S4). Furthermore, when our data was merged with the data from Bik *et al.*¹⁷, similar CCA clustering patterns were observed (despite the low number OTUs common to both studies), in particular regarding the development stage of the animals (p-value = 0.001; Supplementary Fig. S3), in which the samples remained ordered in 3 groups of adult, sub-adult and juvenile animals. Therefore, the approach used may facilitate data comparisons between different research studies on the microbiome composition of several cetacean populations and species.

Moreover, this study identifies a total of 15 bacterial genera and 27 species to be further explored as microbiota fingerprints within the oral cavity of cetaceans.

Therefore, the present study contributes to setting the proper knowledge ground to (i) develop fast bio-monitoring molecular diagnostic assays and tools for microbiome illnesses to predict outbreaks in these populations, as well as to (ii) develop epidemiological models based on high-throughput approaches to assess population health status of coastal cetacean species, and (iii) to identifying impacts, their probable causes, and propose solutions towards marine conservation by using cetacean microbiomes as indicators of marine ecosystem health. Furthermore, the proposed microbiome profiling approach of the oral cavity of cetaceans may become a valuable tool in emergency scenarios like in massive stranding events. In fact, sample collection for microbiome profiling is a fast and straight forward procedure and it will allow to rapidly retrieve important information of large numbers of stranded individuals (e.g. health status, development stage), without resorting to laborious necropsy procedures.

Methods

Sample collection. The samples analysed in this work were obtained in collaboration with the Portuguese Marine Animal Stranding Network, which is coordinated by the Instituto para a Conservação da Natureza e Florestas from Portugal (<http://www.icnf.pt/portal/icnf>), in cooperation with the Portuguese Wildlife Society (SPVS), and by the Spanish Coordinadora para o Estudo dos Mamíferos Mariños (CEMMA), which is responsible for the marine mammal stranding network in Galician waters. All sampled animals were found stranded along the coast. In total, 38 cetaceans (some recently dead and others still alive), were assessed along the northern (11 specimens) and western (27 specimens) Atlantic Iberian coast, by the Marine Animal Stranding Network teams in Portugal and by CEMMA in Galicia, Spain. In the case of the stranded animals found initially alive on the beach, refloating was not a possibility due to their critical clinical state, and despite the efforts of the rescuing teams, animals died before reaching the rehabilitation centre.

Samples from the oral cavity were collected by swabbing the gingival sulcus of the lower and upper jaws with sterile nylon fiber swabs (FLOQSwab™, Copan), according to the procedure established previously in Godoy-Vitorino *et al.*¹², from those animals that had died during rescuing or within less than 24 h prior to sampling (estimated from *rigor mortis* and organ temperature). Additional information regarding gender, development stage and sexual maturity (using the total length as a proxy for age class), occurrence of gross pathologies, and cause of death of the specimens was registered when possible during necropsy procedures^{70,71}. Briefly, the sampled animals belonged to three different species of Odontoceti cetaceans, including 18 common dolphins (*Delphinus delphis*), 10 striped dolphins (*Stenella coeruleoalba*) and 10 harbour porpoises (*Phocoena phocoena*). After having been found stranded on the coast, the animals' necropsies indicated that they had either been incidentally captured in fisheries (27 occurrences as bycatch; animals showing normal body condition and typical cuts caused by fishing gear) or that they had died due to disease (11 occurrences; animals showing emaciation, thin blubber layer, empty stomachs, gross lesions most frequently indicating forms of pneumonia and hepatitis, and high parasite intensity). These cetaceans were classified according to their development stages (14 adults, 12 sub-adults and 12 juveniles) and to their sexual maturity (25 immature and 13 mature), from which 25 were females and 13 males (Supplementary Table S1).

Genomic DNA extraction, PCR amplifications and sequencing. The genomic DNA acquired by the oral cavity swabs, was extracted using the PureLink™ Genomic DNA Mini Kit, according to the manufacturer's instructions. PCR amplification of the 16S rRNA gene targeting the hypervariable region V4 (forward primer: 5' GTGCCAGCMGCCGCGGTAA 3'; reverse primer: 5' GGACTACHVGGGTWTCTAATCC 3'), according to Kozich and co-workers⁷². The generated amplicons covered a region of 251 bp. The DNA was processed according to Illumina instructions to generate Nextera XT paired-end libraries (2 × 250 bp).

Sequence processing and data analysis. Read pairs were trimmed with Sickle, including a minimum Phred score of 30 and a sliding window of 10% of the read length⁷³, to remove adapter and primers sequences, as well as, nucleotides corresponding to low quality base calls. Read pairs were overlapped with Vsearch v2.3.2⁷⁴ and prepared with strict quality and size filtering (minimum length of overlap between reads = 20 bp, minimum length of the merged sequence = 200 bp, maximum expected error of 0.5, maximum number of different bases in the overlap = 2) into uniform error-free sequences. Vsearch⁷⁴ was used for the de-replication, removal of chimeric sequences and clustering with an identity threshold of 97%. The taxonomic classification was assigned with the SILVA ribosomal RNA gene database, version 128⁷⁵, by using the non-redundant representative set of reference sequences clustered at 97% nucleotide identity. Before the analysis of the microbial profiles, the unclassified sequences and low frequency counts (singletons) were removed from the operational taxonomic unit (OTU) table.

The alpha-diversity metrics (richness - as number of OTUs and Shannon diversity index) were estimated with the *amp_alphadiv* function available in the *ampvis2* package version 2.3.11⁷⁶ from the "R" software, and the differences observed between groups of samples was tested with the Kruskal-Wallis chi-squared test, followed by pairwise Wilcoxon test between groups (*kruskal.test* and *pairwise.wilcox.test* functions).

The analysis of the community structure between samples was carried out after performing a total sum scaling normalisation (transforming abundances into relative frequency) and rarefaction of the OTU tables to an even sampling depth of 36760 sequences per sample, thus eliminating any bias due to differing sampling depth during the sequencing process⁷⁷. Constrained canonical analyses (CCA) were performed according to each variable of the metadata collected, by using the *cca* and *anova.cca* functions from the *vegan* package version 2.5.2⁷⁸, with 999 permutation testing. The comparisons of the microbial communities by CCA were performed at the species, genus and family level, using OTU tables with a square root transformation of the relative abundance (Hellinger transformation performed with the *decostand* function) to reduce the range of the data and to make it suitable for analysis by linear methods^{79,80}. Representation of the CCA plot was performed with the *amp_ordinate* function from the *ampvis2* package. The core microbiome and shared OTUs between groups of samples were estimated with the *compute_core_microbiome.py* and *shared_phylotypes.py* scripts from QIIME version 1.9.1⁸¹. BLASTN analysis of the representative sequences of the OTUs was performed against the NCBI non-redundant nucleotide database to identify the closest homologs of relevant taxa⁸².

Comparison with sequencing data available for healthy cetaceans. High-throughput sequencing data regarding the oral microbiome of healthy bottlenose dolphins (*Tursiops truncatus*) was previously published by Bik *et al.*¹⁷. The DNA from these *T. truncatus* samples were acquired from oral swabs and sequenced with the 454 Life Sciences Genome Sequencer FLX Titanium platform, using primers targeting the V3-V4-V5 region of the 16S rRNA gene.

In order to assess if the biological and ecological variables under study could underlie variations in the microbiome of the oral cavity of healthy cetaceans, the approach used in the current work was also applied to the data from 25 animals reported by Bik *et al.*¹⁷. The selected sequence data (identified in Supplementary Table S3 by the SRA accession numbers) derived from two populations inhabiting in the San Diego Bay, San Diego (16 animals) and Sarasota Bay, Florida (9 animals), in the USA. They also included specimens from different gender (13 female and 12 male animals), at different stages of the development (14 adult and 11 juvenile animals) and sexual maturity (11 mature and 14 immature animals).

Initially, BLASTN was used to assess the homology between the representative sequences of the OTUs from the current study and the representative sequences available for the OTUs of Bik *et al.*¹⁷. Common OTUs between studies were identified when (i) 100% nucleotide identity was observed within the 251 bp comprising the V4 region of the 16S rRNA gene, targeted in the amplicon sequencing approach of the present study; (ii) only 1 blast hit with 100% nucleotide identity was available from the data of Bik *et al.*¹⁷. Afterwards, the OTU tables from both

studies were merged as a.csv file, the duplicated common OTUs were manually removed and the final merged OTU table was converted to.biom format using the BIOM project tools⁸³. The CCA and sPLS-DA were performed as described before, at the genus level and using the subsampled tables with 1019 sequences, due to the lower sequencing depth that resulted from the pyrosequencing chemistry used in the Bik *et al.* study¹⁷.

Identification of discriminatory bacterial fingerprints according to animal species, age class, ecology of the animal, or cause of death. The potential signature bacterial genera contributing to the clustering of samples observed in the CCA plots were assessed by two approaches: (i) a supervised Partial Least Squares Discriminant Analysis (sPLS-DA) followed by an Indicator Species Analysis; (ii) the linear discriminant analysis (LDA) effect size algorithm (LEfSe).

The sPLS-DA was carried out with the *plsda* and *tune.plsda* functions from the mixOmics package version 6.3.1⁸⁴ to identify the major bacterial taxa underlying the CCA profile according to the studied variables⁸⁵. In the sPLS-DA plots, a confidence level of 95% was used to draw the confidence ellipses and highlight the sample groups obtained. The significance of the differential abundance of relevant taxa was validated by an Indicator Species Analysis⁸⁶ coupled to a Monte Carlo significance test with 999 permutations, by using the *indval* function from the labdsv package v1.8⁸⁷ and the original OTU table, transformed with total sum scaling. The association of each taxon to a specific metadata variable was measured by the *indval* index, which ranges from a minimum of 0 (not associated) to a maximum of 1 (good indicator), depending on the taxon abundance and fidelity (relative frequency) to a particular group. A bacterial genus was considered a potential signature taxon if the indicator analysis showed (i) a p-value < 0.05, (ii) a relative frequency associated to the grouping variable of at least 0.5, (iii) an *indval* index of at least 0.5 associated to the grouping variable, and if iv) any of the OTUs related to the potential indicator genus showed a p-value < 0.05 in the indicator analysis of the original OTU table.

The LEfSe from the MicrobiomeAnalyst web-tool⁸⁸ was carried out with the original OTU table, transformed with total sum scaling, to complement and validate the previous approach. This approach involves the non-parametric factorial Kruskal-Wallis sum-rank test to identify taxa with significant differential abundances, according the grouping variables of interest, followed by Linear Discriminant Analysis (LDA) to estimate the effect size of each differentially abundant taxa⁸⁹. The threshold of the LDA score to identify the major bacterial taxa driving the clustering of the samples was set to 2, as described in previous studies^{90–93}.

Only the bacterial taxa showing significant differential abundance at the genus and OTU levels by both methods were considered as potential microbial signatures for the respective grouping variable. The differential abundance was represented by the *heatmap.2* function from the “R” software. For representation and due to a broad range of relative abundances, the values were Z-scaled (standard transformation from the function: $z = (x - \text{mean})/\text{standard deviation}$) to highlight their comparisons between groups.

Data Availability

The sequences datasets supporting the conclusions of this article are available in NCBI, deposited with the BioProject database ID PRJNA494623, associated with the SRA accession numbers ranging from SRR7963801 to SRR7963838. The essential data generated and analysed during this study is included in the Supplementary Dataset. The datasets of the *T. truncatus* animals analyzed during the current study, which were published by Bik *et al.*¹⁷, are available in NCBI (see Supplementary, Table S3 for the corresponding Sequence Read Archive accession) and as Supplementary Material of the corresponding publication (<https://doi.org/10.1038/ncomms10516>).

References

- Lubchenco, J., Palumbi, S. R., Gaines, S. D., Andelman, S. & Plugging, A. Hole In The Ocean: The Emerging Science Of Marine Reserves. *Ecol. Appl.* **13**, 3–7 (2003).
- Fleming, L. E. *et al.* Oceans and human health: Emerging public health risks n the marine environment. *Mar. Pollut. Bull.* **53**, 545–560 (2006).
- Halpern, B. S. *et al.* A Global Map of Human Impact on Marine Ecosystems. *Science*. **319**, 948–952 (2008).
- Rawson, A. J., Patton, G. W., Hofmann, S., Pietra, G. G. & Johns, L. Liver Abnormalities Associated with Chronic Mercury Accumulation in Stranded Atlantic Bottlenosed Dolphins. *Ecotoxicol. Environ. Saf.* **25**, 41–47 (1993).
- Lavery, T. J. *et al.* Heavy metal toxicity of kidney and bone tissues in South Australian adult bottlenose dolphins (*Tursiops aduncus*). *Mar. Environ. Res.* **67**, 1–7 (2009).
- Bossart, G. D. Marine mammals as sentinel species for oceans and human health. *Vet Pathol* **48**, 676–690 (2011).
- Frouin, H., Loseto, L. L., Stern, G. A., Haulena, M. & Ross, P. S. Mercury toxicity in beluga whale lymphocytes: Limited effects of selenium protection. *Aquat. Toxicol.* **109**, 185–193 (2012).
- Hammond, P. S. *et al.* Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. *Biol. Conserv.* **164**, 107–122 (2013).
- Vingada, J. & Eira, C. *Conservation of Cetaceans and Seabirds in Continental Portugal. The LIFE+ MarPro project.* (2017).
- Bento, M. C. *et al.* Herpesvirus infection in marine mammals: A retrospective molecular survey of stranded cetaceans in the Portuguese coastline. *Infect. Genet. Evol.* **67**, 222–233 (2019).
- Bento, M. C. Rd. M. *et al.* New insight into dolphin morbillivirus phylogeny and epidemiology in the northeast Atlantic: opportunistic study in cetaceans stranded along the Portuguese and Galician coasts. *BMC Vet. Res.* **12**, 176 (2016).
- Godoy-Vitorino, F. *et al.* The microbiome of a striped dolphin (*Stenella coeruleoalba*) stranded in Portugal. *Res. Microbiol.* **168**, 85–93 (2017).
- Bahrndorff, S., Alemu, T., Alemneh, T. & Lund Nielsen, J. The Microbiome of Animals: Implications for. *Conservation Biology. Int. J. Genomics* **2016**, 7 (2016).
- Schlesselman, T., Iyer, S. S., Kasper, D. L. & Blumberg, R. S. How colonization by microbiota in early life shapes the immune system. *Science*. **352**, 539 (2016).
- Apprill, A. Marine Animal Microbiomes: Toward Understanding Host–Microbiome Interactions in a Changing Ocean. *Front. Mar. Sci.* **4**, 222 (2017).
- Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature* **449**, 811 (2007).
- Bik, E. M. *et al.* Marine mammals harbor unique microbiotas shaped by and yet distinct from the sea. *Nat. Commun.* **7**, 10516 (2016).

18. Nogales, B., Lanfranconi, M. P., Piña-Villalonga, J. M. & Bosch, R. Anthropogenic perturbations in marine microbial communities. *FEMS Microbiol. Rev.* **35**, 275–298 (2011).
19. Ininbergs, K., Bergman, B., Larsson, J. & Ekman, M. Microbial metagenomics in the Baltic Sea: Recent advancements and prospects for environmental monitoring. *Ambio* **44**, 439–450 (2015).
20. Suzuki, A. *et al.* Comparison of the gut microbiota of captive common bottlenose dolphins *Tursiops truncatus* in three aquaria. *J. Appl. Microbiol.* **126**, 31–39 (2019).
21. Dudek, N. K. *et al.* Novel Microbial Diversity and Functional Potential in the Marine Mammal Oral Microbiome. *Curr. Biol.* **27**, 3752–3762.e6 (2017).
22. Soverini, M. *et al.* The bottlenose dolphin (*Tursiops truncatus*) faecal microbiota. *FEMS Microbiol. Ecol.* **92**, fw055 (2016).
23. Diaz, M. A. *et al.* Identification of Lactobacillus strains with probiotic features from the bottlenose dolphin (*Tursiops truncatus*). *J. Appl. Microbiol.* **115**, 1037–51 (2013).
24. Johnson, W. R. *et al.* Novel diversity of bacterial communities associated with bottlenose dolphin upper respiratory tracts. *Environ. Microbiol. Rep.* **1**, 555–62 (2009).
25. Hooper, R. *et al.* Host-derived population genomics data provides insights into bacterial and diatom composition of the killer whale skin. *Mol. Ecol.* **28**, 484–502 (2019).
26. Apprill, A. *et al.* Extensive Core Microbiome in Drone-Captured Whale Blow Supports a Framework for Health Monitoring. *MSystems* **2**, e00119–17 (2017).
27. Apprill, A. *et al.* Humpback whale populations share a core skin bacterial community: towards a health index for marine mammals? *PLoS One* **9**, e90785 (2014).
28. Erwin, P. M. *et al.* High diversity and unique composition of gut microbiomes in pygmy (*Kogia breviceps*) and dwarf (*K. sima*) sperm whales. *Sci. Rep.* **7**, 7205 (2017).
29. Goldman, C. G. *et al.* Novel gastric helicobacters and oral campylobacters are present in captive and wild cetaceans. *Vet. Microbiol.* **152**, 138–145 (2011).
30. Zaura, E., Keijser, B. J. F., Huse, S. M. & Crielaard, W. Defining the healthy “core microbiome” of oral microbial communities. *BMC Microbiol.* **9**, 259 (2009).
31. Ocon, S. *et al.* Transcription profiling reveals potential mechanisms of dysbiosis in the oral microbiome of rhesus macaques with chronic untreated SIV infection. *PLoS One* **8**, e80863 (2013).
32. Harris, S. *et al.* A pyrosequencing investigation of differences in the feline subgingival microbiota in health, gingivitis and mild periodontitis. *PLoS One* **10**, e0136986 (2015).
33. Rautava, J. *et al.* Oral microbiome composition changes in mouse models of colitis. *J. Gastroenterol. Hepatol.* **30**, 521–527 (2015).
34. Greenhalgh, K., Meyer, K. M., Aagaard, K. M. & Wilmes, P. The human gut microbiome in health: establishment and resilience of microbiota over a lifetime. *Environ. Microbiol.* **18**, 2103–2116 (2016).
35. Gao, L. *et al.* Oral microbiomes: more and more importance in oral cavity and whole body. *Protein Cell* **1–13** (2018).
36. Surkov, A. V., Dubinina, G. A., Lysenko, A. M., Glöckner, F. O. & Kuever, J. *Dethiosulfovibrio russensis* sp. nov., *Dethiosulfovibrio marinus* sp. nov. and *Dethiosulfovibrio acidaminovorans* sp. nov., novel anaerobic, thiosulfate- and sulfur-reducing bacteria isolated from Thiodendron sulfur mats in different saline environments. *Int. J. Syst. Evol. Microbiol.* **51**, 327–337 (2001).
37. Wang, Y., Liu, Y., Zhang, Z., Zheng, Y. & Zhang, X.-H. *Marinicella pacifica* sp. nov., isolated from seawater. *Int. J. Syst. Evol. Microbiol.* **66**, 2313–2318 (2016).
38. Yildirim, S. *et al.* Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities. *PLoS One* **5**, e13963 (2010).
39. Gruninger, R. J., Sensen, C. W., McAllister, T. A. & Forster, R. J. Diversity of rumen bacteria in Canadian cervids. *PLoS One* **9**, e89682 (2014).
40. Østbye, K., Wilson, R. & Rudi, K. Rumen microbiota for wild boreal cervids living in the same habitat. *FEMS Microbiol. Lett.* **363** (2016).
41. Smith, S. C., Chalker, A., Dewar, M. L. & Arnould, J. P. Y. Age-related differences revealed in Australian fur seal *Arctocephalus pusillus doriferus* gut microbiota. *FEMS Microbiol. Ecol.* **86**, 246–255 (2013).
42. Lees, H. *et al.* Age and microenvironment outweigh genetic influence on the Zucker rat microbiome. *PLoS One* **9**, e100916 (2014).
43. Bennett, G. *et al.* Host age, social group, and habitat type influence the gut microbiota of wild ring-tailed lemurs (*Lemur catta*). *Am. J. Primatol.* **78**, 883–892 (2016).
44. Markle, J. G. M. *et al.* Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity. *Science*. **339**, 1084–1088 (2013).
45. Avelo, T., Laakkonen, J. & Jernvall, J. Population and individual level dynamics of intestinal microbiota of a small primate. *Appl. Environ. Microbiol.* **82**, 3537–3545 (2016).
46. Muegge, B. D. *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science*. **332**, 970–974 (2011).
47. Nelson, T. M., Rogers, T. L. & Brown, M. V. The gut bacterial community of mammals from marine and terrestrial habitats. *PLoS One* **8**, e83655 (2013).
48. Archie, E. A. & Tung, J. Social behavior and the microbiome. *Curr. Opin. Behav. Sci.* **6**, 28–34 (2015).
49. Gomez, A. *et al.* Gut microbiome composition and metabolomic profiles of wild western lowland gorillas (*Gorilla gorilla gorilla*) reflect host ecology. *Mol. Ecol.* **24**, 2551–2565 (2015).
50. Delpont, T. C., Power, M. L., Harcourt, R. G., Webster, K. N. & Tetu, S. G. Colony location and captivity influence the gut microbial community composition of the Australian sea lion (*Neophoca cinerea*). *Appl. Environ. Microbiol.* **82**, 3440–3449 (2016).
51. Moeller, A. H. *et al.* Social behavior shapes the chimpanzee pan-microbiome. *Sci. Adv.* **2**, e1500997 (2016).
52. Garcia-Martinez, J., Moya, A., Raga, J. A. & Latorre, A. Genetic differentiation in the striped dolphin *Stenella coeruleoalba* from European waters according to mitochondrial DNA (mtDNA) restriction analysis. *Mol. Ecol.* **8**, 1069–1073 (1999).
53. Natoli, A., Birkun, A., Aguilar, A., Lopez, A. & Hoelzel, A. R. Habitat structure and the dispersal of male and female bottlenose dolphins (*Tursiops truncatus*). *Proc. R. Soc. London B Biol. Sci.* **272**, 1217–1226 (2005).
54. Fontaine, M. C. *et al.* Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in Old World waters. *BMC Biol.* **5**, 30 (2007).
55. Gaspari, S., Azzellino, A., Airoidi, S. & Hoelzel, A. R. Social kin associations and genetic structuring of striped dolphin populations (*Stenella coeruleoalba*) in the Mediterranean Sea. *Mol. Ecol.* **16**, 2922–2933 (2007).
56. Foster, G. *et al.* *Phocoenobacter uteri* gen. nov., sp. nov., a new member of the family Pasteurellaceae Pohl (1979) 1981 isolated from a harbour porpoise (*Phocoena phocoena*). *Int. J. Syst. Evol. Microbiol.* **50**(Pt 1), 135–9 (2000).
57. Lee, W.-H. *et al.* Bacterial alterations in salivary microbiota and their association in oral cancer. *Sci. Rep.* **7**, 16540 (2017).
58. Crielaard, W. *et al.* Exploring the oral microbiota of children at various developmental stages of their dentition in the relation to their oral health. *BMC Med. Genomics* **4**, 22 (2011).
59. Allen-Vercos, E., Strauss, J. & Chadee, K. Fusobacterium nucleatum: an emerging gut pathogen? *Gut Microbes* **2**, 294–298 (2011).
60. Al-hebshi, N. N. *et al.* Inflammatory bacteriome featuring Fusobacterium nucleatum and Pseudomonas aeruginosa identified in association with oral squamous cell carcinoma. *Sci. Rep.* **7**, 1834 (2017).
61. Stoddard, R. A. *et al.* Risk factors for infection with pathogenic and antimicrobial-resistant fecal bacteria in northern elephant seals in California. *Public Health Rep.* **123**, 360–70 (2008).

62. Lee, C., Agidi, S., Marion, J. W. & Lee, J. *Arcobacter* in Lake Erie beach waters: an emerging gastrointestinal pathogen linked with human-associated fecal contamination. *Appl. Environ. Microbiol.* **78**, 5511–9 (2012).
63. Fera, M. T. *et al.* Detection of *Arcobacter* spp. in the coastal environment of the Mediterranean Sea. *Appl. Environ. Microbiol.* **70**, 1271–1276 (2004).
64. Das, M., Badley, A. D., Cockerill, F. R., Steckelberg, J. M. & Wilson, W. R. Infective endocarditis caused by HACEK microorganisms. *Annu Rev Med* **48**, 25–33 (1997).
65. Ormerod, K. L. *et al.* Genomic characterization of the uncultured Bacteroidales family S24-7 inhabiting the guts of homeothermic animals. *Microbiome* **4**, 36 (2016).
66. Frantzi, A. & Herzog, D. L. Mixed-species associations of striped dolphins (*Stenella coeruleoalba*), short-beaked common dolphins (*Delphinus delphis*), and Risso's dolphins (*Grampus griseus*) in the Gulf of Corinth (Greece, Mediterranean Sea). *Aquat. Mamm.* **28**, 188–197 (2002).
67. Quérrouil, S. *et al.* Why do dolphins form mixed-species associations in the Azores? *Ethology* **114**, 1183–1194 (2008).
68. Miyazaki, N. & Nishiwaki, M. School structure of the striped dolphin off the Pacific coast of Japan. *Sci. Reports Whales Res. Institute, Tokyo* **30**, 65–115 (1978).
69. Evans, W. E. Common dolphin, white-bellied porpoise *Delphinus delphis* Linnaeus, 1758. *Handb. Mar. Mamm.* **5**, 191–224 (1994).
70. Kuiken, T. & Hartmann, M. G. Cetacean pathology: dissection techniques and tissue sampling. *Eur. Cetacean Soc. Newsl.* **17**, 43 (1991).
71. Pugliare, K. R. *et al.* *Marine mammal necropsy: an introductory guide for stranding responders and field biologists*, <https://doi.org/10.1575/1912/1823> (Woods Hole Oceanographic Institution, 2007).
72. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **AEM**, 01043–13 (2013).
73. Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)[Software]. Available at, <http://github/najoshi/sickle> (2011).
74. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
75. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–6 (2013).
76. Andersen, K. S. S., Kirkegaard, R. H., Karst, S. M. & Albertsen, M. ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. *bioRxiv* 299537 (2018).
77. Chao, A. & Jost, L. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* **93**, 2533–2547 (2012).
78. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**, 927–930 (2003).
79. Legendre, P. & Gallagher, E. D. Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**, 271–280 (2001).
80. Frade, P. R., Roll, K., Bergauer, K. & Herndl, G. J. Archaeal and bacterial communities associated with the surface mucus of Caribbean corals differ in their degree of host specificity and community turnover over reefs. *PLoS One* **11**, e0144702 (2016).
81. Kuczynski, J. *et al.* Using QIIME to analyze 16S rRNA gene sequences from microbial communities. *Curr. Protoc. Bioinformatics* Chapter 10, Unit 10.7. (2011).
82. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
83. McDonald, D. *et al.* The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *Gigascience* **1**, 7 (2012).
84. Lê Cao, K.-A., Rohart, F., Gonzalez, I. & Dejean, S. mixOmics: Omics Data Integration Project. R package Version 6.1. 1. *PLoS Comput Biol* **13**, e1005752 (2017).
85. Lê Cao, K.-A. *et al.* MixMC: a multivariate statistical framework to gain insight into microbial communities. *PLoS One* **11**, e0160169 (2016).
86. Dufréne, M. & Legendre, P. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol. Monogr.* **67**, 345–366 (1997).
87. Roberts, D. W. labdsv: ordination and multivariate analysis for ecology, R package version 1.5–0. 2012. (2016).
88. Dhariwal, A. *et al.* MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res.* **45**, W180–W188 (2017).
89. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011).
90. Szafranski, S. P. *et al.* High resolution taxonomic profiling of the subgingival microbiome for biomarker discovery and periodontitis diagnosis. *Appl. Environ. Microbiol.* **AEM**, 03534–14 (2014).
91. Kennedy, R. *et al.* The microbiome associated with equine periodontitis and oral health. *Vet. Res.* **47**, 49 (2016).
92. Gomez, D. E., Arroyo, L. G., Costa, M. C., Viel, L. & Weese, J. S. Characterization of the fecal bacterial microbiota of healthy and diarrheic dairy calves. *J. Vet. Intern. Med.* **31**, 928–939 (2017).
93. Dong, L. *et al.* Microbial similarity and preference for specific sites in healthy oral cavity and esophagus. *Front. Microbiol.* **9** (2018).

Acknowledgements

The authors would like to acknowledge to Jörg Becker and João Sobral for the sequencing services provided at the Instituto Gulbenkian de Ciência. This work was supported by the project CetSenti RECI/AAGGLO/0470/2012 (FCOMP-01-0124-FEDER-027472), by the strategic program UID/BIA/04050/2013 (POCI-01-0145-FEDER-007569), by the GenomePT project (POCI-01-0145-FEDER-022184), supported by COMPETE 2020 - Operational Programme for Competitiveness and Internationalisation (POCI), Lisboa Portugal Regional Operational Programme (Lisboa2020), Algarve Portugal Regional Operational Programme (CRESC Algarve2020), under the PORTUGAL 2020 Partnership Agreement, funded by national funds through the Fundação para a Ciência e a Tecnologia (FCT) I.P. and the European Regional Development Fund (ERDF), through the COMPETE2020-Programa Operacional Competitividade e Internacionalização (POCI). The work was supported as well through CESAM UID/AMB/50017/2013 co-funded by the FCT/MEC and FEDER within PT2020 and Compete 2020 (C.E.), FCT grants SFRH/BD/30240/2006 (M.F.), and SFRH/BPD/82407/2011 (A.L.) and NIH RCMI Grant No. 2U54MD007600 (F.G.V.).

Author Contributions

J.V., C.E., P.M.S. designed the study. M.F., P.C., A.L., J.V., C.E. were involved in the animal rescue and swab collection. H.A.-R. extracted the DNA from swabs and prepared samples for high-throughput sequencing. P.S.-C. and P.M.S. analyzed the sequencing data. P.S.-C., F.G.-V., J.V., C.E., P.M.S. drafted the manuscript. All authors read, revised and approved the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-50139-7>.

Competing Interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019