

## Introduction

Comparative studies of terrestrial mammal hosts have identified host phylogeny, diet, and gut morphology as primary drivers of the gut bacterial community composition. Marine mammals have been excluded from these studies, although they represent unique feeding strategies, environments and life histories. Cetaceans are the most diverse mammalian group to have adapted to a fully marine existence. Due to their obligate marine lives in which the intestinal flora is largely inherited from mother to calf, it follows that a significant proportion of their gut microbiomes could have descended from terrestrial elements that have been isolated from land for millions of years. As such, whales may represent one of the only arguably "closed system" with which to study the diversity and evolution of the mammalian gut flora over deep time. Extant cetaceans comprise two groups: baleen whales (Mysticeti) and toothed whales (Odontoceti). In addition to different diet and feeding strategies, the digestive anatomy and physiology also differs between the two cetacean groups. The purpose of this comparative study is to investigate the microbial diversity and functional potential in the whale hindgut. We employed both high-throughput amplicon sequencing of 16S rRNA hypervariable tags and genomic DNA shotgun Illumina sequencing of fecal samples from two representative whale species, the long-finned pilot whale (*Globicephala melas*), which is a toothed whale, and the bowhead whale (*Balaena mysticetus*), which belongs to the baleen whales.

## Objectives

- Is the whale host microbiome a product of the evolutionary origin of the hosts themselves, their surrounding marine environment, or both?
- Do differences in the whale diet/feeding strategies and digestive anatomy/physiology play a role in forming the adult microbial community?

## Materials

Study Species: (lower hindgut fecal samples)



➤ Long finned pilot whale:  
*Globicephala melas*

➤ Bowhead whale:  
*Balaena mysticetus*

## Methods

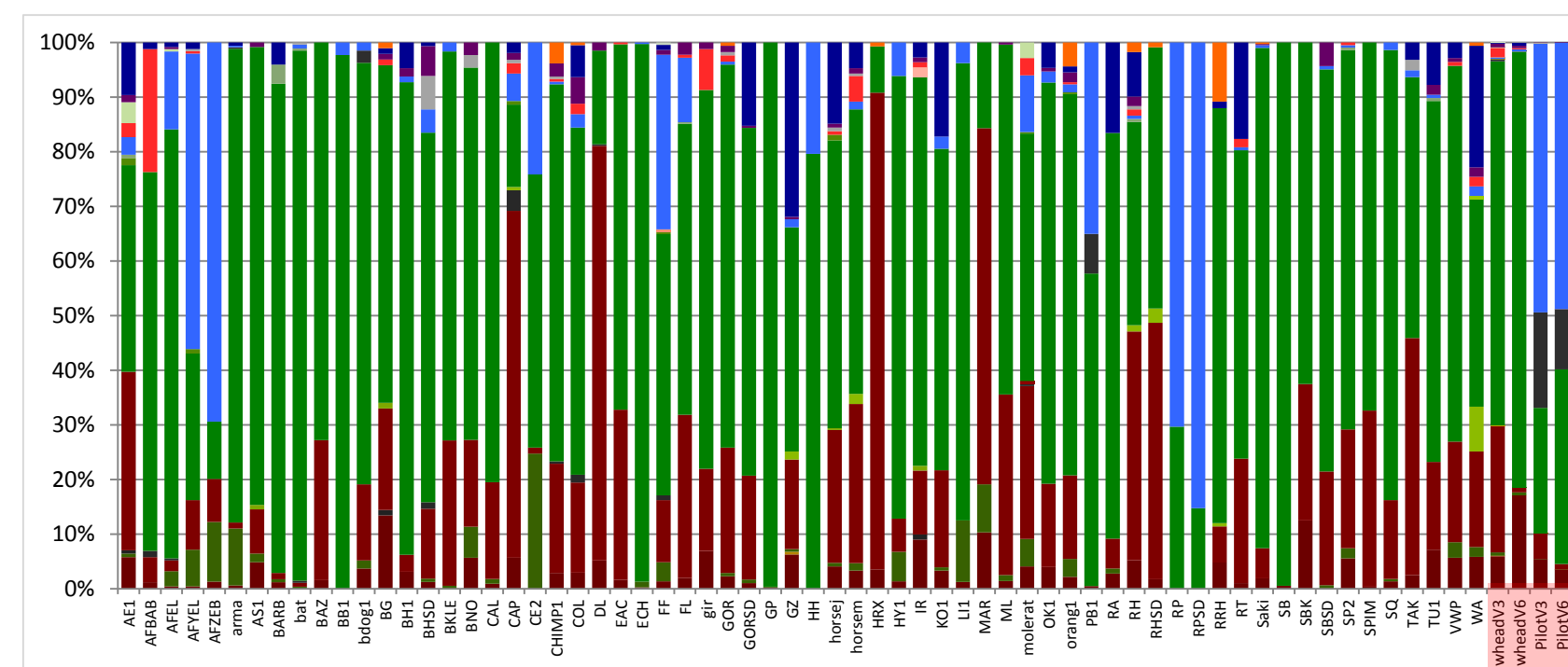
- Collection of fecal samples: Feces from distal guts of six bowhead whales was sampled from legally harvested whales during the subsistence hunt in Barrow, Alaska. Pilot whale feces was collected from six whales during the subsistence hunt at the Faroe Islands.
- Amplicon sequencing of hypervariable 16S rRNA tags V3 and V6 was employed to explore taxonomic abundance.
- Shotgun genomic sequencing on Illumina was employed to explore the functional potential of the microbiota.
- Edena was employed for de novo assembly of Illumina genomic reads.

## Metagenomic Analysis

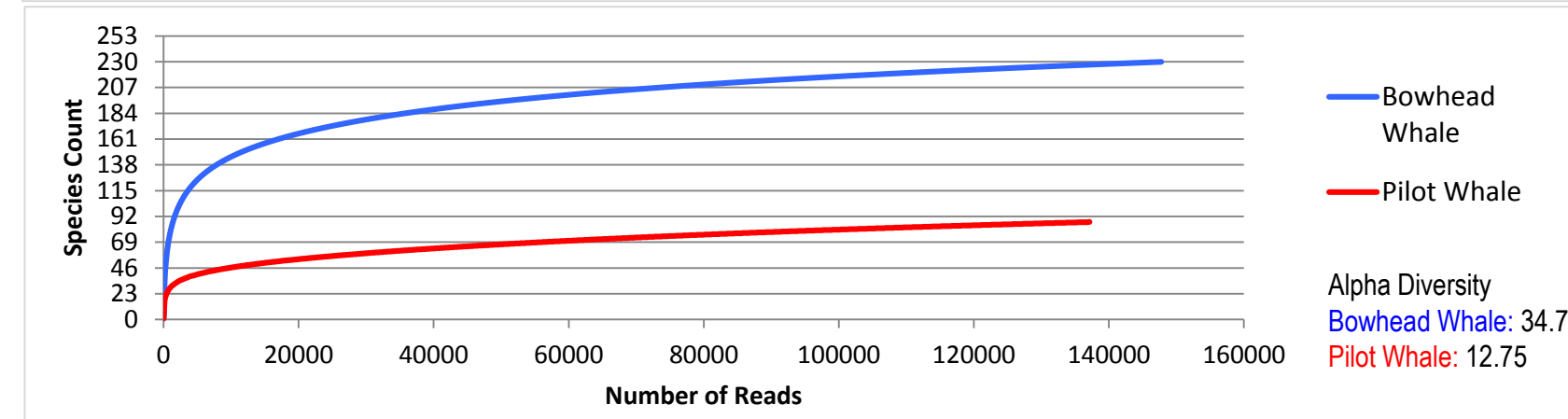
- Qiime, an open-source bioinformatics pipeline for performing microbiome analysis from raw DNA sequencing data, was used to explore taxonomic abundance, determine  $\alpha$  and  $\beta$  diversity of samples, and perform principal coordinate analysis (PCoA).
- MG-RAST, an automated analysis platform for metagenomes that provides insights into microbial populations based on sequence data, was used to annotate Illumina sequence reads and explore the functional potential of the whale gut microbiome.

## Results and Discussion

### Taxonomic Abundance

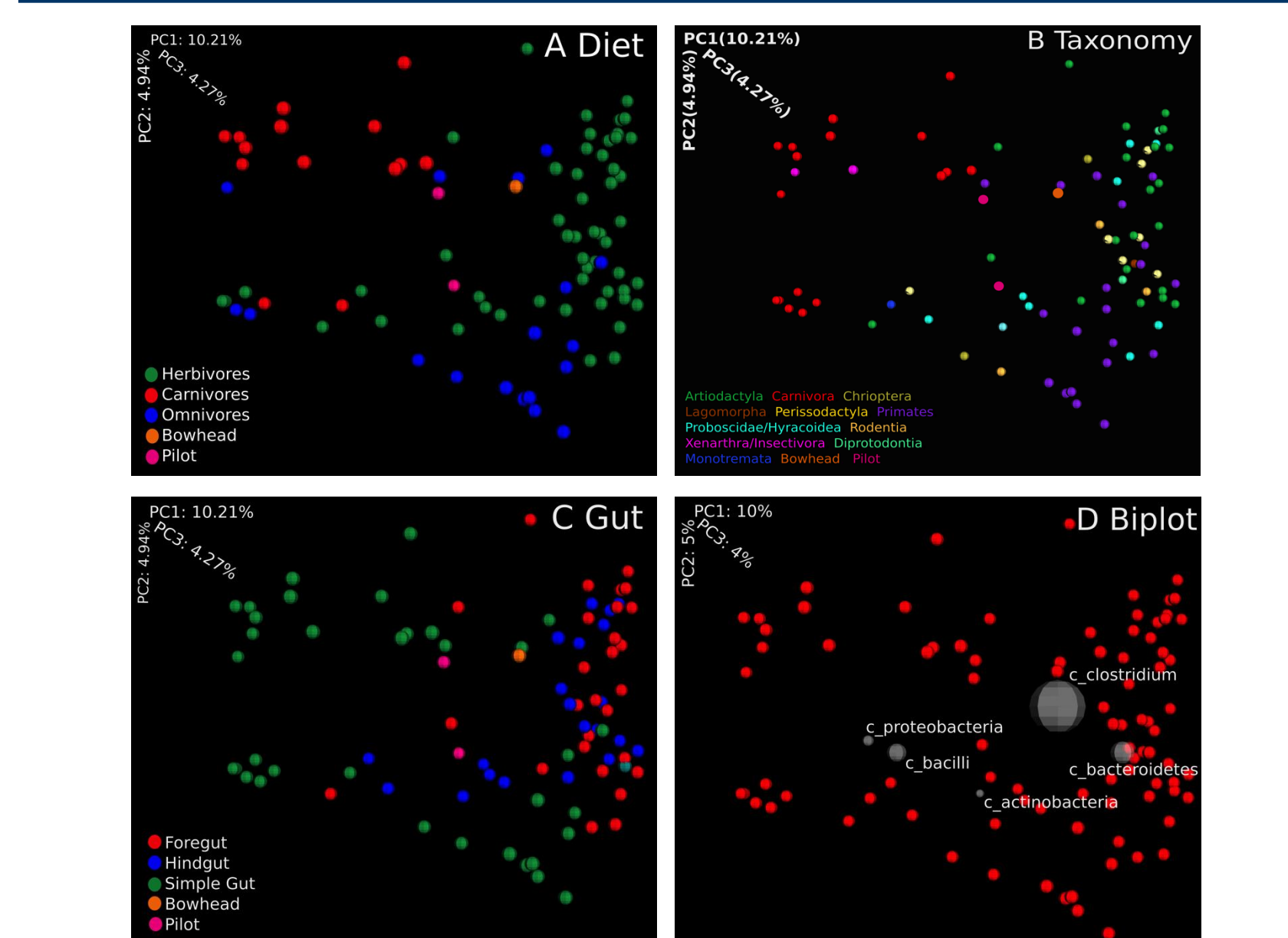


**Figure 1. Taxonomic Abundance.** Percentage of 16S rRNA sequences from each fecal sample assigned to different bacterial phyla. Bowhead and pilot whales (V3 and V6 reads, highlighted) are compared to terrestrial mammal species (Ley et al. 2008).



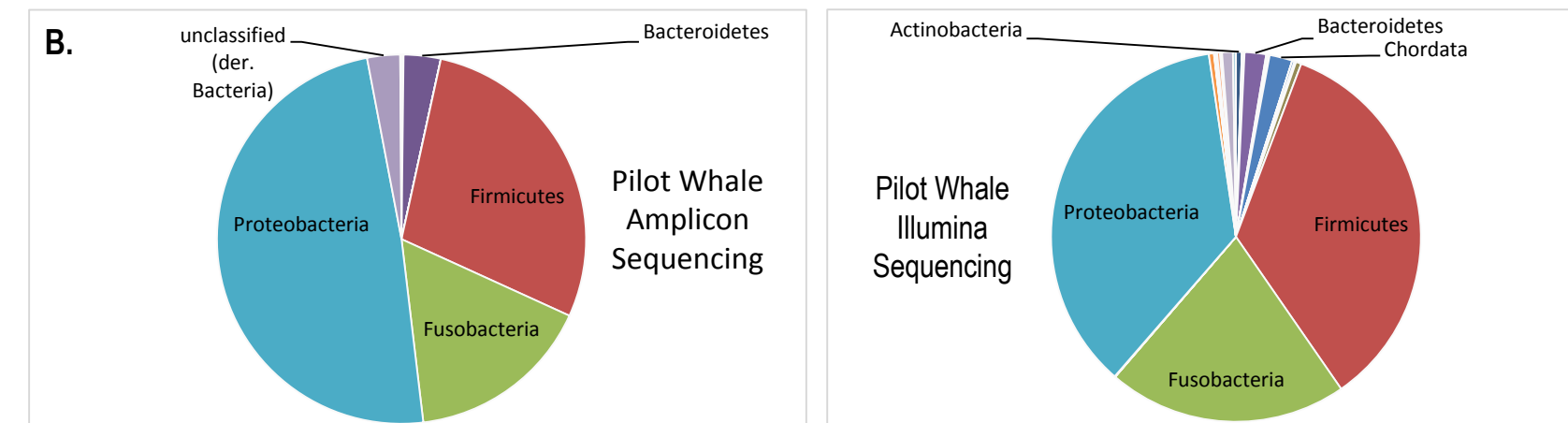
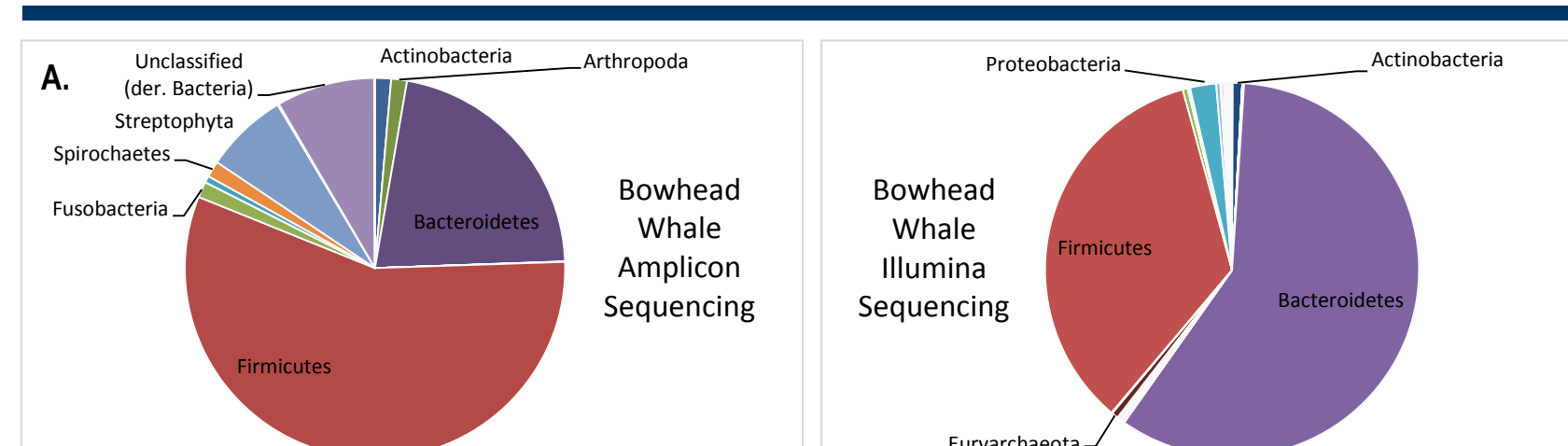
**Figure 2. Rarefaction Plot and Alpha Diversity** is used to determine whether all species in an ecosystem have been observed. The plot shows the total number of distinct species annotations as a function of the number of sequences sampled. The alpha diversity estimate, computed from the antlog of the Shannon diversity metric, summarizes the distribution of species level annotations in each sample.

- The taxon abundance at the phyla level of the pilot whale differs from the majority of the terrestrial mammals in its relatively high concentrations of proteobacteria (~50%) and fusobacteria (~20%). Other mammals with similar compositions included the polar bear (PB1), red panda (RP, RPSD), mountain zebra (AFZEB), african elephant (AFYEL), and flying fox (FF).
- The bowhead whale shows a high degree of similarity to the majority of the terrestrial mammals, potentially representing a "typical" mammalian microbiome, with major abundance of bacteroidetes (over 30%) and firmicutes (over 70%).
- Combined for the 61 mammalian species, a total of 41 bacterial phyla were detected, with a total of 36 detected between the two whale species.
- The increasing, non convergent rarefaction curves of the two whales indicates a fraction of the species diversity remains to be discovered. The alpha diversity indicates that the gut microbiome of the bowhead whale is more diverse than that of the pilot whale.



**Figure 3. Beta Diversity** represents explicit comparison of microbial communities based on their composition. Mammalian fecal communities clustered using PCoA from the unweighted Unifrac metric. PC1, PC2, and PC3 are plotted on the x, y, and z axes, respectively. Each circle represents a fecal sample, which are organized by (A) Diet, (B) Taxonomy, and (C) Gut Type. Plot (D) is a 3D Biplot exhibiting the relative abundances of the five most abundant classes of bacteria present among all samples. Results from bowhead and pilot whale V3 and V6 reads are represented by orange and pink colors.

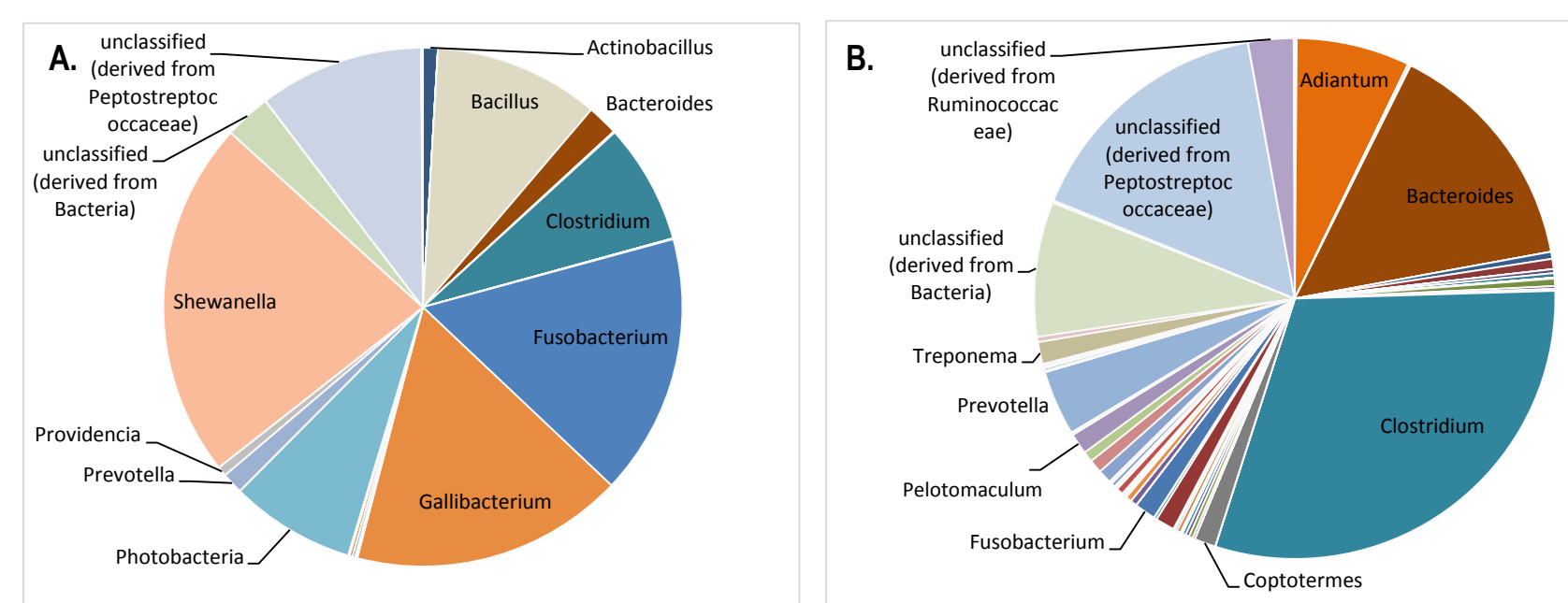
- Unifrac distances within diet classes are smaller than those between diet classes. Bowhead and pilot whales clustered closest with carnivores, exhibiting lower unifrac distances between the two, indicating a higher degree of similarity in their gut microbial community.
- Clustering was more apparent between foregut and hindgut fermentation classes than with simple gut fermentation. Unifrac distance between the bowhead and pilot whales was lowest between themselves and hindgut and foregut fermenters. This clustering is correlated to similarities in the gut microbial community including a shared abundance of clostridium, proteobacteria, and bacilli.
- Plots according to taxonomy showed more significant clustering within mammalian orders than between them. Clustering between the two whales was relatively high, and therefore consistent with the clustering of the rest of mammalian orders. Bowhead and Pilot whales, of the cetacean order, exhibited lower unifrac distances between themselves and the majority of artiodactyla and primates than the majority of any other order.
- The size of the sphere representing a taxon was proportional to its mean relative abundance across all samples.



**Figure 4. Taxonomic annotations** at the phylum level from amplicon and Illumina sequencing for A. bowhead and B. pilot whale samples.

A: amplicon I: Illumina	Firmicutes	Bacteroidetes	Fusobacteria	Proteobacteria
	Clostridium	Bacteroides, Prevotella	Fusobacteria	Photobacterium, Gallibacterium, Pasteurella, Haemophilus
<b>Bowhead</b>	A(57%) (43%)	A(22%) (59%)	A(1%) (0%)	A(1%) (2%)
Amplicon	30.4%	14.8%	4%	1.3%
Illumina	14.8%	40.2%	10.1%	1.3%
<b>Pilot</b>	A(28%) (43%)	A(3%) (2%)	A(16%) (21%)	A(49%) (36%)
Amplicon	7.5%	1.9%	1.3%	16.3%
Illumina	22.6%	0.8%	0.2%	15.7%

- Comparison of taxonomic diversity from amplicon vs. shotgun sequencing is used to explore possible bias introduced by primer specific (amplicon) methods.
- Although the overall phyla representation was comparable between the two methods of sequencing, some major differences were observed.
- In the bowhead whale, firmicutes constituted the most abundant amplicon reads (over 57%) while it was represented by about 35% in the genomic reads. Bacteroidetes represented ~59% of the genomic reads, but was found among 22% of amplicon reads.



**Figure 5. Taxonomic annotations** at the genus level from amplicon sequencing for A. pilot and B. bowhead whale.

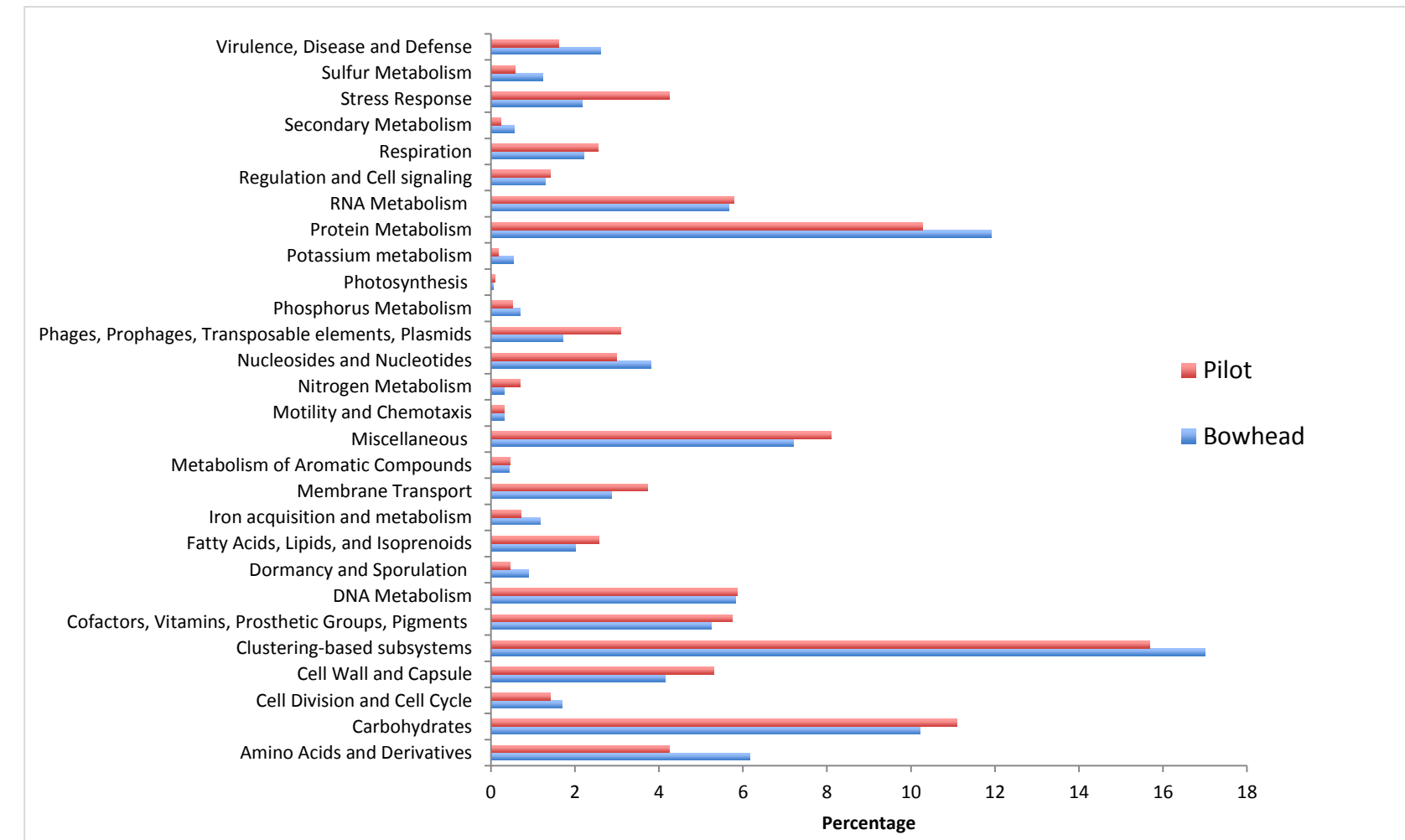
	Firmicutes	Bacteroidetes	Spirochaetes
	Clostridium	Bacteroides, Prevotella	Treponema
<b>Bowhead</b>	30.4%	14.8%	4%
<b>Pilot</b>	7.5%	1.9%	1.3%

	Fusobacteria	Proteobacteria
	Fusobacterium, Gallibacterium, Pasteurella, Haemophilus, Photobacterium	
<b>Bowhead</b>	1.3%	<0.1%
<b>Pilot</b>	16.3%	17%

- Bacteria abundant or commonly found in the GI tract of other vertebrates, including firmicutes, bacteroidetes, and fusobacteria (the latter at low frequency in bowhead whale), were also found in the both whales; however, at the genus level anaerobacter, cetobacterium, oscillibacter, sporacetigenum, and sarcina had negligible contributions to the overall whale gut microbiota suggesting an environmental influence on the whale gut microbiome.
- An abundance of bacteria not commonly found in the GI tract of vertebrates from the class Pasteurellales (primary or opportunistic pathogens), and *Photobacteria* and *Vibrio* were found in the pilot whale, but not the bowhead whale, suggesting a correlation between their gut microbiomes and different species of prey.
- *Photobacteria* and *Vibrio* are bacteria commonly found to maintain a symbiotic relationship with squid, which contributes over 90% of the pilot whale's diet. Illumina shotgun sequencing (below) can provide information on whether these bacteria, not found in the bowhead whale, have some functional potential.
- The pilot whale also displayed a greater average abundance of *Fusobacterium* than the bowhead whale and its terrestrial mammal counterparts (Fig. 1).

## Functional Potential



**Figure 6. Contribution of functional categories** of annotated and assembled Illumina sequences reads.

- Bias in assembled Illumina reads is accounted for by using raw values.
- Clustering-based subsystems (one in which there is functional coupling evidence that genes belong together, but function is unknown) account for the highest section of annotated proteins.
- The majority of the functional categories contributed equally to overall annotated proteins in both whales suggesting diverse bacteria filling similar or the same niches in the gut microbiome of both whales.
- Further breakdown indicated negligible amounts of methanogens (high in terrestrial ruminants), and high amounts of chitinases and trehalases used in the digestion of krill, shrimp, and squid.

- Bowhead whales feed at all levels of the water column, engulfing large volumes of water, including prey that are trapped on the baleen as water flows into the mouth. The preferred prey are copepods, euphausiids, and shrimp.
- Pilot whales forage at a wide variety of water depths. Although small fish and crustaceans are relatively common, its principal food source is cephalopods, more specifically luminescent squid.

Chitin Breakdown	Pilot Whale	Bowhead Whale
Chitinase A	<i>Vibrio shilonii</i> AK1 (chitinivorous)	<i>Clostridium botulinum</i> B str. Eklund 17B
Chitinase B		<i>Clostridium botulinum</i> B str. Eklund 17B
Predicted xylanase/chitin deacetylase		<i>Clostridium acetobutylicum</i> ATCC 824
Putative chitinase		<i>Clostridium botulinum</i> A str. Hall
Xylanase/chitin deacetylase, NodB family		<i>Clostridium acetobutylicum</i> ATCC 824
Trehalose Breakdown	Pilot Whale	Bowhead Whale
Alpha, alpha-phosphotrehalase	<i>Brachyspira murdochii</i> DSM 12563	<i>Clostridium botulinum</i> E1
GntR family transcriptional repressor of trehalose operon		<i>Bacillus pseudofirmus</i> OF4
PTS system, trehalose-specific HBC component		<i>Clostridium perfringens</i> NCTC 8239
Lipid Breakdown	Pilot Whale	Bowhead Whale
putative lipase	<i>Clostridium perfringens</i> ATCC 13124	
Lipase		<i>Clostridium tetani</i> E88
Protein Breakdown	Pilot Whale	Bowhead Whale
ATP-dependent protease La	<i>Haemophilus influenzae</i> RdAW	<i>Clostridium botulinum</i> F str. 230613
		<i>Bacteroides vulgatus</i> ATCC 8482
		<i>Clostridium difficile</i> NAP07
		<i>Prevotella marshallii</i> DSM 16973
Protease	<i>Providencia rustigianii</i> DSM 4541	<i>Bacteroides finegoldii</i> DSM 17565
		<i>Clostridium difficile</i> NAP07
		<i>Clostridium symbiosum</i> WAL-14673
Carboxyl-terminal protease	<i>Clostridium perfringens</i> SM101	<i>Clostridium acetobutylicum</i> DSM 1731
ATP-dependent Clp protease proteolytic subunit ClpP	<i>Haemophilus influenzae</i> NT127	<i>Paolubacter propionigenes</i> WB4
ATP-dependent Clp protease proteolytic subunit	<i>Haemophilus influenzae</i> NT127	<i>Prevotella marshallii</i> DSM 16973
ATP-dependent Clp protease ATP-binding subunit ClpX	<i>Haemophilus somnus</i> 129PT	<i>Psychrobacter cryohaloleptis</i> K5
Aminopeptidase I zinc metalloprotease	<i>Fusobacterium mortiferum</i> ATCC 981	<i>Turicibacter sanguinis</i> PC909

**Figure 7. Functional annotations of proteins** with their associated microorganisms.

- Bias incorporated in the functional analysis from assembled sequences may skew quantitative representations of bacterial assignments, however, significance in the difference in bacteria assigned to functions for each whale is pertinent.
- Proteins and enzymes are not equally represented in both whales due to this bias and natural variation in their gut microbiomes.
- Functional annotations of genomic sequences suggest different bacteria in both whales carry out shared functions between the two.
- Genera mostly present in the pilot whale, e.g., *Haemophilus*, *Fusobacterium*, and *Vibrio*, contribute to the same functions carried out in both whales, but are barely present in the bowhead whale.
- *Vibrio* and *Photobacterium*, who live in symbiotic relationships with squid, may have been assimilated into metabolic processes within the pilot whale over time.
- The whale samples contained high amounts of enzymes important to the breakdown of chitin and trehalose, present in high abundance in krill and squid.

## Conclusion

- A combination of 16S rRNA amplicon and genomic shotgun Illumina sequencing provides a more complete analysis of the taxonomic abundance and functional potential of microbiota.
- The bowhead whale shows a high degree of similarity to the majority of terrestrial mammals representing a "typical" mammalian microbiome.
- The pilot whale contains a significantly greater average abundance of fusobacteria and proteobacteria than its terrestrial mammalian counterparts.
- Assimilation of *Vibrio* and *Photobacterium*, possibly acquired from squid, into metabolic processes is present in the pilot whale.
- Functional annotations of genomic sequences suggest diverse bacteria filling similar niches in the gut microbiome of both whales.
- More accurate quantitative representation of functional potential will be achieved by analyzing raw Illumina sequence reads.
- Gut microbiomes of terrestrial and marine mammals are driven by diet, gut type, and taxonomy.
- The whale gut microbiome is a product of evolutionary origin (terrestrial mammalian elements) and influences from its marine environment.

## References

Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., Pena A.G., Goodrich J.K., Gordon J.I. et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods*, 7, 335-336.

Desportes, G. and R. Mouritsen. 1993. Preliminary results on the diet of long-finned pilot whales off the Faroe Islands. In: *Biology of Northern Hemisphere pilot whales* (eds. G.P. Donovan, C.H. Lockyer, A.R. Martin), pp. 305-324. Report of the International Commission Special Issue 14, Cambridge, UK.

Gill, S. R., M. Pop, R. T. Deboy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett, and K. E. Nelson. 2006. Metagenomic analysis of the human distal gut microbiome. *Science* 312:1355-9.

Ley RE1, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Birchler JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. Evolution of mammals and their gut microbes. *Science* 320, 1647 (2008)

Ley, Ruth E. et al. "Worlds within Worlds: Evolution of the Vertebrate Gut Microbiota." *Nature reviews. Microbiology* 6.10 (2008): 776-788. *PMC*. Web. 6 Apr. 2015.

Lowry, L.F. 1993. Foods and feeding ecology. In: *The bowhead whale* (J.J. Burns, J.J. Montague, C.J. Cowles, eds.), Pp. 201-238. Society for Marine Mammalogy, Lawrence, Kansas.

Lowry, L.F., G. Sheffield, G. George, J.C. 2004. Bowhead whale feeding in the Alaskan Beaufort Sea based on stomach content analyses. *Jour. Cetacean Res. Management* 6: 215-223.

Meyer, F., Paarmann D., D'Souza M., Olson R., Glass E.M., Kubal M., Paczian T., Rodriguez A., Stevens R., Wilke A. et al. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 9, 386.