Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 28:1-12
Copyright © September 2023, M&I UBC

Spatial region of habitat drives fish gut microbial diversity, composition, and functionality

Jordon Chen, Vivian Yan, Christopher Yap, Fione Yip

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY The gut microbiome in fish aids in nutrient distribution by mediating the breakdown of food, nutrient absorption, energy homeostasis and waste excretion. The composition of the gut microbiome is largely influenced by the location of the fish as prey consumption and nutrient availability differs from the coast to the open and deep ocean regions. However, there are limited large-scale studies exploring how the gut microbiome differs across ocean regions, and the functional involvement of the microbiome from an inferred metagenomic perspective. Using a dataset of 101 fish species, we investigated the effects of spatial region habitat on the internal and external fish microbiome. We found that spatial region is a driver of internal organ, but not external organ diversity. Only open ocean internal samples contain a core microbiome of six taxa, including representatives from the genera: *Synechococcus* CC9902, *Psychromonas*, *Acinetobacter*, and Sva0081 Sediment Group. We also identified differing abundances of metabolic pathways between the samples, including pathways involved in xenobiotic detoxification, glycosaminoglycan degradation, ABC transporters, and the renin-angiotensin system. Our findings demonstrate that spatial region is a driver of fish gut microbial diversity, composition, and functionality.

INTRODUCTION

The fish microbiome plays an extensive role in maintaining fish health and digestion. The external organ fish microbiome (skin and gill) consists of a complex pattern of gene expression that maintains the epithelial barrier and appropriate innate immune responses to protect the host (1, 2). Mucus present on skin and gill surfaces are colonised by a highly diverse commensal microbial community (3). The microbiome limits the colonisation of pathogenic microorganisms and prevents disease development by competitive exclusion, releasing compounds that impede the growth of pathogenic microorganisms, facilitating waste excretion, and mucus homeostasis (2). Moreover, the gut microbiome influences digestion and absorption of nutrients, as well as nutrient distribution and energy homeostasis (4). Thus, there is an important relationship between the fish microbiome and the host that will likely differ depending on the fish's habitat.

Some factors that have been previously described in the literature as drivers of both gut and external organ microbiome diversity in fish are water quality, disease, and diet. The focus in current research efforts has largely been on the taxonomic diversity across different fish species and their body surfaces (5). However, there is little research on the functional involvement of the microbiome from a metabolomic perspective (5). Research in Hadal snailfish, which inhabit depths between 6,000 to 8,000 m, has shown that the gut microbiome consists primarily of piezophiles, which are highly adapted to growth under high hydrostatic pressure. They also found genera including *Psychromonas, Moritella,* and *Shewanella* associated with snailfish and hypothesised they contribute to chitin degradation and fatty acid production (6). Thus, the fish microbiome is likely different depending on which part of the ocean the fish is located in as the microbial species would be adapted to different nutrient sources and environmental conditions relevant to that ocean region.

Minich *et al*. sampled the microbiome of gill, skin, midgut, and hindgut of 101 species of fish prevalent at various habitat depths. They compared the alpha diversity of samples from each body site and looked at other predictors including habitat depth, collection substrate, dorsal length, fish mass, and swim modes (7). Their study found that the midgut had the

Published Online: September 2023

Citation: Chen, Yan, Yap, Yip. 2023. Spatial region of habitat drives fish gut microbial diversity, composition, and functionality. UJEMI 28:1-12

Editor: Lauren Pugsley and Shruti Sandilya, University of British Columbia

Copyright: © 2023 Undergraduate Journal of Experimental Microbiology and Immunology. All Rights Reserved.

Address correspondence to: https://jemi.microbiology.ubc.ca/

highest diversity while the gills had the lowest diversity (7). They also compared the beta diversity of the fish microbiome of different body samples and found that sample type, habitat depth, and substrata group were strong predictors of beta diversity (7). Since their dataset includes body part samples from different ocean regions, there is an opportunity to investigate how these two factors interact and drive differences in fish microbiota.

We will expand on their findings by grouping fish based on their natural habitat of either coastal or open oceans to examine how the fish microbiome compares when living near the coast and in the open. The microbiome samples will also be grouped into external or internal organ samples to look for general trends, rather than comparing each anatomical region separately. Based on previous findings, we hypothesise that microbial diversity will differ between the two anatomical locations and spatial regions, respectively. Additionally, ocean salinity is variably distributed horizontally and vertically, and studies have shown that it plays a role in shaping the fish gut microbiome (8). Knowing that oceanic regions have heterogeneous salinities that impact marine species distribution and thus ecology, we can hypothesise that the external and internal anatomical fish microbiomes will differ between ocean spatial regions. Since different adaptations allow fish to survive in a particular habitat, we hypothesise that open ocean microbiomes will express more genes that allow the microbes to survive in high pressures, extreme temperatures, unique gas composition, and limited nutrients of the open ocean (9).

METHODS AND MATERIALS

Fish Microbiome Project Dataset. The dataset consists of one gill, skin, midgut, and hindgut samples from a single fish from each of 101 fish species collected near San Diego County, California, USA (7). Skin samples were composed of scraped mucus. Midgut and hindgut samples were digesta material located at the end of the stomach or at the anus, respectively. Gill samples were composed of either the entire gill or three cuts from the top, middle, and bottom of the gill arch if the gill sample was too large. Depending on the vertical range that the fishes inhabit, habitat depth was categorised as intertidal (0-10 m), neritic (0-200 m), mesopelagic (200-1000m), mesopelagic/benthopelagic (largely demersal at 200-1000m), bathypelagic (1000-4000 m) or abyssopelagic (>4000m). Further information on the sample collection and creation of the metadata can be found in the original paper by Minich *et al.* (7). This dataset is publicly available at the European Nucleotide Archive (project number: PRJEB54736) and on Qiita (Qiita ID: 13414).

Preliminary Metadata Filtering and Binning. Metadata processing prior to downstream analysis was performed using the tidyr and dplyr packages in R (Version 4.2.3) with RStudio (Version 2023.03.0) (10–12). Rows containing "NA" or "not applicable" values were first filtered out, reducing the number of samples from 651 to 412. The sample type column (sample_type) was binned into the first new column (anatomical_location) as either "external", consisting of skin and gill samples, or "internal", consisting of midgut and hindgut samples. The fish habitat column (habitat_depth_level3) was binned into a second new column (spatial region) as either "coastal", consisting of neritic and intertidal fish samples, or "open ocean", consisting of the rest. Finally, a third new column (anat_space_combine) was created by combining the values for the first and second new columns. For further QIIME2 analysis, the manifest file was updated to match samples in filtered metadata.

QIIME2 Sequence Preparation and Export. Sequence data was imported into QIIME2 as single end sequences (13). The DADA2 QIIME2 tool was then used to denoise and apply quality control to sequences, truncating reads to a length of 189 base pairs (14). Potential mitochondrial and chloroplast sequences were filtered out of the data using QIIME2's taxa filter table command, and amplicon sequence variants (ASVs) with less than 50 total counts were removed. Samples with less than 10 ASVs were also filtered out. Remaining ASVs were assigned taxonomy using the SILVA 138 SSU Ref NR 99 database, re-trained with the primers (Forward: GTGYCAGCMGCCGCGGTAA, Reverse: GGACTACNVGGGTWTCTAAT) used in the study (15, 16). A rooted phylogeny tree was

built with FastTree following the alignment of representative sequences with MAFFT using the QIIME2 align-to-tree-mafft-fasttree command (17, 18). Finally, a rarefaction depth of 1552 was selected based on alpha rarefaction plots. Upon filtering the data, the metadata file, ASV table, rooted phylogeny tree, and taxonomy table were then imported to R to be converted to a phyloseq object using the phyloseq package (19).

Diversity Analyses. Alpha diversity of the spatial region and anatomical location categories was assessed by generating Shannon-Wiener index plots in R, using the phyloseq object and ggplot2 and phyloseq packages (20–22). The Wilcoxon rank sum test was performed to test for significance using the ggpubr package (23). The phyloseq package was used in R to estimate beta diversity using Weighted UniFrac distances calculated on the phyloseq object. PCoA was then calculated and plotted using the phyloseq and ggplot2 packages. To test for significance, the vegan package was used to perform PERMANOVAs (24).

Core microbiome analysis. Core microbiome analysis was performed using the microbiome package and tidyverse package in R (25, 26). Core microbiome was calculated using a prevalence threshold of 0.5 and detection threshold of 0.01, and visualised as a venn diagram using the ggVennDiagram package (27).

DESeq2. Differential taxa abundance analysis between open ocean internal and coastal internal samples was performed using the DESeq2 package and tidyverse package, with one added to the ASV table, and visualised as a $\log_2$-fold change barplot with the ggplot2 package in R (21, 25, 28).

PICRUSt2 Functional Prediction. Functional abundances were predicted for representative ASV sequences using tools wrapped within the PICRUSt2 software (29). HMMER and SEPP tools optimally placed ASVs into a phylogeny tree based on sequence similarity to reference 16S sequences. SEPP was used as a low-memory alternative to EPA-ng, outputting a tree with incorporated ASVs (30). The castor R package was applied to the tree to infer genomic content of sample sequences, predicting Kegg Ortholog (KO) abundances for each ASV (31). Edge lengths were set to 0 such that they did not influence predictions. Finally, the metagenomes of each sample were predicted.

Pathway Inference and Differential Expression Analysis. Predicted metagenome data was imported into R and downstream analysis were conducted using the ggpicrust2 package (32). Pathway abundances were first inferred by converting KO abundances to KEGG pathway abundances. KEGG pathway abundance and metadata were then filtered such that samples aligned prior to differential expression analysis with the limma voom tool (33–35). Limma voom was chosen to determine if microbial communities in open ocean samples differed functionally from coastal communities as it allows for powerful analyses of RNA-seq data by providing strict false discovery rate control. Subsequent results were filtered for the 30 lowest Benjamin-Hochberg adjusted p-values, and annotated by connecting to the online KEGG database (36–38). Results were visualised in a grouped bar chart. To exclusively compare internal samples across oceanic spatial regions, the metadata was subsetted for internal samples only.

Code Availability. All R (version 4.2.3) code is available at:
https://github.com/jochennn/micb475_project2

RESULTS

Only open ocean internal microbiomes show significantly greater diversity. Alpha diversity metrics on anatomical sample location and spatial region was conducted using the Shannon diversity index. Internal and open ocean samples were more diverse than external ($p = 0.0074$, Wilcoxon Rank Sum test) and coastal samples ($p = 0.0012$, Wilcoxon Rank Sum test), respectively (Figure 1A & B). To identify potential interactions between categories, we faceted the anatomical location graph with spatial region, and vice versa. In the anatomical

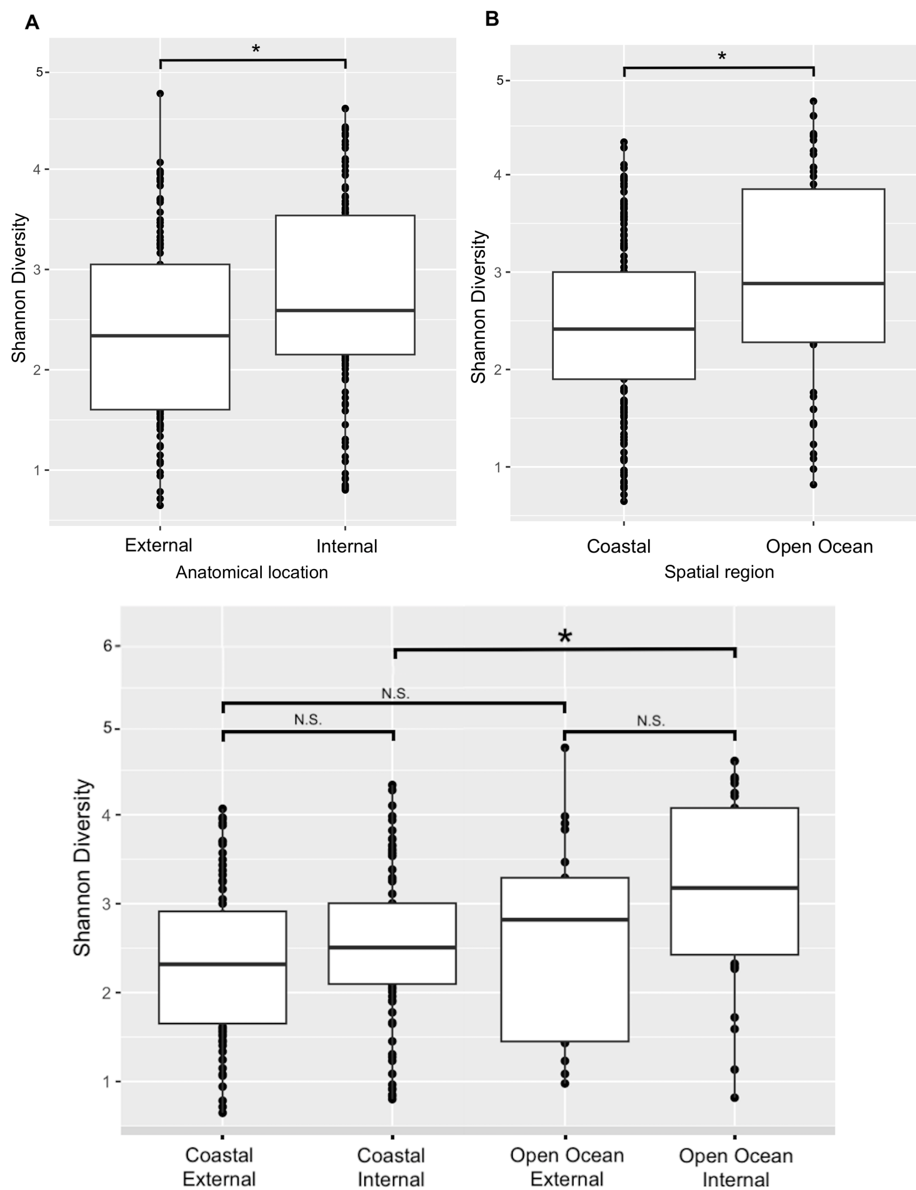


FIG. 1 Open ocean and internal microbiomes have higher alpha diversity. (**A)** Internal samples had significantly higher Shannon diversity compared to external samples ($P = 0.0074$, Wilcoxon Rank Sum test). (**B)** Open ocean samples had significantly higher Shannon diversity compared to coastal samples ($P = 0.0012$, Wilcoxon Rank Sum test). **(C)** Only internal samples had significant differences in alpha diversity between the coastal and open ocean regions. Asterisks indicate significant differences between sample categories calculated with Wilcoxson rank sum tests with Bonferroni correction. $*P < 0.05$, $P_{adjusted} < 0.013$.

location graph faceted by spatial region, we see no statistically significant differences in alpha diversity when comparing coastal or open ocean samples with internal and external samples ($p = 0.084$; $p = 0.067$, Wilcoxon Rank Sum test) (Figure S1). However, when faceting the spatial region graph by anatomical location, open ocean samples were more diverse than coastal samples for internal samples ($p = 0.0012$, Wilcoxon Rank Sum test) (Figure 1C). External sample diversity did not change between spatial regions. To examine whether anatomical location or spatial region are predictors of microbial community composition, we calculated beta diversity using Weighted UniFrac distances. We observed clustering of anatomical samples into 2 clusters, external and internal samples, while no clustering was observed for spatial region ($p = 0.030$; $p = 0.138$, PERMANOVA) (Figure S2A). To assess whether patterns existed on finer resolutions, we repeated the analysis using the original

sample type and habitat depth levels as predictors, which included gill, skin, midgut, and hindgut for different sample types from the intertidal, neritic, abyssalpelagic, meso/benthopelagic and bathypelagic regions (Figure S2B). We found that anatomical location (pseudo-F = 2.7509, $p = 0.030$) and its strata (pseudo-F = 2.6302, $p = 0.007$) were strong predictors while spatial region (pseudo-F = 1.7049, $p = 0.138$) and its strata (pseudo-F = 1.6112, $p = 0.095$) were not (Figure S2C). Overall, while the open ocean internal samples were more diverse, there was no significant difference in microbial community composition between open ocean and coastal samples.

A small core microbiome is associated with open ocean internal samples. Next, to further characterise the differences between open ocean external, open ocean internal, coastal external, and coastal internal groups, we decided to see how the composition between coastal and open ocean internal microbiomes differ. We filtered samples to include core members only, which we defined as having a prevalence of at least 0.5 and a detection threshold of 0.01, and saw no core microbiome for the coastal internal samples. However, we identified a small core microbiome for the open ocean internal group, which consisted of six taxa (Figure 2). Four of the six taxa belonged to the genera: *Synechococcus* CC9902, *Psychromonas*, *Acinetobacter*, and Sva0081 Sediment Group. Using DESeq2, we then confirmed that *Psychromonas*, *Acinetobacter*, and the Sva0081 Sediment Group were differentially abundant between open ocean internal and coastal internal groups (Figure S3), validating the results from the core microbiome analysis.

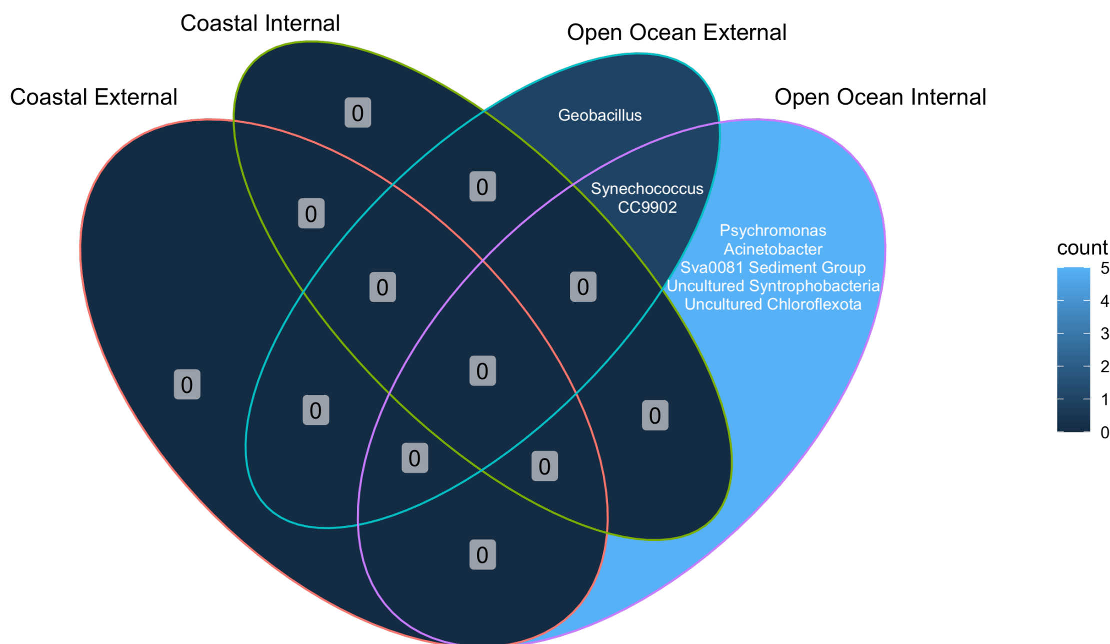
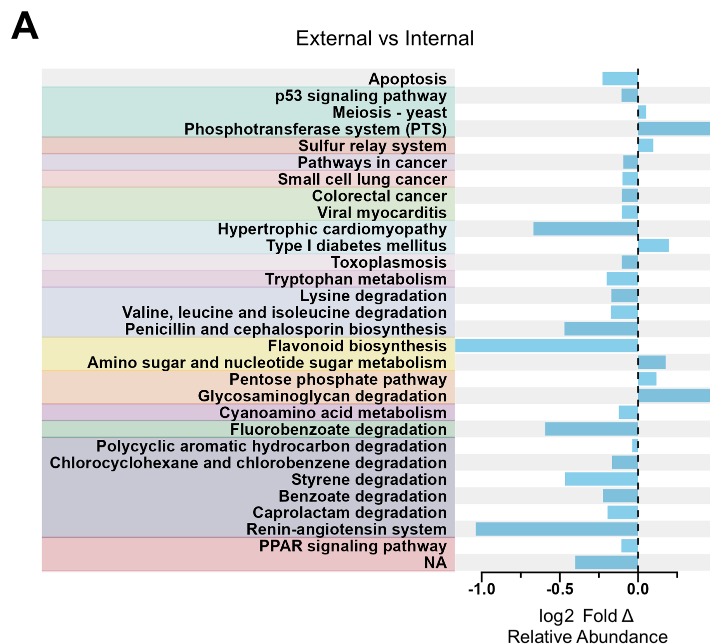


FIG. 2 Six taxa associated with the open ocean internal core microbiome. Four taxa in the open ocean internal microbiome are resolved to the genus level: *Psychromonas*, *Acinetobacter*, and Sva0081 Sediment Group, with one shared taxon: *Synechococcus* CC9902, part of both the open ocean internal and external core microbiomes. No core microbiome of the coastal samples was seen. Prevalence threshold of 0.5 and detection threshold of 0.01.

Significantly enriched catabolic and anabolic pathways are primarily found in the internal fish microbiome. To further understand how microbial communities differ between anatomical regions, we identified differences in the abundance of metabolic pathways to infer functional differences. Using the ggpicrust2 package in R, we converted KO abundances, inferred with PICRUSt2, to KEGG pathway abundances. Upon conducting a limma voom differential expression analysis, we found statistically significant differences in relative pathway abundances between the external and internal fish microbiomes in our study. We decided to focus only on the 30 pathways with the lowest Benjamin-Hochberg adjusted p-values, which were all significantly different, and found catabolic and anabolic pathways to

be primarily enriched in the internal microbial community (Figure 3A). Of the metabolic pathways retained, it was interesting to note that the glycosaminoglycan degradation and amino sugar and nucleotide sugar metabolism pathways were enriched in the external microbial community relative to the internal microbial community (Figure 3A).

ABC transporter pathway significantly enriched in the internal microbiome of coastal fish compared to open ocean fish. Lastly, we wanted to examine whether the internal microbiome of open ocean fish was functionally different from that of coastal fish. Once again, we performed a limma voom differential expression analysis on KEGG pathway abundances, this time for internal samples only. Of the top 30 KEGG pathways with the lowest adjusted p-values, we only observed the ABC transporter pathway to have marginally higher relative abundance in the internal coastal microbial community, while the rest showed increased relative abundance in the open ocean microbiome (Figure 3B).



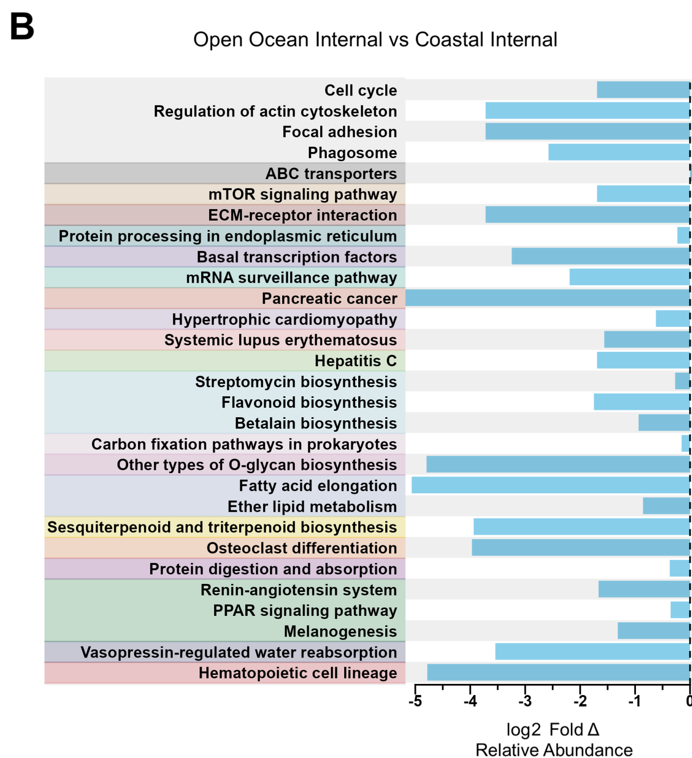


FIG. 3 Significant differences present in relative pathway abundances when comparing external and internal fish microbial communities, and in spatially distinct internal fish microbiome samples. Limma voom differential expression analysis shows log2 fold-change of relative abundance for KEGG pathways with 30 lowest Benjamin-Hochberg adjusted p-values. **(A)** Catabolic and anabolic pathways were primarily decreased in external microbial communities in comparison to internal communities. (**B**) For internal microbiome comparisons across oceanic spatial regions, the open ocean internal samples acted as a baseline. Examination of the grouped bar chart revealed that the ABC transporter pathway was decreased in the internal microbiome of coastal fish. All pathways shown above are significantly different between the compared groups. *$P<0.05$

DISCUSSION

This study expands upon previous research by Minich *et al.* to elucidate the differences between microbiomes in coastal fish species versus offshore fish species. In assessing diversity, core taxa, as well as potential functional differences between taxa, between the internal microbiomes of open ocean and coastal fish, we found that only internal microbiota differed between open ocean and coastal fish. In spite of combining water column depth levels, we saw a trend of differing diversities between anatomical sites sampled, but not between the spatial regions the fish were caught at, highlighting a potential area of further research in microbiomes shared across fish species with different lifestyles.

We first validated whether our binning method, which reduced anatomical location to inner and outer fish, yielded similar results as the preliminary study by Minich *et al*. (7). Minich *et al*. found that midgut samples had the highest diversity across several metrics, relative to gill, skin, and hindgut (7). We also observed greater alpha diversity in internal samples, which includes midgut samples, when comparing with external samples. The high alpha diversity in internal samples is corroborated by other studies, which describe the fish gut as highly dense in bacteria involved in fish development, nutrition, and disease prevention, in contrast to surrounding water (39, 40).

We also looked at how spatial regions affected alpha diversity and found that open ocean samples had higher alpha diversity compared to coastal samples, and that this pattern was driven by patterns in internal fish samples only, and not external. Specifically, internal open ocean samples had higher microbial richness than internal coastal samples, suggesting that, internal samples are sensitive to changes in spatial region. This observation is congruent with findings from Kim et al. who found that host habitat was a major determinant of the gut microbiome composition in fish (41). The researchers looked at how alpha diversity differed in fish from lakes, streams, and seas, all of which differ drastically in environmental conditions, including salinity, temperature, depth, and nutrient availability (41). They found that freshwater fish had significantly higher Shannon index and Observed Species (non-phylogenetic diversity) values compared to seawater fish, but Faith's phylogenetic diversity was comparable (41). Kim et al posited that these patterns were due to host adaptation to the environment, as the lack of cofactors and vitamins in freshwater habitats has been shown to promote the prevalence of diverse microbes involved in the metabolism of those nutrients (41). Since coastal waters are more proximal to river run-off, we might expect the internal coastal fish gut samples to have higher microbial richness compared to internal open ocean samples due to a reduction in water salinity (42). Further research needs to be done to investigate the reasons for differential fish internal microbial diversity, and to elucidate why external samples do not appear to have differences between the regions.

After conducting core microbiome analysis, we found four taxa unique to open ocean internal samples, that resolved to the genus level and were of interest. For example, *Psychromonas* was found in the open ocean internal samples and is a genus with chitin degradation capabilities (6). Due to chitin being a common polysaccharide component of marine crustacean shells, it may have a role in fish digestion. However, whether this genus has an active function in the fish gut has yet to be determined. *Acinetobacter* was shown to be a strong indicator species (data not shown) and the most differentially abundant taxon in DESeq2 of the open ocean internal, as well as part of the core microbiome. *Acinetobacter* is a very diverse genus of bacteria, ranging from human and fish pathogens to commensal organisms in the soil and water. We cannot be certain of its role in the fish gut as roles differ between *Acinetobacter* species (43). Finally, we see a Sva0081 Sediment Group, a genus of anaerobic sulphate reducing bacteria. When running indicator species analysis, we saw that it is an indicator species for both open ocean internal and external categories. Previously, the genus has been shown to have high abundance in sediments but here they have also been observed inside fish (44, 45). However, future studies would need to be performed to see if they are active in the fish gut or skin, or ingested. Additionally, there was one taxon shared between both the open ocean internal and external core microbiomes. *Synechococcus* CC9902, an abundant cyanobacteria. Most blooms of *Synechococcus* occur close to the coast (46). However, when stratifying the data, we saw that it was present in skin and gut samples of a variety of water column strata, including both coastal and open ocean categories.

Synechococcus CC9902 has been shown to alter fish behaviour, and major blooms have been observed off the coast of California, where many of the fish samples for our dataset were taken (46). Interestingly, we did not see the *Synechococcus* CC9902 as differentially abundant in DESeq2 or an indicator species.

Although Minich *et al.* identified that body site differences were the primary contributor to microbial community composition, they did not determine the specific pathways that were differentially abundant between the microbiomes in fish external anatomy and fish internal anatomy. Here we found that pathways associated with catabolism and anabolism were primarily enriched in the internal fish microbiome, and this included pathways involved in amino acid metabolism, as well as styrene, benzoate, and caprolactam degradation pathways. Succinctly, many of the pathways that were enriched in the internal community were associated with dietary processing or serve roles in xenobiotic detoxification and degradation, which is commonly performed by the gut microbiome (47, 48). The few metabolic pathways found to be enriched in the external microbiome included the glycosaminoglycan degradation and amino sugar and nucleotide sugar metabolism pathways. The glycosaminoglycan degradation pathway is known to be a mechanism present in gill microbiota that aids in salt water acclimation and osmoregulation (49). Accordingly, the amino sugar and nucleotide sugar metabolism pathway has also been shown to facilitate osmoregulation, where it is upregulated in the gill upon salinity transfer (50). This can explain the enrichment we see in the external community. Taken together, our findings corroborate the observations made by Minich *et al.*, which indicate that conserved aspects of body sites favour certain microbial communities.

Continuing with our exploration of differences across spatially segregated internal fish microbial communities, we performed an additional pathway prediction functional analysis. We observed that the ABC transporter pathway was the only pathway significantly enriched in the coastal internal community relative to the open ocean internal microbial community. While this pathway is involved in osmotic homeostasis and nutrient uptake, both of which are essential in prokaryotic species, the enrichment that we see in the internal coastal microbiome is likely due to the need for increased xenotoxin resistance and removal (51). As pollutant concentrations have been shown to be the highest in coastal waters, microbes present in these areas fittingly need more ABC transporters to expel foreign substances they encounter (52).

Other patterns observed in differentially abundant pathways corroborated previous work. Particularly, the streptomycin biosynthesis pathway was found to be enriched in open ocean internal samples. This supports literature showing that microbiomes of animals fed zinc restricted diets exhibited increased streptomycin biosynthesis (53). Considering that zinc concentrations are lower in the ocean as opposed to coastal waters, we might expect to see this pathway enriched in the open ocean (54). The renin-angiotensin system was another pathway that showed enrichment in the open ocean internal microbiome. This system performs functions associated with osmoregularity and saltwater acclimation to accommodate shifts in osmolality (55). Although this pathway has only been shown to be reduced in the gut microbiota of fish transferred from saltwater to freshwater, it is plausible that fish gut microbes present in oceanic regions of less salt concentration could exhibit a decreased abundance of the pathway (55). Taking into account that most of our samples were obtained off the coast of California, the differences in relative abundances of this pathway could therefore be attributed to the lesser salt concentrations of the waters along the west coast of North America in comparison to the open ocean (56).

Limitations There are some concerns with treating the open ocean as a single spatial region (57). With greater water column stratification offshore than coastal waters, the reduced mixing events occurring in deep waters may lead to differences in fish diets and lifestyles at different water column depths. However, when analysing the unbinned data (original water column strata), we saw similar results to when combining them. We hypothesise that since open water fish species are known to migrate vertically through the water column on a diurnal basis, there may be a high degree of homogeneity in microbiomes between the different depths (43).

Further, we found that differences in microbial diversity between groups was reduced when binning sample types together into anatomical locations, suggesting that the midgut and

hindgut samples, as well as skin and gill samples, are different from each other (Figure S1C), with a stronger difference than between open ocean and coastal fish. These results would match findings from Minich *et al.* where they found that the beta diversity of the hindgut and midgut of fish differed as fish increased in trophic level (7).

When looking at the taxa present, we were only able to resolve to the genus level. Many taxa are likely unstudied and uncharacterised. There are also likely various species under the same genus, such as within *Acinetobacter,* that were identified as core microbes in the open ocean internal microbiome but had different functions and niches (43). While the genus may be considered part of the core microbiome, if resolved to the species level, we may see otherwise. As well, when using 16S data, we are examining the microbes that are present and are unable to determine whether they are alive or dead, which limits our ability to draw conclusions about the ecological roles of taxa in the microbiome.

Similar limitations apply to the metabolic data from the PICRUSt2 analysis. It is important to note that PICRUSt2 only infers the gene family and pathway abundances from 16S sequences, and while this may be useful in forming hypotheses for future studies, the presence of such pathways and their ecological roles cannot be confirmed. As well, taxa in a microbiome does not always maintain the same functional microbiome as previous studies have shown (58). Predictions from the analysis are limited by the way our study sequences get placed in the reference tree. In other words, the process isn't entirely reproducible, as placement of the sequences are dependent on each other and whether or not they are based on single ASVs or functions. One important limitation of PICRUSt2 is that the reference genome we used may not fully represent the 16S sequences of microbes in the fish of Southern California coastal and oceanic waters, resulting in poor predictions of pathways.

Conclusions This study finds that spatial region is a driver of fish gut microbial diversity, composition, and functionality. Compared to coastal fish internal microbiomes, this study identifies how fish across a diverse range of species comprising the open ocean region share a small core microbiome and yet are more diverse than coastal fish microbiomes. Additionally, looking at inferred metagenomic data, we find novel differences between the metagenomic content in internal microbiomes of coastal and open ocean fish. Patterns found in our study can guide future studies in characterising the importance of the fish microbiome in fish health, and help lead fish management strategies and conservation efforts.

Future Directions To further explore how spatial regions impact microbial diversity, separating the spatial regions into their individual strata will give a more nuanced picture on how depth levels affect diversity. It may identify water depths that have particular environmental factors that influence fish microbiota, which may inform ocean preservation strategies. The balance of the gut microbial community is sensitive to factors like nutritional status, stress, antibiotics, and infection (39). While our study provides a snapshot of the level of diversity currently in Southern California, it will be important to monitor this population to potentially identify the effects of ocean pollution in the area, especially after having found metabolic pathways involved in xenobiotic detoxification, which serves as an environmental warning sign. The coast of Southern California was heavily used in the last century as a dumping ground for chemical waste, including dichlorodiphenyltrichloroethane (DDT), a now banned pesticide which remains on the ocean floor in dangerously high concentrations (59). Understanding how pollution and environmental changes affect diversity over time would help give us a better idea of how fish microbiomes are affected by said changes. Learning the warning indicators for concerning environmental changes can help prompt conservation efforts.

Additional studies that look at more fish species and an increased sampling size of more than one fish per species would allow for stronger characterization of microbiome differences between open ocean and coastal as well as internal and external microbiomes. This will help to confirm and increase confidence in our findings. A very specific scope that would be of interest to investigate is to target specific species that are found between water column regions, to further characterise a wider range of fish that are more migratory between habitat regions and perhaps allow for more stringent thresholds for core microbiome analysis. To

parallel our study, investigation of the metagenomics of the new dataset can help confirm our picrust analysis.

Finally, it will be meaningful to conduct a deeper dive on the specific taxa revealed to be core microbes in the different ocean habitats and determine whether they actually have an active function in the fish gut microbiome. It is unclear if the microbes we identified were alive and actively making significant contributions to metabolic processes in the fish gut.

ACKNOWLEDGEMENTS

The authors wish to express their appreciation to Dr. Evelyn Sun and Dr. Melissa Chen for their advice, support and time meeting to discuss the project, Dr. Jeremiah Minich and the research group at the University of California San Diego for providing the metadata, and the UBC MBIM department for funding this work.

CONTRIBUTIONS

All team members contributed to the completion of this draft manuscript, as well as planning, research, and analysis for this study. J.C. and C.Y. were responsible for preliminary metadata filtering and binning, as well as QIIME 2 sequence preparation and export. V.Y. was responsible for alpha diversity analysis in R, Figure 1 and Supplemental Figure 1 generation. F.Y. was responsible for describing the meta dataset, beta diversity analysis in R, and Supplemental Figure 2 generation. C.Y. was responsible for core microbiome analysis in R, Figure 2 and Supplemental Figure 3 generation. J.C. was responsible for PICRUSt2 analysis and Figure 3 generation. The abstract was written by F.Y. The introduction was written by F.Y. and V.Y. Methods, results, and discussion sections were divided equally amongst all team members depending on which analyses they were responsible for. The conclusion was written by C.Y. and F.Y. All members collaborated on the study limitations. C.Y. and V.Y. collaborated on the future directions. All members participated in manuscript editing. Co-authorship should be considered equal for all 4 authors.

REFERENCES

1. Elsheshtawy A, Clokie BGJ, Albalat A, Beveridge A, Hamza A, Ibrahim A, MacKenzie S. 2021. Characterization of External Mucosal Microbiomes of Nile Tilapia and Grey Mullet Co-cultured in Semi-Intensive Pond Systems. Front Microbiol 12:773860.
2. Pérez T, Balcázar JL, Ruiz-Zarzuela I, Halaihel N, Vendrell D, Blas I de, Múzquiz JL. 2010. Host–microbiota interactions within the fish intestinal ecosystem. Mucosal Immunol 3:355–360.
3. Merrifield DL, Rodiles A. 2015. 10 - The fish microbiome and its interactions with mucosal tissues, p. 273–295. *In* Beck, BH, Peatman, E (eds.), Mucosal Health in Aquaculture. Academic Press, San Diego.
4. Nayak SK. 2010. Role of gastrointestinal microbiota in fish. Aquac Res 41:1553–1573.
5. Legrand TPRA, Wynne JW, Weyrich LS, Oxley APA. 2020. A microbial sea of possibilities: current knowledge and prospects for an improved understanding of the fish microbiome. Rev Aquac 12:1101–1134.
6. Blanton JM, Peoples LM, Gerringer ME, Iacuaniello CM, Gallo ND, Linley TD, Jamieson AJ, Drazen JC, Bartlett DH, Allen EE. 2022. Microbiomes of Hadal Fishes across Trench Habitats Contain Similar Taxa and Known Piezophiles. mSphere 7:e00032-22.
7. Minich JJ, Härer A, Vechinski J, Frable BW, Skelton ZR, Kunselman E, Shane MA, Perry DS, Gonzalez A, McDonald D, Knight R, Michael TP, Allen EE. 2022. Host biology, ecology and the environment influence microbial biomass and diversity in 101 marine fish species. Nat Commun 13:6978.
8. Sullam KE, Essinger SD, Lozupone CA, O'Connor MP, Rosen GL, Knight R, Kilham SS, Russell JA. 2012. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. Mol Ecol 21:10.1111/j.1365-294X.2012.05552.x.
9. Layers of the Ocean. https://www.noaa.gov/jetstream/ocean/layers-of-ocean. Retrieved 13 March 2023.
10. Wickham H, François R, Henry L, Müller K, Vaughan D, Software P, PBC. 2023. dplyr: A Grammar of Data Manipulation (1.1.1).
11. Wickham H, Vaughan D, Girlich M, Ushey K, Posit, PBC. 2023. tidyr: Tidy Messy Data (1.3.0).
12. Posit. Posit. https://www.posit.co/. Retrieved 17 April 2023.
13. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S,

Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. 8. Nat Biotechnol 37:852–857.

14. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. 7. Nat Methods 13:581–583.
15. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596.
16. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, Schweer T, Peplies J, Ludwig W, Glöckner FO. 2014. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucleic Acids Res 42:D643–D648.
17. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE 5:e9490.
18. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066.
19. McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217.
20. Strong WL. 2016. Biased richness and evenness relationships within Shannon–Wiener index values. Ecol Indic 67:703–713.
21. Wickham H. 2016. Data Analysis, p. 189–201. *In* Wickham, H (ed.), ggplot2: Elegant Graphics for Data Analysis. Springer International Publishing, Cham.
22. Kim B-R, Shin J, Guevarra RB, Lee JH, Kim DW, Seol K-H, Lee J-H, Kim HB, Isaacson RE. 2017. Deciphering Diversity Indices for a Better Understanding of Microbial Communities 27:2089–2093.
23. Kassambara A. 2023. ggpubr: "ggplot2" Based Publication Ready Plots (0.6.0).
24. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Solymos P, Stevens MHH, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, Caceres MD, Durand S, Evangelista HBA, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill MO, Lahti L, McGlinn D, Ouellette M-H, Cunha ER, Smith T, Stier A, Braak CJFT, Weedon J. 2022. vegan: Community Ecology Package (2.6-4).
25. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the Tidyverse. J Open Source Softw 4:1686.
26. Huang R, Soneson C, Ernst FGM, Rue-Albrecht KC, Yu G, Hicks SC, Robinson MD. 2021. TreeSummarizedExperiment: a S4 class for data with hierarchical structure. F1000Research 9:1246.
27. Gao C-H, Yu G, Dusa A. 2022. ggVennDiagram: A "ggplot2" Implement of Venn Diagram (1.2.2).
28. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.
29. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2020. PICRUSt2 for prediction of metagenome functions. 6. Nat Biotechnol 38:685–688.
30. Barbera P, Kozlov AM, Czech L, Morel B, Darriba D, Flouri T, Stamatakis A. 2019. EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. Syst Biol 68:365–369.
31. Efficient comparative phylogenetics on large trees | Bioinformatics | Oxford Academic. https://academic.oup.com/bioinformatics/article/34/6/1053/4582279?login=false. Retrieved 17 April 2023.
32. Yang C, Mai J, Cao X, Burberry A, Cominelli F, Zhang L. 2023. ggpicrust2: an R package for PICRUSt2 predicted functional profile analysis and visualization. arXiv:2303.10388. arXiv https://doi.org/10.48550/arXiv.2303.10388.
33. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47.
34. Phipson B, Lee S, Majewski IJ, Alexander WS, Smyth GK. 2016. Robust hyperparameter estimation protects against hypervariable genes and improves power to detect differential expression. Ann Appl Stat 10:946–963.
35. Law CW, Chen Y, Shi W, Smyth GK. 2014. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 15:R29.

36. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27–30.
37. Kanehisa M. 2019. Toward understanding the origin and evolution of cellular organisms. Protein Sci 28:1947–1951.
38. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. 2023. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res 51:D587–D592.
39. Banerjee G, Ray AK. 2017. Bacterial symbiosis in the fish gut and its role in health and metabolism. Symbiosis 72:1–11.
40. Tarnecki A m., Burgos F a., Ray C l., Arias C r. 2017. Fish intestinal microbiome: diversity and symbiosis unravelled by metagenomics. J Appl Microbiol 123:2–17.
41. Kim PS, Shin N-R, Lee J-B, Kim M-S, Whon TW, Hyun D-W, Yun J-H, Jung M-J, Kim JY, Bae J-W. 2021. Host habitat is the major determinant of the gut microbiome of fish. Microbiome 9:166.
42. Salinity | Science Mission Directorate. https://science.nasa.gov/earth-science/oceanography/physical-ocean/salinity. Retrieved 17 April 2023.
43. Touchon M, Cury J, Yoon E-J, Krizova L, Cerqueira GC, Murphy C, Feldgarden M, Wortman J, Clermont D, Lambert T, Grillot-Courvalin C, Nemec A, Courvalin P, Rocha EPC. 2014. The Genomic Diversification of the Whole Acinetobacter Genus: Origins, Mechanisms, and Consequences. Genome Biol Evol 6:2866–2882.
44. Coskun ÖK, Özen V, Wankel SD, Orsi WD. 2019. Quantifying population-specific growth in benthic bacterial communities under low oxygen using H218O. 6. ISME J 13:1546–1559.
45. Fonseca A, Espinoza C, Nielsen LP, Marshall IPG, Gallardo VA. 2022. Bacterial community of sediments under the Eastern Boundary Current System shows high microdiversity and a latitudinal spatial pattern. Front Microbiol 13.
46. Hamilton TJ, Paz-Yepes J, Morrison RA, Palenik B, Tresguerres M. 2014. Exposure to bloom-like concentrations of two marine Synechococcus cyanobacteria (strains CC9311 and CC9902) differentially alters fish behaviour. Conserv Physiol 2:cou020.
47. Talwar C, Nagar S, Lal R, Negi RK. 2018. Fish Gut Microbiome: Current Approaches and Future Perspectives. Indian J Microbiol 58:397–414.
48. Walter JM, Bagi A, Pampanin DM. 2019. Insights into the Potential of the Atlantic Cod Gut Microbiome as Biomarker of Oil Contamination in the Marine Environment. 7. Microorganisms 7:209.
49. Lai KP, Zhu P, Boncan DAT, Yang L, Leung CCT, Ho JCH, Lin X, Chan TF, Kong RYC, Tse WKF. 2022. Integrated Omics Approaches Revealed the Osmotic Stress-Responsive Genes and Microbiota in Gill of Marine Medaka. mSystems 7:e00047-22.
50. Liang P, Saqib HSA, Lin Z, Zheng R, Qiu Y, Xie Y, Ma D, Shen Y. 2021. RNA-seq analyses of Marine Medaka (Oryzias melastigma) reveals salinity responsive transcriptomes in the gills and livers. Aquat Toxicol 240:105970.
51. Kashinskaya EN, Simonov EP, Izvekova GI, Parshukov AN, Andree KB, Solovyev MM. 2020. Composition of the microbial communities in the gastrointestinal tract of perch (Perca fluviatilis L. 1758) and cestodes parasitizing the perch digestive tract. J Fish Dis 43:23–38.
52. New Study Finds Ocean Pollution a Threat to Human Health. Glob Environ Health Newsl. https://www.niehs.nih.gov/research/programs/geh/geh_newsletter/2021/2/articles/new_study_finds_ocean_pollution_a_threat_to_human_health.cfm. Retrieved 17 April 2023.
53. Gaulke CA, Beaver LM, Armour CR, Humphreys IR, Barton CL, Tanguay RL, Ho E, Sharpton TJ. 2020. An integrated gene catalog of the zebrafish gut microbiome reveals significant homology with mammalian microbiomes. bioRxiv https://doi.org/10.1101/2020.06.15.153924.
54. Neff JM. 2002. Chapter 10 - Zinc in the Ocean, p. 175–189. *In* Neff, JM (ed.), Bioaccumulation in Marine Organisms. Elsevier, Oxford.
55. Lai KP, Lin X, Tam N, Ho JCH, Wong MK-S, Gu J, Chan TF, Tse WKF. 2020. Osmotic stress induces gut microbiota community shift in fish. Environ Microbiol 22:3784–3802.
56. 2012. A Measure of Salt. Text.Article. NASA Earth Observatory. https://earthobservatory.nasa.gov/images/78250/a-measure-of-salt. Retrieved 17 April 2023.
57. Burchard H, Craig PD, Gemmrich JR, van Haren H, Mathieu P-P, Meier HEM, Smith WAMN, Prandke H, Rippeth TP, Skyllingstad ED, Smyth WD, Welsh DJS, Wijesekera HW. 2008. Observational and numerical modeling methods for quantifying coastal ocean turbulence and mixing. Prog Oceanogr 76:399–442.
58. Eng A, Borenstein E. 2018. Taxa-function robustness in microbial communities. Microbiome 6:45.
59. Anguiano D. 2023. High concentrations of DDT found across vast swath of California seafloor. The Guardian.