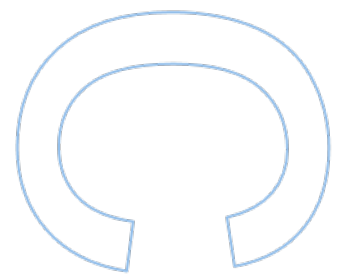
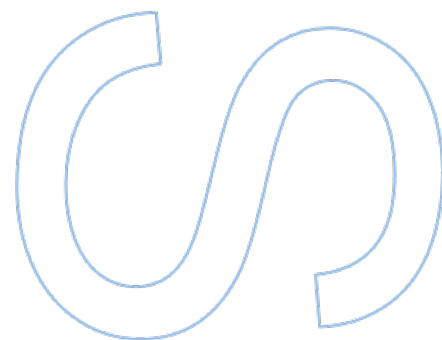
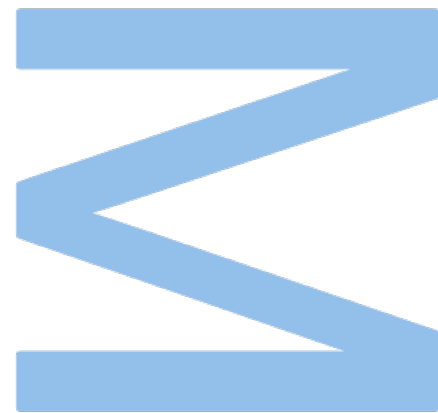




Dive into the blow microbiome: exploring microbial diversity of the respiratory tract in free-ranging cetaceans



Beatriz Santos Silva

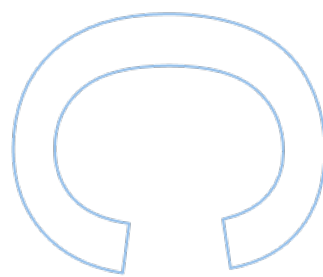
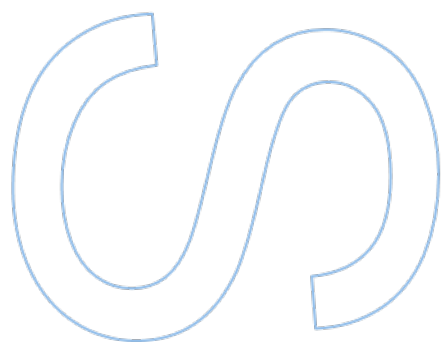
Aquatic Biological Resources
Biology Department
2022

Supervisor

Maria Paola Tomasino, Postdoc Researcher, CIIMAR

Co-supervisor

Ana Mafalda Correia, Postdoc Researcher, CIIMAR
Catarina Magalhães, Invited Auxiliar Professor, CIIMAR



Declaração de Honra

Eu, Beatriz Santos Silva, inscrito(a) no Mestrado em Recursos Biológicos Aquáticos da Faculdade de Ciências da Universidade do Porto declaro, nos termos do disposto na alínea a) do artigo 14.º do Código Ético de Conduta Académica da U.Porto, que o conteúdo da presente dissertação reflete as perspetivas, o trabalho de investigação e as minhas interpretações no momento da sua entrega.

Ao entregar esta dissertação, declaro, ainda, que a mesma é resultado do meu próprio trabalho de investigação e contém contributos que não foram utilizados previamente noutros trabalhos apresentados a esta ou outra instituição.

Mais declaro que todas as referências a outros autores respeitam escrupulosamente as regras da atribuição, encontrando-se devidamente citadas no corpo do texto e identificadas na secção de referências bibliográficas. Não são divulgados na presente dissertação quaisquer conteúdos cuja reprodução esteja vedada por direitos de autor.

Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

Beatriz Santos Silva

Porto, 30 de Setembro de 2022

Acknowledgments

Primeiro de tudo, gostaria de fazer um agradecimento especial e sincero à minha orientadora Maria Paola, por toda a ajuda e confiança depositada no meu trabalho e pela disponibilidade. Também gostaria de agradecer às minhas co-orientadoras Mafalda Correia e Catarina Magalhães, pelo apoio e guia dados durante a preparação desta dissertação. O meu trabalho nunca teria sido possível sem a vossa ajuda.

Quero agradecer também aos membros da equipa do EcoBioTec pela simpatia e boa disposição. Além disso, um grande obrigada a todos os meus colegas e amigos de mestrado de Recursos Biológicos Aquáticos pela entreajuda e motivação.

Um grande obrigada a todos os meus amigos (e companheiro de vida) que me acompanham e me apoiam em todas as minhas decisões. Às minhas amigas de há muitos anos, obrigada por de uma forma ou de outra, estarem presentes em todas as etapas da minha vida. Sem enumerar nomes, vocês são muitos, mas sabem quem são... quero agradecer ao eterno e maravilhoso BDL. de onde guardo tão maravilhosos momentos. Ainda bem que apareceram no meu caminho, definitivamente este percurso não seria o mesmo sem a vossa presença e ajuda.

Finalmente, quero agradecer à minha família por estarem sempre ao meu lado em todas as minhas escolhas e apoiarem-me a seguir em frente. Obrigada por terem sempre acreditado nas minhas capacidades. Não poderia deixar de agradecer aos meus cães por toda a companhia, em especial ao Piruças que acompanhou muito da escrita deste documento no meu colo.

Todos guardam um lugar especial no meu coração. A todos vocês um enorme obrigada por tudo!

Agradeço também aos projetos que financiaram este projeto, nomeadamente o projeto ATLANTIDA (ref. NORTE-01-0145- FEDER-000040) e o projeto OCEAN3R (ref.NORTE-01-0145-FEDER-000064), suportado pelo Programa Operacional Regional do Norte (NORTE 2020), no âmbito do Acordo de Parceria PORTUGAL 2020 e através do Fundo Europeu de Desenvolvimento Regional (FEDER).



Resumo

O clima da Terra está a mudar a uma taxa sem precedentes, em parte devido a impactos relacionados com o clima e a impactos antropogénicos. Os cetáceos são espécies-chave e sentinelas dos ecossistemas marinhos e da saúde dos oceanos, considerados espécies estratégicas para avaliar o bem-estar dos ecossistemas aquáticos e identificar tendências ambientais prejudiciais. As doenças respiratórias são uma das principais causas de morte nos cetáceos. Caracterizar os microbiomas do trato respiratório pode constituir uma indicação adequada da saúde destes animais. Para isso é necessário estabelecer uma base de dados dos microrganismos comumente associados a mamíferos marinhos a fim de detetar anomalias. A análise do microbioma presente no sopro tem sido proposta como um biomarcador para análises do estado de saúde respiratório de cetáceos.

No presente estudo, baleias-piloto de barbatana curta (*Globicephala macrorhynchus*), residentes e visitantes regulares da Ilha da Madeira, foram usadas como espécie modelo para validar o conjunto de primers mais adequado para explorar a diversidade microbiana do trato respiratório. As amostras de sopro e de controlo de água foram recolhidas em 2018, o DNA foi extraído das amostras e depois sequenciado para amplificar as regiões hipervariáveis V3-V4/V4-V5 do gene 16S.

Os resultados demonstraram que, independentemente do conjunto de primers utilizado, todas as baleias-piloto de barbatana curta amostradas compartilham na sua composição os filos *Actinobacteria*, *Bacteroidetes*, *Firmicutes* e *Proteobacteria*. O conjunto de dados V4-V5 capturou a maior riqueza microbiana (biosfera microbiana rara e táxons potencialmente patogénicos), enquanto que o V3-V4 teve mais diversidade. A análise V4-V5 das amostras de sopro e de água do mar demonstrou que as amostras de sopro apresentaram maior diversidade do que a água do mar, no entanto as diferenças não foram significativas. 17,6% das espécies encontradas são cosmopolitas, estando presentes tanto no sopro como na água, enquanto que 58,9% das espécies são específicas do sopro.

Este estudo comprovou que a caracterização das comunidades microbianas associadas à respiração pode ser usada como um futuro biomarcador para o estado fisiológico das vias aéreas de cetáceos, no entanto, esta técnica ainda requer consideração e desenvolvimento futuro para uma maior profundidade.

Palavras-chave: Cetáceos, sopro, infecções respiratórias, amostragem não invasiva, metabarcoding

Abstract

The climate of the earth is changing at an unprecedented rate, in part due to climate-related and anthropogenic-induced impacts. Cetaceans are key species and sentinels of the marine ecosystems and ocean health, being a strategic group to evaluate the well-being of aquatic ecosystems and to detect harmful environmental trends. Respiratory disease is one of the main causes of death in these animals. Characterizing microbiomes related with their respiratory tract may constitute a suitable indication of health status. For that, it's necessary to establish baseline data on microorganisms commonly associated with marine mammals to detect anomalies. The study of microbiomes present in their exhaled breath condensates (EBCs), or blows, have been proposed as a biomarker for assessing respiratory health.

In the present study, the short-finned pilot whales (*Globicephala macrorhynchus*), both residents and regular visitors of Madeira Island, were used as a model species to validate the most suitable primer set to explore the microbial diversity of the respiratory tract. Blow and control water samples were collected in 2018, DNA was extracted from samples and then sequenced to amplify V3-V4/V4-V5 hypervariable regions of 16S rRNA gene.

Results showed that independently of the primer set used all short-finned pilot whales sampled share in their composition *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, and *Proteobacteria* phyla. The V4-V5 dataset captured the higher microbial richness (microbial rare biosphere and potentially pathogenic taxa), whereas V3-V4 had more diversity. The V4-V5 analysis of blow and seawater samples showed that blow samples had higher diversity than seawater, however differences were not significant. 17,6% of species are cosmopolitan, being present either in the blow or in the water. However, 58,9% of species are specific to blow.

This study showed that the characterization of the respiratory-associated microbial communities can be used as a future biomarker for the physiological state of the airways in cetaceans, however, several aspects require consideration and future development to further advance this technique.

Keywords: Cetaceans, blow, respiratory infections, non-invasive sampling, metabarcoding

Table of Contents

List of Tables.....	ix
List of Figures.....	xi
List of Abbreviations.....	xiii
1. Introduction	1
1.1. Threats to Marine Environment.....	1
1.2. From micro to macro in the ocean: host-microbiome interactions.....	2
1.2.1. Core taxa	4
1.3. Microbial communities' identification	5
1.4. 16S rRNA Metabarcoding.....	6
1.5. Microbial communities in cetaceans in a changing ocean.....	7
1.5.1. Respiratory infections in cetaceans	8
1.6. Infraorder Cetacea.....	10
1.6.1. The ecological importance of Cetaceans.....	11
1.6.2. The Exhale Breath Condensate of Cetaceans	13
1.7. Sampling Cetaceans	16
1.7.1. Blow sampling.....	18
1.8. Target Species: Short-finned pilot whale.....	20
1.8.1. Short-finned pilot whale off Madeira Archipelago	21
1.8.2. Conservation status and threats to the short-finned pilot whale	22
2. Objectives	23
3. Material and Methods	24
3.1. Sampling.....	24
3.2. DNA Extraction	25
3.3. PCR amplification, Library Preparation, and Quality Control - Optimization of the PCR Amplification protocol.....	26
3.4. 16S rRNA Amplicon Sequencing.....	29
3.5. Bioinformatics analyses.....	29

3.5.1.	Upstream sequence analysis.....	29
3.5.2.	Downstream sequence analysis.....	30
3.5.3.	Pathogen database and screening.....	31
4.	Results.....	32
4.1.	Comparison of 16S rRNA primers for studies of blow prokaryotic communities	32
4.1.1.	Alpha diversity.....	34
4.1.2.	Beta diversity.....	36
4.1.3.	Characterization of the prokaryotic community in V3-V4 and V4-V5 dataset	38
4.1.4.	Cetacean respiratory pathogens detected in the short-finned pilot whale blow	44
4.2.	Comparison between the blow of short-finned pilot whales and seawater samples.....	45
4.2.1.	Alpha diversity.....	47
4.2.2.	Beta diversity.....	48
4.2.3.	Characterization of the prokaryotic community in short-finned pilot whale blow and seawater samples.....	49
4.2.4.	Venn Diagram Analysis.....	51
5.	Discussion.....	52
6.	Conclusions and Futures Perspectives.....	60
7.	Scientific Communications (based on the work of this thesis).....	62
8.	References.....	63
9.	Supplementary information.....	80

List of Tables

Table 1 - Number of cetacean species, subspecies, and total in the different IUCN (International Union for Conservation of Nature) categories: critically endangered, endangered, vulnerable, near threatened, least concern, data deficient.	11
Table 2 - State of the art of studies focused on the use of Exhale Breath Condensate (blow) for analyzing microorganisms associated with the respiratory tract of cetaceans.	14
Table 3 - Different primers tested in the optimization of the PCR (polymerase chain reaction) Amplification protocol.	26
Table 4 - Different Taq polymerase tested in the optimization of the PCR (polymerase chain reaction) Amplification protocol.	27
Table 5 - Metadata of the blow and seawater samples analyzed.	28
Table 6 - Cleaning Process of the 16S rRNA sequences for each sample of the blow of short-finned pilot whales.	33
Table 7 - Minimum, maximum, median, mean, and standard deviation of the different alpha diversity indices in V3-V4 and V4-V5 datasets, of short-finned pilot whale blow samples.	35
Table 8 - Output of a Wilcoxon Signed-Rank Test, comparing the indexes of the short-finned pilot whale blow community richness obtained with two different primer sets (V3-V4 and V4-V5).	36
Table 9 - Output of the PERMANOVA analysis (Df=degrees of freedom, SS=sum of squares, R ² = coefficient of determination, PR(>F)=p-value) to test for the influence of the primer sets (V3-V4 and V4-V5) on the distribution of prokaryotic communities of the blows of short-finned pilot whales.	37
Table 10 - Cleaning process of the 16S rRNA sequences for each sample of blow of short-finned pilot whale and seawater.	46
Table 11 - Output of a Wilcoxon rank sum Test, comparing indexes of alpha diversity for microbial community, between the blows of short-finned pilot whale and seawater samples (using the primer set V4-V5).	47
Table 12 - Minimum, maximum, median, mean, and standard deviation of the different measures of alpha diversity in the prokaryotic communities of the short-finned pilot whale blow and seawater samples (using the primer set V4-V5).	48
Table 13 - Output of the PERMANOVA (Df=degrees of freedom, SS=sum of squares, R ² = coefficient of determination, PR(>F)=p-value)analysis to test for the influence of	

different types of samples (blow of short-finned pilot whales/seawater samples) on the distribution of prokaryotic communities (using the primer set V4-V5). 49

Table S 1 – DNA concentration (ng/μl), for each sample of the short-finned pilot whale blows and sterivex seawater control samples used in the analysis. 80

Table S 2 - Calculated alpha diversity indices, for each sample of the short-finned pilot whale blows, of the comparison of V3-V4 and V4-V5 primers sets. 81

Table S 3 - The proportion of all phyla present in the V4-V5 and V3-V4 datasets, for samples of the short-finned pilot whale blows.....81

Table S 4 - The proportion of all classes present in the V4-V5 and V3-V4 datasets, for samples of the short-finned pilot whale blows.....82

Table S 5 - The proportion of all families present in the V4-V5 and V3-V4 datasets, for samples of the short-finned pilot whale blows.....83

Table S 6 - The proportion of all genera present in the V4-V5 and V3-V4 datasets, for samples of the short-finned pilot whale blows.....86

Table S 7 - Overview of ASV composition of bacterial families in both datasets. The table consists of all bacterial families that were observed in both V3-V4 and V4-V5 datasets, for samples of the short-finned pilot whale blows, and had a sequence proportion of at least 1%. 90

Table S 8 - Calculated alpha diversity indices for each sample of short-finned pilot whale blows and respective control seawater samples, using the primer set V4-V5. 91

List of Figures

Figure 1 - Illustration of the classification of the interactions between animal and microbiome, retrieved from Apprill et al., 2017.....	3
Figure 2 - Illustration of male (top) and female (bottom) short-finned pilot whale. Retrieved from Jefferson, T. A., Webber, M. A., & Pitman, (2015).	20
Figure 3 - Map with the location of Madeira Island, Portugal (in red) (Google maps)...	24
Figure 4 - Photograph of blow collection using the Performagene sample collection kit, Ágatha Gil.	25
Figure 5 - Rarefaction curves of the V3-V4 dataset (on the left) and the V4-V5 dataset (on the right), for short-finned pilot whale blow samples.	34
Figure 6 - Observed richness, Chao1, Shannon, and Inverse Simpson indices of alpha diversity among blow microbiome. Different primer sets were represented by different shapes.....	35
Figure 7- Non-metric multidimensional scaling (nMDS) (stress=0,07) of the microbiota found in short-finned pilot whale blow samples, in merged V3-V4 and V4-V5 datasets, based on Bray-Curtis Dissimilarity. The primer sets used are represented by different shapes and the days of sampling by different colors.	37
Figure 8 - Relative abundance of prokaryotic phyla identified across the short-finned pilot whale blow samples, by V3-V4 (on the left) and V4-V5 (on the right) datasets.	38
Figure 9 - Relative abundance of the top 10 prokaryotes classes identified across the short-finned pilot whale blow samples (V3-V4 dataset on the top and V4-V5 on the bottom).	40
Figure 10 - Relative abundance of the top 10 prokaryotes families identified across the short-finned whale blow samples (V3-V4 dataset on the top and V4-V5 on the bottom).	41
Figure 11 - Major taxonomic families obtained for the short-finned pilot whale blow samples (in V3–V4 and V4–V5 datasets). The total sequence proportion of each family in the V3–V4 (left side) and V4–V5 (right side) datasets is represented in the x-axis. The numbers in each column represent the number of observed ASVs affiliated with each taxonomic family. Color coding represents the different taxonomic classes. Only families that comprised at least 1% of sequences, in at least one of the datasets, were included in the figure.	42
Figure 12 - Relative abundance of the top 10 prokaryotic genera identified across the short-finned pilot whale blow community (V3-V4 dataset on the top and V4-V5 on the bottom).	43

Figure 13 - Venn diagram of sequences abundances of all ASVs (Amplicon Sequence Variants) of V3-V4 (red) and V4-V5 (yellow) datasets at genus rank, for short-finned pilot whale blow samples.	44
Figure 14 - Venn diagram analysis of all potential pathogenic genera detected in V4-V5 (blue) and V3-V4 (red) datasets, for short-finned pilot whale blow samples.	45
Figure 15 - Rarefaction curves of the V4-V5 dataset, for short-finned pilot whale blow and seawater samples.	46
Figure 16 - Observed richness, Shannon, and Simpson indices of alpha diversity among blow and seawater microbiomes, amplified with the V4-V5 primer set. Blow samples (brown) are represented on the left and seawater control samples (orange) are on the right.	47
Figure 17 - Non-metric multidimensional scale (nMDS) (stress=0,04) of the microbiota found in short-finned pilot whale blow and sterivex samples, based on Bray-Curtis dissimilarity. The different types of samples are represented by colours, red for blow and blue for seawater.	48
Figure 18 - Relative abundance of prokaryotes phyla identified across the short-finned pilot whale blow and seawater samples separated by three sampling days (11/09/2018; 11/20/2018; 12/12/2018) – using the primer set V4-V5.	50
Figure 19 - Relative abundance of the top 10 prokaryotic genera identified across the short-finned pilot whale blow and seawater samples separated by three sampling days (11/09/2018; 11/20/2018; 12/12/2018) – using the primer set V4-V5.	51
Figure 20 - Venn Diagram analysis of short-finned pilot whale blows (red) and seawater sterivex samples (blue).	51

List of Abbreviations

16S rRNA	16S ribosomal ribonucleic acid
18S rRNA	18S ribosomal ribonucleic acid
ASV	Amplicon Sequence Variant
BOLD	Barcode of Life Data System
Bp	Base pair(s)
CIIMAR	Interdisciplinary Center for Marine and Environmental Research
DNA	Deoxyribonucleic acid
EBC	Exhaled breath condensate
EEZ	Exclusive Economic Zone
EU	European Union
GTDB	Genome Taxonomy Database
IUCN	International Union for Conservation of Nature
n/a	Not available
NGS	Next Generation Sequencing
nMDS	Non-metric multidimensional scale
OTU	Operational taxonomic unit
PCBs	Polychlorinated biphenyls
PCR	Polymerase chain reaction
RDP	Ribosomal Database Project
UAVs	Unmanned aerial vehicles
URT	Upper respiratory tract

1. Introduction

1.1. Threats to Marine Environment

Anthropogenic-induced and climate-related impacts are changing the climate of the earth at an unprecedented rate (Doney et al., 2012; Halpern et al., 2008). In general, the temperature of the planet is increasing, sea ice and glaciers are receding, sea levels are rising, and pollutants are accumulating in the environment and within organisms (Moore, 2008). These physical global changes, in addition to habitat degradation, overexploitation, overfishing coastal sprawl, species redistributions with the introduction of exotic species, ocean acidification, pollution, and eutrophication, among other impacts, affect and alter the marine ecosystem worldwide. The sustainability, resilience, and biodiversity of marine environments are becoming deteriorated (Halpern et al., 2007).

Cetaceans are sentinel species highly affected by these ecosystem changes, and, as top predators, they suffer the consequences of bioaccumulation and biomagnification (Roman et al., 2014). These impacts deteriorate their health, compromise their immune system, and make individuals more susceptible to diseases and pathogens, which can then terminate in substantial mortality and morbidity rates (Boonstra et al., 2015; Bossart, 2011; Duarte, 2014; Halpern et al., 2007; Sehnaal et al., 2021; Worm et al., 2006). Often, causes of disease and death are correlated with secondary infections in the respiratory system (Raverty et al., 2017). Lately, reports of mass mortalities caused by bacterial pathogens and disease outbreaks have increased, leading to population declines (Harkonen et al., 2008; Jepson et al., 2013; Nelson et al., 2015; Omata et al., 2006; Pamer, 2007). Respiratory diseases like pneumonia are one of the principal causes of mortality in both captive and wild cetaceans (Nabi et al., 2020; Venn-Watson et al., 2012).

Cetacean species are ideal sentinels for human coastal populations and ocean health, as they have similarities in the bioaccumulation of pollutants and susceptibility to infectious disease (Bossart, 2007; Roman et al., 2014). The monitoring of free-ranging cetaceans provides a valuable method for assessing the well-being of aquatic ecosystems and identifying adverse environmental trends, enabling the comparison of stressors between geographic locations and related changes in individual and population health status (Barratclough et al., 2019; Bossart, 2011). As a result, these alerts can

assist specialists to describe and manage the impact of these trends on the ocean or human health, in order to provide solutions to mitigate threats and protect these species, as well as know populations' long-term survival and reproductive success (Bik et al., 2016; Bossart, 2007, 2011; Moore, 2008). Therefore, the monitoring of these animals can act as a biological primary warning system, alerting humans to hazards earlier than outbreak of human diseases (Stewart et al., 2014). Furthermore, conservation needs and objectives highlighted in important EU (European Union) directives enhance the requirement of expanding the information about the biology and ecology of cetaceans (Centelleghé et al., 2020).

1.2. From micro to macro in the ocean: host-microbiome interactions

The ocean is inhabited by millions of organisms, ranging from marine mammals to microorganisms. Microbial communities have a key role in animal health and provide crucial functions to the earth and ocean (oxygen production, nutrient cycling, and organic matter degradation), being very important to ecosystem stability (Ainsworth & Gates, 2016; Apprill, 2017; Fischbach & Segre, 2016). The composition of the microbial communities present in water is variable over time, geography, and under different environmental conditions. Nevertheless, their various functions are conserved globally (Sehna et al., 2021).

The microbiome is composed of the genetic information and inferred physicochemical features of the gene products of a microbiota (membership of microbial communities linked with a specific habitat) (Fañer et al., 2017). Microorganisms live in diverse anatomical areas of cetaceans, colonizing their skin, mucosa, and intestines. Each of these host environments supports a particular microbial community, with only a few microorganisms being privileged above others (Tlaskalová-Hogenová et al., 2004). The relationship between the microorganisms present in water and aquatic organisms is not considerably known (Sehna et al., 2021). There is evidence that the composition of environmental microbial communities can influence the health of the animals, acting as a vector for several infectious microbes. However, the host-associated microbiome is under various environmental and biological selection pressures (nutritional status, stress, disease), hence the aquatic animal microbiomes is not a direct consideration of the near water microbiome (Ainsworth & Gates, 2016; Sehna et al., 2021).

Although the composition of the microbial community of an individual host is relatively steady, the relative proportion of each microorganism in each host is dynamic (Cardona et al., 2018). It is suggested that cetaceans and their microbial symbionts have formed close bonds during millions of years of co-evolution (Johnson et al., 2009; Nelson et al., 2015). The microbial community can straightly influence disease progression through infection, as well as adjust the regulation and response of the host's immune system (Nelson et al., 2015). Cardona and colleagues (2018) suggested that the interactions of cetaceans' respiratory tract and gastrointestinal epithelium with air- and food-associated microorganisms have the greatest impact on the host-microbiota dynamics.

Host-microbiome dynamics are divided into two classifications (Figure 1): 1. dysbiosis, in which relationships or interactions are disrupted and may be linked to large stress or infection events; 2. symbiosis, in which organisms participate in regular metabolic and immunological signal interactions (Aprill, 2017).

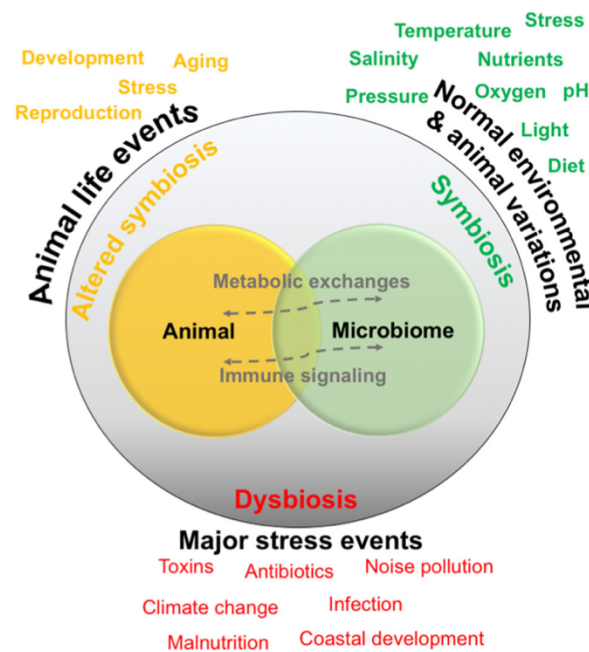


Figure 1 - Illustration of the classification of the interactions between animal and microbiome, retrieved from Aprill et al., 2017.

The normal microbiota, also called as resident or indigenous, includes archaea, bacteria, viruses, protozoa, and fungi. It is founded on the interaction between the host and the environment and is estimated that over 1000 bacterial species could be identified. Microbiota can be transitory or permanent. In the microbiota are considered microorganisms from the respiratory, urogenital, and gastrointestinal tract, oral cavity, and skin of healthy animals (Avalos-Téllez et al., 2010; Sørum & Sunde, 2001;

Taskalová-Hogenová et al., 2004). The host-microbiota composition is determined by factors like specie, age, diet, sexual comportment, stress, and hormones of the animals. In the case of captive individuals of cetaceans, the use of antibiotics and cleaning approaches also influences the microbiota (Avalos-Téllez et al., 2010; Taskalová-Hogenová et al., 2004). Microorganisms have evolved to coexist with their hosts without inflicting negative impacts, and their coexistence is usually beneficial. They are non-pathogens as residents of a specific area, however, when animals are exposed to altered conditions (physiological, immunological, environmental, or nutritional), they may become opportunistic pathogens as the microorganisms move to another area and can cause problems to the animals. This often happens in stressed, immunosuppressed, or in individuals under antimicrobial medical treatment (Avalos-Téllez et al., 2010; Sørum & Sunde, 2001; Taskalová-Hogenová et al., 2004).

1.2.1. Core taxa

The core taxa are microbiota members that exist in a single or most hosts of a particular species and have remained persistent and abundant taxa, across time or multiple habitats, hosts, or geographic areas (Bäckhed et al., 2005; Hooper et al., 2002; Nelson et al., 2015; C. Vendl et al., 2019). The determination of the core taxa is fundamental to portray the importance of the complex and diverse microbial communities in the host (Shade & Handelsman, 2012). Risely and colleagues (2020) considered five types of core microbiome: common, temporal, ecological, functional, and host-adapted. The common core is found through a huge amount of hosts belonging to a specific host population or species. The temporal core is stable or predictable over time. The ecological core is the part of the microbiome that has a big importance in the organization and shaping of the ecological community. The functional core has essential biological functions for the host (physiological, biochemical, and ecological services). Finally, the host-adapted core is a microbe collection that has co-evolved with the host species or sub-population and whose presence improves the host's fitness (Risely, 2020).

A persistent abundance or highly prevalent core taxa is proof of their beneficial nature towards their host and represents a stable and healthy microbial community as these bacteria interact and benefit, and can even be crucial, to the immune system of the host. Identifying a bacterial core may be a significant step in characterizing a healthy microbial

community (Apprill et al., 2017; Hernandez-Agreda et al., 2017; Nelson et al., 2015; Shade & Handelsman, 2012; C. Vendl et al., 2019; Waltzek et al., 2012).

1.3. Microbial communities' identification

Until recently, the majority of worldwide microbial diversity was unavailable or underestimated by isolation and culture-dependent methods, which are neither universal nor unsuitable for the growth of most microorganisms. Traditional media can only be used to cultivate less than 1% of bacteria, and isolation techniques are very time-consuming (Armougom, 2009; Callewaert et al., 2020; Johnson et al., 2009). Bacterial communities associated with the respiratory tract of cetaceans remain to be described, with current understanding almost entirely derived from clinical isolation investigations. Molecular analyses often reveal higher diversity and novel taxa of bacterial communities than culture-based methods (Johnson et al., 2009).

Thanks to multidisciplinary approaches, such as the discovery of cultivation-independent, and advances in high-throughput technology provided by next-generation sequencing (NGS) technologies and their commercialization, bioinformatics, and analytics, the number of studies investigating the diversity of whole microbial communities has increased. NGS has yielded powerful new tools in terms of time-saving, cost-effectiveness, and data production capability (Armougom, 2009; Ranjan et al., 2016).

The most commonly used sequencing method for analyzing the microbiome is amplicon analysis of the 16S rRNA (16S ribosomal ribonucleic acid) gene. In this approach, a 16S rRNA region is amplified by PCR (polymerase chain reaction), using primers that target highly conserved regions of the gene, and then it is sequenced using NGS platforms like 454 pyrosequencing, Ion Torrent, and Illumina (HiSeq and MiSeq) (Ranjan et al., 2016). It has been shown that the various sequencing platforms give distinctive results, especially due to the specific variable region of the 16S gene chosen, primers used, and the length of the amplicons retrieved by the different platforms. Depending on the type of sequencing technology, if longer reads are used, sequence errors will be reduced (Faner et al., 2017). The selection of the PCR primers and the hypervariable region targeted for DNA (deoxyribonucleic acid) amplification is an important and crucial step (Armougom, 2009; Wang & Qian, 2009). Each hypervariable region produces different

taxonomic resolutions and behaves differently in different niches (Graspeuntner et al., 2018; Willis et al., 2019). Primers (15-20 base-pair [bp] nucleotide sequences) are located in conservative regions and are utilized to create amplicons of variant regions that are illuminating for taxonomic assignment (Wang & Qian, 2009). Efforts have been made to design more universal 16S rRNA gene primer sets that can cover almost the full diversity of natural microbial communities (e.g Earth Microbiome Project; Caporaso et al., 2012; Gilbert et al., 2014). For the investigation of marine microbiomes, the primer set 515YF/Y926R-jed was designed by Caporaso et al., (2011, 2012) to target the V4–V5 hypervariable regions of the 16S rRNA gene, which can capture the diversity of the archaeal communities. It was later modified by Apprill et al., (2015) and Parada et al., (2016). The aforementioned primers have a degeneracy to cover a broad spectrum of diversity, specifically the *Crenarchaeota/Thaumarchaeota* (degeneracy at 515YF) phylum and the clade SAR11 (alphaproteobacterial class; degeneracy at Y906R-jed). The V3-V4 hypervariable regions is suggested for marine bacteria (Klindworth et al., 2013).

1.4. 16S rRNA Metabarcoding

Metabarcoding is a powerful tool for the identification of multiple taxa from a mixed environmental sample (Kress & Erickson, 2008). This method relies on a short section of DNA (called DNA barcode) from a specific gene marker (18S rRNA gene for eukaryotes and 16S rRNA for prokaryotes), on techniques such as NGS, PCR (Polymerase chain reaction), and bioinformatic tools (Armougom, 2009; Bourlat et al., 2013). Bioinformatics pipelines are used to perform quality checks, remove low-quality reads, merge paired-sequence ends, and check for chimeras. Depending on the pipeline and clustering algorithms being used, sequence reads are grouped into clusters named amplicon sequence variants (ASVs) or operational taxonomic units (OTUs). Once sequencing clusters are produced, they can be investigated directly (often done for bacterial or fungal diversity), or they can be associated with taxonomic identifiers utilizing relevant reference databases (Hebert et al., 2003). The DNA sequencing reference databases contain the DNA barcode sequences for different species. The most commonly used databases are Barcode of Life Data System (BOLD), SILVA, Genome Taxonomy Database (GTDB), Ribosomal Database Project (RDP) and NCBI. The metabarcoding approach can be used to assign unknown individuals to species or to

improve the discovery of new species by performing large-scale screening of one or a few reference genes (Moritz & Cicero, 2004). Species detection using DNA metabarcoding is considered trustworthy (Yang et al., 2014), however, it is subjected to some biases, like the incompleteness of reference libraries; choice of DNA barcode (marker); selection and efficacy of PCR primers (Pawlowski et al., 2016).

The 16S rRNA gene is present in all bacteria and archaea; even eukaryotes have it in mitochondria and chloroplasts. This barcode region is the component of a 30S small subunit of prokaryotic ribosomes. The bacterial 16S rRNA gene sequence is about 1,550 bp long and is composed of 9 variable (V1 to V9) and 10 conserved regions, with different evolutionary rates. Hypervariable regions display substantial sequence diversity among different bacterial species and can be used to design species-specific hybridization probes or PCR primers, suitable for microbial detection and identification and essential for distinguishing one microbial taxon from another (taxonomic assignment) (Chakravorty et al., 2007; Van de Peer et al., 1996). On the other hand, conserved regions are very useful for finding all 16S rRNA genes in complex samples. Also, they are very similar among organisms, and are important for sequence alignment and search of homologous regions in sequences from various organisms (Kirchman, 2018; Van de Peer et al., 1996). The 16S gene is the most widely used molecular marker for bacteria taxonomic purposes and for phylogenetic studies (Case et al., 2007). It can also be used for determining the genetic diversity of environmental samples (Baker et al., 2003; Chakravorty et al., 2007). The best definition of species based on the 16S rRNA is that two bacteria or two archaea belong to the same species if their 16S rRNA genes are $\geq$ 97% identical (Kirchman, 2018).

1.5. Microbial communities in cetaceans in a changing ocean

Microbial and host dynamics in aquatic individuals are being affected by rising changes in the marine environment, due to climate-related and anthropogenic impacts (Sehna et al., 2021). Currently, the role of host-microbe interaction as a sentinel of a changing host is being investigated, in order to better understand the interactions and exchanges between animals, microbiomes, and the marine environment (Ainsworth & Gates, 2016; Lee et al., 2019; Nelson et al., 2015; Webster et al., 2016). Overall, the composition of bacterial communities in marine mammals is poorly known, particularly the respiratory microbiome and specific functions of airway bacteria (Apprill et al., 2017; Faner et al.,

2017; Lima et al., 2012; Ridgway et al., 1969; C. Vendl et al., 2019). The understanding of the degree of natural variety inside and between healthy cetaceans, host-microbe interactions, and the establishment of baseline data on microorganisms commonly within and between them, will potentially allow the quick and non-invasively identification of anomalies present in individuals and bacterial disease markers, like the identification of known and potentially emerging pathogens associated with them. Such knowledge may also allow for the determination of health risks like respiratory diseases and alerts for the presence of endangered populations, helping in the creation of correct management decisions (Christiansen et al., 2020; Lima et al., 2012; Raverty et al., 2017; Sehnal et al., 2021).

1.5.1. Respiratory infections in cetaceans

Airway microbiota display complex interactions with their host organism. A close linkage amid the airways' physiological state and the composition of microbial communities has already been identified, influencing the respiratory health in marine mammals (Lima et al., 2012; Catharina Vendl, Slavich, Wemheuer, et al., 2020). Once the rate of bacterial immigration, elimination, and growth conditions in the airways is altered, a state of sickness or even a modest shift in the host immunological capacity can affect the diversity and total microbial composition of the airways (Dickson et al., 2016, 2018; Einarsson et al., 2016; C. Vendl et al., 2021). Numerous factors, such as the health condition or different levels of social interaction, may relate to the bacterial composition in the respiratory tract of cetaceans (Dickson et al., 2018; Teo et al., 2015). The factors that affect the airway control the type of modification that occurs in the microbial community (Catharina Vendl, Slavich, Wemheuer, et al., 2020). Some recent analyses of the exhale breath condensate (EBC) of cetaceans, also called blow, reveal an extraordinary diversity and richness of the respiratory microbiota of the animals' (Apprill et al., 2017; Geoghegan et al., 2018; Pirodda et al., 2017). The EBC is formed by the exhaled droplets of condensed respiratory vapor produced when animals exhale huge tidal volumes at the water surface.

The respiratory system of cetaceans is the main connection between the internal body and the external environment. The principal role of their respiratory system is to transport air to the alveoli of the lungs and promote the gas exchange (Johnson et al., 2009). Due to certain anatomic features of cetaceans, that contribute to their survival in aquatic

conditions and diving efficiency, breath analysis methods are potentially suitable to provide important physiological insights (Burgess et al., 2018; Johnson et al., 2009).

The respiratory tract of mammals is composed by the upper (nasal cavity, pharynx, larynx) and the lower (trachea, bronchi, bronchioles, and alveolar structure) airways (Catharina Vendl, Slavich, Nelson, et al., 2020). The upper respiratory tract (URT) of cetaceans finishes in the blowhole, a special nasal opening on the dorsal surface of the head, and an exceptional adaptation to life in the aquatic environment (Johnson et al., 2009; Nelson et al., 2015, 2019). The seawater is almost entirely excluded from the airways during the dive, with the blowhole being protected by a muscle flap, and due to the exhalation force. Despite that isolation, this opening makes the respiratory system directly exposed to the microbiome of the surface water and air when animals take a breath, which allows exogenous variables to impact the microbiota composition (Sehnal et al., 2021). The environment around the blowhole is warm, moist, and rich in proteins and lipids, which makes it an appropriate habitat for bacteria. However, the growth of most organisms can be limited by the rapid changes in pressure and oxygen that occur in cetaceans, and by the different host defenses, like antimicrobial surfactants produced by these animals, and also by the expulsion of mucus through the blowhole (Foot et al., 2006; Johnson et al., 2009).

The microbiota of the respiratory tract is restricted to the upper structures, with the lower structures hypothesized to be sterile in healthy individuals (Avalos-Téllez et al., 2010; Shukla et al., 2017). In terrestrial mammals like humans, at the back of the nose and mouth, the pharynx connects the nasal cavity and the mouth. The pharynx is part of the digestive and respiratory systems, transporting both food and air, so this common pathway allows breath chemicals to be accidentally contaminated by odors from the digestive system (Aksenov et al., 2014; Burgess et al., 2018). However, cetaceans have these systems disconnected by the close of the palatopharyngeal sphincter. In these animals, the larynx extends up into the nasal cavity rather than opening into the throat (Aksenov et al., 2014; Macleod et al., 2007). This separation avoids water from entering the respiratory tract during feeding, limiting the exchange of microorganisms between both tracts, and making microbiota overlap improbable (Bik et al., 2016; Johnson et al., 2009; C. Vendl et al., 2021). This unique respiratory anatomy of cetaceans may harbor rare microbes (Apprill, 2017). Furthermore, human lungs contain fewer alveoli than cetaceans' lungs do. The efficiency of gas exchange is higher in cetaceans, because their lungs have two layers of capillaries, whereas most mammals have only one (Aksenov et al., 2014). These two layers make cetaceans breathe in enormous tidal

volumes with exceptional efficiency because it maximizes the exposure of lung air to blood circulation (Burgess et al., 2018). Therefore, compared with the human respiratory microbiota, the blow may be a better bacterial representative of the cetacean respiratory microbiota (C. Vendl et al., 2021).

Nonetheless, these animals established a more powerful breathing mechanism than humans, due to their marine lifestyle, their trachea is short, with a large diameter. Deep dives of cetaceans have evolved to include prolonged breath-holding, followed by rapid breaths. Cetaceans can inhale a tidal volume of 70–85%, and exhale of up to 90% of their total lung volume in 0.3s for smaller cetaceans and 1-2 seconds for larger whales, whereas a physically active human can exhale up to only 50% (Aksenov et al., 2014; Johnson et al., 2009; Nelson et al., 2015; Raverty et al., 2017; Ridgway et al., 1969; C. Vendl et al., 2021). This may result in a higher turnover rate of cetacean respiratory bacteria, possibly making them more transient than in humans, enabling deep lung exposure to airborne threats at the marine surface. Additionally, cetaceans lack upper airway filters and turbinates, which filter inhaled particles (C. Vendl et al., 2021; Venn-Watson et al., 2012). The rapid respiratory exchange rates and the large tidal volumes make the deposition of particles and pathogens deep into the respiratory tract in these animals easier than in terrestrial mammals. As a result, respiratory illness is a considerably more common cause of morbidity and mortality among cetaceans (Leger et al., 2018).

1.6. Infraorder Cetacea

Cetaceans (Chordata, Cetacea) belong to a highly specialized infraorder of mammals and play a very important role in the balance of the marine ecosystems (Katona & Whitehead, 1988). The International Union for Conservation of Nature (IUCN) Red List provides the status evaluation of the 90 species of cetaceans, including whales, dolphins, and porpoises, and 40 subspecies or subpopulations. The IUCN presently classifies 60 out of 130 cetacean species/subspecies as endangered or vulnerable (Table 1).

Table 1 - Number of cetacean species, subspecies, and total in the different IUCN (International Union for Conservation of Nature) categories: critically endangered, endangered, vulnerable, near threatened, least concern, data deficient.

Category	Species	Subspecies/populations	Total
Critically Endangered	4	18	22
Endangered	11	11	22
Vulnerable	7	9	16
Near Threatened	10	0	10
Least Concern	49	0	49
Data Deficient	9	2	11
Total	90	40	130

The Cetacea infraorder is divided into 2 superfamilies: Mysticeti and Odontoceti. The Mysticeti superfamily, also commonly known as baleen whales, includes 5 families, 6 genera, and 16 species described (WoRMS, 2022). Baleen whales lack functional teeth and feed on very small organisms like small fish or euphausiid crustaceans such as krill, through a highly specialized filtering apparatus, the “beards”. These are constituted by brush-like keratin structures named “baleen plates”, also called “whalebone”. Also, they have a symmetrical skull with two blowholes (Bannister, 2009; Carwardine, 2020; Jefferson, T. A., Webber, M. A., & Pitman, 2015; Thewissen et al., 2009). Odontocetes, or toothed cetaceans, are made up of at least 76 species from 37 genera and 10 families, including marine and freshwater dolphins, porpoises, beaked, and sperm whales. They are characterized by their teeth that allow them to catch individual prey, and their asymmetrical skull with one single blowhole (Carwardine, 2020; Jefferson, T. A., Webber, M. A., & Pitman, 2015). Odontocetes and Mysticetes differ fundamentally in three main points: the placement/structure of the skull, the method of obtaining food, and the anatomical changes (such as the specialization in echolocation of Odontocetes) (Hooker, 2009). Cetaceans live in diverse ecological niches, have different ways of life and social systems. Despite their ecological differences, they share important features, like being live-bearing homeotherms and air-breathing animals (Ballance, 2018).

1.6.1. The ecological importance of Cetaceans

Cetaceans play a significant role in marine ecosystems, being responsible for important changes like material fluxes and species diversity (Roman et al., 2014). These animals

have functional and structural impacts on their ecosystems in different ways, including top-down (consumers' influence on their prey) and bottom-up (factors affecting primary production or the transfer of energy up the food web) effects (Ballance, 2018). A strong pressure on marine communities can be applied by cetaceans through indirect food-web interactions and predation. Their large size, high metabolic rate, and position as top predators, result in the intake of tons of prey. Despite being important consumers, cetaceans are also important prey to other marine animals like sharks or killer whales. Furthermore, these marine mammals are also important nutrient vectors, releasing nitrogen, iron, and other elements through their feces and urine. Since many of them forage at depth and defecate at the surface, and migrate between productive feeding and oligotrophic breeding grounds, these nutrients are transferred vertically and horizontally through urea, carcasses, and placenta. When cetaceans die, they sink to the bottom of the ocean, providing food and habitat for deep-sea water communities. Also, by pulling their prey to the surface, cetaceans provide feeding opportunities for other species, such as seabirds. Moreover, cetaceans possess a variety of symbiotic or parasitic species, in some cases (whale cyamids), which are completely dependent on cetaceans for all their life stages (Ballance, 2018; Roman et al., 2014).

Cetaceans' position on the trophic ladder, as top predators, makes them act as umbrella species, exerting control over lower trophic levels. Therefore, is expected that natural co-occurring species benefit (i.e. are protected) by their conservation (Roberge & Angelstam, 2004; Segio et al., 2008; Sergio et al., 2006). Cetaceans are considered flagship species, given their charismatic nature, attracting the public's attention, and helping to trigger important behavioral responses among humans, including increasing awareness about ocean health issues (Mazzoldi et al., 2019). They also contribute to the blue economy by drawing key sources of financial profit through their roles in tourism, culture, and conservation. These animals act as sentinel species as most of them are long-lived, slow-breeding animals, with few offspring, feed at a high trophic level and their sensitivity to environmental changes makes them reflect and incorporate ecological changes across high temporal and spatial scales (Moore, 2008). Cetaceans are in contact with contaminants and toxins, and since they possess large fat reservoirs, this enables the bioaccumulation of these products, being considered as biological "sinks" for many persistent pollutants (Bossart, 2011; Halpern et al., 2008; Murphy et al., 2009). Essentially, the overall health and wellness of marine mammals is a reflection of the ecosystems on which they rely (Moore, 2008). Furthermore, cetaceans usually select

areas with high biodiversity richness and diversity, therefore also acting as biodiversity indicators (Bik et al., 2016; Bossart, 2011; Segio et al., 2008).

1.6.2. The Exhale Breath Condensate of Cetaceans

Blow is a relatively unexplored biological sample with a poorly known composition. Hormones, as well as respiratory microbes, a range of metabolites, and a number of inflammation- and immune-related substances are found in blow samples (K. E. Hunt et al., 2013). The EBC contains mucus and moisture that, when liberated into the external air, which is comparably cooler, condenses to create a detectable mass of vapor, prone to be collected (Aprill et al., 2017). Furthermore, due to the high flow rates that cetaceans exhale, some respiratory fluid and lung cells will probably be exhaled. The primary biological fluid in the blow is likely to be lung surfactant (Berta et al., 2006). It is supposed that the composition of blow and lung surfactants is similar, with 97% of lipids, 2.3% of protein, and <1% of carbohydrates. It may also contain airway secretions, supposed to be originated from the alveolar surfactants and used for airway stabilization (Frère et al., 2010).

It has been observed that microbial communities present in the blow of cetaceans are distinct from those present in the surrounding seawater, air, prey, other body sites, and from the hands and noses of their human caregivers in the case of captive cetaceans, thus being considered specific of cetaceans respiratory tract (Geoghegan et al., 2018; Pirodda et al., 2017; C. Vendl et al., 2021). Within this material, specific microbial pathogens have been identified (Acevedo-Whitehouse et al., 2010), which emphasizes the potential of using the EBC collection and characterization as a method for monitoring microbial communities associated with the respiratory system of cetaceans (Lima et al., 2012).

Several investigations studying the airway microbiota of cetaceans resorted to the analysis of EBC samples (Table 2). They applied different methodologies for blow collection, DNA extraction, and amplification, targeting different hypervariable gene regions. Some of these studies used the metabarcoding approach. The majority of data accessible on cetaceans pathogens, diseases, and parasites (cetacean-associated microbiome) come from captive, stranded, sick, or injured individuals which cannot be considered representative of the free-ranging populations (Acevedo-Whitehouse et al., 2010; Johnson et al., 2009; Lima et al., 2012). Knowledge of the cetaceans' respiratory microbiome from free-ranging species is still limited.

Table 2 - State of the art of studies focused on the use of Exhale Breath Condensate (blow) for analyzing microorganisms associated with the respiratory tract of cetaceans.

Specie	Wild/Captive	Local	Blow Collection	Blow DNA extraction	Target Amplification region	Primers	Sequencing	Reference
Bottlenose dolphin (<i>Tursiops aduncus</i>)	Captive	n/a	sterile bag or container	n/a	n/a	n/a	n/a	(Avalos-Téllez et al., 2010)
Bottlenose dolphin (<i>Tursiops aduncus</i>)	Captive	Sea World Marine Park Gold Coast, Australia	directly to a "yellow cap" container	FastDNA Spin Kit for Soil (Vendl, C. et al., 2020)a	V1-V3 16S rRNA	27F and 519R	Illumina MiSeq platform	(C. Vendl et al., 2021)
Bottlenose Dolphins (<i>Tursiops truncatus</i> ; <i>Tursiops aduncus</i>); Hybrid	Captive	SeaWorld, Australia	70 ml polypropylene tubes	QIAamp DNA Micro Kit	V1-V3 16S rRNA	27F and 519R	ribosomal tag prosequencing	(Lima et al., 2012)
Bottlenose Dolphins (<i>Tursiops truncatus</i>)	Live capture and release	Charleston, South Carolina and River Lagoon, Florida	sterile swabs with ejectable heads	Qiagen Dneasy Bood and Tissue Kit	nearly the full 16S gene	27F and 1492R	pGEM-T cloning vectors	(Johnson et al., 2009)
Common dolphin (<i>Delphinus sp.</i>), Bottlenose dolphin (<i>Tursiops truncatus</i>), Bryde's whale (<i>Balaenoptera edeni</i>), Fin (<i>Balaenoptera physalus</i>), Sperm (<i>Physeter macrocephalus</i>), Humpback (<i>Megaptera novaeangliae</i>), Gray (<i>Eschrichtius robustus</i>) and Blue whale (<i>Balaenoptera musculus</i>)	Wild	Gulf of California and along the Pacific coast of Baja California (Mexico)	sterile petri dishes attached to the Perspex plate and a long pole/ an helicopter with 2 sterile Petri dishes attached	Qiagen Dneasy	1400 bp fragment of the eubacterial 16S rRNA	3F and 1319R	n/a	(Acevedo-Whitehouse et al., 2010)
Humpback whales (<i>Megaptera novaeangliae</i>)	Wild	Hervey Bay, Southern Queensland, Australia/ Coast of Sydney, New South Wales, Australia	6 blank sterile petri dishes attached to the Perspex plate and a long pole (Acevedo-Whitehouse et al., 2010; Hogg et al., 2009)	FastDNA Spin Kit for Soil	V1-V3 16S rRNA	27F and 519R	Illumina MiSeq platform	(Catharina Vendl, Slavich, Wemheuer, et al., 2020)
Blue (<i>Balaenoptera musculus</i>), Grey (<i>Eschrichtius robustus</i>), Humpback (<i>Megaptera novaeangliae</i>) and Long-finned pilot whales (PWs) (<i>Globicephala melas</i>)	Wild	Hervey Bay, Queensland, Australia/ Gulf of California / along the Pacific coast of Baja California (Mexico) / Alboran Sea, western part of the Mediterranean	small remote-controlled helicopter with Petri dishes attached to its skids (Acevedo-Whitehouse et al., 2010) / telescope pole with 6 blank sterile Petri dishes (Vendl et al., 2019)	FastDNA Spin Kit for Soil	V4 16S rRNA	515F and 806R	Illumina MiSeq platform	(Catharina Vendl, Slavich, Nelson, et al., 2020)
Humpback whales (<i>Megaptera novaeangliae</i>)	Wild	Hervey Bay, Queensland, Australia	6 blank sterile petri dishes attached to the Perspex plate and a long pole (Acevedo-Whitehouse et al., 2010)	FastDNA Spin Kit for Soil	V1-V3 16S rRNA	27F and 519R	Illumina MiSeq platform	(C. Vendl et al., 2019)
Humpback whales (<i>Megaptera novaeangliae</i>)	Wild	East coast of Sydney, Australia	UAV (4-motor electric multirotor)	QuickDNA Fungal/Bacterial Miniprep kit	16S rRNA	27F and 519R	Sanger	(Pirrotta et al., 2017)
Bottlenose dolphin (<i>Tursiops aduncus</i>); Humpback dolphin (<i>Sousa sahulensis</i>)	Wild	n/a	custom-built waterproof 4-motor electric multirotor (Pirrotta et al., 2017)	QIAamp DNA Investigator Kit	600bp mitochondrial D-Loop	H00034 and D-Loop	n/a	(Raudino et al., 2019)
Bottlenose Dolphins (<i>Tursiops truncatus</i>); Sperm whale (<i>Physeter macrocephalus</i>)	Captive/Wild	Acquario di Genova/ Gulf of Ambracia (Greece)/ southern Tyrrhenian Sea (Italy)	drone with 6-well polystyrene sterile Petri plate	Qiagen Dneasy Bood and Tissue Kit	V3-V4 16S rRNA bacterial and archaeal gene	341F and 806R	Illumina MiSeq platform	(Centelleghé et al., 2020)
Humpback whales (<i>Megaptera novaeangliae</i>)	Wild	Massachusetts coastal waters off Cape Cod / coastal waters around Vancouver Island	unmanned hexacopter drone	PowerBiofilm DNA isolation kit	V4 SSU rRNA	515FY and 806RB	MiSeq and NanoSeq	(Apprill et al., 2017)
Bottlenose dolphin (<i>Tursiops truncatus</i>); Sea lions (<i>Zalophus californianus</i>)	Captive/Wild	US Navy Marine Mammal Program (MMP) in San Diego Bay, San Diego, CA/ Sarasota Bay, Florida	n/a	n/a	V3-V4-V5 16S rRNA	n/a	Sanger	(Bik et al., 2016)
Killer Whale (<i>Orcinus orca</i>)	Wild	Salish Sea around the San Juan Islands of Washington State	pole with petri plates with selective or non-selective agars	QIAamp DNA Mini kit	Bacterial 16S rRNA and Fungal 18S rRNA; culturable vs non-culturable methods	Uni-C and Uni-D primers	n/a	(Raverty et al., 2017)

Humpback (<i>Megaptera novaeangliae</i>), blue (<i>Balaenoptera musculus</i>), and killer whale (<i>Orcinus orca</i>)	Wild	Frederick Sound in southeast Alaska, USA; off Loreto and Bahia La Paz, Mexico (herein referred to as the Gulf of California); Gulf of Guinea, Gabon; Caribbean Sea, Dominican Republic; and Stellwagen Bank, Massachusetts, USA	UAS	0.2µm filter followed by phenol- chloroform- isoamyl alcohol (PCI) extraction or using the QIAGEN DNAeasy blood and tissue spin column kit	V4 16S SSU rRNA	V4_515F and V4_806R-B	Illumina MiSeq v2	(Atkinson et al., 2021)
Bottlenose dolphins (<i>Tursiops aduncus</i> ; <i>Tursiops truncatus</i>)	Wild/Captive	Shark Bay, Australia/ National Aquarium Baltimore, USA	Funnel connected to a 50ml sterile polypropylene falcon centrifuge tube/directly into 50 ml sterile polypropylene	Qiagen Dneasy Bood and Tissue Kit	V4 16S rRNA	515F and 806R	Ion Torrent	(Nelson et al., 2019)
Killer Whale (<i>Orcinus orca</i>)	Wild	outer US Pacific Northwest coast and into the Salish Sea	pole-mounted sterile Petri dishes covered with sterile 100-µm mesh nitex	n/a	16S rRNA	Nextera XT primers	Illumina MiSeq v3	(Rhodes et al., 2022)
Bottlenose dolphins (<i>Tursiops truncatus</i>), Short-beaked common dolphin (<i>Delphinus delphis</i>)	Wild	New Quay, Aberystwyth, and Pwllheli	telescopic pole with a single sterile Petri dish affixed to the end via a suction cup	Qiagen DNeasy Blood and Tissue Kit	144 bp product of the 12S mtDNA gene.	Cetacea12S_F and Cetacea12S_R	n/a	(Robinson & Nuuttila, 2020)
Bottlenose dolphins (<i>Tursiops truncatus</i>)	Wild	Sarasota, Florida	sterile petri dish	QIAamp Mini body	V3-V4 of 16S rRNA	F: 5'-TCGTCGGCA GCGTCAGAT GTGTATAAG AGACAGCCT ACGG GNGGCWGC AG - 3' R: 5' - GTCTCGTGG GCTCGGAG ATGTG TATAAGAGA CAGGACTAC HVGGGTATC TAATCC - 3'	Illumina MiSeq platform	(Robles-Malagamba et al., 2020)

1.7. Sampling Cetaceans

Cetaceans face multiple pressures, with their conservation and management relying, partially, on the physiological information obtained with health monitoring of individuals and populations (Apprill, 2017; K. E. Hunt et al., 2013). The lack of information and the challenges in the study of these species is owed to their high speed and mobility, and also legal and political restrictions on working on protected species, which include the majority species of cetaceans (Ballance, 2018). These animals spend most of their time underwater and out of human sight, with the capacity to cover big areas in a short period, inhabiting a huge environment both in terms of the use of the water column and their geographical range. Often, in wildlife's studies it is difficult to address biological questions without direct interaction with animals, which leads to unpleasant capture and restraint techniques (Geoghegan et al., 2018). Furthermore, sampling has logistical, and cost implications for researchers, and also welfare implications for animals - some methods are invasive with a strong impact on the animals (Acevedo-Whitehouse et al., 2010). When taking biological samples directly from free-ranging cetaceans, it is important to take into account a variety of things, such as the group composition, the behavior displayed, proximity to the vessel, and environmental factors like the sea state and wind direction (Burgess et al., 2018; Geoghegan et al., 2018).

The development of appropriate tools to perform sampling protocols that are logistically feasible and non- or minimally- invasive is required to increase the collection of important ecological data, including data on the marine mammal microbiome, and to monitor the emergence of infectious diseases, particularly for threatened or vulnerable populations (Acevedo-Whitehouse et al., 2010; K. E. Hunt et al., 2013; C. Vendl et al., 2019). These tools and methods should provide biological material suitable for a variety of investigations, be easily reproducible (Johnson et al., 2009; Nelson et al., 2015; Raudino et al., 2019), and enable investigators to gather samples over time, seasons, and life history (Burgess et al., 2018; K. E. Hunt et al., 2013).

Blood sampling is considered the gold standard method for endocrine assessment in captive animals (Thompson et al., 2014). However, blood sampling is not possible with animals that cannot be restrained or handled. In large whales, it is nearly impossible to take blood from them non-invasively (K. E. Hunt et al., 2013; Thompson et al., 2014). Less invasive and, optimally, non-invasive techniques, such as the collection of fecal plumes and blubber/skin biopsies have been explored to study cetacean physiology

(Hudson et al., 2021). Fecal samples have been collected successfully from diverse species of cetaceans and are being used to retrieve information specific to digestive physiology (diet analysis and intestinal parasitology), and also to study other physiological measures like steroid hormones (K. E. Hunt et al., 2013). However this is a challenging technique with a risk of environmental contamination of the samples (potential loss of polar hormone metabolites to seawater) and a low sampling rate. Moreover, it is difficult to target and identify individuals because focal animals may not defecate while under observation, and furthermore they cannot always be genotyped due to DNA degradation (K. E. Hunt et al., 2013; Thompson et al., 2014). Dart-biopsy is the most frequently used method of collecting genetic data from free-ranging cetaceans. This technique is also used for analyses of contaminants and fatty-acid, and endocrine studies (Frère et al., 2010; K. E. Hunt et al., 2013). Although minimally invasive, it still requires direct contact and impact on the animal, and it is challenging to employ on species that are wary of vessels and reluctant to approach (Raudino et al., 2019). It is also considered unsuitable for young animals or to apply in groups with calves (Frère et al., 2010).

Visual assessment using photographic techniques of the external appearance of animals is a time-honored method for analyze the health and nutritional status of individuals, such as skin condition, ectoparasite load, along with wounds, and scars from ship strikes and fishing gear entanglement (Christiansen et al., 2020; K. E. Hunt et al., 2013).

It has recently been shown that sampling wildlife individuals and populations can be made using remotely piloted aircrafts, for example, Unmanned Aerial Vehicles (UAVs). These tools have huge potential to modernize and transform how studies are being conducted, improving the efficiency and safety of sample collection from wild individuals, permitting the collection of samples from inaccessible and unsafe environments (Christie et al., 2016; Domínguez-Sánchez et al., 2018; Geoghegan et al., 2018; Pirotta et al., 2017). Furthermore, they may have a broad application in the cost-effective monitoring and management of whale populations worldwide (Pirotta et al., 2017). In fact, drones are potentially the premium and emerging non-invasive tool with minimal impact on animals (Christiansen et al., 2016). They have been successfully used in photogrammetric, physiological, and behavioral studies, microbiome research, and infrared heat detection analyses (Atkinson et al., 2021).

A non-invasive collection technique used to obtain information on cetaceans' respiratory health is the sampling of the EBC. This technique is frequently used for diagnosing

respiratory disorders in humans, but its use in cetaceans is less well established (Acevedo-Whitehouse et al., 2010). This procedure takes advantage of the natural breathing behavior of the animals, can be applied to individuals from all age classes, involves no harmful contact, and can be achieved when dolphins voluntarily bow-ride (Frère et al., 2010; Raverty et al., 2017). Usually, during a single surfacing phase, cetaceans blow many times and, some species, frequently allow for a close approach of the vessels during this period (K. E. Hunt et al., 2013). The EBC can have a chemical composition that reflects the circulating components of blood. It represents a novel approach to explore animal physiology, allowing the analysis of a wide range of metabolites simultaneously and overcoming the logistic challenges of blood sampling from free-swimming individuals by conventional invasive methods (Aksenov et al., 2014; Burgess et al., 2016). Moreover, during at-sea blow sampling, researchers can collect samples from target individuals and obtain repeated samples over time (Burgess et al., 2016). The application of this non-invasive approach represents an important tool for wildlife investigations and conservation, as well as respiratory medicine (Thompson et al., 2014).

1.7.1. Blow sampling

Blow sampling was first attempted by Hogg et al., (2005) to examine hormonal and reproductive states in bottlenose dolphins (*Tursiops truncatus*). The use of the cetacean blow as a biomarker to study animal health parameters has been one of the most promising tools to emerge in recent years, especially to study the respiratory tract, including the identification of potential fungal, bacterial, and viral respiratory pathogens and monitor spatiotemporal fluctuations, aiming to inform conservation management (Acevedo-Whitehouse et al., 2010; Geoghegan et al., 2018; Raverty et al., 2017; C. Vendl et al., 2019). Blow samples can also be used to analyze host DNA for population genetic investigations (Frère et al., 2010), stress-related and reproductive hormones (Burgess et al., 2018; Hogg et al., 2005, 2009; Thompson et al., 2014), energetic cost of reproduction (Christiansen et al., 2016), screen for epizootics (Pirotta et al., 2017), viruses (Geoghegan et al., 2018), and respiratory associated microorganisms (Acevedo-Whitehouse et al., 2010; K. E. Hunt et al., 2013; Johnson et al., 2009; Lima et al., 2012).

Blow sampling is challenging and requires a skilled team, principally in studies of free-ranging cetaceans, that are difficult to approach closely in a consistent and easy way

(Raudino et al., 2019). In order to collect good respiratory samples representative of the microbial communities in the respiratory tract of cetaceans, water contamination needs to be avoided (Avalos-Téllez et al., 2010). In the case of studies with animals in captivity, blow samples have been collected by holding an inverted tube or another device directly over the animal's blowhole (Frère et al., 2010; K. E. Hunt et al., 2013; Lima et al., 2012). In this case, to complete a normal microbiota study, it is important that, at least, in the first three weeks before the sampling, cetaceans must not show any clinical pathology, infection, or received antibiotic treatment (Avalos-Téllez et al., 2010).

In the wild, a variety of sampling devices have been used to gather blow droplets, like pole-based methods, and more recently, using remote-controlled helicopters and UAVs (Table 2) (Acevedo-Whitehouse et al., 2010; Hogg et al., 2005, 2009; Kathleen E. Hunt et al., 2014). Using pole-based methods has the advantage to collect a larger sample volume because the EBC is collected more closely to the blowhole, however, it has the disadvantage that it can be more behaviorally intrusive to individuals (swimming speed, increased breathing synchrony, and direction changes). Moreover, it requires that the vessels maintain a stable sailing close to the animals, with the same speed as them during the sampling. It is also very important that the behavior of the animals is propitious, without sudden movements, allowing themselves to be approached, and without very erratic blows (Burgess et al., 2018; Hastie et al., 2003; Tosi & Ferreira, 2009).

Using UAVs is minimally intrusive and noninvasive, once the sampling can be done while maintaining the boat at a safe distance from the animal's (Nelson et al., 2019; Pirotta et al., 2017; Raverty et al., 2017; C. Vendl et al., 2019). However, it is more difficult to apply to smaller individuals, due to their swimming and diving behavior (Centelleghe et al., 2020). It also requires a bigger vessel with space for a landing/takeoff platform, as well as a trained pilot. Furthermore, the equipment has a higher cost in comparison to the pole sampling (Acevedo-Whitehouse et al., 2010).

Blow sampling have been using various collection materials, including nylon mesh, nylon stocking, nylon veil, cotton gauze, polypropylene containers, and Petri dishes (Table 2) (Acevedo-Whitehouse et al., 2010; Hogg et al., 2005, 2009; Kathleen E. Hunt et al., 2014; Thompson et al., 2014). By sampling with more absorbent materials, it is possible to increase the blow sample volume (Raudino et al., 2019).

1.8. Target Species: Short-finned pilot whale

The genus *Globicephala* Lesson 1828 (superfamily Odontoceti) contains two species, with a slight overlap in distribution, the short-finned pilot whale, *G. macrorhynchus* Gray (1846) (van Bree, 1971); and the long-finned pilot whale, *G. melas* Traill (1809). This study is focused on the first one.

The pectoral fins of the short-finned pilot whale *G. macrorhynchus* are shorter and thicker than those of its identical cool-water relative, *G. melas* (Olson, 2009). They also differ in the number of teeth, as the short-finned whale has 7-9 pairs in both the upper and lower jaws, while the long-finned has 9-12 (Servidio, 2014).

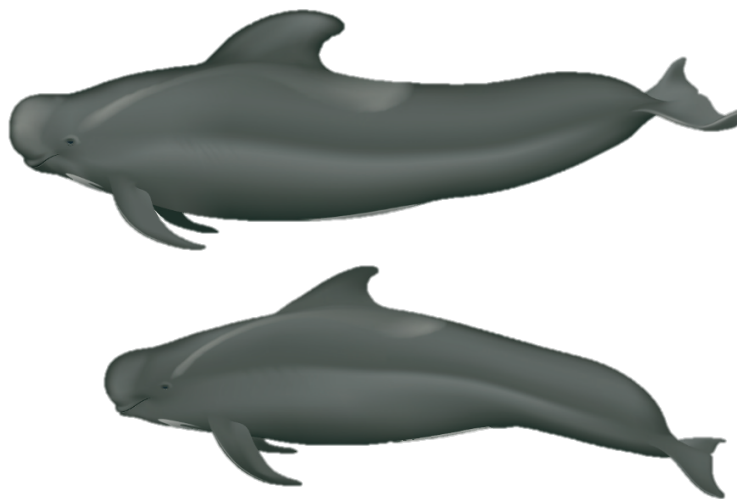


Figure 2 - Illustration of male (top) and female (bottom) short-finned pilot whale. Retrieved from Jefferson, T. A., Webber, M. A., & Pitman, (2015).

Although commonly referred to as a whale, *G. macrorhynchus* belongs to the delphinids (Family Delphinidae) (WoRMS, 2022). These animals are deep-divers: Alves and colleagues (2013) showed that they can dove to 1000 meters depth in only 20 minutes. They have long and robust bodies, with adults reaching an average length of approximately 6 meters. They have a dark grey-black color, bulbous heads with a pronounced melon, dramatically up-curved mouthlines, and small or inexistent beaks. Depending on their age and sex, the shape of the head varies significantly, becoming more globose in adult males (Jefferson, T. A., Webber, M. A., & Pitman, 2015). In terms of weight and size, sexual dimorphism is present, with males being heavier and larger.

Females are sexually mature at 9 years with a capacity to produce one calf every 3 to 5 years. The life span of females is higher, 63 years, when compared to their male counterparts. Males only live till 45 years old and their sexually mature age ranges from 13 to 16 years. In the Northern Hemisphere, the mating season of these animals usually takes place in spring or early summer, and parturition in summer or early autumn (Olson, 2009).

Short-finned pilot whales are among the most known cetacean species due to their distribution, global abundance, and propensity for mass strandings (Filipe Alves, 2013). They inhabit circum-tropical and warm temperate waters, ranging within 50°N and 40°S (Minton et al., 2018), including the biogeographic region of Macaronesia (NE Atlantic), which contains four archipelagos: Azores, Madeira, Canaries and Cape Verde. The preferred habitat of these animals includes places of great topographic relief, island slopes, and continental shelf breaks, and they are found primarily in deep waters (Filipe Alves, 2013). These animals are very social and usually form stable groups called pods composed of an average of 20–90 individuals with close matrilinear associations (Olson, 2009). Individuals in a pod can travel together, alone, or, often, along with other pods (Hauser et al., 2007).

1.8.1. Short-finned pilot whale off Madeira Archipelago

The Madeira Archipelago, Portugal, offers favorable conditions for spotting short-finned pilot whales, with the species being between the four most encountered odontocete in these waters. Madeira is located in a warm-temperate latitude and its pelagic waters are among the most isolated and oligotrophic oceanic environments in the North Atlantic. A small continental shelf, steep submarine canyons, and deep waters make this a unique habitat (F. Alves et al., 2020; Filipe Alves, Dinis, et al., 2013; Filipe Alves, Quérrouil, et al., 2013).

Since 1985, *G. macrorhynchus* is legally protected in the Exclusive Economic Zone (EEZ) of the archipelago of Madeira by the Dec.-Leg. Regional 6/86/M. The status of the species in this area is considered 'Least Concern' (Cabral et al., 2005). The species' population structure has been investigated, combining genetic analysis and photo-id methodologies (Filipe Alves, 2013; Sambolino et al., 2022). Their dive characteristics (Filipe Alves, Dinis, et al., 2013), survival, and abundance, along with distribution and habitat preference, have also been deeply studied (Filipe Alves et al., 2015).

The study of Alves and colleagues (2013) showed that the degree of site fidelity displayed by these individuals is quite variable, with individuals classified as residents (individuals that have been captured more than 5 times in 3 years), others as transients of different regions (individuals captured only once), or even as visitors (individuals that stay between these thresholds). It is proposed that all short-finned pilot whales encountered in Madeira genetically belong to the same population that may include three clans of island-related whales, and other transients, interacting for mating purposes when they meet (Filipe Alves, Dinis, et al., 2013).

1.8.2. Conservation status and threats to the short-finned pilot whale

Worldwide estimations of short-finned pilot whale population parameters are limited in the literature, and as of 2008, the species was classified as “Data Deficient” on the IUCN Red List. In 2018, the species was evaluated with the status of “Least Concern”, once there is currently no evidence of mass deaths or population decline. The species is currently listed in Appendix II of CITES, Annex II of Bern Convention, and Annex IV of Habitats Directive. The species protected under the aforementioned annex should have a strict protection and regulation regime applied across their entire natural range within the EU (within and outside Natura 2000 sites). It is estimated that the worldwide abundance of short-finned pilot whales consists of a minimum of 700.000 individuals, however, there is no data on worldwide population trends available (Minton et al., 2018).

These animals are subject to multiple types of threats like human interactions, including direct exploitation, and hunting in populations of Sri Lanka, Indonesia, Japan, and the eastern Caribbean (Jefferson, T. A., Webber, M. A., & Pitman, 2015). The species is also bycaught by different fishing activities like longline, gillnet, or trawls fisheries (Servidio, 2014). Another threat is the exposure to chemical contaminants, like PCBs (polychlorinated biphenyl) pesticides, or mercury, that accumulate in the muscles and blubber tissues. Moreover, collisions with vessels, often occurring where their habitat overlaps areas of intense traffic, as well as sounds generated by commercial shipping, military operations, and fisheries can endanger their populations (Minton et al., 2018; Olson, 2009). Physiological responses, as a result of these conservation pressures, can be identified at the individual-level before population-level impacts on health, mortality, or fecundity become notable (K. E. Hunt et al., 2013).

2. Objectives

Until now, no studies were found in the literature focusing on short-finned pilot whales (*Globicephala macrorhynchus*) respiratory microbiome. Besides, none of the previous studies with cetaceans has compared the amplification targeting different hypervariable gene regions, using the same methodology on the same samples. For these reasons, the present study will build on the poor knowledge on the respiratory microbiome of free-ranging cetaceans, using the short-finned pilot whale as a model species. The main goal is to provide a network of consistent core taxa from the EBC short-finned pilot whale microbiome, that will serve as a baseline to set a working and optimized methodology for cetacean health monitoring, from sample collection to laboratory analysis. This will be accomplished by achieving the following specific objectives:

- a) Optimize the protocol for extraction of DNA from cetacean blow samples.
- b) Perform a comprehensive comparison of the primer sets performance (regarding amplification [V3-V4 and V4-V5 of the 16S rRNA gene] and taxa coverage) to select the most suitable primer sets for describing blow and seawater microbiomes.
- c) Characterize the microbiome of the tract respiratory of free-ranging cetaceans and seawater controls, showing commonalities and differences between the samples, by metabarcoding 16S rRNA amplicon analyses.
- d) Highlight potential pathogens in the EBC microbiota.

3. Material and Methods

3.1. Sampling

The sampling activities were conducted in 2018, during an at-sea campaign, in the framework of the MARINFO and MARCET projects, by a multidisciplinary team composed of CIIMAR (Interdisciplinary Center for Marine and Environmental Research) and MARE (Marine and Environmental Sciences Centre, ARDITI, Madeira, Portugal) researchers. Fieldwork was performed in the southern waters of Madeira Island, Portugal (Figure 3).



Figure 3 - Map with the location of Madeira Island, Portugal (in red) (Google maps).

Blow samples targeted free-ranging short-finned pilot whales (*Globicephala macrorhynchus*). During the sampling activities, the sampled animals' natural behavior (traveling, resting, or socializing), their class age, as well as the number of individuals in the group, were also recorded.

The Performagene sample collection kit (Performagene™ PG-100 swab collection kit, DNA Genotek, Ontario, Canada) was attached to an extendable 4m carbon pole and used as a sampling device to collect EBCs (Figure 4). This was the first study using this kit to study the respiratory tract of cetaceans. When the animals surfaced to exhale near

the boat, the blow collector device was positioned about 40 cm above the blowhole into the exhaled plume to collect the droplets in the blow. Whenever possible, before attempting the collection of blow samples, the animals were photo-identified to crosscheck with the Madeira's photo-ID catalogue of the species, and also to avoid replication of an individual's EBC.



Figure 4 - Photograph of blow collection using the Performagene sample collection kit, Ágatha Gil.

Control samples of seawater were taken once for each day of sampling to ensure that any DNA results were obtained from blow samples rather than seawater contamination. A sterile bucket was used to collect seawater control samples. A total of 1L of seawater per control sample was filtered on-site through a 0.22 μm pore size Sterivex® filter cartridge combined with a 50-ml syringe, following Kopf et al. (2015). The collected filters were stored on land at -20°C and sent to CIIMAR laboratories in frozen conditions.

A total of 22 blow samples and 4 seawater control samples were collected (detailed sampling list is reported in Table 5).

3.2. DNA Extraction

DNA extraction of blow samples was performed using the QIAmp DNA Mini Kit (Qiagen®), following the manufacturer's instructions. The seawater control samples (Sterivex™ Filters) were defrosted, at room temperature, to extract the total DNA, using the PowerWater® Sterivex™ DNA Isolation Kit protocol (MO BIO Laboratories, Inc.,

Portugal), following the manufacturer's instructions (Table 5). A concentration step in a concentrator plus (Eppendorf) was added to enhance the quantity of the extracted DNA in both EBC and seawater samples. To quantify the DNA extracted, a DeNovix® DS-11 Spectrophotometer was used, and also a Qubit® 3.0 Fluorometer with a Qubit® dsDNA High Sensibility Assay Kit (Thermo Fisher Scientific, Inc., Waltham, MA, USA Thermo Fisher Scientific) (Table S1).

3.3. PCR amplification, Library Preparation, and Quality Control - Optimization of the PCR Amplification protocol

Samples were then prepared for the amplification of the 16S rRNA gene hypervariable V4-V5 region (≈412 bp), using specific primers 515F-Y/926R, (Table 3) recommended by the Earth Microbiome Project (Parada et al., 2016). The V3-V4 region of the 16S rRNA bacterial and archaeal gene was also amplified using the 341F and 806R primers (Takahashi et al., 2014). The PCR reaction was prepared in a volume of 10 µL containing: one of the following three Taq polymerases (DreamTaq™ PCR Master Mix, Bioline MyTaq™, HotStarTaq® Plus), the primer forward and primer reverse, ultrapure H₂O, and the template DNA. Negative and positive controls were included in all PCR amplification procedures.

Table 3 - Different primers tested in the optimization of the PCR (polymerase chain reaction) Amplification protocol.

Primers	Nucleotide sequence (5' - 3')	Target 16S rRNA gene region	References
515YF (Forward)	5'-GTGYCAGCMGCCGCGGTAA-3'	V4-V5	(Parada et al., 2016)
Y926R (Reverse)	5'-CCGYCAATTYMTTTRAGTTT-3'	V4-V5	
341F (Forward)	5' - CCTACGGGAGGCAGCAG-3'	V3-V4	(Takahashi et al., 2014)
806R (Reverse)	5' - GGACTACHVGGGTWTCTAAT-3'	V3-V4	

PCR reactions in CIIMAR labs were carried out in Veriti® 96-Well Thermal Cycler (ThermoFisher Scientific), featuring different cycles and minutes of denaturation, annealing, and extension accordingly to three different Taq DNA polymerases used

(Table 4). Due to the low quantity of DNA extracted in each sample were used these different Taq polymerases with different sensibilities (Table S1).

At the end of this process, the PCR products from each sample were ran in an electrophoresis gel 1,5% (1.5 g of agarose in 100mL of TAE, Tris-acetate-Ethylenediamine tetraacetic acid, with 0.5 µl of SYBR® Safe DNA Gel Stain (ThermoFisher Scientific), exposed to an electric flow of 150V for 30 min. The results were visualized on a UV transilluminator at Gel Doc XR-System (BIO-RAD) and analyzed with Image Lab™ software (BIO-RAD) to verify the success of extraction and the relative amount of DNA available.

Table 4 - Different Taq polymerase tested in the optimization of the PCR (polymerase chain reaction) Amplification protocol.

TAQ polymerase	Step 1 (1x)	Step 2 (30x)	Step 3 (1x)	Step 4
DreamTaq™ PCR Master Mix (Fermentas GmbH)	95°C / 15'	95°C / 40'' 52°C / 40'' 72°C / 1'	72°C / 5'	4°C / ∞
Bioline MyTaq™ (BIOPORTUGAL)	95°C / 2'	95°C / 30'' 55°C / 30'' 72°C / 1'	72°C / 1'	12°C / ∞
HotStarTaq® Plus (QIAGEN)	95°C / 2'	95°C / 30'' 55°C / 30'' 72°C / 1'	72°C / 7'	12°C / ∞

Table 5 - Metadata of the blow and seawater samples analyzed.

Sample ID	Codes used in the analyses	Sample type	Sampling KIT	Collection Date	DNA extraction kit	Latitude	Longitude	Natural behavior	No. of animals sampled
K_blow_Gma1	Gma_1	Blow	PERFORMAgene	09/28/18	QIAamp® DNA Mini and Blood Mini Handbook - QIAGEN	32,5941	-16,969517	Travelling	1
K_blow_Gma2	Gma_2			09/28/18		32,5941	-16,969517	Travelling	1
K_blow_Gma3	**			09/29/18		32,5902	-16,90541667	Travelling	1
K_blow_Gma4	Gma_4			09/29/18		32,5902	-16,905417	Travelling	1
K_blow_Gma5	Gma_5			09/29/18		32,5902	-16,905417	Travelling	1
K_blow_Gma6	Gma_6			10/02/18		32,6109	-17,06	Travelling	1
K_blow_Gma7	Gma_7			10/02/18		32,6109	-17,06	Travelling	1
K_blow_Gma8	Gma_8			10/05/18		32,5975	-16,698017	Travelling	1
K_blow_Gma9	Gma_9			10/06/18		32,5485	-16,875317	Travelling	1
K_blow_Gma10	Gma_10			10/06/18		32,5485	-16,875317	Travelling	1
K_blow_Gma11	Gma_11			10/06/18		32,5485	-16,875317	Travelling	Pool
K_blow_Gma12	Gma_12			10/06/18		32,5485	-16,875317	Travelling	Pool
K_blow_Gma13	Gma_13			10/10/18		32,5875	-16,98825	Travelling	Pool
K_blow_Gma14	**			11/06/18		32,6214	-16,76168333	Travelling	1
K_blow_Gma15	Blow_15			11/09/18		32,5922	-16,873167	Resting	Pool
K_blow_Gma16	Blow_16			11/09/18		32,5922	-16,873167	Travelling	Pool
K_blow_Gma17	Blow_17			11/09/18		32,5392	-16,905417	Travelling	1
K_blow_Gma18	Blow_18			11/30/18		32,594	-16,905417	Travelling	1
K_blow_Gma19	Blow_19			12/12/18		32,5814	-16,9329	Travelling	Pool
K_blow_Gma20	Blow_20			12/12/18		32,5814	-16,89608	Travelling	Pool
K_blow_Gma21	Blow_21			12/12/18		32,5897	-16,802	Travelling	Pool
K_blow_Gma22	Blow_22			12/12/18		32,5897	-16,802	Travelling	Pool
SW_C_Gma14	**	Seawater control	SW 1L filtered troughst sterivex filter	11/09/18	Power water sterivex extraction kit - MO BIO Laboratories	*	*	-	-
SW_C_Gma16-17	A_St			11/09/18		*	*	-	-
SW_C_Gma18	B_St			11/30/18		*	*	-	-
SW_C_Gma19-20-22	C_St			12/12/18		*	*	-	-

(*seawater samples were taken one per sampling day away from the blow collection point; ** Samples K_blow_Gma_3, K_blow_Gma14, and Sw_C_Gma14 were not included in the subsequent analysis due to the following reasons: Sample K_blow_Gma_3 was submerged in water during the blow sampling, Sample K_blow_Gma14 failed to sequence; seawater control sample for K_blow_GMA14, the SW_C_Gma14, was also, consequently, not included in the analysi

3.4. 16S rRNA Amplicon Sequencing

Samples were prepared for the amplification of V3-V4 and V4-V5 of the 16S rRNA gene by high-throughput sequencing. Detailed description of the protocol is reported in Ribeiro et al., (2018). Briefly, KAPA HiFi HotStart PCR Kit was used in the first PCR reaction following manufacturer suggestions (using 0.3 μ M of each PCR primer and 12.5 ng of template DNA in a total volume of 25 μ L; PCR conditions: 3 min denaturation at 95 °C, followed by 25 cycles of 98 °C for 20 s, 50 °C for 30 s and 72 °C for 30 s and finally 5 min at 72 °C). Indexes and sequencing adapters were added to the target amplicon in the second PCR reactions. The amplified products obtained were purified and normalized with SequalPrep 96-well plate kit (ThermoFisher Scientific, USA). Pair-end sequencing was carried out in the Illumina MiSeq® sequencer with the V3 chemistry (Illumina, Inc., San Diego, CA, USA) at Genoinseq laboratories (Cantanhede, Portugal).

3.5. Bioinformatics analyses

The bioinformatics analyses were divided into two parts:

Part 1. The first part was a comprehensive comparison of the primer sets performance to select the most suitable primer set for describing blow microbial communities. This comparison was made with 12 samples of blow that were sequenced to amplify the V3-V4 and V4-V5 regions of the 16S rRNA gene.

Part 2. The second part used 8 blow and 3 seawater sterivex samples sequenced to amplify the V4-V5 region of the 16S rRNA gene. This comparison was made with blow and seawater control samples collected on the same day for a total of 3 days. The principal aim was to investigate the commonalities and the differences between different types of samples and between the different days of sampling.

3.5.1. Upstream sequence analysis

The “DADA2” software package (1.20 version) (Callahan et al., 2016) on R studio (4.1.1 version) was used to process the fastq files obtained after Illumina MiSeq sequencing.

Initially, the raw sequences were quality filtered. For the Part 1, i.e. the primer set performance comparison pipeline with blow samples, raw reads were truncated at 275 bp and 215 bp for forward and reverse sequences, respectively – the position where the final sequence numbers were higher. Afterwards, the trimmed sequences were dereplicated, denoised, and merged. Sequences were then processed to obtain the ASV table, instead of OTUs, for taxonomic resolution, increased precision, and reproducibility. In ASVs, reads represent differences of one nucleotide (Callahan et al., 2017). Additionally, chimeras sequences were identified and excluded. To determine the taxonomic classification of each ASV, the Naïve Bayes classifier was used against the 132 version of the SILVA database (Quast et al., 2013). The upstream analysis performed for the primer set comparison pipeline followed by Fadeev et al., (2021).

In Part 2, i.e. the comparative analysis between the blow and seawater samples, reads were truncated at 250 bp and 200 bp for forward and reverse sequences, respectively. Then, the same upstream steps and parameters reported above for the Part 1 analysis were followed.

3.5.2. Downstream sequence analysis

The ASV counts and taxonomy tables from the upstream analysis, together with the metadata table containing the sample information (type of sample, species, primer set, sample day), were used as input for the “phyloseq” R package (v.1.36) (a tool to import, store, analyze, and graphically display complex phylogenetic sequencing data that has already been clustered into ASVs) to create a phyloseq object and perform the downstream analysis (McMurdie & Holmes, 2013). The ASVs that were taxonomically unclassified at phylum rank or were not assigned to bacterial or archaeal lineages, and the undesirable lineages, such as “Chloroplast”, “Eukaryota” and “Mitochondria”, were excluded from further analysis.

The distribution and diversity of the prokaryotic community across the different primer sets used, and between the blow and seawater samples were investigated. First, rarefaction curves of the samples were created and plotted using the R-package “vegan” (v. 2.6.2; Oksanen et al., 2019). These curves reported the numbers of ASVs in function to sampling depth.

Secondly, the alpha diversity was analyzed for different subsets of samples by calculating four different indexes: Observed (count of unique ASVs in each sample – number of observed samples), Chao 1 (abundance-based estimator of species richness), Shannon (estimator of species richness and species evenness), and Inverse Simpson (estimator of species evenness), using the “phyloseq” package. Differences in alpha diversity were tested using the Wilcoxon signed rank test (part 1) and Wilcoxon rank-sum test (part 2), considering a level of significance of 0.05. Moreover, it was analyzed the beta diversity (community differentiation along habitat gradients) using the “vegan” package and the multidimensional plot with the “phyloseq” package. To test if the primers set used (Part 1) had a significant effect on the prokaryotic communities present in the different blow samples, a PERMANOVA statistical test was implemented. Therefore, two hypotheses were generated with $p\text{-max}=0,05$: H0: The use of different primer set does not influence the distribution of prokaryotic communities; H1: The use of different primer set influences the distribution of prokaryotic communities. The same was done in Part 2 to verify if there were significant differences between the prokaryotic communities detected in each type of samples (blow/seawater).

To calculate and see possible relations between shared and exclusive ASVs between the different types of samples, and different primer sets employed, Venn diagrams were constructed, using the “vegan” and “Venn” packages of R (Dusa, 2020). Lastly, the microbial communities’ taxonomic composition was investigated by creating taxonomy bar plots. The distribution of prokaryotes taxa across the samples was analyzed at six taxonomic levels: kingdom, phylum, class, order, family, and genus. Taxonomic distribution plots were performed using several R-packages, namely, “phyloseq”, “ggplot2” (v.3.3.6), “tidyverse” (v.1.32) and “scales” (v.1.20) (McMurdie & Holmes, 2013; Wickham, 2009; Wickham et al., 2019).

3.5.3. Pathogen database and screening

Due to the interest in studying the cetacean respiratory microbiome as an indicator of cetacean health, the taxonomy of blow samples at the phyla and genus level was compared to a pathogen database previously compiled by Apprill et al., (2017). This database is composed of human and animal pathogens that are potentially pathogenic to marine mammals. These pathogens were recognized by the American Biological Safety Association, and by previously published papers, like Raverty et al., (2016) that

identified putative pathogens of any marine mammal body site (Table S2 – Apprill et al., 2017). Additionally, to highlight differences between the shared and exclusive pathogens within the different primers sets used (V3-V4 and V4-V5) a Venn Diagram was made.

4. Results

4.1. Comparison of 16S rRNA primers for studies of blow prokaryotic communities

Environmental DNA was extracted from the blow of the short-finned pilot whale (Gma_1 – Gma_13). The two hypervariable regions of the 16S rRNA gene V3–V4 and V4–V5, were amplified and compared using Fadeev 2021 pipeline.

After the sequencing process, a total of 685 692 reads (with a mean of $57\,141 \pm 7\,920.8$ reads per sample) was obtained for the V3-V4 dataset; and 862 763 for the V4-V5 (with a mean of $71\,896.9 \pm 12\,990.5$ reads per sample). After the quality filter steps, the V3-V4 dataset showed a lower decrease in the number of sequences throughout the workflow compared to V4-V5, with $60.4 \pm 5.7\%$ and $48.5 \pm 7.7\%$ of sequences retained per sample, respectively. Therefore, the non-target sequences were eliminated. Consequently, the final output was 415 340 sequences in the V3-V4 dataset that were assigned to 2 764 ASVs (with a mean of $34\,611.7 \pm 6\,228.9$ reads per sample) and 417 733 sequences in the V4-V5 dataset that were assigned to 3 665 ASVs (with a mean of $34\,811.1 \pm 7\,554.7$ reads per sample). The cleaning process for each sample is indicated in Table 6.

Table 6 - Cleaning Process of the 16S rRNA sequences for each sample of the blow of short-finned pilot whales.

Sample ID	Sample type	Primer set	Raw sequences	Sequences clustered to ASVs	Filtered sequences	Cleaned sequences
Gma_1	blow	V3-V4	48 404	29 551 (61%)	666 (1%)	28 885 (60%)
Gma_2			64 270	44 316 (69%)	412 (1%)	43 904 (68%)
Gma_4			53 851	31 649 (59%)	363 (1%)	31 286 (58%)
Gma_5			56 840	30 079 (53%)	297 (1%)	29 782 (52%)
Gma_6			65 986	37 781 (57%)	189 (0%)	37 592 (57%)
Gma_7			59 640	31 587 (53%)	357 (1%)	31 230 (52%)
Gma_8			46 679	26 351 (57%)	146 (1%)	26 205 (56%)
Gma_9			68 874	46 654 (68%)	519 (1%)	46 135 (67%)
Gma_10			51 715	33 370 (65%)	1 205 (2%)	32 165 (62%)
Gma_11			67 625	43 739 (65%)	2 722 (4%)	41 017 (61%)
Gma_12			52 187	36 069 (69%)	2 885 (6%)	33 184 (64%)
Gma_13			49 621	35 365 (71%)	1 410 (3%)	33 955 (68%)
Gma_1	blow	V4-V5	74 101	47 918 (65%)	14 286 (19%)	33 632 (45%)
Gma_2			68 863	44 893 (65%)	6 732 (10%)	38 161 (55%)
Gma_4			74 979	45 795 (61%)	4 456 (6%)	41 339 (55%)
Gma_5			79 609	53 570 (67%)	6 207 (8%)	47 363 (59%)
Gma_6			67 448	44 101 (65%)	6 028 (9%)	38 073 (56%)
Gma_7			87 967	53 676 (61%)	8 183 (9%)	45 493 (52%)
Gma_8			61 857	38 462 (62%)	8 160 (13%)	30 302 (49%)
Gma_9			65 556	40 706 (62%)	13 534 (21%)	27 172 (41%)
Gma_10			58 244	35 598 (61%)	10 698 (18%)	24 900 (43%)
Gma_11			102 328	74 022 (72%)	38 923 (38%)	35 099 (34%)
Gma_12			60 659	36 693 (61%)	11 999 (20%)	24 694 (41%)
Gma_13			61 152	40 555 (66%)	9 050 (15%)	31 505 (52%)

In the rarefaction curves (Figure 5) is possible to see that the two sequence libraries, showed by the majority of the blows, were sequenced to complete saturation. The number of the reads obtained after sequencing, for each sample, contained most of the potential community richness covered by both primer sets.

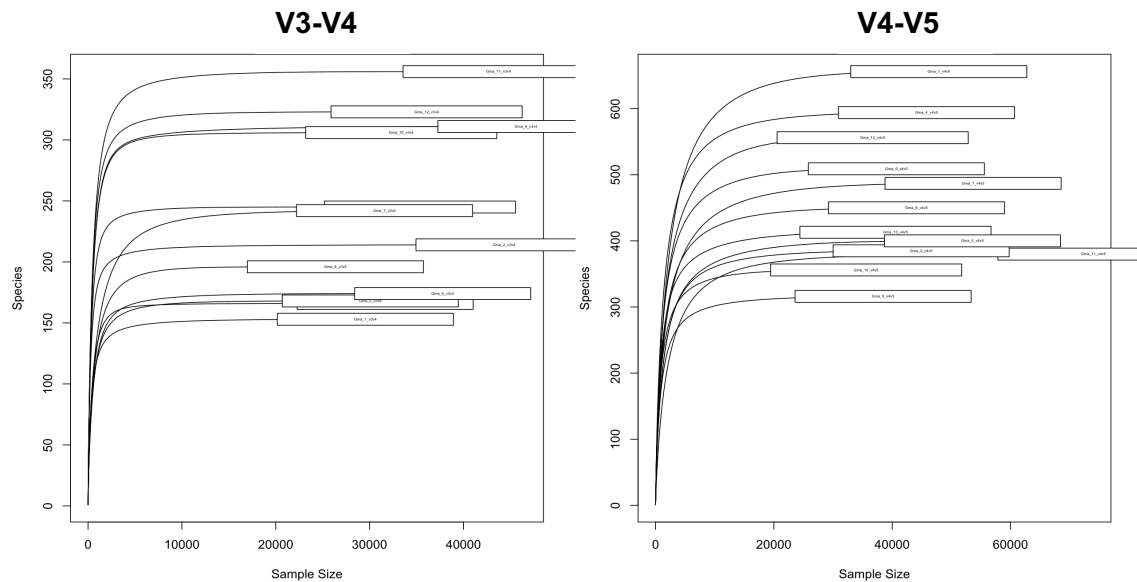


Figure 5 - Rarefaction curves of the V3-V4 dataset (on the left) and the V4-V5 dataset (on the right), for short-finned pilot whale blow samples.

4.1.1. Alpha diversity

Alpha diversity measures varied with the primer set (Figure 6). The Observed ASVs and Chao1 had lower values in the V3-V4 dataset compared to values of V4-V5 dataset. In the V4-V5 dataset, alpha indices values were higher and more agglomerate. In the V4-V5 dataset, sample Gma_4 was highlighted as an outlier, presenting highest values in Observed and Chao1 indexes. The Shannon and Inverse Simpson indexes demonstrated a more considerable ASVs diversity in the V3-V4 region. Sample Gma_12

of the V4-V5 dataset represented the lowest values among all the samples in Shannon and Inverse Simpson index.

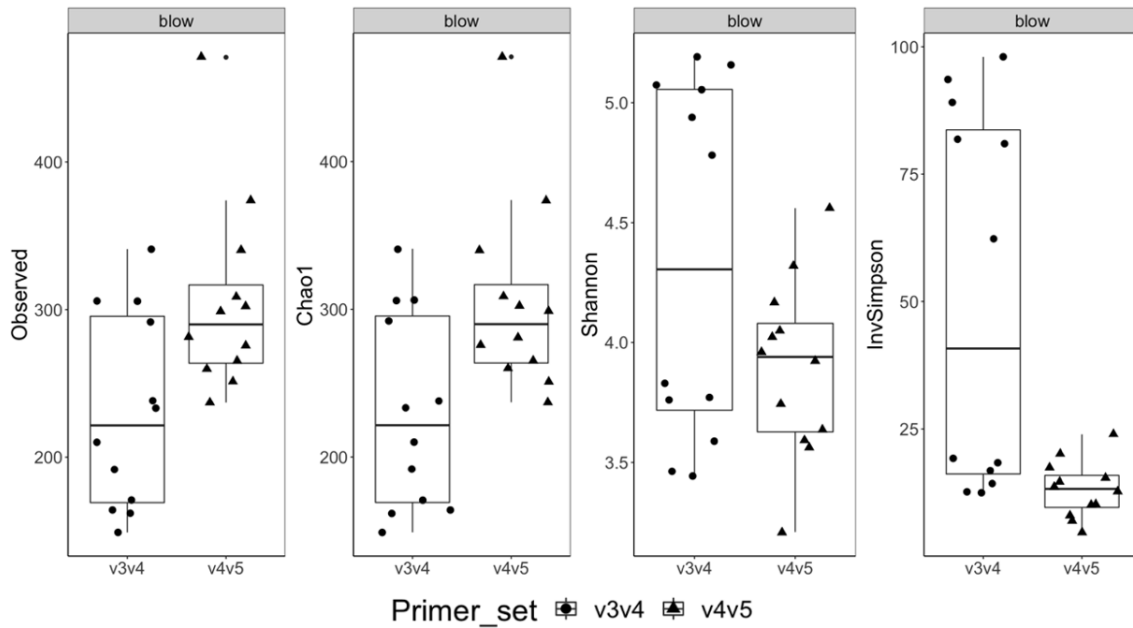


Figure 6 - Observed richness, Chao1, Shannon, and Inverse Simpson indices of alpha diversity among blow microbiome. Different primer sets were represented by different shapes.

Table 7 - Minimum, maximum, median, mean, and standard deviation of the different alpha diversity indices in V3-V4 and V4-V5 datasets, of short-finned pilot whale blow samples.

	Index	V3-V4				V4-V5			
		min	max	median	mean±sd	min	max	median	mean±sd
Blow	Observed	149	341	222	(230±67)	237	374	281	(290±40)
	Chao1	149	341	222	(230±67)	237	374	281	(290±40)
	Shannon	3.44	5.19	4.3	(4.3±0.7)	3.21	4.56	3.9	(3.9±0.4)
	Inverse Simpson	12.45	97.99	40.8	(50±37)	4.71	23.96	13.3	(13±6)

* the observed and Chao1 measures did not include sample Gma_4 of V4-V5 dataset

To verify if the primer set influenced or not the richness and the diversity of prokaryotic communities, a Wilcoxon Signed-Rank Test was performed.

Table 8 - Output of a Wilcoxon Signed-Rank Test, comparing the indexes of the short-finned pilot whale blow community richness obtained with two different primer sets (V3-V4 and V4-V5).

	Index	group1	group2	n1	n2	statistic	p	p.signif
Blow	Observed	V3-V4	V4-V5	12	12	33	0.0262	*
	Chao1							
	Shannon					91.5	0.273	ns
	Inverse Simpson					118	0.00681	**

Considering a level of significance of 0.05, the blow community richness for the Observed, Chao1, and Inverse Simpson measures was significantly different between the two primers sets (Table 8), with ca. 32% more bacterial ASVs in the V4-V5. Statistically significant differences were not found for the Shannon index (Table 8).

4.1.2. Beta diversity

The beta diversity of microbial communities was visualized by creating non-metric multidimensional scaling (nMDS) plots, based on Bray-Curtis dissimilarity (Figure 7).

The nMDS demonstrated an obvious separation of the samples based on the primer set used, being possible to visualize 4 distinct main clusters, two of each primer set. One of the clusters of the V3-V4 dataset was composed of samples Gma_1,4,5,6,7, and 8, and the other cluster with samples Gma_2,9,10,11,12, and 13. In the V4-V5 dataset, one of the clusters was composed of Gma_1,2,4,5,6,7, and 8 samples and the other one of Gma_8,9,10,11,12, and 13. In the V4-V5 dataset, the individuals that were traveling together when they were sampled, appeared in the same cluster and had more approximate values in the nMDS than in the V3-V4 dataset, as can be specifically seen for the individuals Gma_6 and 7, which had nearly the same values.

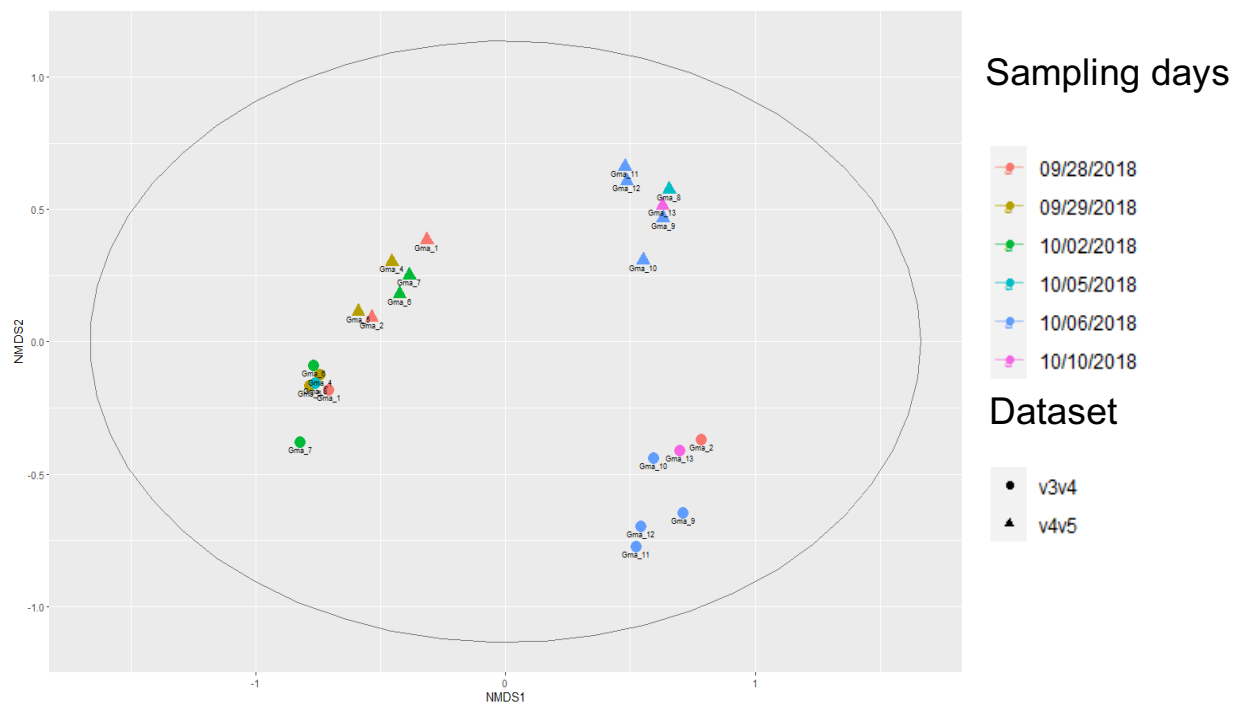


Figure 7- Non-metric multidimensional scaling (nMDS) (stress=0,07) of the microbiota found in short-finned pilot whale blow samples, in merged V3-V4 and V4-V5 datasets, based on Bray-Curtis Dissimilarity. The primer sets used are represented by different shapes and the days of sampling by different colors.

To verify if the primer set had a significant effect on the prokaryotic communities present in the different blow samples, a PERMANOVA Test was performed.

Table 9 - Output of the PERMANOVA analysis (Df=degrees of freedom, SS=sum of squares, R²= coefficient of determination, PR(>F)=p-value) to test for the influence of the primer sets (V3-V4 and V4-V5) on the distribution of prokaryotic communities of the blows of short-finned pilot whales.

	Df	SS	R ²	F	Pr(>F)
Primer_set	1	0,38948	0,24864	7,2802	0,006
Residual	22	1,17697	0,75136		
Total	23	1,56645	1		

The primer set used influenced the distribution of prokaryotic communities ($p=0.006$, lower than the level of significance of 0.05). Specifically, approximately 24% ($R^2=0.24864$) of the variation of the prokaryotic community distribution was explained by the different primers (Table 9).

4.1.3. Characterization of the prokaryotic community in V3-V4 and V4-V5 dataset

The composition of the blow prokaryotic communities was investigated between the two datasets, V3-V4 and V4-V5, to analyze the differences and commonalities between the use of the different primer sets. Considering the richness, a total of 1 890 bacterial and 38 archaeal taxa were recorded in V3-V4 dataset, whereas V4-V5 recorded 2 327 and 30 taxa, respectively.

Phylum Level

Taxonomic analysis at phylum level revealed that a core of 4 most abundant phyla was identified by the two datasets (Figure 8): *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, and *Proteobacteria*.

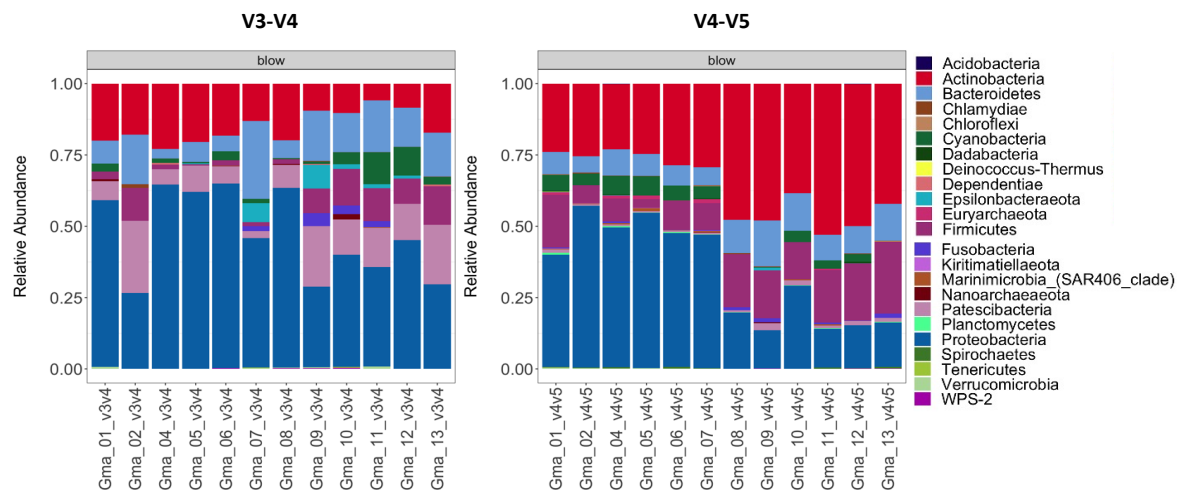


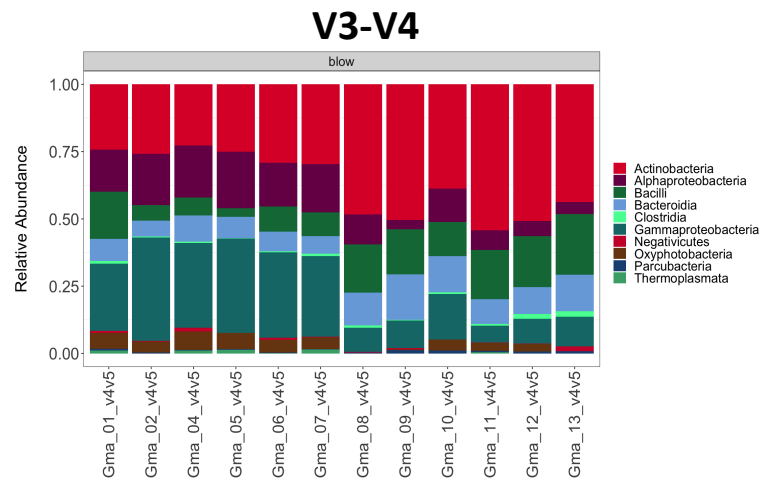
Figure 8 - Relative abundance of prokaryotic phyla identified across the short-finned pilot whale blow samples, by V3-V4 (on the left) and V4-V5 (on the right) datasets.

In the V3-V4 dataset, the prokaryotic community was distributed within 15 different phyla, 7 of them had more than 1% of the total proportion of all the ASVs sequences of this dataset. This dataset was dominated by sequences of the phylum *Proteobacteria* (47%). On the other hand, V4-V5 was distributed within 21 phyla, with only 5 phyla with more than 1% of the total proportion of the V4-V5 dataset. This dataset was dominated by *Proteobacteria* (37%) and *Actinobacteria* (34%). *Cyanobacteria* was present in more samples of the V4-V5 dataset.

Class Level

The V3-V4 dataset included 37 different classes (11 of them were highly abundant, with >1% of the sequence reads), and it was dominated by *Gammaproteobacteria* (31%), followed by the same proportion of *Actinobacteria* and *Alphaproteobacteria* (15%). The V4-V5 dataset was dominated by the *Actinobacteria* class (34%), with low abundance of the majority of the sequence reads (representing <1%) (30 out of 36). The V4-V5 dataset was mainly composed of 6 classes (*Actinobacteria*, *Gammaproteobacteria*, *Alphaproteobacteria*, *Bacilli*, *Bacteroidia*, and *Oxyphoyobacteria*), whereas the V3-V4 dataset had in its composition the same taxa plus *Gracilibacteria*, *Campylobacteria*, *Saccharimonadia*, and *Clostridia* (Figure 9).

Within the classes *Actinobacteria*, *Alphaproteobacteria*, *Bacteroidia*, *Clostridia*, *Gammaproteobacteria*, *Gracilibacteria*, *Parcubacteria*, and *Saccharimonadia*, there were many differences between datasets in the number of observed ASVs. These classes belong to the top 10 of both datasets, except *Clostridia* (position 11 of V3-V4 dataset), *Gracilibacteria* (position 13 of V4-V5 dataset), and *Saccharimonadia* (position 27 of V4-V5 dataset).



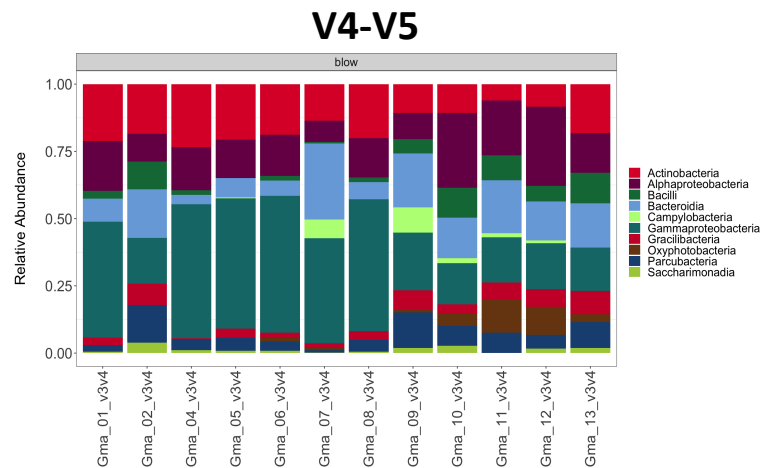
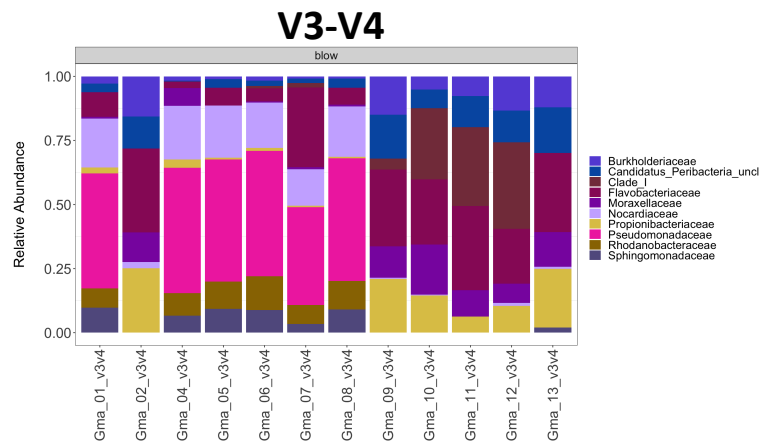


Figure 9 - Relative abundance of the top 10 prokaryotes classes identified across the short-finned pilot whale blow samples (V3-V4 dataset on the top and V4-V5 on the bottom).

Family Level

In the V3-V4 dataset, the prokaryotic community was distributed within 108 different families, and 24 families had more than 1% of the total proportion. This dataset was dominated by sequences of the family *Pseudomonadaceae* (18%). Whereas V4-V5 was distributed by 153 families, and only 17 families with more than 1% of the total proportion. Despite V4-V5 had more different families, most of them had a relative abundance <1%. *Propionibacteriaceae* (25%) was the predominant family of this dataset (Table S4). In this taxon it was possible to visualize that some samples had similar compositions (Figure 10).



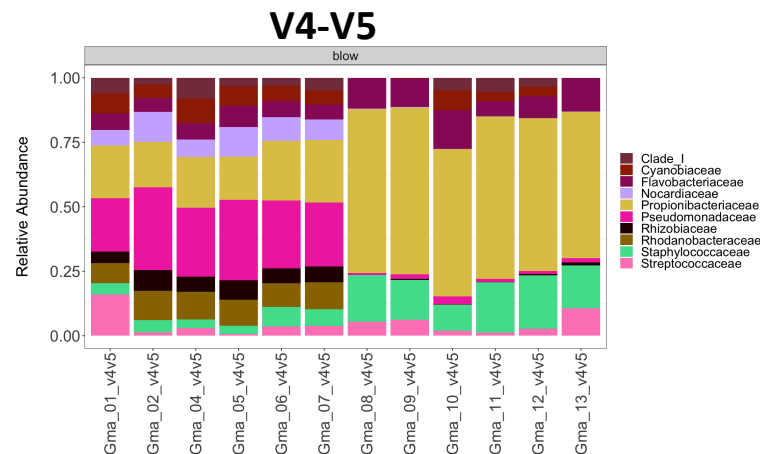


Figure 10 - Relative abundance of the top 10 prokaryotes families identified across the short-finned whale blow samples (V3-V4 dataset on the top and V4-V5 on the bottom).

For the V3-V4 dataset, the similarities in taxonomic composition among samples Gma_1,4,5,6,7, and 8; and between the samples Gma_2,9,10,1,12, and 13. As for the V4-V5 dataset, the similarities were between the samples Gma_1,2,4,5,6,7, and between the samples Gma_8,9,10,11,12, and 13 (Table 9). These clusters of samples were the same as those identified in the nMDS analysis (Figure 7). In both datasets, some families of the top 10 were common: *Nocardiaceae*, *Propionibacteriaceae*, and *Rhodanobacteraceae* (Figure 10).

The taxa *Clade_I* (included in the class *Alphaproteobacteria*) comprised 3% of all sequences in both datasets, however, the number of ASVs belonging to this taxa in V3-V4 was less than half compared to the V4-V5 dataset (total of 10 and 20 ASVs, respectively). The sequence proportion of the family *Flavobacteriaceae* (class *Bacteroidia*) was higher in the V3-V4 dataset (9%), in relation to 6% of the V4-V5 dataset, but it consisted of only one-third ASVs in the V3-V4 dataset, compared to the V4-V5 dataset (total of 10 and 32, respectively, Figure 11).

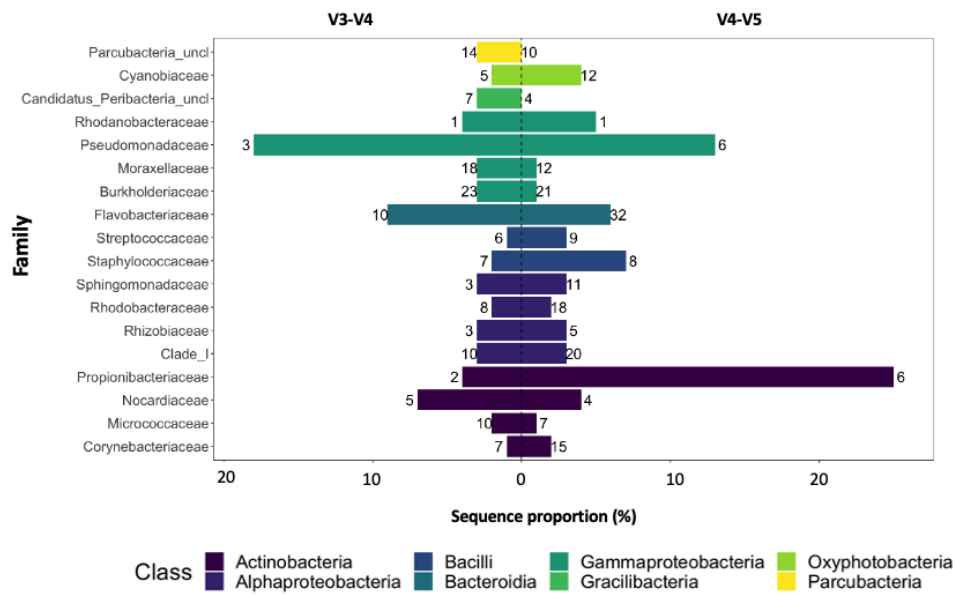


Figure 11 - Major taxonomic families obtained for the short-finned pilot whale blow samples (in V3–V4 and V4–V5 datasets). The total sequence proportion of each family in the V3–V4 (left side) and V4–V5 (right side) datasets is represented in the x-axis. The numbers in each column represent the number of observed ASVs affiliated with each taxonomic family. Color coding represents the different taxonomic classes. Only families that comprised at least 1% of sequences, in at least one of the datasets, were included in the figure.

Genus Level

Concerning the lower taxonomic level, the V3-V4 dataset detected 24 genera (with more than 1% of the total proportion of the sequences) associated with 128 different ASVs. This dataset was dominated by sequences of the genus *Pseudomonas* (18%) with only 3 ASVs associated with this genus. Whereas V4-V5 had only 14 genera (with more than 1% of the total proportion) associated with 79 different ASVs. They were dominated by *Cutibacterium* (25%). The same pattern observed in the families in terms of the composition of the samples, occurred in the analysis of the genus. The top10 genera included some taxa that are common to both datasets: *Cutibacterium*, *Dyella*, *Flavobacterium*, *Prochlorococcus_MIT9313*, *Pseudomonas*, *Rhodococcus*, and *Shingomonas*.

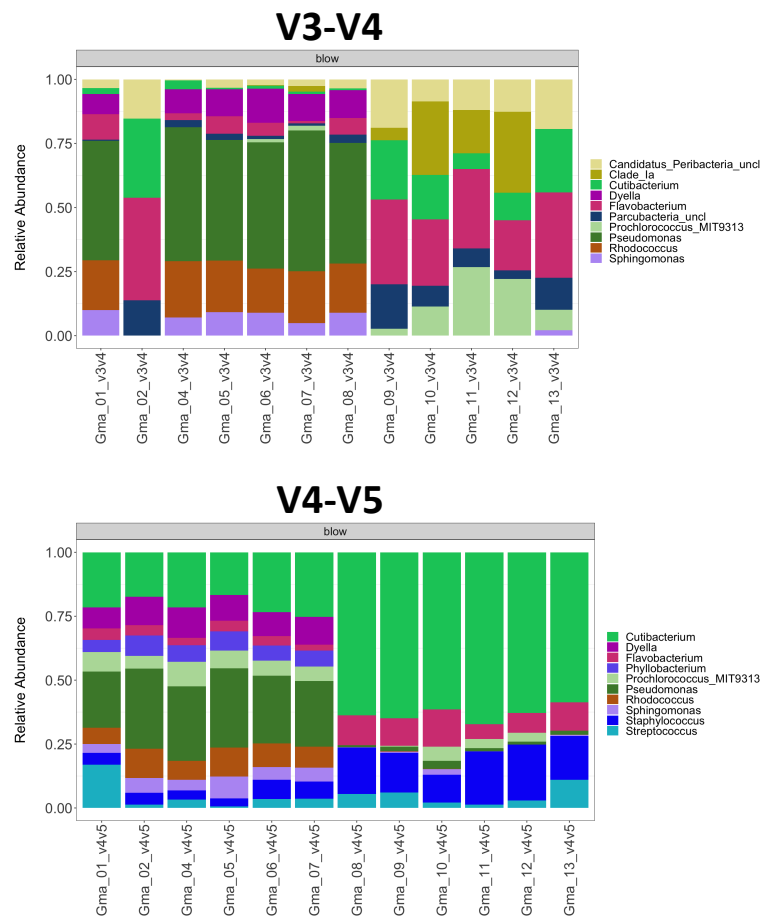


Figure 12 - Relative abundance of the top 10 prokaryotic genera identified across the short-finned pilot whale blow community (V3-V4 dataset on the top and V4-V5 on the bottom).

In the V3–V4 dataset, the ASVs were merged into 104 different genera and 53 lineages that were affiliated to higher taxonomic ranks. In the V4–V5 dataset, the ASVs were merged into 184 different genera and 68 lineages that were affiliated with higher taxonomic ranks. Overall, at this level 110 (36.8% of the total) lineages were observed in both datasets. Considering all the taxa, in the V3–V4 dataset 47 (15.7% of the total) lineages were absent from the V4–V5 dataset. Moreover, in the V4–V5 dataset there were 142 (47.5% of the total) lineages that were absent from the V3–V4 dataset (Figure 13). This result highlighted the presence of a rare biosphere in the V4-V5 dataset, with a highly diverse pool of low relative abundance microbial populations reported in samples, this is also important to the detection of potential pathogenic taxa.

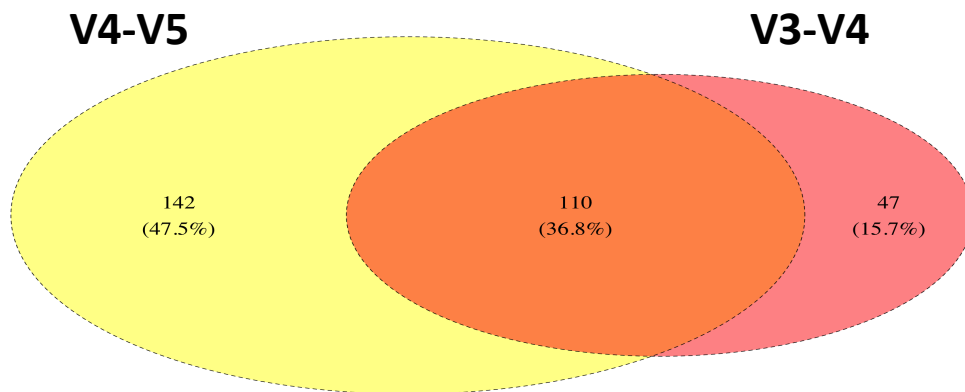


Figure 13 - Venn diagram of sequences abundances of all ASVs (Amplicon Sequence Variants) of V3-V4 (red) and V4-V5 (yellow) datasets at genus rank, for short-finned pilot whale blow samples.

4.1.4. Cetacean respiratory pathogens detected in the short-finned pilot whale blow

A deeper analysis was performed to highlight the presence of potential pathogenic species within the blow microbiome of pilot whales by the two primer sets. Several ASVs belonging to different taxa that comprise potential pathogenic species were identified (377 ASVs in V3-V4 and 416 in V4-V5 dataset) and some of them belonged to taxa linked to infectious diseases in marine mammals.

At a genus level, with the V3-V4 dataset, 59 different potential pathogenic genera were detected, and with the V4-V5, there were 86 different potential pathogenic genera. The most abundant potential pathogenic genus in both datasets was *Pseudomonas*. In the V3-V4 dataset, the more abundant potential pathogenic genera were followed by *Flavobacterium*, *Rhodococcus*, *Staphylococcus*, *Porphyromonas*, *Sphingomonas*, *Streptococcus*, and *Acinetobacter*. Whereas in the V4-V5 were *Staphylococcus*, *Flavobacterium*, *Streptococcus*, *Rhodococcus*, *Sphingomonas*, *Porphyromonas*, and *Acinetobacter* (Figure 14).

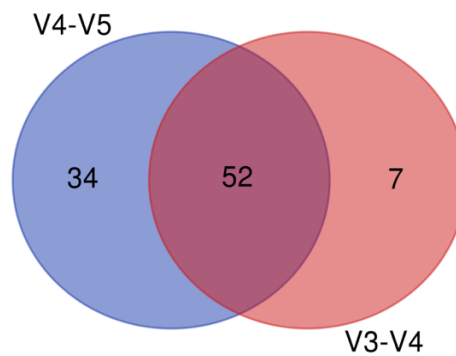


Figure 14 - Venn diagram analysis of all potential pathogenic genera detected in V4-V5 (blue) and V3-V4 (red) datasets, for short-finned pilot whale blow samples.

4.2. Comparison between the blow of short-finned pilot whales and seawater samples

For the comparison between the blows and the seawater, 8 blows and 3 seawater samples were used, targeting V4-V5 hypervariable region of the 16S rRNA gene.

After the sequencing process, a total of 140 891 sequences were obtained: 102 838 for blow samples (with a mean of $12\,855 \pm 6\,333$ reads per sample) and 38 053 for seawater samples (with a mean of $12\,684 \pm 6\,472$ reads per sample). Data processing was performed by DADA 2 bioinformatic pipeline.

In blow and seawater sterivex samples, ca. 7% and 20% of sequences respectively, were non-target taxa and were removed from the dataset. Blow samples showed a lower decrease in the number of sequences throughout the preprocessing workflow in relation to sterivex samples, with $43 \pm 8.5\%$ and $31 \pm 4.6\%$ of sequences retained per sample, respectively. Consequently, the final output was 56 137 sequences, with 44 621 sequences in the blow samples that were assigned to 1 116 ASVs (with a mean of $5\,578 \pm 2\,881$ reads per sample) and 11 516 sequences in the seawater sterivex samples that were assigned to 338 ASVs (with a mean of $3\,839 \pm 1\,891$ reads per sample). The cleaning process for each sample can be seen in Table 10.

Table 10 - Cleaning process of the 16S rRNA sequences for each sample of blow of short-finned pilot whale and seawater.

Sample Name	Sample type	Primer set	Raw sequences	Sequences clustered to ASVs	Filtered sequences	Cleaned sequences
Blow_15	blow	V4-V5	5 583	3 059 (55%)	1 607 (29%)	1 452 (26%)
Blow_16			8 735	5 302 (61%)	492 (6%)	4 810 (55%)
Blow_17			8 076	4 265 (53%)	586 (7%)	3 696 (46%)
Blow_18			15 359	6 586 (43%)	378 (3%)	6 208 (40%)
Blow_19			24 629	12 031 (49%)	1 218 (5%)	10 813 (44%)
Blow_20			18 550	8 554 (47%)	649 (4%)	7 905 (43%)
Blow_21			9 636	3 746 (39%)	47 (0.5%)	3 699 (38%)
Blow_22			12 270	6 436 (53%)	398 (3.2%)	6 038 (49%)
A_St	water	V4-V5	20 008	10 949 (55%)	4 927 (25%)	6 022 (30%)
B_St			7 735	3 083 (40%)	323 (4%)	2 760 (36%)
C_St			10 310	5 815 (57%)	3 081 (30%)	2 734 (27%)

In the rarefaction curves (Figure 15), it was possible to see that the majority of samples reach the plateau, therefore the number of reads obtained after the sequencing, for each sample, was sufficient to describe the entirety of the diversity harbored in the samples.

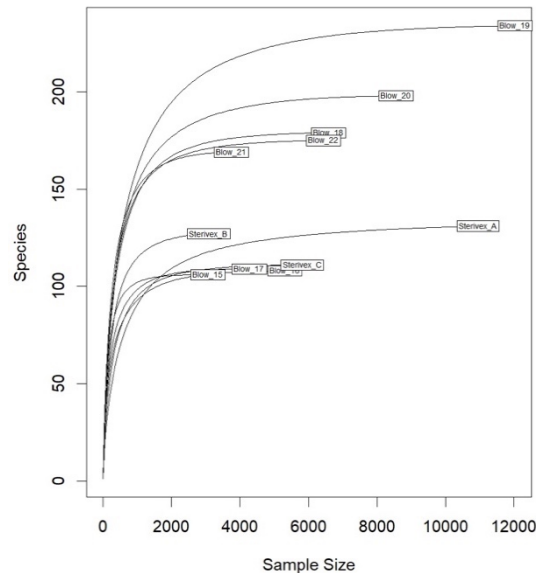


Figure 15 - Rarefaction curves of the V4-V5 dataset, for short-finned pilot whale blow and seawater samples.

4.2.1. Alpha diversity

To study the diversity of prokaryotic communities across different types of samples (blow and seawater), the Shannon, and Simpson indexes, along with the number of observed ASVs were calculated for each sample (Table S7 and Figure 16).

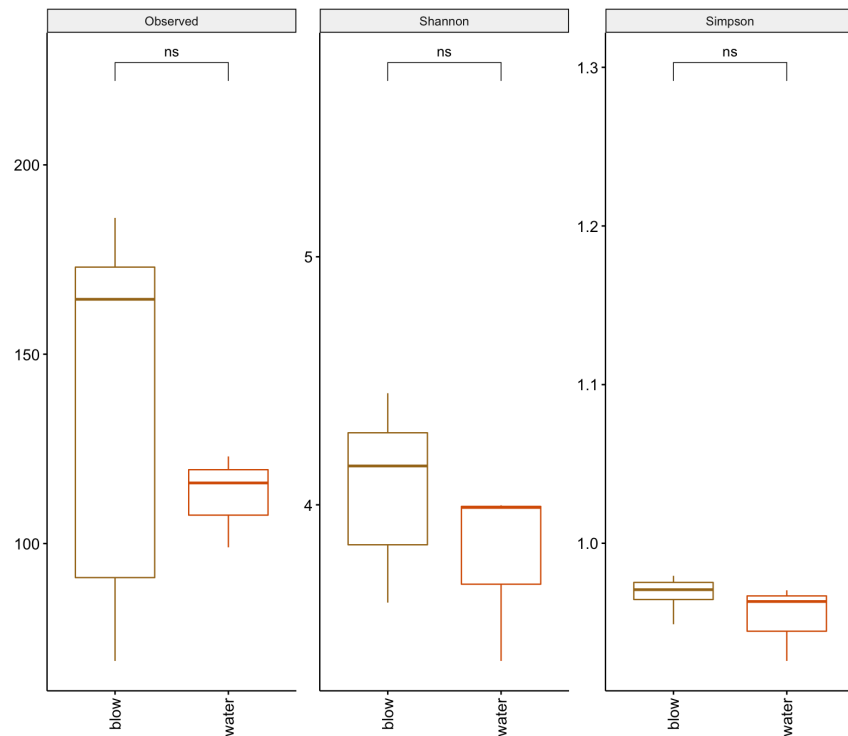


Figure 16 - Observed richness, Shannon, and Simpson indices of alpha diversity among blow and seawater microbiomes, amplified with the V4-V5 primer set. Blow samples (brown) are represented on the left and seawater control samples (orange) are on the right.

A Wilcoxon rank sum Test was performed to test for significant differences between the richness and the diversity of prokaryotic communities among blow and seawater samples (Table 11).

Table 11 - Output of a Wilcoxon rank sum Test, comparing indexes of alpha diversity for microbial community, between the blows of short-finned pilot whale and seawater samples (using the primer set V4-V5).

measure	group1	group2	n1	n2	statistic	p	p.signif
Observed	blow	sterivex	8	3	15	0.63	ns
Chao1					18	0.279	ns
Shannon					18	0.279	ns
Inverse Simpson					18	0.279	ns

Considering a level of significance of 0.05, the blow and seawater community richness and diversity were not significantly different (Table 11). However, blow samples have a higher richness and diversity than water samples. Blow_19 which was a sample retrieved from a pool of individuals was the one with a higher richness (Table 5).

Table 12 - Minimum, maximum, median, mean, and standard deviation of the different measures of alpha diversity in the prokaryotic communities of the short-finned pilot whale blow and seawater samples (using the primer set V4-V5).

		blow				seawater			
		min	max	median	mean±sd	min	max	median	mean±sd
Measures	Observed	69	186	164.5	(139.5±49)	99	123	116	(112±12)
	Chao1	69	186.33	164.5	(139.5±49)	99	123.1	117.4	(113±13)
	Shannon	3.6	4.5	4.2	(4.1±0.3)	3.4	3.99	3.9	(3.8±0.4)
	Inverse Simpson	19.6	48.7	34.4	(34.2±10)	13.5	33.7	27.2	(24±10)

4.2.2. Beta diversity

The beta diversity of both blows and seawater was visualized through a nMDS plot, based on Bray-Curtis dissimilarity (Figure 17). The nMDS demonstrated a separation of the samples based on the type of sample (blow/seawater). Sample Blow_16 and 17, as well as Blow 20 and 22 were more approximate in the nMDS, and these individuals were traveling together when they were sampled.

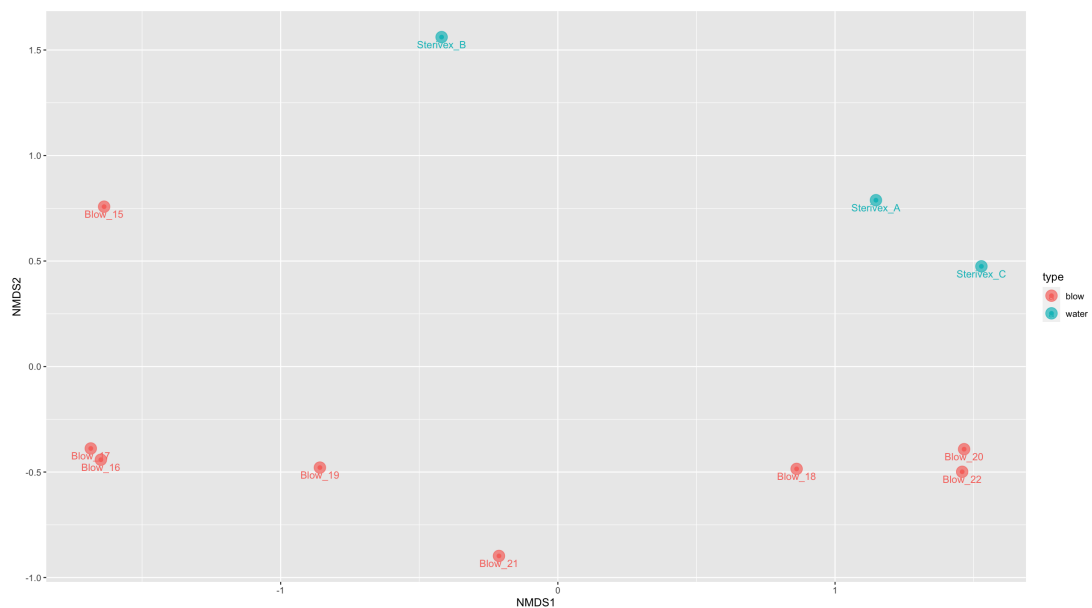


Figure 17 - Non-metric multidimensional scale (nMDS) (stress=0,04) of the microbiota found in short-finned pilot whale blow and sterivex samples, based on Bray-Curtis dissimilarity. The different types of samples are represented by colours, red for blow and blue for seawater.

To verify if different type of samples had a significant effect on the prokaryotic communities present among blow and seawater samples, a PERMANOVA Test was performed.

Table 13 - Output of the PERMANOVA (Df=degrees of freedom, SS=sum of squares, R²= coefficient of determination, PR(>F)=p-value)analysis to test for the influence of different types of samples (blow of short-finned pilot whales/seawater samples) on the distribution of prokaryotic communities (using the primer set V4-V5).

	Df	SS	R ²	F	Pr(>F)
type	1	0.6672	0.19732	2.2124	0.05
Residual	9	2.7143	0.80268		
Total	10	3.3815	1		

Almost significant differences were found between the prokaryotic communities found in blow and sweater samples ($p=0.05$, $\alpha=0.05$). Specifically, approximately 19% ($R^2=0.19732$) of the variation of the prokaryotic community distribution was related to the different type of samples (Table 13).

4.2.3. Characterization of the prokaryotic community in short-finned pilot whale blow and seawater samples

The prokaryotic composition was investigated between the two types of samples, blow and seawater, and the outcome was compared across the different samples collected and between the different days of sampling to analyze the differences and commonalities between them.

Phylum level

Metabarcoding dataset from blow and seawater samples showed that prokaryotic communities were distributed within 16 phyla, and the *Bacteroidetes* was the most abundant phylum. Taxonomic analysis, at this level, revealed that there is no clear differentiation between blow and seawater samples. However, some taxa are more representative of blow samples which is the case of *Cyanobacteria*. All samples were composed of *Actinobacteria*, *Bacteroidetes*, and *Proteobacteria* (Figure 18).

On day 11/09/2018, *Cyanobacteria* and *Thaumarchaeota* were unique to short-finned pilot whale respiratory vapor and not found in the matching seawater. Phyla

Euryarchaeota and *Marinimicrobia* (SAR406_clade) mark more presence in the Blow_15 sample - this was also the sample with a higher number of representative phyla. On day 11/30/2018, *Acidobacteria*, *Actinobacteria*, and *Cyanobacteria* were more representative of blow samples, whereas *Planctomycetes* and *Thaumarchaeota* were more representative of seawater samples. On day 12/12/2018, the Blow_19 had the highest percentage of *Proteobacteria*.

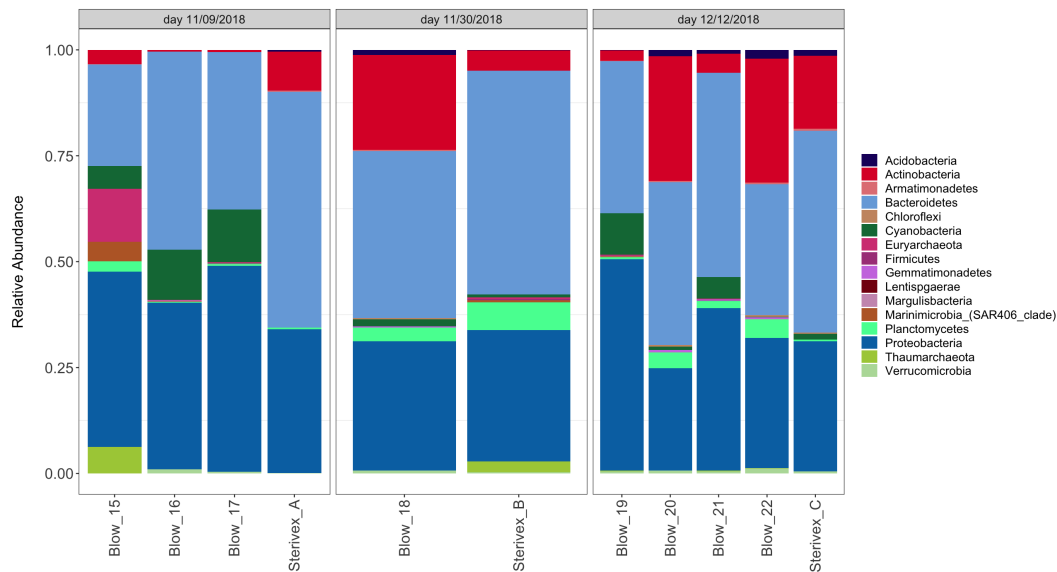


Figure 18 - Relative abundance of prokaryotes phyla identified across the short-finned pilot whale blow and seawater samples separated by three sampling days (11/09/2018; 11/20/2018; 12/12/2018) – using the primer set V4-V5.

Genus level

At the Genus level, the blow and seawater samples contained 108 taxa. Although a significant portion of these taxa was found in both seawater and blow, a difference between the genera detected in each type of samples was evident. Some taxa are more representative of blow, while others have a higher proportion of abundance in seawater. This was the case of the genera NS3a_marine_group (family *Flavobacteriaceae*) that was most dominant in seawater samples, and *Planktomarina* was most dominant in blow samples. Overall, the most dominant taxon across all samples was the Unclassified genus of Bacteroidetes. *Synechococcus_CC9902* was only present in the blow samples (Figure 19).

On day 11/09/2018, *Amylibacter* and *Fluviicola* have higher abundance in blow samples. In seawater samples, *Limnohabitans*, *NS3a_marine_group*, and *Sediminibacterium* were unique to them.

On day 11/30/2018, *hgcl_clade* and *Sediminibacterium* were more representative of blow samples, whereas *NS3a_marine_group* was more representative of seawater samples.

On day 12/12/2018, *Amylibacter*, *Planktomarina*, and *Synechococcus_CC9902* were unique to the two blow samples (Blow_19 and Blow_21).

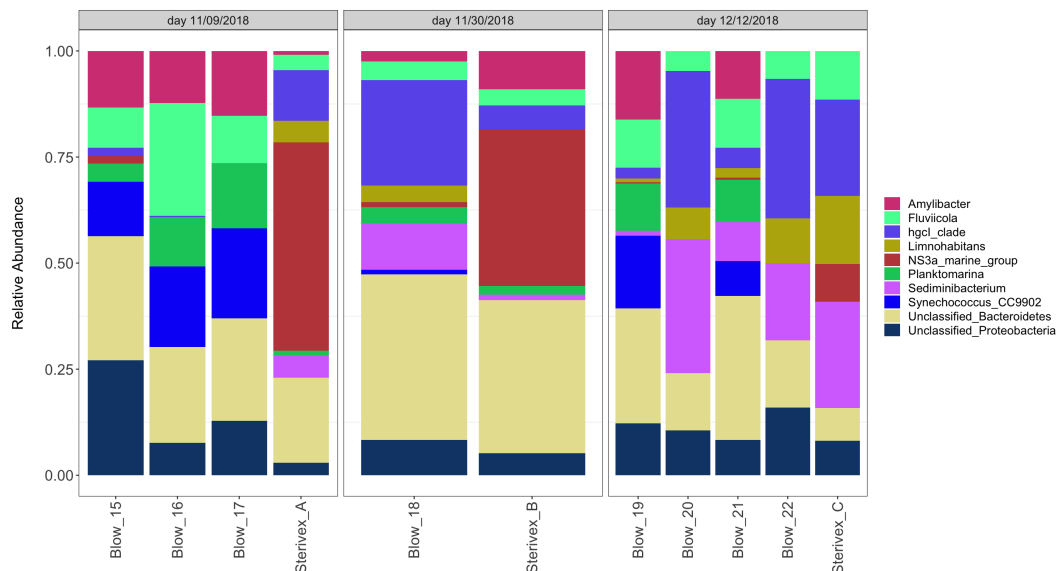


Figure 19 - Relative abundance of the top 10 prokaryotic genera identified across the short-finned pilot whale blow and seawater samples separated by three sampling days (11/09/2018; 11/20/2018; 12/12/2018) – using the primer set V4-V5.

4.2.4. Venn Diagram Analysis

Blow and seawater control samples shared a total of 107 ASVs, with the blows showing the highest number of unique taxa (359) in relation to seawater samples (143) (Figure 20).

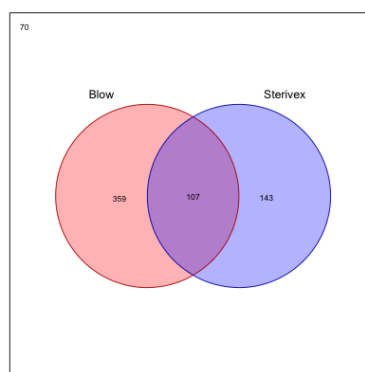


Figure 20 - Venn Diagram analysis of short-finned pilot whale blows (red) and seawater sterivex samples (blue).

5. Discussion

Sampling

The main goal of non-invasive sampling techniques for cetaceans is to avoid disturbing, injuring, or negatively influencing the target species while collecting samples (Robinson & Nuuttila, 2020). During the fieldwork activities reported in the present research, none of the short-finned pilot whales appeared to be negatively impacted by the sampling procedure used, with no evident changes in dolphin swimming patterns or anomalous behavior, suggesting that our blow-sampling method caused minimal distress to the individuals.

The sampling kit used here (PERFORMAgene), has been employed in many studies on different animals such as livestock (bovine/ovine/caprine) (Neary et al., 2014) and companion animals (canine/feline/equine) (Colpitts et al., 2022; Sacco et al., 2017). This kit is suitable for genotyping, sequencing, parent-child appraisal, and biobanking. Until now, no studies in the literature reported the use of the PERFORMAgene kit for sampling cetacean blow. In the present work, this kit was tested for the first time for the collection of short-finned pilot whales EBC. Environmental DNA was extracted from the swab of the PERFORMAgene kit, and although low concentrations of DNA were obtained, the method allowed for the blow metabarcoding analysis.

Primers set comparison

In this study, the prokaryotic community harbored in the respiratory tract of short-finned pilot whale individuals was described for the first time, reporting also a comprehensive comparison of the performance of different 16S primer sets' (V3-V4 and V4-V5). Both hypervariable 16S regions are recommended in the literature for assessing marine microbial diversity (Klindworth et al., 2013; Walters et al., 2015).

The use of different hypervariable regions, as well as different types of samples, storage, methods of DNA extraction, and 16S databases, can influence the obtained data and the interpretation of the results. In the study of McNichol et al. (2021), comparing commonly-used primers, with >300 million rRNA gene sequences retrieved from marine metagenomes around the world, the best-performing primers, when comparing predicted median coverage to Bacteria and Archaea of 16S rRNA, were 515Y/926R (regions V4-V5) and 515Y/806RB (region V4). The V4-V5 regions are currently recommended to

target marine microbes (including both bacteria and archaea) by the Earth Microbiome Project (Walters et al., 2015). On the other end, the primers 341F and 806R used in this study to amplify the V3-V4 regions of the 16S rRNA bacterial and archaeal genes were already applied in a previous blow study (Centelleghé et al., 2020). Other prior studies analyzing the blow microbiota have amplified the regions V3-V4, but used different primers in relation to the present study (Bik et al., 2016; Robles-Malagamba et al., 2020).

One of the aims of the present research was to assess the advantages and limitations in the amplification of both V3-V4 and V4-V5 regions, in regard to the obtained diversity and richness, without the influence of technical biases (same experimental and bioinformatic processes on the same samples). Moreover, with this work we also intended to select of an optimally DNA extraction method and a single PCR primer set (or different applications of each primer set), that can be applied to the study of the microbiome of blow samples. The results denote that the different hypervariable regions analyzed provided different degrees of resolution in taxonomic identification resulting in different estimates of microbial community composition.

The metabarcoding analyses confirmed and extended previous studies demonstrating that the respiratory microbiome of cetaceans can be determined using EBC (Apprill et al., 2017; Nelson et al., 2015, 2019; C. Vendl et al., 2019, 2021; Catharina Vendl, Slavich, Wemheuer, et al., 2020). In previous studies, it was already demonstrated that blow samples are a close approximation to the microbiome of blowhole swabs, thereby removing the need for more invasive blowhole swabbing procedures (Bik et al., 2016; Lima et al., 2012; Robles-Malagamba et al., 2020). This study focuses on prokaryotic community, however, besides bacteria and archaea, also protists, viruses, and fungi may be key indicators of respiratory diseases (Geoghegan et al., 2018), and these organisms can be detected in larger-volume samples using a metagenomic approach. Despite that, with the examination of blow microbial diversity, it was possible to see a diverse and rich community of microorganisms, with several features that may be useful for monitoring health.

(A) Diversity

Alpha diversity measures differed between primer pairs. On an ASV level, our results showed that the V4-V5 dataset captured more richness (higher values in the Observed ASVs and Chao1 measures), whereas the V3-V4 resulted in more diversity (higher values in Shannon and Inverse Simpson measures).

Beta diversity was analyzed using one of the most frequently used measures of community similarity, the Bray-Curtis dissimilarity. This measure is a statistic index used to quantify the compositional dissimilarity between two different sites, depending on the two communities' counts of shared and non-shared specimen (0 means the two sites have the same composition - sharing all the species; and 1 means the two sites do not share any species). In this case, this analysis revealed that samples were grouped according to the different primer sets used. Despite the separation of the samples in different clusters depending on the primer set used, the V4-V5 dataset appeared to be a better representation of the distribution of prokaryotic communities. In this dataset, individuals that were traveling together when they were sampled appeared within the same cluster, and had more approximate values in the nMDS, than in the V3-V4 dataset. There is some evidence that sociality affects microbes in the respiratory tract, and this seems to be the case of short-finned pilot whales, contrary to a previous study on long-finned pilot whales (Catharina Vendl, Slavich, Nelson, et al., 2020). Nevertheless, this hypothesis needs further evidence, since at this moment it is uncertain if individuals who were traveling together when they were sampled belong to the same pod or not. Cetaceans exhibit behaviors, such as surfacing and breathing near each other, or feeding cooperatively, which could facilitate the transfer of microbes and the spread of pathogens between individuals (Apprill et al., 2017; Bogomolni et al., 2008). This process has been recognized as an exclusive and important aspect of social living providing health benefits to animals, and access to associated microbes may be a driving force in the evolution of sociality (Lombardo, 2008).

Bacterial diseases, such as morbillivirus infection (*Paramyxoviridae* family) and fitotoxins, have been notorious to be the main causes of large-scale mass mortality in cetaceans across different geographic areas (Groch et al., 2021). Previous studies reported that short-finned pilot whales could be affected by the morbilliviruses in a non-virulent form, which may then become virulent and lethal in other marine mammals (Di Guardo & Mazzariol, 2016; Sierra et al., 2016). Morbillivirus is thought to be transmitted through blow exchange (Van Bresse et al., 2014). In some areas, short-finned pilot whales associate frequently with bottlenose dolphins, acting as a vector for morbilliviruses to these animals (Olson, 2009).

(B) Taxonomic profiles

This study provides baseline information on microbiota found in the blow of short-finned pilot whales.

All samples from both datasets shared a main core microbiota in their blow, composed by *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, and *Proteobacteria* phyla. Nevertheless, the dominant ASVs were not the same between the results obtained from the amplification of V3-V4 and V4-V5 regions. The analysis of the blow core microbiome could provide useful features for the health monitoring of cetaceans worldwide. Lima et al. (2012) described a blow temporal analysis of captive dolphins, suggesting that microbial community composition in healthy animals is quite stable and that individual dolphins harbor consistently unique microbial communities. Several studies provided preliminary evidence that cetaceans host a core group of bacteria associated with the respiratory system (Apprill et al., 2017; Bik et al., 2016; Johnson et al., 2009; Lima et al., 2012). Apprill and colleagues (2017) detected, on average, 25 sequence clusters that were found in 100% of humpback whales, accounting for 36% of the microbiota, being one of the most extensive core microbiotas found in any mammal to date. This large core microbiome was shared across individual whales from populations separated geographically into two ocean basins. Out of these core microbes, 20 closest phylogenetic relatives have been previously found in marine mammals (normally in the mouths and blowholes of bottlenose dolphins), indicating that this core microbiome assemblage is unique to marine mammals and may denote a healthy, non-infected pulmonary system. Conversely, Vendl et al. (2019) found low microbiota richness and a small core microbiome in humpback whales off the coast of Australia. The animals sampled in Vendl et al. (2019) were migrating for four months, which is the period of fasting; whereas, the animals sampled in the study by Apprill et al. (2017) were at their foraging sites and early stages of migration. The authors of the study of Vendl et al. (2019) hypothesized that the lack of a core microbiome might be related to the animals' physiological state at the sampling time. Also, none of the five core genera detected matched with any core found in other studies that investigated cetacean blow (Apprill et al., 2017; Johnson et al., 2009; Lima et al., 2012; Pirota et al., 2017).

In this study, the most dominant phylum recovered from both datasets was *Proteobacteria*. This result agrees with previous studies made for other species of cetaceans, such as Atkinson. (2021), April et al. (2017), Centelleghes et al. (2018), Nelson et al. (2019), and Vendl et al. (2019). Lima and colleagues (2012) study showed

that the aforementioned phyla, here present in both datasets (*Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, and *Firmicutes*), plus the *Fusobacteria* and unclassified *Bacteria*, represented 98% of the total community composition in bottlenose dolphins. In the present work, *Fusobacteria* was detected, but as a low abundant phylum (<1% of the total sequence reads) in both datasets. It also contrasts with the study of Nelson et al., 2019, where this taxon accounted for 7,2% of the total community. In the study of Pirotta and colleagues (2017), the most common bacteria identified at the phylum level were *Proteobacteria*, *Firmicutes*, and *Actinobacteria*.

The presence of *Cyanobacteria*, a seawater species, in blow samples is not surprising. It has already been detected in previous studies of blow microbiota (Pirotta et al., 2017). *Epsilonbacteraeota*, a phylum with a marked presence in the V3-V4 dataset, has been already identified as part of the core microbiota of bottlenose dolphins' respiratory system (Lima et al., 2012). The phyla *Euryarchaeota*, *Dadabacteria*, *Kiritimatiellaeota*, *Acidobacteria*, and *Deinococcus-Thermus*, in the present research only found in the V4-V5, were not reported in previous studies investigating the blow microbiota of different cetacean species. The same occurs with *Dependentiae* of the V3-V4 dataset. Contrary, *Planctomycetes*, *Chloroflexi* (Bik et al., 2016), and *Spirochaetes* (Bik et al., 2016; Lima et al., 2012; C. Vendl et al., 2021), here only found in the V4-V5 dataset, were described in previous studies of blow microbiota. This also happens with the *Tenericutes* phylum, here exclusive to the V3-V4 dataset, which, in a previous work, was the most dominant phylum in samples of dolphin forestomach (Bik et al., 2016).

Regarding the class level, the *Gammaproteobacteria*, *Actinobacteria*, and *Alphaproteobacteria* were the classes that dominated both datasets (with different proportions). The *Gammaproteobacteria* class is typically correlated with dominating classes in blow samples (Pirotta et al., 2017). In the study of Johnson et al., (2009), this class accounts for an average of 55% of all sequences. Similar findings were described in Lima et al., (2012), where this taxon accounts for an average of 75% of sequences across all blow samples. This class was also encountered in the mouth of bottlenose dolphins (Bik et al., 2016; Catharina Vendl, Slavich, Wemheuer, et al., 2020). *Alphaproteobacteria* was described to be much more common in seawater samples (Bik et al., 2016; Pirotta et al., 2017). However, in the study of Nelson et al., (2019), *Alphaproteobacteria* (26.1%) and *Gammaproteobacteria* (25.8%) had nearly the same percentages within the blow microbiome community. In contrast to other studies, the blow microbiota of short-finned pilot whales, assessed in the present study, did not have

the order *Cardibacteriales* (*Gammaproteobacteria*) as an abundant bacterial taxon (Bik et al., 2016; Johnson et al., 2009; Lima et al., 2012).

The *Flavobacteriaceae* family, here detected in both datasets, was already identified in the blow of bottlenose dolphins (C. Vendl et al., 2021) and humpback whales (Catharina Vendl, Slavich, Wemheuer, et al., 2020). This family has been reported as a significant contributor to the skin microbiotas of whales (Apprill et al., 2011, 2020).

The *Cutibacterium* genus (dominant in the V4-V5 dataset) is a typical dominant microbial community in the nasal microbiota of middle-aged adults (Kumpitsch et al., 2019). The genus *Prochlorococcus*, detected in both datasets, is one of the most abundant bacteria in the ocean (Zehr et al., 2017), and it is probably a seawater-associated bacteria within the short-finned pilot whale blow. Four of the most abundant genera detected in both datasets (*Pseudomonas*, *Flavobacterium*, *Rhodococcus*, and *Sphingomonas*) are potential pathogenic genera within the blow microbiome.

The V4-V5 dataset detected more taxa of less abundance, which can be important to detect the “rare biosphere” in the blow, especially considering the low represented taxa that can potential have a pathogenic role in causing respiratory diseases. In the study of Nelson et al., (2019), different genera, with a potential pathogenic role, were less abundant taxa in blow samples (*Staphylococcus* with 0.01% relative abundance of identified genera in blow samples; and *Streptococcus* with 0.09%).

(C) Potential pathogens in blow samples

Microbiomes can be coarsely screened for pathogens using the amplicon sequencing approach. There is limited data on cetacean pathogens, and the prevalence and progression of respiratory diseases in wild populations. In this study, the V4-V5 dataset captured a higher richness of potential pathogenic genera compared to the V3-V4 dataset. However, the most abundant genera were the same in both datasets, with different relative abundances.

With this type of analysis, it is difficult to assess the potential pathogenic role of the microorganisms found in the blow of these animals that can cause respiratory diseases. In the blow of short-finned pilot whales were found genera closely related to certain respiratory pathogens in cetaceans. These genera include few representative species capable of causing disease in cetaceans, however, there are a big number of non-pathogenic species belonging to these genera, represented by them. Some of these

genera were already been reported in the blow of killer whales (Raverty et al., 2017), bottlenose dolphins (Nelson et al., 2019; Bik et al., 2016), humpback whales (Apprill et al., 2017; Pirrota et al., 2017). In the study of Apprill et al., (2017), despite the presence of these potential pathogenic genera, whales appeared healthy, therefore these genera probably were not acting as pathogens at the time. However, using screening methods could quickly identify samples that need to be examined for pathogens with finer resolution methods.

The most common potential pathogenic genus in both datasets was *Pseudomonas* (12.62% in the V4-V5 dataset, 17.80% in the V3-V4 dataset), which includes the species *P. aeruginosa*, a commensal and opportunistic pathogen when host defenses are impaired in the lung environment, responsible for causing pneumonia, septicemia, and abscesses in marine mammals (Nelson et al., 2019). Finding evidence of this genus in cetacean blow samples, as well as the genera *Staphylococcus* (6.81% in the V4-V5 dataset, 1.93% in the V3-V4 dataset) and *Streptococcus* (3.10% in the V4-V5 dataset, 1.48% in the V3-V4 dataset), warrants attention in terms of the potential impact on the health of their populations. Members of these 3 genera have been recovered from dead, sick, but also healthy cetaceans, being related to the cause of severe pneumonia and also mortality, suggesting that commensal agents can have pathogenic behavior (Acevedo-Whitehouse et al., 2010). Another genus present in this study was the *Porphyromonas* (1.01% in the V4-V5 dataset, 1.02% in the V3-V4 dataset), which is known to be part of the salivary microbiome and linked with oral diseases, such as periodontitis (Falagas & Siakavellas, 2000). This genus was previously detected on the blow core microbiome of humpback whales (Apprill et al., 2017), and was also detected in the skin of this species (Apprill et al., 2014).

In 2011, in Florida, a live mass stranding of 26 short-finned pilot whales occurred. Some of them were transferred for rehabilitation, and their bacteria cultured during the routine clinical care, and necropsy of a deceased whale included methicillin-sensitive and methicillin-resistant to *Staphylococcus aureus*. These studies support that *S. aureus* may be shed into an environment by humans or pilot whales and subsequently colonize or infect an exposed new host (Hower et al., 2013).

Staphylococcus and *Streptococcus* have also been described as a constant component of the external skin microbiota near the blowhole external skin (Morris et al., 2011). These taxa have been isolated also in feces and blow fluids from individuals that exhibited no signs of disease (Johnson et al., 2009). This suggests that these organisms

may exist as part of the normal microbiota, or as low-abundance transients, only capable of infection after disruption of host defenses (immunosuppression or alterations in the normal microbiota), that remove barriers to colonization in the bronchi and lungs (Acevedo-Whitehouse et al., 2010; Johnson et al., 2009). Understanding the composition and ecology of the bacterial communities associated with the healthy URT is a critical first step in elucidating the role of the normal microbiota in maintaining respiratory health (Johnson et al., 2009). However, infer population health only based on bacterial detection at a single time point from individuals blow is naïve - it is necessary a long-term monitoring of the prevalence of these microorganisms, together with additional assays (e.g. mucosal inflammatory responses and diagnostic cytology).

Comparison of blow and seawater samples

The comparison between blow and seawater samples was focused on 3 days of sampling. This limitation along with a low sequencing depth makes it difficult to interpret the results unambiguously. The blow and seawater microbial community richness and diversity were not significantly different. However, the distribution of prokaryotic communities was influenced by the different types of samples (blow and seawater). Although a significant portion of taxa was found in both seawater and blow samples, some of them were present only in the blow, suggesting they are effectively members of the respiratory microbiome. Blow samples presented a higher number of ASVs compared with seawater samples. There were more unique blow ASVs than the sum of blow ASVs shared with water and unique water ASVs.

In some studies (Apprill et al. 2017; Pirodda et al. 2017, Raverty et al. 2017, and Vendl et al. 2019), the microbial communities in seawater were significantly different from those detected in cetaceans' blow, although with some similarities between communities. These similarities can be due to a mixture of seawater sequences with the blow when the blow is expired by the whale, making the seawater ASVs contaminants. It can also be due to the often entry of seawater into the upper airways of cetaceans, after exhalation. In this last case, these seawater-related ASVs can be true airways' inhabitants, either transient or with a more permanent nature (Apprill et al., 2017).

6. Conclusions and Futures Perspectives

This study is a first step to optimize a standardized pipeline for the analysis of the microbial communities harbored in the respiratory tract of short-finned pilot whales and seawater.

In the comparison between the primers sets, all samples from both datasets (V3-V4 and V4-V5), shared a main core microbiome, composed by *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, and *Proteobacteria* phyla.

The selection of the primer set to use for the assessment of microbiomes in cetacean blow samples depends on the main goal of the analyses. If the main goal is to capture more diversity, the V3-V4 primer set demonstrated to have a better performance. Whereas if the target is to capture the higher microbial richness, targeting the microbial rare biosphere, it is proposed to use the primer set targeting the hypervariable regions V4-V5. Indeed, despite the V4-V5 dataset detected a higher number of taxa, the majority of them had a relative abundance <1%. Higher richness in V4-V5 dataset was detected also in the deeper analysis focused on the potentially pathogenic genera.

Concerning the comparison between seawater and blow samples performed by the V4-V5 primers, blow samples showed higher diversity and richness than seawater (although not significantly different). This analysis allowed to highlight some phyla (e.g. *Actinobacteria*, *Bacteroidetes*, and *Proteobacteria*) that are cosmopolitan, being present either in the blow or in the water, and other that are more specific to blow (e.g. *Euryarchaeota*).

Moreover, with this investigation, several bacterial genera known to be primary or opportunistic pathogens in cetaceans were detected; however, their potential role in the disease of short-finned pilot whales remains to be determined. In order to understand the connection between the respiratory tract microbiota and health-related factors of short-finned pilot whales, it is important to screen for changes to the signatures of blow microbiotas (presence of non-typical microbes, potential pathogens, and changes to the diversity of the overall microbiotas and/or the composition and abundance of core bacteria). This can be achieved by doing studies with a bigger sample size, repeating individual sampling over a longer period of time to determine if the microbial composition evolves, and studies analyzing the respiratory microbiomes of captive/managed populations of cetaceans with diagnosed respiratory diseases. It will be necessary to try

to examine blow microbiomes from unhealthy/diseased animals to understand the stability of the core microbiome with different health states. The data shown in this study could also serve as a baseline for more in-depth studies focused on the role that short-finned pilot whales play in assessing the health of marine ecosystems and the human coastal population.

Several aspects require consideration and future development to further advance the blow microbiome as a health monitoring tool for cetaceans. First of all, it is important to optimize the amount of EBC necessary to collect by the PERFORMAgene kit, in order to obtain higher DNA concentration, which could affect the sequencing efficiency and the downstream procedures. Besides this, different sampling techniques should be taken into consideration, such as the use of drone-enabled photogrammetry, making it possible to detect poor body conditions of animals, therefore facilitating the detection of diseased individuals. Moreover, in future samplings, it will be useful to collect air samples, accounting for potential contamination from the atmospheric environment.

The next steps of this study will include the verification of the resident pattern of individuals sampled, using photographs captured during the sampling activities, to verify if individuals that were traveling together when they were sampled belong to the same pod. Also, a comparison between pathogens present in seawater and blow will be undertaken, to understand the extent of pathogen exposure and the recruitment within the host.

Studying the impact on blow microbiota of sex and age of the animals would also be interesting. This is a poorly studied topic, and while a previous study found evidence that these factors have an impact on the microbiome of the animal's (C. Vendl et al., 2021), others did not (Bik et al., 2016; Lima et al., 2012) – influence of sex and age could therefore be dependent on the species and even on the populations. Future studies would benefit from deeper sequencing to the level of species to identify the infectious potential of commensals and their transmission pathways. To fully define the total respiratory tract microbiota of cetaceans and to further improve this type of analysis could be done a whole-genome analysis approach (shotgun sequencing), which would help to explore the functional differences provided by different regions. Moreover, it will be remarkable to use a metagenomic approach, with more investigations about the association between viruses, fungus, and bacteria, present in the respiratory tract of cetaceans.

7. Scientific Communications (based on the work of this thesis)

Santos Silva, B., Correia, M., Magalhães, C., Tomasino, M. P., 2022. Dive into the blow microbiome: exploring microbial diversity of the respiratory tract in free-ranging cetaceans. In: Book of Abstracts of the 15th Meeting of Young Researchers at U. Porto (IJUP22). 04-06/05/2022, Porto, Portugal. (Oral communication)

Santos Silva, B., Alves, F., Dinis, A., Ferreira, R., Correia, M., Valente, R., Gil, A., Castro, F., Sousa-Pinto, I., Rosso, Massimiliano, Centelleghé, C., Mazzariol, S., Magalhaes, C., Tomasino, M. P., 2022. Investigation of respiratory prokaryotic communities of free-ranging short-finned pilot whales using non-invasive sampling. In Book of Abstracts of the 3th Blue Think Conference 2022: Share Science, Spread Knowledge. 27/09/2022, Matosinhos, Portugal. (Poster presentation)

Tomasino, M. P., Santos Silva, B., Alves, F., Dinis, A., Ferreira, R., Correia, M., Valente, R., Gil, A., Castro, F., Sousa-Pinto, I., Rosso, Massimiliano, Centelleghé, C., Mazzariol, S., Magalhaes, C. 2022. Dive into the blow microbiome of free-ranging short-finned pilot whales: protocol optimization and preliminary results. 24th Biennial Conference on the Biology of Marine Mammal. 01-05/08/2022, Palm Beach, Florida, USA. (Poster presentation)

8. References

- Acevedo-Whitehouse, K., Rocha-Gosselin, A., & Gendron, D. (2010). A novel non-invasive tool for disease surveillance of free-ranging whales and its relevance to conservation programs. *Animal Conservation*, 13(2), 217–225.
- Ainsworth, T. D., & Gates, R. D. (2016). Corals' microbial sentinels. *Science*, 352(6293), 1518–1519.
- Aksenov, A. A., Yeates, L., Pasamontes, A., Siebe, C., Zrodnikov, Y., Simmons, J., McCartney, M. M., Deplanque, J. P., Wells, R. S., & Davis, C. E. (2014). Metabolite content profiling of bottlenose dolphin exhaled breath. *Analytical Chemistry*, 86(21), 10616–10624.
- Alves, F., Dromby, M., Baptista, V., Ferreira, R., Correia, A. M., Weyn, M., Valente, R., Froufe, E., Rosso, M., Sousa-Pinto, I., Dinis, A., Dias, E., & Teodósio, M. A. (2020). Ecophysiological traits of highly mobile large marine predators inferred from nucleic acid derived indices. *Scientific Reports*, 10(1), 1–10.
- Alves, Filipe. (2013). *Population Structure, Habitat Use and Conservation of Short-Finned Pilot Whales Globicephala macrorhynchus in the Archipelago of Madeira*. Doctoral thesis (PhD) in Biological Sciences, Sciences Faculty of Madeira University, Madeira, Portugal. 186 pp.
- Alves, Filipe, Dinis, A., Nicolau, C., Ribeiro, C., Kaufmann, M., Fortuna, C., & Freitas, L. (2015). Survival and abundance of short-finned pilot whales in the archipelago of Madeira, NE Atlantic. *Marine Mammal Science*, 31(1), 106–121.
- Alves, Filipe, Dinis, A., Ribeiro, C., Nicolau, C., Kaufmann, M., Fortuna, C., & Freitas, L. (2013). Daytime dive characteristics from six short-finned pilot whales *Globicephala macrorhynchus* off Madeira Island. *Arquipélago. Life and Marine Science*, 31, 1–8.
- Alves, Filipe, Quérrouil, S., Dinis, A., Nicolau, C., Ribeiro, C., Freitas, L., Kaufmann, M., & Fortuna, C. (2013). Population structure of short-finned pilot whales in the oceanic archipelago of Madeira based on photo-identification and genetic analyses: Implications for conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 23(5), 758–776.
- Apprill, A. (2017). Marine animal microbiomes: Toward understanding host-microbiome

- p>interactions in a changing ocean.
- Frontiers in Marine Science*
- , 4(JUL), 1–9.
- Apprill, A., McNally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, 75(2), 129–137.
- Apprill, A., Miller, C. A., Moore, M. J., Durban, J. W., Fearnbach, H., & Barrett-Lennard, L. G. (2017). Extensive Core Microbiome in Drone-Captured Whale Blow Supports a Framework for Health Monitoring. *MSystems*, 2(5), e00119-17.
- 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. *Royal Society Open Science*, 7(5), 192046.
- Apprill, A., Mooney, T. A., Lyman, E., Stimpert, A. K., & Rappé, M. S. (2011). Humpback whales harbour a combination of specific and variable skin bacteria. *Environmental Microbiology Reports*, 3(2), 223–232.
- Apprill, A., Robbins, J., Eren, A. M., Pack, A. A., Reveillaud, J., Mattila, D., Moore, M., Niemeyer, M., Moore, K. M. T., & Mincer, T. J. (2014). Humpback whale populations share a core skin bacterial community: Towards a health index for marine mammals? *PLoS ONE*, 9(3), e90785.
- Armougom, F. (2009). Exploring Microbial Diversity Using 16S rRNA High-Throughput Methods. *Journal of Computer Science & Systems Biology*, 2(1), 74–92.
- Atkinson, S., Rogan, A., Baker, C. S., Dagdag, R., Redlinger, M., Polinski, J., Urban, J., Sremba, A., Branson, M., Mashburn, K., Pallin, L., Klink, A., Steel, D., Bortz, E., & Kerr, I. (2021). Genetic, Endocrine, and Microbiological Assessments of Blue, Humpback and Killer Whale Health using Unoccupied Aerial Systems. *Wildlife Society Bulletin*, 45(4), 654–669.
- Avalos-Téllez, R., Suárez-Guemes, F., Carrilo-Casas, E. M., & Hernández-Castro, R. (2010). Bacteria and yeast normal microbiota from respiratory tract and genital area of bottlenose dolphins (*Tursiops truncatus*). *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, 2, 666–673.
- Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., & Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *Science*, 307(5717), 1915–1920.
- Baker, G. C., Smith, J. J., & Cowan, D. A. (2003). Review and re-analysis of domain-

- specific 16S primers. *Journal of Microbiological Methods*, 55(3), 541–555.
- Ballance, L. T. (2018). Cetacean Ecology. *Encyclopedia of Marine Mammals, 1991*, 172–180.
- Bannister, J. L. (2009). Baleen whales (mysticetes). In eds. Perrin WF, Würsig B, Thewissen JGM (Ed.), *Encyclopedia of marine mammals* Academic Press, Amsterdam, 2, 80–89.
- Barratclough, A., Wells, R. S., Schwacke, L. H., Rowles, T. K., Gomez, F. M., Fauquier, D. A., Sweeney, J. C., Townsend, F. I., Hansen, L. J., Zolman, E. S., Balmer, B. C., & Smith, C. R. (2019). Health Assessments of Common Bottlenose Dolphins (*Tursiops truncatus*): Past, Present, and Potential Conservation Applications. *Frontiers in Veterinary Science*, 6(Dec), 1–18.
- Berta, A., Sumich, J. L., & Kovacs, K. M. (2006). *Marine Mammals Evolutionary Biology*. Academic Press.
- 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. *Nature Communications*, 7, 1–13.
- Bogomolni, A. L., Gast, R. J., Ellis, J. C., Dennett, M., Pugliares, K. R., Lentell, B. J., & Moore, M. J. (2008). Victims or vectors: A survey of marine vertebrate zoonoses from coastal waters of the Northwest Atlantic. *Diseases of Aquatic Organisms*, 81(1), 13–38.
- Boonstra, W. J., Ottosen, K. M., Ferreira, A. S. A., Richter, A., Rogers, L. A., Pedersen, M. W., Kokkalis, A., Bardarson, H., Bonanomi, S., Butler, W., Diekert, F. K., Fouzai, N., Holma, M., Holt, R. E., Kvile, K., Malanski, E., Macdonald, J. I., Nieminen, E., Romagnoni, G., ... Whittington, J. D. (2015). What are the major global threats and impacts in marine environments? Investigating the contours of a shared perception among marine scientists from the bottom-up. *Marine Policy*, 60, 197–201.
- Bossart, G. D. (2007). Emerging Diseases in Marine Mammals: from Dolphins to Manatees-FAU Institutional Repository. *Microbe*, 2(97), 106–112.
- Bossart, G. D. (2011). Marine mammals as sentinel species for oceans and human health. *Veterinary Pathology*, 48(3), 676–690.
- Bourlat, S. J., Borja, A., Gilbert, J., Taylor, M. I., Davies, N., Weisberg, S. B., Griffith, J.

- F., Lettieri, T., Field, D., Benzie, J., Glöckner, F. O., Rodríguez-Ezpeleta, N., Faith, D. P., Bean, T. P., & Obst, M. (2013). Genomics in marine monitoring: New opportunities for assessing marine health status. *Marine Pollution Bulletin*, 74(1), 19–31.
- Burgess, E. A., Hunt, K. E., Kraus, S. D., & Rolland, R. M. (2016). Get the most out of blow hormones: Validation of sampling materials, field storage and extraction techniques for whale respiratory vapour samples. *Conservation Physiology*, 4(1), 1–11.
- Burgess, E. A., Hunt, K. E., Kraus, S. D., & Rolland, R. M. (2018). Quantifying hormones in exhaled breath for physiological assessment of large whales at sea. *Scientific Reports*, 8(1), 1–14.
- Cabral, M., Almeida, J., Almeida, P., Dellinger, T., Almeida, N., Oliveira, M., Palmeirim, J., Queiroz, A., Rogado, L., & Santos-Reis, M. (2005). Livro Vermelho dos Vertebrados de Portugal. *Livro Vermelho dos Vertebrados de Portugal*.
- Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal*, 11(12), 2639–2643.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583.
- Callewaert, C., Ravard Helffer, K., & Lebaron, P. (2020). Skin Microbiome and its Interplay with the Environment. *American Journal of Clinical Dermatology*, 21(1), 4–11.
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., & Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME Journal*, 6(8), 1621–1624.
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America*, 108(1), 4516–4522.
- Cardona, C., Henry, C., Bonn, B. Van, & Gilbert, A. (2018). Environmental Sources of

- Bacteria Differentially Influence. *MSystems*, 3(3), 1–18.
- Carwardine, M. (2020). *Handbook of Whales, Dolphins and Porpoises*. Bloomsbury Publishing.
- Case, R. J., Boucher, Y., Dahllöf, I., Holmström, C., Doolittle, W. F., & Kjelleberg, S. (2007). Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Applied and Environmental Microbiology*, 73(1), 278–288.
- Centelleghes, C., Carraro, L., Gonzalvo, J., Rosso, M., Esposti, E., Gili, C., Bonato, M., Pedrotti, D., Cardazzo, B., Povinelli, M., & Mazzario, S. (2020). The use of Unmanned Aerial Vehicles (UAVs) to sample the blow microbiome of small cetaceans. *PLoS ONE*, 15(Jul), 1–14.
- Chakravorty, S., Helb, D., Burday, M., Connel, N., & Alland, D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J Microbiol Methods*, 69(2), 330–339.
- Christiansen, F., Dawson, S. M., Durban, J. W., Fearnbach, H., Miller, C. A., Bejder, L., Uhart, M., Sironi, M., Corkeron, P., Rayment, W., Leunissen, E., Haria, E., Ward, R., Warick, H. A., Kerr, I., Lynn, M. S., Pettis, H. M., & Moore, M. J. (2020). Population comparison of right whale body condition reveals poor state of the North Atlantic right whale. *Marine Ecology Progress Series*, 640, 1–16.
- Christiansen, F., Dujon, A. M., Sprogis, K. R., Arnould, J. P. Y., & Bejder, L. (2016). Noninvasive unmanned aerial vehicle provides estimates of the energetic cost of reproduction in humpback whales. *Ecosphere*, 7(10), 1–18.
- Christie, K. S., Gilbert, S. L., Brown, C. L., Hatfield, M., & Hanson, L. (2016). Unmanned aircraft systems in wildlife research: Current and future applications of a transformative technology. *Frontiers in Ecology and the Environment*, 14(5), 241–251.
- Colpitts, J., McLoughlin, P. D., & Poissant, J. (2022). Runs of homozygosity in Sable Island feral horses reveal the genomic consequences of inbreeding and divergence from domestic breeds. *BMC genomics*, 23(1), 1–17.
- Di Guardo, G., & Mazzariol, S. (2016). Cetacean Morbillivirus-associated pathology: Knowns and unknowns. *Frontiers in Microbiology*, 7(Feb), 1–5.
- Dickson, R. P., Erb-Downward, J. R., Falkowski, N. R., Hunter, E. M., Ashley, S. L., & Huffnagle, G. B. (2018). The lung microbiota of healthy mice are highly variable,

- cluster by environment, and reflect variation in baseline lung innate immunity. *American Journal of Respiratory and Critical Care Medicine*, 198(4), 497–508.
- Dickson, R. P., Erb-Downward, J. R., Martinez, F. J., & Huffnagle, G. B. (2016). The Microbiome and the Respiratory Tract. *Annual Review of Physiology*, 78(Oct), 481–504.
- Domínguez-Sánchez, C. A., Acevedo-Whitehouse, K. A., & Gendron, D. (2018). Effect of drone-based blow sampling on blue whale (*Balaenoptera musculus*) behavior. *Marine Mammal Science*, 34(3), 841–850.
- Doney, S. C., Ruckelshaus, M., Emmett Duffy, J., Barry, J. P., Chan, F., English, C. A., Galindo, H. M., Grebmeier, J. M., Hollowed, A. B., Knowlton, N., Polovina, J., Rabalais, N. N., Sydeman, W. J., & Talley, L. D. (2012). Climate change impacts on marine ecosystems. *Annual Review of Marine Science*, 4, 11–37.
- Duarte, C. M. (2014). Global change and the future ocean: A grand challenge for marine sciences. *Frontiers in Marine Science*, 1(Dec), 1–16.
- Dusa, A. (2020). Draw Venn Diagrams. *Programa*. <https://cran.r-project.org/web/packages/venn/>
- Einarsson, G. G., Comer, D. M., McIlreavey, L., Parkhill, J., Ennis, M., Tunney, M. M., & Elborn, J. S. (2016). Community dynamics and the lower airway microbiota in stable chronic obstructive pulmonary disease, smokers and healthy non-smokers. *Thorax*, 71(9), 795–803.
- Fadeev, E., Cardozo-Mino, M. G., Rapp, J. Z., Bienhold, C., Salter, I., Salman-Carvalho, V., ... & Boetius, A. (2021). Comparison of two 16S rRNA primers (V3–V4 and V4–V5) for studies of arctic microbial communities. *Frontiers in microbiology*, 12, 637526.
- Falagas, M. E., & Siakavellas, E. (2000). Bacteroides, Prevotella, and Porphyromonas species: A review of antibiotic resistance and therapeutic options. *International Journal of Antimicrobial Agents*, 15(1), 1–9.
- Faner, R., Sibila, O., Agustí, A., Bernasconi, E., Chalmers, J. D., Huffnagle, G. B., Manichanh, C., Molyneaux, P. L., Paredes, R., Brocal, V. P., Ponomarenko, J., Sethi, S., Dorca, J., & Monsó, E. (2017). The microbiome in respiratory medicine: Current challenges and future perspectives. *European Respiratory Journal*, 49(4), 1–12.

- Fischbach, M. A., & Segre, J. A. (2016). Signaling in Host-Associated Microbial Communities. *Cell*, 164(6), 1288–1300.
- Foot, N. J., Orgeig, S., & Daniels, C. B. (2006). The evolution of a physiological system: The pulmonary surfactant system in diving mammals. *Respiratory Physiology and Neurobiology*, 154(1–2), 118–138.
- Frère, C. H., Krzyszczyk, E., Patterson, E. M., Hunter, S., Ginsburg, A., & Mann, J. (2010). Thar she blows! A novel method for DNA collection from cetacean blow. *PLoS ONE*, 5(8), 1–5.
- Geoghegan, J. L., Pirodda, V., Harvey, E., Smith, A., Buchmann, J. P., Ostrowski, M., Eden, J. S., Harcourt, R., & Holmes, E. C. (2018). Virological sampling of inaccessible wildlife with drones. *Viruses*, 10(6), 1–7.
- Gilbert, J. A., Jansson, J. K., & Knight, R. (2014). The Earth Microbiome project: Successes and aspirations. *BMC Biology*, 12(1), 1–4.
- Graspeuntner, S., Loeper, N., Künzel, S., Baines, J. F., & Rupp, J. (2018). Selection of validated hypervariable regions is crucial in 16S-based microbiota studies of the female genital tract. *Scientific Reports*, 8(1), 4–10.
- Halpern, B. S., Selkoe, K. A., Micheli, F., & Kappel, C. V. (2007). Evaluating and ranking the vulnerability of global marine ecosystems to anthropogenic threats. *Conservation Biology*, 21(5), 1301–1315.
- Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D'Agrosa, C., Bruno, J. F., Casey, K. S., Ebert, C., Fox, H. E., Fujita, R., Heinemann, D., Lenihan, H. S., Madin, E. M. P., Perry, M. T., Selig, E. R., Spalding, M., Steneck, R., & Watson, R. (2008). A global map of human impact on marine ecosystems. *Science*, 319(5865), 948–952.
- Harkonen, T., Bäcklin, B. M., Barrett, T., Bergman, A., Corteyn, M., Dietz, R., Harding, K. C., Malmsten, J., Roos, A., & Teilmann, J. (2008). Mass mortality in harbour seals and harbour porpoises caused by an unknown pathogen. *Veterinary Record*, 162(17), 555–556.
- Hastie, G. D., Wilson, B., Tufft, L. H., & Thompson, P. M. (2003). Bottlenose dolphins increase breathing synchrony in response to boat traffic. *Marine Mammal Science*, 19(1), 74–084.
- Hauser, D. D., Logsdon, M. G., Holmes, E. E., VanBlaricom, G. R., & Osborne, R. W.

- (2007). Summer distribution patterns of southern resident killer whales *Orcinus orca*: core areas and spatial segregation of social groups. *Marine Ecology Progress Series*, 351, 301–310.
- Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270(1512), 313–321.
- Hernandez-Agreda, A., Gates, R. D., & Ainsworth, T. D. (2017). Defining the Core Microbiome in Corals' Microbial Soup. *Trends in Microbiology*, 25(2), 125–140.
- Hogg, C. J., Rogers, T. L., Shorter, A., Barton, K., Miller, P. J. O., & Nowacek, D. (2009). Determination of steroid hormones in whale blow: It is possible. *Marine Mammal Science*, 25(3), 605–618.
- Hogg, C. J., Vickers, E. R., & Rogers, T. L. (2005). Determination of testosterone in saliva and blow of bottlenose dolphins (*Tursiops truncatus*) using liquid chromatography-mass spectrometry. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 814(2), 339–346.
- Hooker, S. K. (2009). Toothed whales, overview. In eds Perrin WF, Würsig B, Thewissen JGM (Ed.), *Encyclopedia of marine mammals* (2nd ed.). Academic Press, Amsterdam, 1173–1179.
- Hooper, L. V., Midwedt, T., & Gordon, J. I. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annual Review of Nutrition*, 22, 283–307.
- Hower, S., Phillips, M. C., Brodsky, M., Dameron, A., Tamargo, M. A., Salazar, N. C., Jackson, C. R., Barrett, J. B., Davidson, M., Davis, J., Mukherjee, S., Ewing, R. Y., Gidley, M. L., Sinigalliano, C. D., Johns, L., Johnson, F. E., Adebajo, O., & Plano, L. R. W. (2013). Clonally Related Methicillin-Resistant *Staphylococcus aureus* Isolated from Short-Finned Pilot Whales (*Globicephala macrorhynchus*), Human Volunteers, and a Bayfront Cetacean Rehabilitation Facility. *Microbial Ecology*, 65(4), 1024–1038.
- Hudson, J. M., Anderson, W. G., & Marcoux, M. (2021). Measurement of cortisol in blow samples collected from free-swimming beluga whales (*Delphinapterus leucas*). *Marine Mammal Science*, 37(3), 888–900.
- Hunt, K. E., Moore, M. J., Rolland, R. M., Kellar, N. M., Hall, A. J., Kershaw, J., Raverty,

- S. A., Davis, C. E., Yeates, L. C., Fauquier, D. A., Rowles, T. K., & Kraus, S. D. (2013). Overcoming the challenges of studying conservation physiology in large whales: a review of available methods. *Conservation Physiology*, 1(1).
- Hunt, Kathleen E., Rolland, R. M., & Kraus, S. D. (2014). Detection of steroid and thyroid hormones via immunoassay of North Atlantic right whale (*Eubalaena glacialis*) respiratory vapor. *Marine Mammal Science*, 30(2), 796–809.
- Jefferson, T. A., Webber, M. A., & Pitman, R. L. (2015). *Marine Mammals of the World: A Comprehensive Guide to their Identification*. Academic Press.
- Jepson, P. D., Deaville, R., Acevedo-Whitehouse, K., Barnett, J., Brownlow, A., Brownell, R. L., Clare, F. C., Davison, N., Law, R. J., Loveridge, J., Macgregor, S. K., Morris, S., Murphy, S., Penrose, R., Perkins, M. W., Pinn, E., Seibel, H., Siebert, U., Sierra, E., ... Fernández, A. (2013). What Caused the UK's Largest Common Dolphin (*Delphinus delphis*) Mass Stranding Event? *PLoS ONE*, 8(4), e60953.
- Johnson, W. R., Torralba, M., Fair, P. A., Bossart, G. D., Nelson, K. E., & Morris, P. J. (2009). Novel diversity of bacterial communities associated with bottlenose dolphin upper respiratory tracts. *Environmental Microbiology Reports*, 1(6), 555–562.
- Katona, S., & Whitehead, H. (1988). Are Cetacea ecologically important? *Oceanography and Marine Biology Annual Review*, 26, 553–568.
- Kirchman, D. L. (2018). *Processes in Microbial Ecology*. Oxford University Press
- 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 Research*, 41(1), 1–11.
- Kopf, A., Bicak, M., Kottmann, R., Schnetzer, J., Kostadinov, I., Lehmann, K., ... & Ducluzeau, A. L. (2015). The ocean sampling day consortium. *Gigascience*, 4(1), 1-5.
- Kress, W. J., & Erickson, D. L. (2008). *DNA barcodes: Genes, genomics, and bioinformatics*. 105(8), 2761–2762.
- Kumpitsch, C., Koskinen, K., Schöpf, V., & Moissl-Eichinger, C. (2019). The microbiome of the upper respiratory tract in health and disease. *BMC Biology*, 17(1), 1–20.
- Lee, K. H., Gordon, A., Shedden, K., Kuan, G., Ng, S., Balmaseda, A., & Foxman, B.

- (2019). The respiratory microbiome and susceptibility to influenza virus infection. *PLoS ONE*, 14(1), 1–19.
- Leger, J. S., Raverty, S., & Mena, A. (2018). *Cetacea. Pathology of wildlife and zoo animals*, 533–568.
- Lima, N., Rogers, T., Acevedo-Whitehouse, K., & Brown, M. V. (2012). Temporal stability and species specificity in bacteria associated with the bottlenose dolphins respiratory system. *Environmental Microbiology Reports*, 4(1), 89–96.
- Lombardo, M. P. (2008). Access to mutualistic endosymbiotic microbes: An underappreciated benefit of group living. *Behavioral Ecology and Sociobiology*, 62(4), 479–497.
- Macleod, C. D., Reidenberg, J. S., Weller, M., Santos, M. B., Herman, J., Goold, J., & Pierce, G. J. (2007). Breaking symmetry: The marine environment, prey size, and the evolution of asymmetry in cetacean skulls. *Anatomical Record*, 290(6), 539–545.
- Mazzoldi, C., Bearzi, G., Brito, C., Carvalho, I., Desiderà, E., Endrizzi, L., Freitas, L., Giacomello, E., Giovos, I., Guidetti, P., Ressurreição, A., Tull, M., & MacDiarmid, A. (2019). From sea monsters to charismatic megafauna: Changes in perception and use of large marine animals. *PLoS ONE*, 14(12), e0226810.
- McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4), e61217.
- McNichol, J., Berube, P. M., Biller, S. J., & Fuhrman, J. A. (2021). *Evaluating and Improving SSU rRNA PCR Primer Coverage for Bacteria, Archaea, and Eukaryotes Using Metagenomes from Global Ocean Surveys*. 4(1), 1–74.
- Minton, G., Braulik, G., & Reeves, R. (2018). *Globicephala macrorhynchus (Short-finned Pilot Whale)*. 8235, 1–8. <http://www.iucnredlist.org/details/summary/9249/0>
- Moore, S. E. (2008). Marine mammals as ecosystem sentinels. *Journal of Mammalogy*, 89(3), 534–540.
- Moritz, C., & Cicero, C. (2004). DNA barcoding: Promise and pitfalls. *PLoS Biology*, 2(10), e354.
- Morris, P. J., Johnson, W. R., Pisani, J., Bossart, G. D., Adams, J., Reif, J. S., & Fair, P.

- A. (2011). Isolation of culturable microorganisms from free-ranging bottlenose dolphins (*Tursiops truncatus*) from the southeastern United States. *Veterinary Microbiology*, 148(2–4), 440–447.
- Murphy, S., Pierce, G. J., Law, R. J., Bersuder, P., Jepson, P. D., Learmonth, J. A., Addink, M., Dabin, W., Santos, M. B., Deaville, R., Zegers, B. N., Mets, A., Rogan, E., Ridoux, V., Reid, R. J., Smeenk, C., Jauniaux, T., López, A., Alonso Farré, J. M., ... Boon, J. P. (2009). Assessing the effect of persistent organic pollutants on reproductive activity in common dolphins and harbour porpoises. *Journal of Northwest Atlantic Fishery Science*, 42(May), 153–173.
- Nabi, G., McLaughlin, R. W., Khan, S., Hao, Y., & Chang, M. X. (2020). Pneumonia in endangered aquatic mammals and the need for developing low-coverage vaccination for their management and conservation. *Animal Health Research Reviews*, 21(2), 122-130.
- Neary, M. T., Neary, J. M., Lund, G. K., Garry, F. B., Holt, T. N., Mohun, T. J., & Breckenridge, R. A. (2014). A comparison of DNA collection methods in cattle and yaks. *Journal of Animal Science*, 92(9), 3811-3815.
- Nelson, T. M., Apprill, A., Mann, J., Rogers, T. L., & Brown, M. V. (2015). The marine mammal microbiome: current knowledge and future directions. *Microbiology Australia*, 36(1), 8.
- Nelson, T. M., Wallen, M. M., Bunce, M., Oskam, C. L., Lima, N., Clayton, L., & Mann, J. (2019). Detecting respiratory bacterial communities of wild dolphins: Implications for animal health. *Marine Ecology Progress Series*, 622, 203–217.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., ... Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-2. Cran R, 1, 2.
- Olson, P. A. (2009). Pilot Whales: *Globicephala melas* and *G. macrorhynchus*. In eds. Perrin WF, Würsig B, thewissen JGM (Ed.), *Encyclopedia of marine mammals* 2nd ed, Academic Press, Amsterdam, 847–852.
- Omata, Y., Umeshita, Y., Watarai, M., Tachibana, M., Sasaki, M., Murata, K., & Yamada, T. K. (2006). Investigation for presence of *Neospora caninum*, *Toxoplasma gondii* and *Brucella*-species infection in killer whales (*Orcinus orca*) mass-stranded on the coast of Shiretoko, Hokkaido, Japan. *Journal of Veterinary Medical Science*, 68(5),

523–526.

- Pamer, E. G. (2007). Immune responses to commensal and environmental microbes. *Nature Immunology*, 8(11), 1173–1178.
- Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, 18(5), 1403–1414.
- Pawlowski, J., Lejzerowicz, F., Apotheloz-Perret-Gentil, L., Visco, J., & Esling, P. (2016). Protist metabarcoding and environmental biomonitoring: Time for change. *European Journal of Protistology*, 55, 12-25.
- Pirotta, V., Smith, A., Ostrowski, M., Russell, D., Jonsen, I. D., Grech, A., & Harcourt, R. (2017). An economical Custom-Built drone for assessing whale health. *Frontiers in Marine Science*, 4(Dec), 1–12.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1), 590–596.
- Ranjan, R., Rani, A., Metwally, A., S.McGee, H., & L.Perkins, D. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochem Biophys Res Commun.*, 469(4), 967–977.
- Raudino, H. C., Tyne, J. A., Smith, A., Ottewell, K., McArthur, S., Kopps, A. M., Chabanne, D., Harcourt, R. G., Pirotta, V., & Waples, K. (2019). Challenges of collecting blow from small cetaceans. *Ecosphere*, 10(10).
- Raverty, S. A., Rhodes, L. D., Zabek, E., Eshghi, A., Cameron, C. E., Hanson, M. B., & Schroeder, J. P. (2017). Respiratory Microbiome of Endangered Southern Resident Killer Whales and Microbiota of Surrounding Sea Surface Microlayer in the Eastern North Pacific. *Scientific Reports*, 7(1), 1–12.
- Rhodes, L. D., Emmons, C. K., Wisswaesser, G. S., Wells, A. H., & Hanson, M. B. (2022). Bacterial microbiomes from mucus and breath of southern resident killer whales (*Orcinus orca*). *Conservation Physiology*, 10(1), 1–11.
- Ribeiro, H., de Sousa, T., Santos, J. P., Sousa, A. G. G., Teixeira, C., Monteiro, M. R., Salgado, P., Mucha, A. P., Almeida, C. M. R., Torgo, L., & Magalhães, C. (2018). Potential of dissimilatory nitrate reduction pathways in polycyclic aromatic

- hydrocarbon degradation. *Chemosphere*, 199, 54–67.
- Ridgway, S. H., Scronce, B. L., & Kanwisher, J. (1969). Respiration and deep diving in the bottlenose porpoise. *Science*, 166(3913), 1651–1654.
- Risely, A. (2020). Applying the core microbiome to understand host–microbe systems. *Journal of Animal Ecology*, 89(7), 1549–1558.
- Roberge, J. M., & Angelstam, P. (2004). Usefulness of the Umbrella Species Concept as a Conservation Tool. *Conservation Biology*, 18(1), 76–85.
- Robinson, C. V., & Nuuttila, H. K. (2020). Don't hold your breath: Limited DNA capture using non-invasive blow sampling for small cetaceans. *Aquatic Mammals*, 46(1), 32–41.
- Robles-Malagamba, M. J., Walsh, M. T., Ahasan, M. S., Thompson, P., Wells, R. S., Jobin, C., Fodor, A. A., Winglee, K., & Waltzek, T. B. (2020). Characterization of the bacterial microbiome among free-ranging bottlenose dolphins (*Tursiops truncatus*). *Heliyon*, 6(6), e03944.
- Roman, J., Estes, J. A., Morissette, L., Smith, C., Costa, D., McCarthy, J., Nation, J. B., Nicol, S., Pershing, A., & Smetacek, V. (2014). Whales as marine ecosystem engineers. *Frontiers in Ecology and the Environment*, 12(7), 377–385.
- Sacco, J., Ruplin, A., Skonieczny, P., & Ohman, M. (2017). Polymorphisms in the canine monoamine oxidase a (MAOA) gene: identification and variation among five broad dog breed groups. *Canine genetics and epidemiology*, 4(1), 1–8.
- Sambolino, A., Alves, F., Fernandez, M., Krakauer, A. B., Ferreira, R., & Dinis, A. (2022). Spatial and temporal characterization of the exposure of island-associated cetacean populations to whale-watching in Madeira Island (NE Atlantic). *Regional Studies in Marine Science*, 49, 102084.
- Segio, F., Caro, T., Brown, D., Cluclas, B., Hunter, J., Ketchum, J., McHugh, K., & Hiraldo, F. (2008). Top Predators as Conservation Tools: Ecological Rationale, Assumptions, and Efficacy. *Annual Review of Ecology, Evolution, and Systematics*, 39, 1–19.
- Sehnal, L., Brammer-Robbins, E., Wormington, A. M., Blaha, L., Bisesi, J., Larkin, I., Martyniuk, C. J., Simonin, M., & Adamovsky, O. (2021). Microbiome Composition and Function in Aquatic Vertebrates: Small Organisms Making Big Impacts on Aquatic Animal Health. *Frontiers in Microbiology*, 12(Mar).

- Sergio, F., Newton, I., Marchesi, L., & Pedrini, P. (2006). Ecologically justified charisma: Preservation of top predators delivers biodiversity conservation. *Journal of Applied Ecology*, 43(6), 1049–1055.
- Servidio, A. (2014). *Distribution, social structure and habitat use of short-finned pilot whale, Globicephala macrorhynchus, in the Canary Islands*. Doctoral thesis (PhD) University of St Andrews. 452 pp.
- Shade, A., & Handelsman, J. (2012). Beyond the Venn diagram: The hunt for a core microbiome. *Environmental Microbiology*, 14(1), 4–12.
- Shukla, S. D., Budden, K. F., Neal, R., & Hansbro, P. M. (2017). Microbiome effects on immunity, health and disease in the lung. *Clinical and Translational Immunology*, 6(3), e133-12.
- Sierra, E., Fernández, A., Suárez-Santana, C., Xuriach, A., Zucca, D., Bernaldo De Quirós, Y., García-Álvarez, N., De La Fuente, J., Sacchini, S., Andrada, M., Díaz-Delgado, J., & Arbelo, M. (2016). Morbillivirus and pilot whale deaths, Canary Islands, Spain, 2015. *Emerging Infectious Diseases*, 22(4), 740–742.
- SØrum, H., & Sunde, M. (2001). Resistance to antibiotics in the normal flora of animals. *Veterinary Research*, 32(3–4), 227–241.
- Stewart, J. R., Townsend, F. I., Lane, S. M., Dyar, E., Hohn, A. A., Rowles, T. K., Staggs, L. A., Wells, R. S., Balmer, B. C., & Schwacke, L. H. (2014). Survey of antibiotic-resistant bacteria isolated from bottlenose dolphins *Tursiops truncatus* in the southeastern USA. *Diseases of Aquatic Organisms*, 108(2), 91–102.
- Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., & Nishijima, M. (2014). Development of a prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-generation sequencing. *PLoS ONE*, 9(8), e105592.
- Teo, S. M., Mok, D., Pham, K., Kusel, M., Serralha, M., Troy, N., Holt, B. J., Hales, B. J., Walker, M. L., Hollams, E., Bochkov, Y. A., Grindle, K., Johnston, S. L., Gern, J. E., Sly, P. D., Holt, P. G., Holt, K. E., & Inouye, M. (2015). The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host and Microbe*, 17(5), 704–715.
- Thewissen, J. G. M., Cooper, L. N., George, J. C., & Bajpai, S. (2009). From Land to Water: The Origin of Whales, Dolphins, and Porpoises. *Evolution: Education and Outreach*, 2(2), 272–288.

- Thompson, L. A., Spoon, T. R., Goertz, C. E. C., Hobbs, R. C., & Romano, T. A. (2014). Blow collection as a non-invasive method for measuring cortisol in the beluga (*Delphinapterus leucas*). *PLoS ONE*, 9(12), 1–22.
- Traskalová-Hogenová, H., Štěpánková, R., Hudcovic, T., Tučková, L., Cukrowska, B., Lodinová-Žádníková, R., Kozáková, H., Rossmann, P., Bártová, J., Sokol, D., Funda, D. P., Borovská, D., Řeháková, Z., Šinkora, J., Hofman, J., Drastich, P., & Kokešová, A. (2004). Commensal bacteria (normal microflora), mucosal immunity and chronic inflammatory and autoimmune diseases. *Immunology Letters*, 93(2–3), 97–108.
- Tosi, C. H., & Ferreira, R. G. (2009). Behavior of estuarine dolphin, *Sotalia guianensis* (Cetacea, Delphinidae), in controlled boat traffic situation at southern coast of Rio Grande do Norte, Brazil. *Biodiversity and Conservation*, 18(1), 67–78.
- van Bree, P. J. H. (1971). On *Globicephala sieboldii* Gray, 1846, and other species of Pilot Whales. *Beaufortia*, 79–87.
- Van Bresse, M. F., Duignan, P. J., Banyard, A., Barbieri, M., Colegrove, K. M., de Guise, S., di Guardo, G., Dobson, A., Domingo, M., Fauquier, D., Fernandez, A., Goldstein, T., Grenfell, B., Groch, K. R., Gulland, F., Jensen, B. A., Jepson, P. D., Hall, A., Kuiken, T., ... Wellehan, J. F. X. (2014). Cetacean morbillivirus: Current knowledge and future directions. *Viruses*, 6(12), 5145–5181.
- Van de Peer, Y., Chapelle, S., & De Wachter, R. (1996). A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Research*, 24(17), 3381–3391.
- Vendl, C., Ferrari, B. C., Thomas, T., Slavich, E., Zhang, E., Nelson, T., & Rogers, T. (2019). Interannual comparison of core taxa and community composition of the blow microbiota from East Australian humpback whales. *FEMS Microbiology Ecology*, 95(8), fiz102.
- Vendl, C., Nelson, T., Ferrari, B., Thomas, T., & Rogers, T. (2021). Highly abundant core taxa in the blow within and across captive bottlenose dolphins provide evidence for a temporally stable airway microbiota. *BMC Microbiology*, 21(1), 1–15.
- Vendl, Catharina, 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? *Environmental Microbiology Reports*, 12(3), 324–333.

- Vendl, Catharina, Slavich, E., Wemheuer, B., Nelson, T., Ferrari, B., Thomas, T., & Rogers, T. (2020). Respiratory microbiota of humpback whales may be reduced in diversity and richness the longer they fast. *Scientific Reports*, 10(1), 1–13.
- Venn-Watson, S., Daniels, R., & Smith, C. (2012). Thirty year retrospective evaluation of pneumonia in a bottlenose dolphin *Tursiops truncatus* population. *Diseases of Aquatic Organisms*, 99(3), 237–242.
- Walters, W., Hyde, E. R., Berg-lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J. a, & Jansson, J. K. (2015). Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *MSystems*, 1(1), e0009-15.
- Waltzek, T. B., Cortés-Hinojosa, G., Wellehan, J. F. X., & Gray, G. C. (2012). Marine mammal zoonoses: A review of disease manifestations. *Zoonoses and Public Health*, 59(8), 521–535.
- Wang, Y., & Qian, P. Y. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS ONE*, 4(10), e7401.
- Webster, N. S., Negri, A. P., Botté, E. S., Laffy, P. W., Flores, F., Noonan, S., Schmidt, C., & Uthicke, S. (2016). Host-associated coral reef microbes respond to the cumulative pressures of ocean warming and ocean acidification. *Scientific Reports*, 6(Dec), 1–9.
- Wickham, H. (2009). ggplot2: elegant graphics for data analysis. *Springer-Verlag New York*.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemond, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686.
- Willis, C., Desai, D., & Laroche, J. (2019). Influence of 16S rRNA variable region on perceived diversity of marine microbial communities of the Northern North Atlantic. *FEMS Microbiology Letters*, 366(13), 1–9.
- Worm, B., Barbier, E. B., Beaumont, N., Duffy, E., Folke, C., Halpern, B. S., Jackson, J. B. C., Lotze, H. K., Micheli, F., Palumbi, S. R., Sala, E., Selkoe, K. A., Stachowicz, J. J., & Watson, R. (2006). Impacts of Biodiversity Loss on Ocean Ecosystem

Services. *Science*, 314(5800), 787–790.

Yang, C., Wang, X., Miller, J. A., De Blécourt, M., Ji, Y., Yang, C., Harrison, R. D., & Yu, D. W. (2014). Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecological Indicators*, 46, 379–389.

Zehr, J. P., Weitz, J. S., & Joint, I. (2017). How microbes survive in the open ocean. *Science*, 357(6352), 646–647.

9. Supplementary information

Table S 1 – DNA concentration (ng/μl), for each sample of the short-finned pilot whale blows and sterivex seawater control samples used in the analysis.

Sample Name	DNA Concentration (ng/μl) Qubit
Gma_1	1,5
Gma_2	1,5
Gma_4	1,4
Gma_5	2,0
Gma_6	1,0
Gma_7	0,5
Gma_8	1,4
Gma_9	1,7
Gma_10	1,4
Gma_11	0,1
Gma_12	2,1
Gma_13	1,2
Blow_15	LOW
Blow_16	1,06
Blow_17	1,11
Blow_18	1,09
Blow_19	LOW
Blow_20	LOW
Blow_21	0,063
Blow_22	0,054
A_St	0,861
B_St	0,21
C_St	0,16

Table S 2 - Calculated alpha diversity indices, for each sample of the short-finned pilot whale blows, of the comparison of V3-V4 and V4-V5 primers sets.

Sample Name	Primer set	No. of ASVs	Chao1 richness	Shannon diversity index (H)	Inverse Simpson's div. index
Gma_1	V3-V4	149	149	3.76	18.34
Gma_2	V3-V4	210	210	4.78	62.33
Gma_4	V3-V4	162	162	3.77	16.77
Gma_5	V3-V4	164	164	3.46	12.67
Gma_6	V3-V4	171	171	3.44	12.45
Gma_7	V3-V4	233	233	3.83	19.24
Gma_8	V3-V4	192	192	3.59	14.29
Gma_9	V3-V4	306	306	5.05	81
Gma_10	V3-V4	292	292	5.07	89.06
Gma_11	V3-V4	341	341	5.19	93.53
Gma_12	V3-V4	306	306	5.16	97.99
Gma_13	V3-V4	238	238	4.94	81.87
Gma_1	V4-V5	374	374	4.17	20.11
Gma_2	V4-V5	276	276	3.74	13.7
Gma_4	V4-V5	471	471	4.56	23.96
Gma_5	V4-V5	251	251	3.64	12.8
Gma_6	V4-V5	309	309	3.92	14.69
Gma_7	V4-V5	340	340	3.96	15.37
Gma_8	V4-V5	237	237	3.59	7.97
Gma_9	V4-V5	299	299	4.05	10.13
Gma_10	V4-V5	260	260	4.32	17.43
Gma_11	V4-V5	302	302	3.21	4.71
Gma_12	V4-V5	265	265	3.56	6.96
Gma_13	V4-V5	281	281	4.02	10.27

Table S 3 - The proportion of all phyla present in the V4-V5 and V3-V4 datasets, for samples of the short-finned pilot whale blows.

V4-V5			V3-V4		
	Phylum	Proportion		Phylum	Proportion
1	Proteobacteria	36.68	1	Proteobacteria	47.11
2	Actinobacteria	34.12	2	Actinobacteria	15.5
3	Firmicutes	12.84	3	Bacteroidetes	12.47
4	Bacteroidetes	9.11	4	Patescibacteria	12.2
5	Cyanobacteria	3.97	5	Firmicutes	6.27
6	Patescibacteria	0.94	6	Cyanobacteria	3.08

7	Euryarchaeota	0.54	7	Epsilonbacteraeota	1.55
8	Fusobacteria	0.42	8	Fusobacteria	0.93
9	Marinimicrobia (SAR406 clade)	0.33	9	Nanoarchaeaeota	0.23
10	Spirochaetes	0.25	10	Verrucomicrobia	0.2
11	Planctomycetes	0.22	11	Chlamydiae	0.19
12	Verrucomicrobia	0.13	12	Dependentiae	0.11
13	Chlamydiae	0.12	13	WPS-2	0.1
14	Dadabacteria	0.09	14	Tenericutes	0.04
15	Epsilonbacteraeota	0.07	15	Marinimicrobia (SAR406 clade)	0.03
16	Kiritimatiellaeota	0.04			
17	Chloroflexi	0.03			
18	Nanoarchaeaeota	0.03			
19	WPS-2	0.03			
20	Acidobacteria	0.02			
21	Deinococcus-Thermus	0.02			

Table S 4 - The proportion of all classes present in the V4-V5 and V3-V4 datasets, for samples of the short-finned pilot whale blows.

V4-V5			V3-V4		
	Class	Proportion		Class	Proportion
1	Actinobacteria	33.56	1	Gammaproteobacteria	31.3
2	Gammaproteobacteria	22.67	2	Actinobacteria	15.41
3	Alphaproteobacteria	13.59	3	Alphaproteobacteria	15.39
4	Bacilli	11.61	4	Bacteroidia	12.47
5	Bacteroidia	9.11	5	Parcubacteria	6.2
6	Oxyphotobacteria	3.73	6	Bacilli	4.76
7	Clostridia	0.69	7	Gracilibacteria	4.19
8	Thermoplasmata	0.54	8	Oxyphotobacteria	2.42
9	Negativicutes	0.48	9	Campylobacteria	1.55
10	Parcubacteria	0.48	10	Saccharimonadia	1.3
11	Fusobacteriia	0.42	11	Clostridia	1.23
12	Acidimicrobiia	0.4	12	Fusobacteriia	0.93
13	Gracilibacteria	0.37	13	Melainabacteria	0.66
14	Deltaproteobacteria	0.37	14	Deltaproteobacteria	0.41
15	Marinimicrobia (SAR406 clade) uncl	0.33	15	Microgenomatia	0.31
16	Spirochaetia	0.25	16	Negativicutes	0.24
17	Melainabacteria	0.23	17	Woeseearchaeia	0.23
18	Coriobacteriia	0.16	18	Verrucomicrobiae	0.2
19	Phycisphaerae	0.14	19	Chlamydiae	0.19
20	Verrucomicrobiae	0.13	20	Babeliae	0.11
21	Chlamydiae	0.12	21	WS6 (Dojkabacteria)	0.11
22	Dadabacteriia	0.09	22	WPS-2 uncl	0.1
23	Campylobacteria	0.07	23	Acidimicrobiia	0.09
24	Planctomycetacia	0.06	24	ABY1	0.08
25	Proteobacteria uncl	0.06	25	Mollicutes	0.04
26	Erysipelotrichia	0.05	26	Erysipelotrichia	0.03
27	Saccharimonadia	0.05	27	Marinimicrobia (SAR406 clade) uncl	0.03
28	Kiritimatiellae	0.04			
29	Microgenomatia	0.04			
30	Dehalococcoidia	0.03			
31	Woeseearchaeia	0.03			
32	WPS-2 uncl	0.03			
33	OM190	0.02			
34	Subgroup_6	0.02			

35	Deinococci	0.02
36	ABY1	0.01

Table S 5 - The proportion of all families present in the V4-V5 and V3-V4 datasets, for samples of the short-finned pilot whale blows.

V4-V5			V3-V4		
	Family	Proportion		Family	Proportion
1	Propionibacteriaceae	25.28	1	Pseudomonadaceae	17.8
2	Pseudomonadaceae	12.62	2	Flavobacteriaceae	9.31
3	Staphylococcaceae	6.81	3	Nocardiaceae	7.37
4	Flavobacteriaceae	5.6	4	Propionibacteriaceae	3.91
5	Rhodanobacteraceae	4.52	5	Rhodanobacteraceae	3.86
6	Nocardiaceae	4.09	6	Candidatus Peribacteria_uncl	3.48
7	Cyanobiaceae	3.73	7	Clade_I	3.25
8	Streptococcaceae	3.15	8	Sphingomonadaceae	3.13
9	Rhizobiaceae	3.01	9	Moraxellaceae	3.04
10	Clade_I	2.7	10	Burkholderiaceae	2.89
11	Sphingomonadaceae	2.65	11	Rhizobiaceae	2.67
12	Rhodobacteraceae	2.04	12	Parcubacteria_uncl	2.53
13	Corynebacteriaceae	1.97	13	Rhodobacteraceae	2.5
14	Burkholderiaceae	1.35	14	Cyanobiaceae	2.42
15	Micrococcaceae	1.05	15	Micrococcaceae	2.39
16	Moraxellaceae	1.04	16	Staphylococcaceae	1.96
17	Porphyromonadaceae	1.01	17	Streptococcaceae	1.48
18	Caulobacteraceae	0.93	18	Arcobacteraceae	1.39
19	Pasteurellaceae	0.86	19	Saccharimonadales_uncl	1.3
20	Neisseriaceae	0.65	20	Pasteurellaceae	1.28
21	Bacillaceae	0.64	21	Lactobacillaceae	1.1
22	Prevotellaceae	0.64	22	Candidatus_Nomurabacteria_uncl	1.07
23	Family_XI	0.62	23	Caulobacteraceae	1.02
24	Marine_Group_II_uncl	0.53	24	Porphyromonadaceae	1.02
25	Clade_II	0.53	25	Family_XI	0.93
26	Veillonellaceae	0.48	26	Leptotrichiaceae	0.93
27	SAR116_clade	0.45	27	Corynebacteriaceae	0.88
28	Crocinitomicaceae	0.44	28	Neisseriaceae	0.79
29	Microbacteriaceae	0.43	29	Beijerinckiaceae	0.77
30	Parcubacteria_uncl	0.39	30	Candidatus_Zambryskibacteria_uncl	0.69
31	Actinomarinaceae	0.39	31	Bacteroidales_uncl	0.67
32	Weeksellaceae	0.39	32	Obscuribacterales_uncl	0.66
33	SAR86_clade_uncl	0.39	33	Gracilibacteria_uncl	0.63
34	Xanthomonadaceae	0.34	34	Candidatus_Kaiserbacteria_uncl	0.6
35	Beijerinckiaceae	0.33	35	Crocinitomicaceae	0.5
36	Marinimicrobia_(SAR406_clade)_uncl	0.33	36	Candidatus_Campbellbacteria_uncl	0.41
37	Lactobacillaceae	0.33	37	Spirosomaceae	0.38
38	Candidatus_Peribacteria_uncl	0.28	38	Candidatus_Yonathbacteria_uncl	0.36
39	NS9_marine_group	0.25	39	Legionellaceae	0.36
40	Spirochaetaceae	0.25	40	Microbacteriaceae	0.35
41	Enterobacteriaceae	0.24	41	SAR86_clade_uncl	0.35
42	S25-593	0.23	42	SAR116_clade	0.32
43	Obscuribacterales_uncl	0.23	43	Weeksellaceae	0.3
44	Chitinophagaceae	0.23	44	Clade_II	0.29
45	AEGEAN-169_marine_group	0.23	45	Chitinophagaceae	0.29
46	Fusobacteriaceae	0.22	46	AEGEAN-169_marine_group	0.28
47	Spirosomaceae	0.21	47	SM2D12	0.27

48	Leptotrichiaceae	0.21	48	Xanthomonadaceae	0.26
49	Actinomycetaceae	0.2	49	Xanthobacteraceae	0.25
50	Carnobacteriaceae	0.2	50	Lachnospiraceae	0.24
51	Atopobiaceae	0.16	51	Veillonellaceae	0.24
52	Bifidobacteriaceae	0.15	52	Cardiobacteriaceae	0.23
53	Peptostreptococcaceae	0.14	53	Candidatus_Woesebacteria_uncl	0.23
54	Nakamurellaceae	0.14	54	Woeseearchaeia_uncl	0.23
55	Phycisphaeraceae	0.14	55	Clade_IV	0.21
56	Clade_IV	0.12	56	0319-6G20	0.21
57	Family_XIII	0.11	57	Rubritaleaceae	0.2
58	Alphaproteobacteria_uncl	0.11	58	Candidatus_Vogelbacteria_uncl	0.2
59	Gammaproteobacteria_uncl	0.11	59	Gammaproteobacteria_uncl	0.19
60	Chlamydiaceae	0.11	60	Candidatus_Azambacteria_uncl	0.17
61	SAR324_clade(Marine_group_B)_uncl	0.1	61	Bifidobacteriaceae	0.17
62	Cyclobacteriaceae	0.1	62	Candidatus_Yanofskybacteria_uncl	0.16
63	0319-6G20	0.1	63	Campylobacteriales_uncl	0.15
64	Puniceicoccaceae	0.1	64	Chlamydiaceae	0.13
65	Lachnospiraceae	0.09	65	Babeliaceae	0.11
66	Dadabacteriales_uncl	0.09	66	WS6_(Dojkabacteria)_uncl	0.11
67	Leuconostocaceae	0.08	67	Rhodospirillales_uncl	0.11
68	Xanthobacteraceae	0.08	68	SAR324_clade(Marine_group_B)_uncl	0.11
69	Chitinophagales_uncl	0.08	69	Aerococcaceae	0.1
70	Rhodocyclaceae	0.07	70	WPS-2_uncl	0.1
71	Brevibacteriaceae	0.07	71	Actinomarinaceae	0.09
72	Gracilibacteria_uncl	0.07	72	Brevibacteriaceae	0.09
73	Frankiales_uncl	0.06	73	Micavibrionales_uncl	0.08
74	NS7_marine_group	0.06	74	Pseudoalteromonadaceae	0.08
75	Ectothiorhodospiraceae	0.06	75	Dermabacteraceae	0.07
76	Proteobacteria_uncl	0.06	76	Oligoflexaceae	0.07
77	Arcobacteraceae	0.05	77	Oceanospirillales_uncl	0.07
78	PB19_uncl	0.05	78	Carnobacteriaceae	0.06
79	Dermabacteraceae	0.05	79	Candidatus_Pacebacteria_uncl	0.06
80	Erysipelotrichaceae	0.05	80	Candidatus_Magasanikbacteria_uncl	0.06
81	Oligoflexaceae	0.05	81	Peptostreptococcaceae	0.06
82	Candidatus_Nomurabacteria_uncl	0.05	82	Family_XII	0.05
83	Dysgonomonadaceae	0.05	83	Rhizobiales_Incertae_Sedis	0.05
84	Candidatus_Zambryskibacteria_uncl	0.04	84	Mycobacteriaceae	0.05
85	Enterococcaceae	0.04	85	Actinomycetaceae	0.05
86	Porticoccaceae	0.04	86	Hyphomicrobiaceae	0.05
87	Kiritimatiellaceae	0.04	87	Dietziaceae	0.04
88	Family_X	0.04	88	Intrasporangiaceae	0.04
89	Saccharimonadales_uncl	0.04	89	Clade_III	0.04
90	Halomonadaceae	0.04	90	Absconditabacteriales_(SR1)_uncl	0.04
91	KI89A_clade_uncl	0.04	91	Alphaproteobacteria_uncl	0.04
92	SAR202_clade_uncl	0.03	92	Mycoplasmataceae	0.04
93	Aeromonadaceae	0.03	93	Candidatus_Peregrinibacteria_uncl	0.04
94	Woeseearchaeia_uncl	0.03	94	OCS116_clade	0.04
95	Gemmataceae	0.03	95	bac2nit3	0.03
96	Thiotrichaceae	0.03	96	Simkaniaceae	0.03
97	Bdellovibrionaceae	0.03	97	Erysipelotrichaceae	0.03
98	Pseudonocardiaceae	0.03	98	Marinimicrobia_(SAR406_clade)_uncl	0.03
99	Pirellulaceae	0.03	99	Solimonadaceae	0.03
100	env.OPS_17	0.03	100	Enterobacteriaceae	0.03
101	WPS-2_uncl	0.03	101	Aeromonadaceae	0.03
102	Rubritaleaceae	0.03	102	Candidatus_Adlerbacteria_uncl	0.03
103	Hydrogenophilaceae	0.02	103	cvE6	0.03
104	Candidatus_Woesebacteria_uncl	0.02	104	Bacteriovoracaceae	0.03

105	Nocardioidaceae	0.02	105	Candidatus Falkowbacteria_uncl	0.02
106	OCS116_clade	0.02	106	Candidatus Collierbacteria_uncl	0.02
107	Legionellaceae	0.02	107	Hydrogenophilaceae	0.02
108	Saccharospirillaceae	0.02	108	Caenarcaniphilales_uncl	0.004
109	OM190_uncl	0.02			
110	Paracaedibacteraceae	0.02			
111	SM2D12	0.02			
112	Peptococcaceae	0.02			
113	Halieaceae	0.02			
114	Unknown_Family	0.02			
115	Ga0077536_uncl	0.02			
116	Micavibrionales_uncl	0.02			
117	OM182_clade_uncl	0.02			
118	Parvibaculaceae	0.02			
119	Rickettsiales_uncl	0.02			
120	Deltaproteobacteria_uncl	0.02			
121	Subgroup_6_uncl	0.02			
122	Campylobacteraceae	0.02			
123	Nisaeaceae	0.02			
124	Deinococcaceae	0.02			
125	Absconditabacteriales_(SR1)_uncl	0.02			
126	Aerococcaceae	0.01			
127	B1-7BS	0.01			
128	Listeriaceae	0.01			
129	Solimonadaceae	0.01			
130	Hyphomicrobiaceae	0.01			
131	Spongiibacteraceae	0.01			
132	Diplorickettsiaceae	0.01			
133	Methylophilaceae	0.01			
134	Polyangiaceae	0.01			
135	Simkaniaceae	0.01			
136	Microtrichaceae	0.01			
137	Candidatus Amesbacteria_uncl	0.01			
138	Pleomorphomonadaceae	0.01			
139	Clade III	0.01			
140	Saccharimonadaceae	0.01			
141	Sphingobacteriales_uncl	0.01			
142	NS11-12_marine_group	0.01			
143	PS1_clade	0.01			
144	Rikenellaceae	0.01			
145	Dietziaceae	0.01			
146	Oceanospirillales_uncl	0.01			
147	Intrasporangiaceae	0.01			
148	Marine_Group_III_uncl	0.01			
149	Candidatus Kerfeldbacteria_uncl	0.004			
150	Candidatus Collierbacteria_uncl	0.003			
151	Streptomycetaceae	0.003			
152	Pseudohongiellaceae	0.003			
153	Candidatus Magasanikbacteria_uncl	0.002			

Table S 6 - The proportion of all genera present in the V4-V5 and V3-V4 datasets, for samples of the short-finned pilot whale blows.

V4-V5			V3-V4		
	Genus	Proportion		Genus	Proportion
1	Cutibacterium	25.27	1	Pseudomonas	17.8
2	Pseudomonas	12.62	2	Flavobacterium	7.59
3	Staphylococcus	6.81	3	Rhodococcus	7.06
4	Dyella	4.52	4	Cutibacterium	3.91
5	Rhodococcus	4.07	5	Dyella	3.86
6	Flavobacterium	4.04	6	Candidatus_Peribacteria_uncl	3.48
7	Prochlorococcus_MIT9313	3.46	7	Sphingomonas	3.13
8	Streptococcus	3.1	8	Parcubacteria_uncl	2.53
9	Phyllobacterium	2.89	9	Clade_la	2.51
10	Sphingomonas	2.58	10	Prochlorococcus_MIT9313	2.33
11	Clade_la	1.86	11	Phyllobacterium	2.25
12	Paracoccus	1.56	12	Paracoccus	2.22
13	Lawsonella	1.13	13	Staphylococcus	1.96
14	Porphyromonas	1.01	14	Coenonia	1.61
15	Clade_lb	0.84	15	Streptococcus	1.48
16	Haemophilus	0.76	16	Enhydrobacter	1.39
17	Acinetobacter	0.67	17	Arcobacter	1.39
18	Corynebacterium_1	0.65	18	Saccharimonadales_uncl	1.3
19	Brevundimonas	0.65	19	Phocoenobacter	1.2
20	Marine_Group_II_uncl	0.53	20	Acinetobacter	1.2
21	Clade_II_uncl	0.53	21	Paenarthrobacter	1.15
22	Anoxybacillus	0.51	22	Lactobacillus	1.1
23	NS2b_marine_group	0.49	23	Candidatus_Nomurabacteria_uncl	1.07
24	NS5_marine_group	0.48	24	Porphyromonas	1.02
25	NS4_marine_group	0.47	25	Micrococcus	0.81
26	Fluviicola	0.44	26	Neisseriaceae_uncl	0.79
27	SAR116_clade_uncl	0.41	27	Clade_lb	0.75
28	Parcubacteria_uncl	0.39	28	Candidatus_Zambryskibacteria_uncl	0.69
29	Paenarthrobacter	0.39	29	Methylobacterium	0.67
30	Candidatus_Actinomarina	0.39	30	Bacteroidales_uncl	0.67
31	SAR86_clade_uncl	0.39	31	Obscuribacterales_uncl	0.66
32	Veillonella	0.37	32	Gracilibacteria_uncl	0.63
33	Rothia	0.36	33	Candidatus_Kaiserbacteria_uncl	0.6
34	Microbacterium	0.36	34	Oceanivirga	0.56
35	Cloacibacterium	0.36	35	Limnohabitans	0.54
36	Enhydrobacter	0.36	36	Anaerococcus	0.51
37	Burkholderiaceae_uncl	0.36	37	Fluviicola	0.5
38	Marinimicrobia_(SAR406_clade)_uncl	0.33	38	Lawsonella	0.49
39	Lactobacillus	0.33	39	Ralstonia	0.49
40	Neisseriaceae_uncl	0.3	40	Alcaligenes	0.45
41	Gemella	0.3	41	Delftia	0.44
42	Candidatus_Peribacteria_uncl	0.28	42	Brevundimonas	0.44
43	Tepidimonas	0.28	43	Candidatus_Campbellbacteria_uncl	0.41
44	Synechococcus_CC9902	0.27	44	Caulobacteraceae_uncl	0.41
45	Neisseria	0.26	45	Corynebacterium_1	0.38
46	NS9_marine_group_uncl	0.25	46	Pseudarcicella	0.38
47	Treponema_2	0.25	47	Mesorhizobium	0.38
48	Prevotella	0.25	48	Candidatus_Yonathbacteria_uncl	0.36
49	Caulobacteraceae_uncl	0.24	49	Legionella	0.36
50	Prevotella_7	0.24	50	Kocuria	0.36
51	Micrococcus	0.24	51	SAR86_clade_uncl	0.35
52	S25-593_uncl	0.23	52	SAR116_clade_uncl	0.32

53	Obscuribacterales_uncl	0.23	53	Moraxella	0.32
54	Sediminibacterium	0.23	54	Gordonia	0.3
55	AEGEAN-169_marine_group_uncl	0.23	55	Clade_II_uncl	0.29
56	Rhodobacteraceae_uncl	0.22	56	Sediminibacterium	0.29
57	Fusobacterium	0.22	57	AEGEAN-169_marine_group_uncl	0.28
58	Pseudarcicella	0.2	58	Leptotrichiaceae_uncl	0.28
59	Methylobacterium	0.19	59	SM2D12_uncl	0.27
60	Actinomyces	0.19	60	Finegoldia	0.26
61	Corynebacterium	0.19	61	Bradyrhizobium	0.25
62	Xanthomonas	0.18	62	Pseudoxanthomonas	0.25
63	Anaerococcus	0.16	63	Lachnospiraceae_AC2044_group	0.24
64	Atopobium	0.16	64	Veillonella	0.24
65	Gardnerella	0.15	65	Cardiobacteriaceae_uncl	0.23
66	Nakamurella	0.14	66	Candidatus_Woesebacteria_uncl	0.23
67	Alloprevotella	0.14	67	Woeseearchaeia_uncl	0.23
68	Bosea	0.13	68	Clade_IV_uncl	0.21
69	Granulicatella	0.13	69	0319-6G20_uncl	0.21
70	Yersinia	0.12	70	Candidatus_Vogelbacteria_uncl	0.2
71	Roseovarius	0.12	71	Gammaproteobacteria_uncl	0.19
72	Clade_IV_uncl	0.12	72	Tabrizicola	0.18
73	Urania-1B-19_marine_sediment_group	0.12	73	Chryseobacterium	0.18
74	Leptotrichiaceae_uncl	0.12	74	Candidatus_Azambacteria_uncl	0.17
75	Capnocytophaga	0.11	75	Gardnerella	0.17
76	Alphaproteobacteria_uncl	0.11	76	Paucibacter	0.17
77	Gammaproteobacteria_uncl	0.11	77	Polynucleobacter	0.17
78	Pseudoxanthomonas	0.11	78	Microbacterium	0.16
79	Chlamydiaceae_uncl	0.11	79	Peptoniphilus	0.16
80	Aquabacterium	0.11	80	Candidatus_Yanofskybacteria_uncl	0.16
81	Anaerobacillus	0.11	81	Campylobacteriales_uncl	0.15
82	SAR324_clade(Marine_group_B)_uncl	0.1	82	Ottowia	0.15
83	Family_XIII_uncl	0.1	83	Psychrobacter	0.14
84	Marinoscillum	0.1	84	Chlamydiaceae_uncl	0.13
85	0319-6G20_uncl	0.1	85	Cloacibacterium	0.13
86	Oceanivirga	0.09	86	Diaphorobacter	0.12
87	Kingella	0.09	87	Babeliaceae_uncl	0.11
88	Dadabacteriales_uncl	0.09	88	WS6_(Dojkabacteria)_uncl	0.11
89	Leuconostoc	0.08	89	Luteolibacter	0.11
90	Polynucleobacter	0.08	90	Rhodospirillales_uncl	0.11
91	Filifactor	0.08	91	SAR324_clade(Marine_group_B)_uncl	0.11
92	Chitinophagales_uncl	0.08	92	Aquabacterium	0.1
93	Tabrizicola	0.08	93	Aerococcus	0.1
94	Lautropia	0.07	94	Leucobacter	0.1
95	Methyloversatilis	0.07	95	WPS-2_uncl	0.1
96	Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium	0.07	96	Bosea	0.09
97	Brevibacterium	0.07	97	Caulobacter	0.09
98	Comamonas	0.07	98	Synechococcus_CC9902	0.09
99	Dolosigranulum	0.07	99	Candidatus_Actinomarina	0.09
100	Gracilibacteria_uncl	0.07	100	Brevibacterium	0.09
101	Parvimonas	0.07	101	Roseibacillus	0.09
102	Peptoniphilus	0.06	102	Streptobacillus	0.09
103	Ottowia	0.06	103	Phenylobacterium	0.09
104	Frankiales_uncl	0.06	104	Micavibrionales_uncl	0.08
105	NS7_marine_group_uncl	0.06	105	Haemophilus	0.08
106	Ectothiorhodospiraceae_uncl	0.06	106	Pseudoalteromonas	0.08
107	Burkholderia-Caballeronia-Paraburkholderia	0.06	107	Pseudorhodobacter	0.08
108	Escherichia/Shigella	0.06	108	Brachybacterium	0.07
109	Lachnoanaerobaculum	0.06	109	Rothia	0.07

110	Proteobacteria_uncl	0.06	110	Oligoflexaceae_uncl	0.07
111	Arcobacter	0.05	111	Burkholderia-Caballeronia-Paraburkholderia	0.07
112	Mesorhizobium	0.05	112	Oceanospirillales_uncl	0.07
113	Dialister	0.05	113	Granulicatella	0.06
114	PB19_uncl	0.05	114	Candidatus_Pacebacteria_uncl	0.06
115	Lentimonas	0.05	115	Rhodoluna	0.06
116	Bisgaardia	0.05	116	Candidatus_Magasanikbacteria_uncl	0.06
117	Brachybacterium	0.05	117	Romboutsia	0.06
118	Pelistega	0.05	118	Acidovorax	0.06
119	Delftia	0.05	119	NS5_marine_group	0.05
120	Candidatus_Nomurabacteria_uncl	0.05	120	Exiguobacterium	0.05
121	Dysgonomonas	0.05	121	Phreatobacter	0.05
122	Lactococcus	0.05	122	Tenacibaculum	0.05
123	Massilia	0.05	123	Mycobacterium	0.05
124	Pantoea	0.04	124	Actinomyces	0.05
125	Candidatus_Zambryskibacteria_uncl	0.04	125	Hyphomicrobium	0.05
126	Oligoflexus	0.04	126	Dietzia	0.04
127	SAR92_clade	0.04	127	Janibacter	0.04
128	Caulobacter	0.04	128	Clade_III_uncl	0.04
129	Afipia	0.04	129	Absconditabacteriales_(SR1)_uncl	0.04
130	Erysipelothrix	0.04	130	Alphaproteobacteria_uncl	0.04
131	R76-B128	0.04	131	Mycoplasma	0.04
132	Ascidiaehabitans	0.04	132	Candidatus_Peregrinibacteria_uncl	0.04
133	Thermicanus	0.04	133	OCS116_clade_uncl	0.04
134	Candidatus_Puniceispirillum	0.04	134	Hydrogenophaga	0.04
135	Duganella	0.04	135	Rhodoferrax	0.03
136	Stenotrophomonas	0.04	136	bac2nit3_uncl	0.03
137	Bradyrhizobium	0.04	137	Ga0074140	0.03
138	Saccharimonadales_uncl	0.04	138	Aurantimicrobium	0.03
139	Halomonas	0.04	139	Dielma	0.03
140	K189A_clade_uncl	0.04	140	Marinimicrobia_(SAR406_clade)_uncl	0.03
141	Diaphorobacter	0.04	141	Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium	0.03
142	Peptostreptococcus	0.03	142	Solimonadaceae_uncl	0.03
143	SAR202_clade_uncl	0.03	143	Aeromonas	0.03
144	Aeromonas	0.03	144	Candidatus_Adlerbacteria_uncl	0.03
145	Finegoldia	0.03	145	cvE6_uncl	0.03
146	MB11C04_marine_group	0.03	146	Peredibacter	0.03
147	Woeseearchaeia_uncl	0.03	147	Candidatus_Falkowbacteria_uncl	0.02
148	Gemmata	0.03	148	Escherichia/Shigella	0.02
149	Thiotrichaceae_uncl	0.03	149	Candidatus_Collierbacteria_uncl	0.02
150	Bdellovibrio	0.03	150	Schlegelella	0.02
151	Megasphaera	0.03	151	Comamonas	0.02
152	Pirellula	0.03	152	Tepidiphilus	0.02
153	Phocoenobacter	0.03	153	Pelomonas	0.01
154	env.OPS_17_uncl	0.03	154	Falsirhodobacter	0.01
155	WPS-2_uncl	0.03	155	Stenotrophomonas	0.01
156	Romboutsia	0.03	156	Serratia	0.01
157	Kocuria	0.03	157	Caenarcarniphilales_uncl	0.004
158	Roseibacillus	0.03			
159	Sphingobium	0.02			
160	Bacillus	0.02			
161	Hydrogenophilus	0.02			
162	Weeksellaceae_uncl	0.02			
163	Candidatus_Woesebacteria_uncl	0.02			
164	Nocardioides	0.02			
165	Pseudorhodobacter	0.02			
166	OCS116_clade_uncl	0.02			

167	Arthrobacter	0.02
168	Legionella	0.02
169	Thalassolituus	0.02
170	Vagococcus	0.02
171	OM190_uncl	0.02
172	Candidatus_Finniella	0.02
173	SM2D12_uncl	0.02
174	Peptococcus	0.02
175	SM1A02	0.02
176	Acidibacter	0.02
177	OM60(NOR5)_clade	0.02
178	Janthinobacterium	0.02
179	Aggregatibacter	0.02
180	Altererythrobacter	0.02
181	Enterococcus	0.02
182	Ga0077536_uncl	0.02
183	Micavibrionales_uncl	0.02
184	OM182_clade_uncl	0.02
185	Parvibaculaceae_uncl	0.02
186	Proteus	0.02
187	Gordonia	0.02
188	Saccharopolyspora	0.02
189	Shuttleworthia	0.02
190	Rickettsiales_uncl	0.02
191	Selenomonas	0.02
192	Coralimargarita	0.02
193	Deltaproteobacteria_uncl	0.02
194	Subgroup_6_uncl	0.02
195	Xylophilus	0.02
196	Frigoribacterium	0.02
197	Rhodoluna	0.02
198	Sphingopyxis	0.02
199	Campylobacter	0.02
200	OM75_clade	0.02
201	Deinococcus	0.02
202	Psychrobacter	0.02
203	Absconditabacteriales_(SR1)_uncl	0.02
204	Stomatobaculum	0.02
205	Abiotrophia	0.01
206	Actinotignum	0.01
207	B1-7BS_uncl	0.01
208	Brachymonas	0.01
209	Luteimonas	0.01
210	Brochothrix	0.01
211	Solimonadaceae_uncl	0.01
212	Rathayibacter	0.01
213	Candidatus_Aquiluna	0.01
214	BD1-7_clade	0.01
215	Methylothera	0.01
216	Mogibacterium	0.01
217	Rickettsiella	0.01
218	Dyadobacter	0.01
219	Ga0074140	0.01
220	Pajaroellobacter	0.01
221	Pseudonocardiaceae_uncl	0.01
222	Sva0996_marine_group	0.01
223	Prevotella_2	0.01

224	Candidatus_Amesbacteria_uncl	0.01
225	Pleomorphomonas	0.01
226	Selenomonas_3	0.01
227	Clade_III_uncl	0.01
228	Methylocella	0.01
229	Candidatus_Saccharimonas	0.01
230	Sphingobacteriales_uncl	0.01
231	NS11-12_marine_group_uncl	0.01
232	Pedomicrobium	0.01
233	PS1_clade_uncl	0.01
234	Faecalibaculum	0.01
235	Rikenellaceae_RC9_gut_group	0.01
236	Sphingomonadaceae_uncl	0.01
237	Dietzia	0.01
238	Oceanospirillales_uncl	0.01
239	Rhizorhapis	0.01
240	Chryseobacterium	0.01
241	Janibacter	0.01
242	Aurantimicrobium	0.01
243	Marine_Group_III_uncl	0.01
244	Propioniciclava	0.01
245	Oligoflexaceae_uncl	0.01
246	Hyphomicrobium	0.005
247	Olsenella	0.005
248	Candidatus_Kerfeldbacteria_uncl	0.004
249	Candidatus_Collierbacteria_uncl	0.003
250	Streptomyces	0.003
251	Pseudohongiella	0.003
252	Candidatus_Magasanikbacteria_uncl	0.002

Table S 7 - Overview of ASV composition of bacterial families in both datasets. The table consists of all bacterial families that were observed in both V3-V4 and V4-V5 datasets, for samples of the short-finned pilot whale blows, and had a sequence proportion of at least 1%.

Class	Family	Seq.proportion (% of total V3- V4)	Seq.proportion (% of total V4- V5)	Num. ASVs (V3-V4)	Num. ASVs (V4-V5)	ASVs ratio (V3-V4/ V4-V5)
Actinobacteria	Corynebacteriaceae	1	2	7	15	0.47
Actinobacteria	Micrococcaceae	2	1	10	7	1.43
Actinobacteria	Nocardiaceae	7	4	5	4	1.25
Actinobacteria	Propionibacteriaceae	4	25	2	6	0.33
Alphaproteobacteria	Beijerinckiaceae	1	0	4	7	0.57
Alphaproteobacteria	Caulobacteraceae	1	1	6	5	1.20
Alphaproteobacteria	Clade_I	3	3	10	20	0.50
Alphaproteobacteria	Clade_II	0	1	2	5	0.40
Alphaproteobacteria	Rhizobiaceae	3	3	3	5	0.60
Alphaproteobacteria	Rhodobacteraceae	2	2	8	18	0.44
Alphaproteobacteria	Sphingomonadaceae	3	3	3	11	0.27
Bacilli	Lactobacillaceae	1	0	3	4	0.75
Bacilli	Staphylococcaceae	2	7	7	8	0.88
Bacilli	Streptococcaceae	1	3	6	9	0.67
Bacteroidia	Crocinitomicaceae	1	0	3	3	1.00
Bacteroidia	Flavobacteriaceae	9	6	10	32	0.31
Bacteroidia	Porphyromonadaceae	1	1	4	7	0.57

Campylobacteria	Arcobacteraceae	1	0	2	1	2.00
Clostridia	Family_XI	1	0	9	9	1.00
Gammaproteobacteria	Burkholderiaceae	3	1	23	21	1.10
Gammaproteobacteria	Moraxellaceae	3	1	18	12	1.50
Gammaproteobacteria	Neisseriaceae	1	1	3	10	0.30
Gammaproteobacteria	Pasteurellaceae	1	1	3	11	0.27
Gammaproteobacteria	Pseudomonadaceae	18	13	3	6	0.50
Gammaproteobacteria	Rhodanobacteraceae	4	5	1	1	1.00
Gracilibacteria	Candidatus_Peribacteria_uncl	3	0	7	4	1.75
Gracilibacteria	Gracilibacteria_uncl	1	0	2	1	2.00
Oxyphotobacteria	Cyanobiaceae	2	4	5	12	0.42
Parcubacteria	Candidatus_Nomurabacteria_uncl	1	0	17	4	4.25
Parcubacteria	Candidatus_Zambryskibacteria_uncl	1	0	7	1	7.00
Parcubacteria	Parcubacteria_uncl	3	0	14	10	1.40
Saccharimonadia	Saccharimonadales_uncl	1	0	13	3	4.33

Table S 8 - Calculated alpha diversity indices for each sample of short-finned pilot whale blows and respective control seawater samples, using the primer set V4-V5.

Sample Name	Primer set	No. of ASVs	Chao1 richness	Shannon diversity index (H)	Inverse Simpson's div. index
Blow_15	V4-V5	69	69	3.89	37.18
Blow_16	V4-V5	93	93	3.61	19.60
Blow_17	V4-V5	85	85	3.70	25.01
Blow_18	V4-V5	169	169	4.27	40.14
Blow_19	V4-V5	186	186.33	4.12	31.64
Blow_20	V4-V5	185	185.11	4.19	29.46
Blow_21	V4-V5	163	163	4.45	48.75
Blow_22	V4-V5	166	166	4.34	41.74
Sterivex_A	V4-V5	116	117.43	3.37	13.49
Sterivex_B	V4-V5	123	123.08	4.00	27.24
Sterivex_C	V4-V5	99	99	3.99	33.73