water

MDPI

Article

Metagenomic Analysis of Microbial Diversity in the Moroccan Coastal Water of the Gibraltar Strait

Manal Chrairi [1], Said Barrijal [2], Antonio Castellano-Hinojosa [3,*], Youssra Boumait [1], Chahrazade El Hamouti [1], Afaf Lamzouri [1,4], Hassan Ghazal [5,6,7] and Rajaa Chahboune [1]

1 Molecular Biology Unit, Life and Health Sciences Laboratory, Faculty of Medicine and Pharmacy of Tangier, Abdelmalek Essaadi University, Tangier 90000, Morocco; chrairi.manal@gmail.com (M.C.); boumait.youssra@gmail.com (Y.B.); chahrazade.elhamouti@gmail.com (C.E.H.); lamzouriafaf@hotmail.com (A.L.); rchahboune@uae.ac.ma (R.C.)
2 Faculty of Sciences and Techniques of Tangier, Abdelmalek Essaadi University, Tangier 90000, Morocco; barrijal@yahoo.fr
3 Department of Microbiology, Institute of Water Research, University of Granada, C/Ramon y Cajal, 4, 18071 Granada, Spain
4 Department of Medical Genetics and Oncogenetics, Mohammed VI University Hospital, Tangier 90000, Morocco
5 Laboratory of Precision Medicine and One Health, Faculty of Medicine, Mohammed VI University of Sciences and Health, Casablanca 82403, Morocco; hassan.ghazal@fulbrightmail.org
6 Royal Institute of Executive Management for Youth and Sport (IRFC), Department of Sports Sciences, Laboratory of Sports Sciences and Performance Optimization, Salé 10102, Morocco
7 National Center for Scientific and Technical Research (CNRST), Rabat 10102, Morocco
* Correspondence: ach@ugr.es

Abstract: Coastal waters are known for higher productivity and organic matter levels, which support a high diversity and abundance of microorganisms compared to some aquatic environments. The characterization of marine microbiomes can provide valuable information for evaluating the sustainability of coastal waters that are increasingly subjected to environmental and human impacts. Our study is the first metagenomic study realized on Moroccan Mediterranean coastal seawater. We analyzed and described the Gibraltar Detroit marine microbiome taxonomic and functional profiling using MG-RAST software. Shotgun sequencing revealed a predominance of bacterial taxa, particularly the Proteobacteria (57.29%) and Bacteroidetes (27.06%) phyla, alongside notable populations of eukaryotes, viruses, and archaea. Alphaproteobacteria and Gammaproteobacteria emerged as the dominant bacterial classes, while Flavobacteria represented a significant portion of Bacteroidetes. Functional profiling of the microbial community highlighted a wide array of metabolic pathways, emphasizing genes related to carbohydrate metabolism, amino acid synthesis, and protein processing. The marine microbiome exhibited essential biogeochemical activities, particularly in nitrogen, sulfur, and carbon cycles, with notable pathways including denitrification, thiosulfate oxidation, and carbon fixation. This functional diversity underlines the microbiome's role in sustaining ecosystem health through nutrient cycling and organic matter degradation. Our findings offer a crucial baseline for understanding microbial community structure and resilience in Mediterranean coastal ecosystems, with implications for assessing future environmental and anthropogenic impacts on these microbial dynamics.

Keywords: Mediterranean Sea; bacterial diversity; marine microbiome; bioinformatics; whole-genome sequencing

Citation: Chrairi, M.; Barrijal, S.; Castellano-Hinojosa, A.; Boumait, Y.; El Hamouti, C.; Lamzouri, A.; Ghazal, H.; Chahboune, R. Metagenomic Analysis of Microbial Diversity in the Moroccan Coastal Water of the Gibraltar Strait. *Water* **2024**, *16*, 3202. https://doi.org/10.3390/w16223202

Academic Editor: Abasiofiok Mark Ibekwe

Received: 26 September 2024
Revised: 31 October 2024
Accepted: 6 November 2024
Published: 8 November 2024

1. Introduction

Seawater contains more than a billion microorganisms per liter including bacteria and other microorganisms, which are essential for maintaining and organizing marine

ecosystem sustainability. Microbial communities in marine ecosystems actively participate in various biogeochemical processes including carbon cycling, impacting how carbon moves through the ocean and shaping pathways for this crucial element. Moreover, they are pivotal for biogeochemical primary production in marine environments [1], maintaining ocean fertility and marine food chain production [2,3]. The estimated number of prokaryotes on Earth is 1.2×10^{30} cells, with the most occurring in the deep oceanic subsurface (4×10^{29} cells), soil (3×10^{29} cells), deep continental subsurface (3×10^{29} cells), oceanic subsurface (1×10^{29} cells), and upper oceanic sediment (5×10^{28} cells) [4]. Interestingly, ~70% of the prokaryotic biomass in the oceanic water column harbors the epipelagic waters [5]. Despite their important roles in different life cycles, scientists have not yet identified and understood the diversity and role of microbial communities in marine ecosystems, but bacteria seem to be the predominant life form in the ocean, exhibiting high levels of diversity, abundance, and metabolic activity [1]. While marine microbes play crucial ecological roles, much remains unknown about their diversity, functional roles, and interactions within marine ecosystems. Notably, only a small fraction of bacteria in environmental samples can be isolated and cultured, which has historically limited the characterization of marine microbial communities [6].

The advent of high-throughput sequencing and whole-metagenomic shotgun (WGS) sequencing has markedly expanded our ability to analyze microbial diversity in marine environments by enabling culture-independent studies of entire microbial communities [7,8]. These approaches allow scientists to generate comprehensive genetic profiles that reveal the immense microbial diversity and functional complexity within marine ecosystems, illuminating previously hidden metabolic pathways and ecological interactions [9,10]. WGS sequencing is especially transformative, as it not only enables taxonomic profiling via 16S rRNA gene analysis but also facilitates the examination of functional genes within microbial communities, providing a nuanced understanding of their ecological roles and adaptive strategies [11].

Recent advancements in sequencing technology have propelled a wave of marine microbiome studies, leading to significant data accumulation on microbial community dynamics, diversity, and functionality across various oceanic regions. For instance, metagenomics studies have highlighted shifts in microbial community structures in response to environmental changes, contributing to a more detailed picture of microbial adaptability and resilience in marine ecosystems [3,9,12,13]. As a result, there is now a robust dataset that reflects the diversity and functional potential of marine microbiomes, underscoring the value of high-throughput sequencing as an indispensable tool for marine microbial ecology. Over the past decade, these approaches have not only broadened our understanding of microbial diversity but have also paved the way for innovative ecological insights, revealing complex microbial interactions that sustain marine ecosystem health and productivity.

The Mediterranean Sea is an enclosed basin [14], and its coasts have been a cradle of many civilizations and industries, thus supporting a high density of inhabitants and representing an important transport network [15]. This oligotrophic basin is characterized by highly variable environmental conditions; thus, the salinity, nutrient gradient availability, and temperature are affected. Its coastal areas are enriched by regional features through changes in some climatic conditions and several types of human impacts. Interestingly, the Strait of Gibraltar is one of the most principal canals in the Mediterranean and the world, as it is the only connection between the Mediterranean Sea and the Atlantic open ocean [16]. It is known as a complex system of contractions, submarine sills, and daily climatic conditions and plays a key role in environmental changes. However, it is described as an interesting mixture of water types coming from the Mediterranean and the eastern North Atlantic Ocean [16,17]. As a result, the Strait of Gibraltar represents an important hotspot for marine microbial biodiversity, unique microbial habitats, and the production of primary ecological cycles. However, few studies have used metagenomics analyses to explore the microbial communities of the Gibraltar Strait.

One of the important issues in marine ecology investigations is to acquire insights into the links between the ecological cycle processes and the microbial composition data of the studied area. Microorganisms facilitate elemental cycling in the oceans and modulate biogeochemical processes that are essential for sustaining planetary habitability [18,19]. Therefore, the study of marine microbiomes constitutes a critical step towards describing the genetic diversity and metabolic potential of marine ecosystems. In this context, the objective of this study was to characterize the community structure, composition, and functional traits of the marine microbiome of the Gibraltar Strait using metagenomics. To our knowledge, this is the first metagenomics study that has been applied in Morocco to study a Mediterranean marine coastal microbial ecosystem.

2. Materials and Methods

2.1. Sample Site and Sample Processing

Fresh seawater was collected from the southern section of the Strait of Gibraltar site (35°49′27.6″ N 5°43′22.6″ W) (Tangier OSD26), which is located on the northern coast of Morocco (Figure 1A,B). The sampling was conducted on 21 June 2014, in accordance with the Ocean Sampling Day (OSD) project's standard sampling protocol (https://github.com/MicroB3-IS/osd-analysis/wiki/Guide-to-OSD-2014-data) (accessed on 10 June 2014 for sampling processing). The OSD is a global research sampling campaign that involves marine biologists from around the world collecting surface water samples from coastal areas to study marine microbial complexity and interactions with the surrounding environment. This provides a snapshot of the world's oceans, enabling the assessment of the composition and function of prokaryotic and eukaryotic microbial communities. Approximately 25 L of surface seawater (1 m) was collected from the shore from five different points and pooled together. The environmental parameters (metadata) (Supplementary Table S1) were measured in situ during the sampling process using multivariate appliances and probes. After sampling, the water samples were kept refrigerated before being filtered (approximately 1.800 L of filtered water per filter unit) through GV Sterivex cartridge filters (Millipore) (Merck KGaA, Darmstadt, Germany) (4–5 filter units with a 0.22 µm pore size) using a hand pump. The filtered samples were stored at −80 °C until they were shipped for DNA extraction and sequencing.

2.2. DNA Extraction and Sequencing

The Sterivex filters used in the sampling process were subjected to a DNA extraction process as described in the OSD protocol. The detailed protocol can be found under this link: https://github.com/MicroB3-IS/osd-analysis/wiki/Guide-to-OSD-2014-data. DNA extraction was carried out using the Power Water Isolation Kit (MoBio, Carlsbad, CA, USA). The microbial-extracted DNAs were shotgun-sequenced using Illumina MiSeq V3 (2 × 300 bp; Illumina, San Diego, CA, USA). A total of 1,252,579 raw reads were generated after sequencing.

2.3. Bioinformatic Analysis

2.3.1. Preprocessing, Quality Control, Trimming, and Filtering of Reads

The raw data were obtained already demultiplexed and trimmed of adapters and primers before starting the analysis. The preprocessing steps of the fastq raw data, including quality control, filtering, and trimming by quality and length, were conducted using bash scripting. Low-quality reads were filtered out and sequences were trimmed to a consistent length using Trimmomatic [20] (v0.39). A total of 1,177,617 high-quality sequences were obtained, representing 94.02% of the total raw sequences.

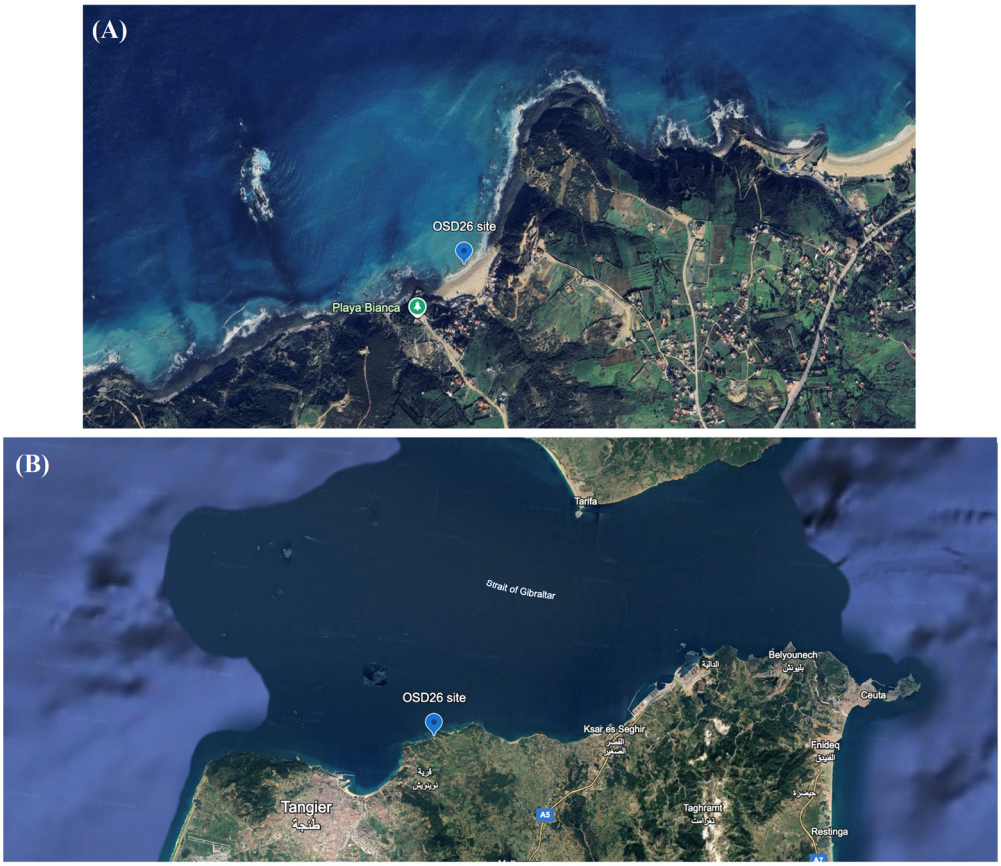

Figure 1. Tangier site, OSD26, north coast of Morocco (35°49′27.6″ N 5°43′22.6″ W), zoom-in on the site (**A**), global view of the site within the Strait of Gibraltar area (**B**) (image from Google Earth, 2024).

2.3.2. Assembly

MetaSpades [21] (v3.13.1) software was used for the metagenomic assembly. MetaSPAdes is designed to assemble shotgun metagenomic reads and can assemble sequenced data from a mixture of bacteria of different abundances and closely related species. It uses paired reads to identify and resolve chimeric contigs resulting from the mis-assembly of different genomes [21,22]. The k-mer sizes were set to (21, 33, 55). The assembly quality of the metagenome was evaluated using MetaQuast [23], for evaluation of the genome assemblies based on contig alignment to a reference.

2.3.3. MG-RAST Pipeline and Analyses

The MG-RAST pipeline (Metagenomics Rapid Annotation using Subsystems Technology) [24,25] was utilized to assess the diversity of organisms at different taxonomic levels. Although whole-genome sequencing was employed, taxonomic profiling was conducted at the genus level to ensure consistency and accuracy in classification, given the limitations of current taxonomic databases and assignment algorithms in reliably distinguishing species in complex environmental samples. The MG-RAST was used to predict the percentage of known annotated proteins in the studied metagenome. The distribution of functional categories and subsystems present in the metagenome was annotated against three different databases: COG, NOG, and KO. COG is the database of Clusters of Orthologous Groups of proteins [26]; NOG is the database of Orthology Relationships, Functional Annotation, and Gene Evolution [27]; and KO is the KEGG Orthology database, which represents molecular functions in terms of functional orthologs of genes and proteins [28]. These databases were used to compare the annotation results of the sequencing data.

2.4. Microbiome-Level Differences in Biogeochemical Processes

Besides the bioinformatic analyses, the studied metagenomic data were assessed using the DiTing tool [29] to decipher the involved genes in the carbon, nitrogen, and sulfur biogeochemical cycles in this marine microbiome. The input clean reads were returned to DiTing that uses MetaSPAdes [21] for assembly (related to DiTing analysis) (with default parameters), as it is recommended for complex metagenomes [29], and Prodigal [30] for recovering the Open Reading Frames (ORFs) by their prediction and translation from the assembled contigs and their annotation with KofamScan [31] using KEGG Orthologs (KOs). The presence or absence of genes involved in each of the steps of the carbon, nitrogen, and sulfur biogeochemical cycles as well as their relative was determined. Heat maps were produced to better illustrate the results.

3. Results

3.1. Bioinformatic Analysis of the Obtained Raw Reads

The contigs in the studied dataset consist of 270,114 high-quality sequences and 122,752,517 base pairs, with an average length of 454 base pairs. The dataset underwent a rigorous QC pipeline and no artificial duplicate reads were identified, which is indicative of good quality. The majority (99.24%) of the analyzed sequences had predicted features, with (54.39%) of the sequences (145,807) containing predicted proteins with known functions and (45.40%) (121,708) containing predicted proteins with unknown functions. A minimal number of sequences (554) contain ribosomal RNA genes. All these produced data culminate in the interpretation of biologically insightful information.

3.2. Microbial Taxonomy

The results provided by the MG-RAST pipeline provide an overview of the diversity and composition of the microbial communities present in the studied site. The pipeline enabled the identification and categorization of the taxonomic composition of the sample, with bacteria representing the most dominant domain (96.06%), followed by eukaryotes (2.16%), viruses (1.02%), and archaea (0.54%) (Figure 2A). The analysis revealed the presence of more than seven phyla. The two most abundant phyla were Proteobacteria (57.29%) (Figure 2B), which mainly belong to the Alphaproteobacteria (30.72%) and Gammaproteobacteria (21.03%) classes (Figure 2C). The second most abundant phylum was Bacteroidetes (27.06%), with the class Flavobacteria (18.02%) being the most abundant. The remaining phyla, including Planctomycetes, Verrucomicrobia, Firmicutes, Cyanobacteria, Actinobacteria, Chlorophyta, Lentisphaerae, and Chloroflexi, were identified but were poorly represented, with minimal percentages ranging from 0.39% to 2%.

Regarding the relative orders (Supplementary Figure S1), Rhodobacterales was the most abundant (20.66%), which belongs to the most abundant class Alphaproteobacteria, followed by Flavobacteriales (14.88%), which belongs to the Flavobacteria class and other orders, with smaller proportions not exceeding 7%. At the family level (Supplementary Figure S2), Rhodobacteraceae had the highest abundance (15.02%), followed by Flavobacteriaceae (12.53%). At the genus level (Figure 2D), unclassified genera derived from Gammaproteobacteria were the most abundant genera in this sample (6%), followed by *Candidatus* Pelagibacter (4.95%), unclassified genera derived from the Flavobacteria class (4.14%), and *Candidatus* Puniceispirillum (3.07%). The remaining genera were represented by very small percentages (less than 2%).

Finally, MG-RAST summarized the diversity of organisms present in the sample by calculating the number of species as a proxy of alpha diversity. The analysis revealed that our dataset contained 475 species, without specifying the exact species present in this studied microbiome.

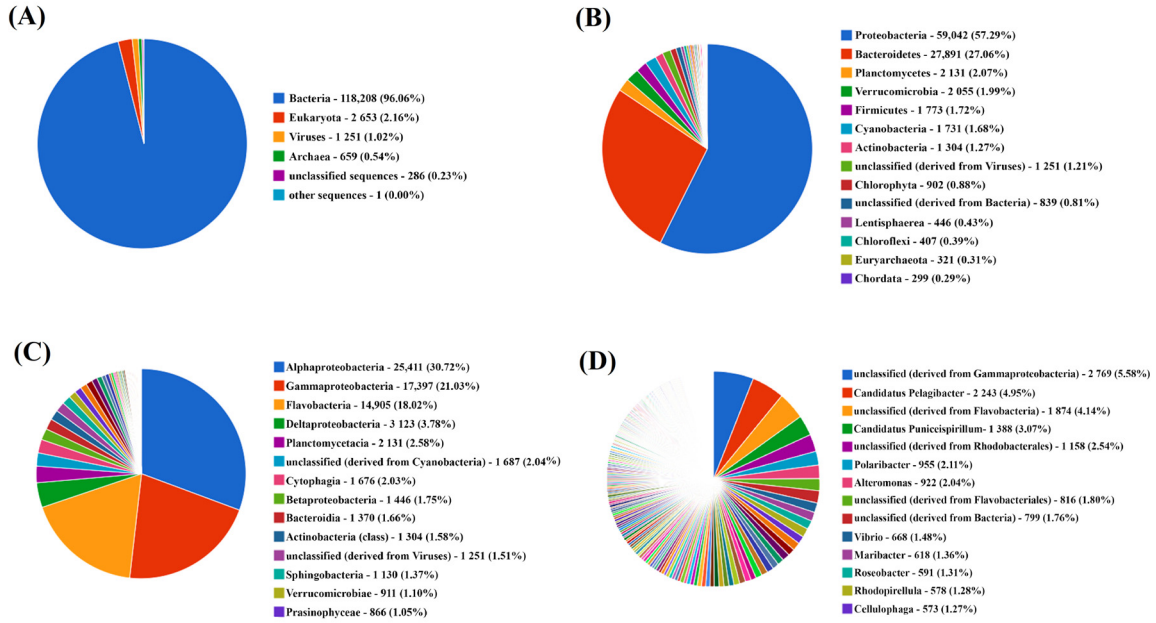

Figure 2. Domain-level (**A**), phylum-level (**B**), class-level (**C**), and genus-level (**D**) classifications of obtained metagenomic reads in the Gibraltar sample OSD26.

3.3. Microbial Community Functional Profiling

The results in Supplementary Figure S3 show that 54.39% of the proteins were known, 45.40% were unknown, and rare ribosomal RNA was present at approximately 0.21%. For the annotation of functional genes (Figure 3), it was observed that clustering-based subsystems (14.49%), carbohydrates (12.86%), amino acids and derivatives (10.54%), and protein metabolism (7.53%) were the most common categories. Additionally, a significant number of genes were assigned to other subsystem categories, such as RNA metabolism (5.97%), DNA metabolism (3.66%), respiration (2.40%), stress response (2.35%), and membrane transport (2.35%).

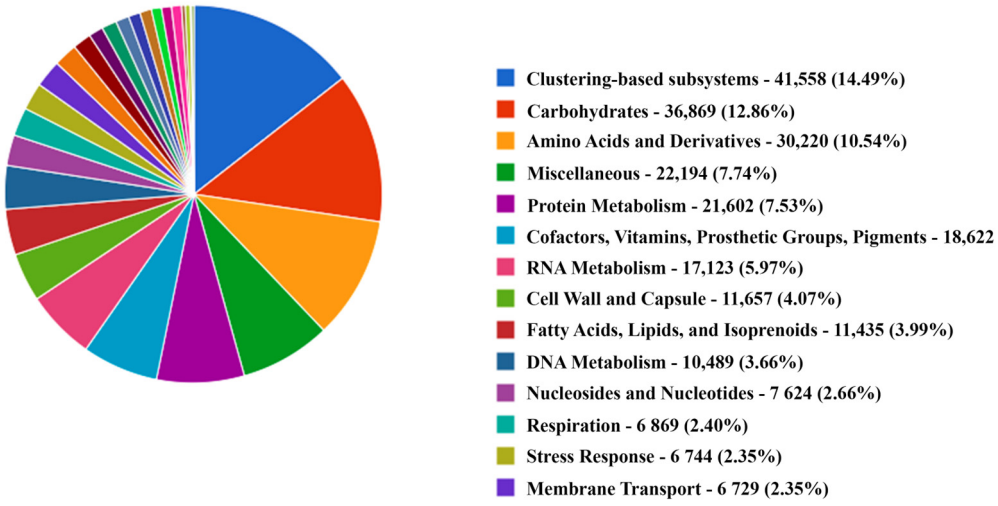

Figure 3. Pie chart summarizing subsystem terms inferred from MG-RAST annotation.

Figure 4 shows the percentage of sequences with predicted protein functions annotated with each functional category based on the respective database used for annotation. For instance, in the case of metabolic functions, COG annotated 45.56% of our sample genes, while KO annotated 61.22% of the total genes, and NOG annotated only 10.03%. For cellular processes and signaling, and information storage and processing, each database provided

its own annotations and percentages of sequences with predicted functions. Notably, COG annotated a larger percentage of our sample sequences with predicted functions, totaling 83.25% of annotated sequences, while NOG annotated fewer sequences, representing a total of 20.22%. Moreover, the KOs provided more predicted protein functions related to genetic information processing (19.52%), environmental information processing (12.80%), and human diseases (1.39%).

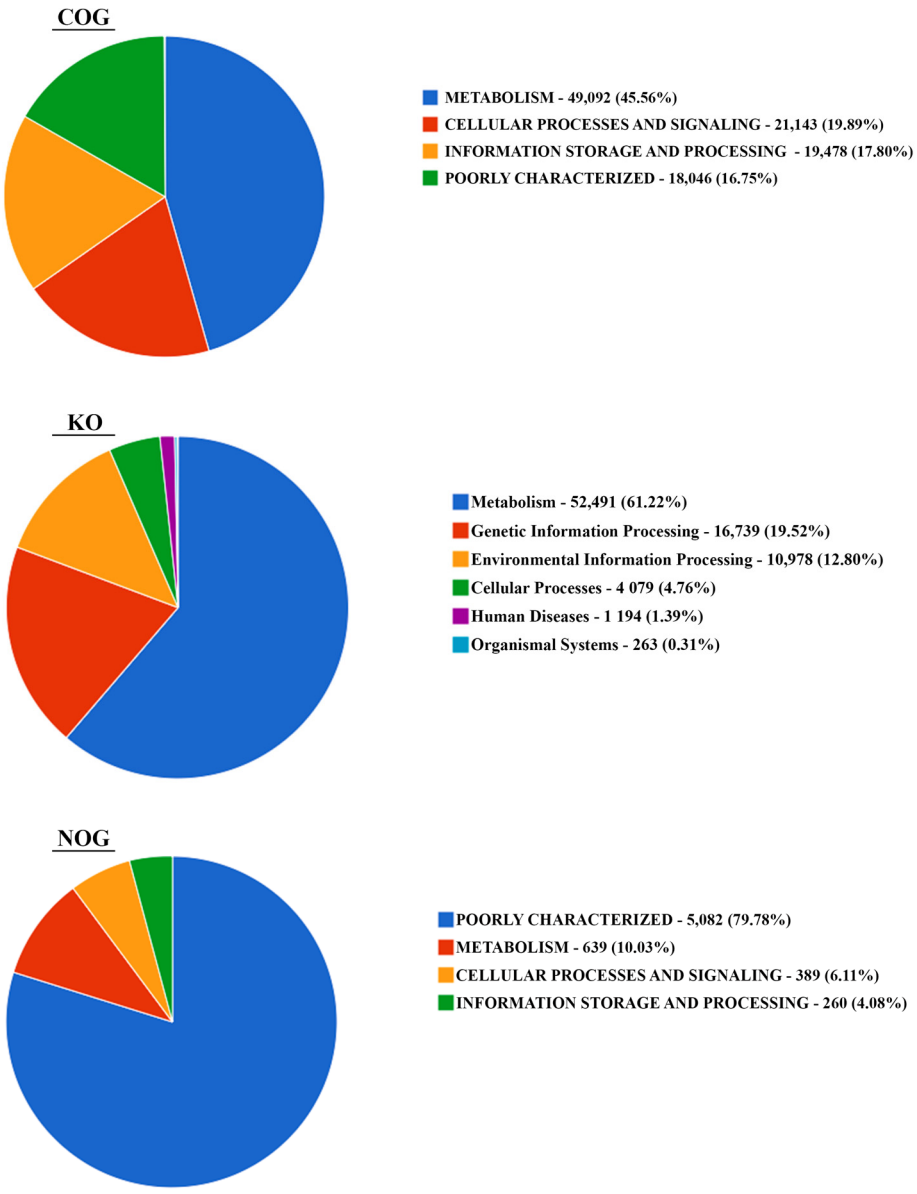

Figure 4. Pie charts illustrating the distribution of functional categories for the COG, KO, and NOG databases.

3.4. Biogeochemical Processes at the Sample Level

The relative abundances of genes involved in the nitrogen, sulfur, and carbon cycles that were present in the marine microbiome are shown in Figure 5A, Figure 5B, and Figure 5C, respectively. For the nitrogen cycle, we observed that only three pathways were present in the marine microbiome (nitrate reduction to ammonium [DNRA], reduction of NO_3^- to NO_2^-, and reduction of NO to N_2O) (Figure 5A). Denitrification was the dominant pathway, and genes such as *narGHI* (denitrification through cytoplasmic nitrate reductase), *napAB* (periplasmic nitrate reductase) pathways, and *norBC* (nitric oxide reductase) were present in the microbiome with relative abundances greater than 5%. The DNRA was

present with a relative abundance of more than >=3% and <5% through *nirBD* (nitrite reductase) and *nrfAH* (ammonia-forming dissimilatory nitrite reductase) pathways.

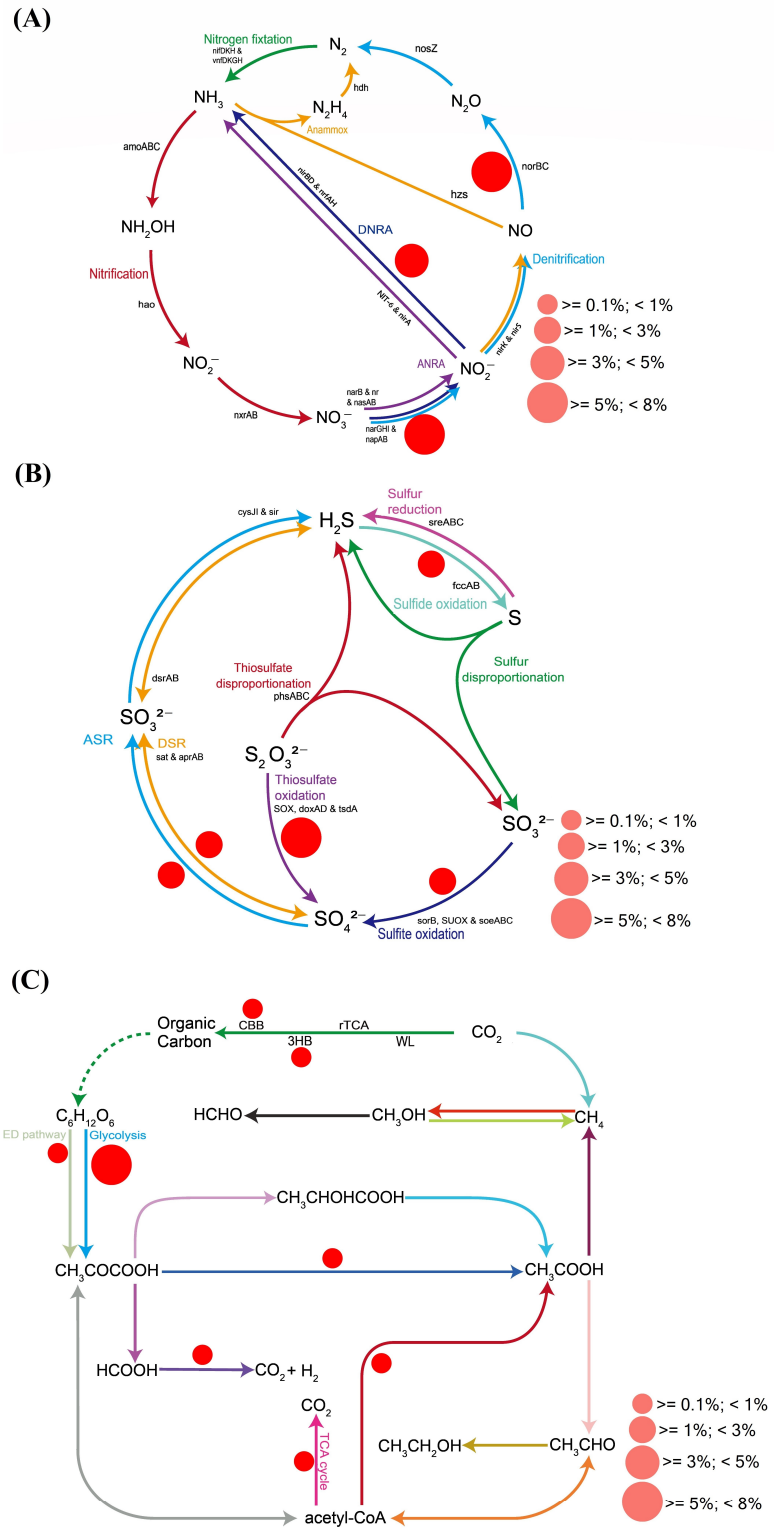

Figure 5. Relative abundances of the involved pathways in the nitrogen (**A**), sulfur (**B**), and carbon (**C**) cycles.

For the sulfur cycle, thiosulfate oxidation was the most abundant pathway (via SOX, *doxAD*, and *tsdA*) with a relative abundance greater than 3% in the marine microbiome.

Moreover, the sulfide oxidation and sulfite oxidation, the ASR (assimilatory sulfate reduction), and the DSR (dissimilatory sulfate reduction) pathway abilities were also present with a relative abundance of >=1% and <3% (Figure 5B).

Regarding the carbon cycle, we found that the marine microbiome contained genes related to eight functional pathways (e.g., Entner–Doudoroff pathway, glycolysis, TCA cycle, fixation of CO_2 to organic carbon). Among them, genes coding for glycolysis were the most abundant with values of relative abundance in the range of >=3% and <5% (Figure 5C).

4. Discussion

This study provides novel insights into the diversity, composition, and functional traits of microbial communities living in the surface seawaters of the Gibraltar Strait between Morocco and Spain. This information contributes to understanding the relationships between the microbial communities and their environment in this fragile and specific Mediterranean ecosystem, and to our knowledge, this is the first metagenomic analysis of the Moroccan marine microbiome of this marine area of the Gibraltar Strait. Our findings also provide valuable information to evaluate the sustainability of this marine area that is subjected to environmental and human impacts. While the scope of this study is limited by the use of a single pooled sample, this approach is consistent with exploratory metagenomic studies aimed at establishing initial characterizations of microbiomes in ecologically significant but under-studied areas. Future studies should consider a more comprehensive sampling strategy to allow for spatial and temporal comparisons. However, our findings provide pioneering insight into the microbial ecology of the Gibraltar Strait, highlighting key microbial players and metabolic pathways that could be foundational for assessing anthropogenic and natural impacts in Mediterranean coastal waters. Despite nearly a decade since sample collection, the relative stability of key environmental parameters in this region (e.g., salinity, nutrient levels) underscores the continued relevance of these findings as indicators of microbial structure in this coastal ecosystem.

Our results showed that bacteria dominate (95% of the sequences) the seawater microbiome on this Mediterranean sampling site. This finding is consistent with results from other studies across the Mediterranean Sea [12]. Like in most marine environments, we found that the microbial community was dominated by the Proteobacteria (especially the Alphaproteobacteria and Gammaproteobacteria classes) and Bacteroidetes (mainly the Flavobacteria class) [3,9,12,32]. Proteobacteria members are often linked to denitrification and amino acid biosynthesis in marine environments [33], while Bacteroidetes are known to be widespread degraders of particulate matter, often growing attached to particles, surfaces, or algal cells [34]. Additionally, Bacteroidetes include photoautotrophic members that produce energy via photosynthesis, utilizing light and carbon dioxide. Within Bacteroidetes, Flavobacteria are notable for generating significant metabolites and include species pathogenic to fish and algae [35]. These taxa are abundant in coastal waters and display high diversity [36]. Interestingly, the dominance of Proteobacteria and Bacteroidetes is not exclusive to marine environments. Studies have shown that these phyla also dominate in various freshwater ecosystems, underscoring their adaptive roles in diverse aquatic environments. For instance, Proteobacteria, particularly Alphaproteobacteria and Gammaproteobacteria, have been frequently identified in freshwater systems, where they fulfill essential roles in organic matter degradation and nutrient cycling, often linked to areas with high anthropogenic activity [37,38]. Similarly, Bacteroidetes in freshwater systems play crucial roles in breaking down complex organic compounds, suggesting a shared functional role with their marine counterparts [37,38].

In line with other studies, the Proteobacteria phylum (mainly Alphaproteobacteria) is the most abundant taxon across various Mediterranean marine sampling sites [12,39], followed by Gammaproteobacteria, Deltaproteobacteria, and Betaproteobacteria, which were also detected in our study. Marine Alphaproteobacteria are especially prevalent in Mediterranean marine regions [39], particularly in the epipelagic zone of coastal waters [40].

Gammaproteobacteria play significant roles in nutrient cycling, with prominent orders like Alteromonadales and Vibrionales thriving in the nutrient-rich waters of coastal zones, including Tangier [41]. Gammaproteobacteria are also present in productive or polluted freshwater environments, serving as potential biomarkers for pollution due to human activities [42]. Additionally, small proportions of other phyla, such as Planctomycetes, Verrucomicrobia, Cyanobacteria, Actinobacteria, and Firmicutes, were observed, suggesting a high diversity of microorganisms in the surface waters along the Moroccan coast of the Gibraltar Strait.

We detected taxa of the Rhodobacterales order and Rhodobacteraceae within the Alphaproteobacteria. These taxa are frequently found in marine environments, particularly in the pelagic zone, where they live in symbiosis with other aquatic macro- and microbial habitats and exhibit significant ecologic and phenotypic diversity, and they are involved in important carbon and sulfur biochemical cycles in their ecosystems [43,44]. The identified genera *Candidatus* Pelagibacter and *Candidatus* Puniceispirillum belong to the Alphaproteobacteria class and have been detected in coastal regions and open oceans [45]. The *Pelagibacter* genus was also observed in our sample and is relatively abundant in the southern Mediterranean waters such as coastal Tunisian seawaters [3]. Members of this genus are known to play an indispensable role in geochemical cycles in ocean systems by recycling unstable dissolved organic matter and have a strong response to nitrogen limitation [46]. Genera derived from Gammaproteobacteria such as *Alteromonas* and *Vibrio* were found in the marine microbiome. Taxa belonging to these two genera are widespread in coastal marine ecosystems where they are involved in nutrient cycling and the biodegradation of organic matter and are associated with marine eukaryotes [3,41]. In particular, the Gammaproteobacteria class is dominant in Mediterranean coastal seawater and has emerged on the Moroccan Mediterranean coast as well as in Tunisian coastal waters [3].

Our analysis identified not only bacteria but also eukaryotes, viruses, and archaea, though at lower abundances. These non-bacterial taxa play significant roles in marine ecosystems, contributing to biogeochemical cycles and ecological interactions. For example, marine viruses are known to influence bacterial population dynamics and nutrient cycling, while certain eukaryotic algae contribute to primary production in coastal environments. This highlights the complex microbial interactions within the Gibraltar marine microbiome, where bacteria form the primary community but are supported by a diverse array of other microorganisms.

The functional activities revealed by metagenomics are a key tool for determining and understanding the interactions between microbial communities living within an ecosystem and its microbiome functionalities. This study showed the overall predicted functions of the Moroccan marine microbiome through their assignments and comparisons to three functional gene databases, COG, NOG, and KO. The results were annotated with various metabolic functions, mainly clustering-based subsystems, carbohydrates, amino acids and derivatives, and proteins. These results suggest that the marine microbiome is enriched with microbes that can perform diverse functions, including the degradation of carbohydrates, which often represent a substantial fraction of marine organic matter and are substrates for heterotrophs [47]. This is important as the Strait of Gibraltar is subjected to increasing environmental and human impacts. The presence of other functions, such as RNA metabolism, DNA metabolism, respiration, membrane transport, and stress response, points toward the plasticity of marine microbiomes to adapt to different stress and environmental conditions. This adaptation is maintained by different microorganisms present in the microbiome and is adjusted through molecular and cellular network systems [48]. This can be attributed to the presence of microbial diversity communities through the relatively abundant members of Proteobacteria and Bacteroidetes, which are known for their enrollment and contribution to crucial environmental life cycles, as well as to metabolic functions and biochemical degradation reactions [15]. Indeed, some members of the Gammaproteobacteria class can produce proteases and extracellular alkaline phosphatases [3], which

can explain the high prevalence of metabolism and membrane transport functions in our studied marine microbiome.

We found that the marine microbiome of the Strait of Gibraltar is enriched in microorganisms performing specific steps of the carbon, nitrogen, and sulfur cycles. The carbon cycle pathways we detected included glycolysis, the TCA cycle, and carbon fixation, with glycolysis genes exhibiting the highest relative abundance. This highlights the community's potential for organic carbon processing and energy production, essential for maintaining ecological balance and supporting higher trophic levels [47–51]. The nitrogen cycle in the studied microbiome was characterized predominantly by denitrification pathways, with genes such as *narGHI* and *napAB* showing notable relative abundances. This suggests that denitrification is a critical process in the region, potentially facilitating the removal of excess nitrogen, which is vital for maintaining water quality and mitigating eutrophication risks [47]. The presence of the dissimilatory nitrate reduction to ammonium (DNRA) pathway further emphasizes the microbial community's role in nitrogen cycling, indicating an adaptive capacity to fluctuating nitrogen conditions. The sulfur cycle also revealed important metabolic activities, with thiosulfate oxidation being the most prevalent pathway. Heterotrophic bacteria, particularly Alphaproteobacteria, play a significant role in mitigating sulfide toxicity through sulfur oxidation processes, thus protecting marine ecosystems [18,52,53]. The interconnectedness of these biogeochemical cycles underscores the complexity of microbial interactions in this environment, where microorganisms contribute not only to their own survival but also to the broader ecological framework.

Due to their diverse metabolic activities in biogeochemical processes, marine microorganisms play a crucial role in interconnecting the sulfur transformation cycles with the carbon and nitrogen cycles. Numerous researchers have shown that dissimilatory sulfur oxidation pathways by heterotrophic bacteria, such as Proteobacteria and especially Alphaproteobacteria, can protect marine ecosystems against sulfide toxicity [52]. The high abundance of denitrification pathways in this marine ecosystem can be explained by the presence of anaerobic microorganisms [54], such as some existing members in this microbiome of Alphaproteobacteria, Gammaproteobacteria, and Flavobacteria, which are facultatively anaerobic.

It was noteworthy to emphasize that comparative genomics has shown that Bacteroidetes, and especially Flavobacteriales, are widely regarded as experts in the destruction of polymers due to their possession of a significant number of peptidases and enzymes that are active on carbohydrates [53]. Furthermore, they play complementary important roles in the carbon cycle of the oceans besides the Proteobacteria, especially the Alphaproteobacteria class [34]. Together, our findings show that the dominant presence of the Proteobacteria and Bacteroidetes members in the studied microbiome can be related to different active pathways of the carbon cycle uncovered in the microbiome.

5. Conclusions

Our metagenomic analysis of the Gibraltar Strait surface water provided the first characterization of the community structure, composition, and functional traits of the marine microbiome of the Moroccan coast of the Gibraltar Strait. We found that this marine ecosystem has a high diversity and functionality, with bacteria being the dominant member of the microbiome. This study has significant environmental implications for understanding the richness of marine microbial communities in surface seawater and the various metabolic functions encoded by the presented coding genes of the microbiome. Our results can be used to evaluate the response of the Gibraltar marine microbiome to respond to future anthropogenic and environmental impacts. In addition, some identified bacterial species or communities can be used as bioindicators for water quality assessment in a strait context, which should be explored in further studies.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/w16223202/s1. Table S1: Physical and chemical data measured in fresh seawater of Gibraltar Detroit site at the time of sampling. Figure S1: Order-level classification of obtained metagenomic reads in the Gibraltar sample OSD26. Figure S2: Family-level classification of obtained metagenomic reads in the Gibraltar sample OSD26. Figure S3: Predicted features of analyzed metagenomic sequences dataset.

Author Contributions: Conceptualization, M.C., S.B., H.G. and R.C.; methodology, M.C. and H.G.; validation, S.B., H.G. and R.C.; formal analysis, M.C.; investigation, M.C.; data curation, M.C.; writing—original draft preparation, M.C.; writing—review and editing, S.B., A.C.-H., H.G. and R.C.; visualization, M.C., S.B., A.C.-H., Y.B., C.E.H., A.L., H.G. and R.C; supervision, S.B. and R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the EU MicroB3 project, funded by the European Union's Seventh Framework Programme (FP7; Joint Call OCEAN.2011-2: Marine microbial diversity-new insights into marine ecosystems functioning and its biotechnological potential) under grant agreement no. 287589 (Ocean Sampling Day (OSD) Consortium). Manal Chrairi is a recipient of the Erasmus internship at the Water Institute, Granada, Spain, for the 2022/23 academic year. Hassan Ghazal is a US NIH grant recipient through the H3abionet/H3africa consortium U24HG006941 and the OSD coordinator for Morocco.

Informed Consent Statement: The necessary and proper licenses for marine water collection were obtained prior to project commencement at the Mediterranean Moroccan studied site.

Data Availability Statement: The bacterial metagenome raw read sequence dataset was deposited in EMBL-EBI, https://www.ebi.ac.uk/ena/browser/view/ERR771019 (accessed on 1 February 2023).

Acknowledgments: This study is part of the Ocean Sampling Day consortium. The bioinformatics analysis has been performed in the MARWAN HPC server at the National Center for Scientific and Technical Research (CNRST).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Azam, F.; Malfatti, F. Microbial Structuring of Marine Ecosystems. *Nat. Rev. Microbiol.* **2007**, *5*, 782–791. [CrossRef] [PubMed]
2. Zehr, J.P.; Kudela, R.M. Nitrogen Cycle of the Open Ocean: From Genes to Ecosystems. *Annu. Rev. Mar. Sci.* **2011**, *3*, 197–225. [CrossRef]
3. Quéméneur, M.; Bel Hassen, M.; Armougom, F.; Khammeri, Y.; Lajnef, R.; Bellaaj-Zouari, A. Prokaryotic Diversity and Distribution Along Physical and Nutrient Gradients in the Tunisian Coastal Waters (South Mediterranean Sea). *Front. Microbiol.* **2020**, *11*, 593540. [CrossRef]
4. Flemming, H.-C.; Wuertz, S. Bacteria and Archaea on Earth and Their Abundance in Biofilms. *Nat. Rev. Microbiol.* **2019**, *17*, 247–260. [CrossRef]
5. Herndl, G.J.; Bayer, B.; Baltar, F.; Reinthaler, T. Prokaryotic Life in the Deep Ocean's Water Column. *Annu. Rev. Mar. Sci.* **2023**, *15*, 461–483. [CrossRef]
6. Rodrigues, C.J.C.; de Carvalho, C.C.C.R. Cultivating Marine Bacteria under Laboratory Conditions: Overcoming the "Unculturable" Dogma. *Front. Bioeng. Biotechnol.* **2022**, *10*, 964589. [CrossRef]
7. Thomas, T.; Gilbert, J.; Meyer, F. Metagenomics—A Guide from Sampling to Data Analysis. *Microb. Inform. Exp.* **2012**, *2*, 3. [CrossRef]
8. Riesenfeld, C.S.; Schloss, P.D.; Handelsman, J. Metagenomics: Genomic Analysis of Microbial Communities. *Annu. Rev. Genet.* **2004**, *38*, 525–552. [CrossRef] [PubMed]
9. Laiolo, E.; Alam, I.; Uludag, M.; Jamil, T.; Agusti, S.; Gojobori, T.; Acinas, S.G.; Gasol, J.M.; Duarte, C.M. Metagenomic Probing Toward an Atlas of the Taxonomic and Metabolic Foundations of the Global Ocean Genome. *Front. Sci.* **2024**, *1*, 1038696. [CrossRef]
10. Pérez-Cobas, A.E.; Gomez-Valero, L.; Buchrieser, C. Metagenomic Approaches in Microbial Ecology: An Update on Whole-Genome and Marker Gene Sequencing Analyses. *Microb. Genom.* **2020**, *6*, mgen000409. [CrossRef]
11. Navgire, G.S.; Goel, N.; Sawhney, G.; Sharma, M.; Kaushik, P.; Mohanta, Y.K.; Mohanta, T.K.; Al-Harrasi, A. Analysis and Interpretation of Metagenomics Data: An Approach. *Biol. Proced. Online* **2022**, *24*, 18. [CrossRef]
12. Tully, B.J.; Sachdeva, R.; Graham, E.D.; Heidelberg, J.F. 290 Metagenome-Assembled Genomes from the Mediterranean Sea: A Resource for Marine Microbiology. *PeerJ* **2017**, *5*, e3558. [CrossRef]
13. Chaouni, B.; Idrissi Azami, A.; Essayeh, S.; Arrafiqui, E.H.; Bailal, A.; Raoui, S.; Amzazi, S.; Twaddle, A.; El Hamouti, C.; Boukhatem, N.; et al. Moroccan Lagoon Microbiomes. *Water* **2022**, *14*, 1715. [CrossRef]

14. Sánchez-Avila, J.; Meyer, J.; Lacorte, S. Spatial Distribution and Sources of Perfluorochemicals in the NW Mediterranean Coastal Waters (Catalonia, Spain). *Environ. Pollut.* **2010**, *158*, 2833–2840. [CrossRef]
15. Coll, M.; Piroddi, C.; Steenbeek, J.; Kaschner, K.; Lasram, F.B.R.; Aguzzi, J.; Ballesteros, E.; Bianchi, C.N.; Corbera, J.; Dailianis, T.; et al. The Biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats. *PLoS ONE* **2010**, *5*, e11842. [CrossRef]
16. Sanchez-Roman, A.; Jorda, G.; Sannino, G.; Gomis, D. Modelling Study of Transformations of the Exchange Flows along the Strait of Gibraltar. *Ocean Sci.* **2018**, *14*, 1547–1566. [CrossRef]
17. Abdulla, A.; Linden, O. Maritime Traffic Effects on Biodiversity in the Mediterranean Sea: Review of Impacts. In *Priority Areas and Mitigation Measures*; IUCN: Gland, Switzerland, 2008.
18. Joye, S.B.; Bowles, M.W.; Ziervogel, K. Marine Biogeochemical Cycles. In *The Marine Microbiome*; Stal, L.J., Cretoiu, M.S., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 623–671. [CrossRef]
19. Cossarini, G.; Feudale, L.; Teruzzi, A.; Bolzon, G.; Coidessa, G.; Solidoro, C.; Di Biagio, V.; Amadio, C.; Lazzari, P.; Brosich, A.; et al. High-Resolution Reanalysis of the Mediterranean Sea Biogeochemistry (1999–2019). *Front. Mar. Sci.* **2021**, *8*, 741486. [CrossRef]
20. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef]
21. Nurk, S.; Meleshko, D.; Korobeynikov, A.; Pevzner, P.A. metaSPAdes: A New Versatile Metagenomic Assembler. *Genome Res.* **2017**, *27*, 824–834. [CrossRef] [PubMed]
22. Ayling, M.; Clark, M.D.; Leggett, R.M. New Approaches for Metagenome Assembly with Short Reads. *Brief. Bioinform.* **2020**, *21*, 584–594. [CrossRef] [PubMed]
23. Mikheenko, A.; Saveliev, V.; Gurevich, A. MetaQUAST: Evaluation of Metagenome Assemblies. *Bioinformatics* **2016**, *32*, 1088–1090. [CrossRef]
24. Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E.; Kubal, M.; Paczian, T.; Rodriguez, A.; Stevens, R.; Wilke, A.; et al. The Metagenomics RAST Server—A Public Resource for the Automatic Phylogenetic and Functional Analysis of Metagenomes. *BMC Bioinform.* **2008**, *9*, 386. [CrossRef]
25. Wilke, A.; Bischof, J.; Gerlach, W.; Glass, E.; Harrison, T.; Keegan, K.P.; Paczian, T.; Trimble, W.L.; Bagchi, S.; Grama, A.; et al. The MG-RAST Metagenomics Database and Portal in 2015. *Nucleic Acids Res.* **2016**, *44*, D590–D594. [CrossRef]
26. Tatusov, R.L.; Galperin, M.Y.; Natale, D.A.; Koonin, E.V. The COG Database: A Tool for Genome-Scale Analysis of Protein Functions and Evolution. *Nucleic Acids Res.* **2000**, *28*, 33–36. [CrossRef]
27. Muller, J.; Szklarczyk, D.; Julien, P.; Letunic, I.; Roth, A.; Kuhn, M.; Powell, S.; von Mering, C.; Doerks, T.; Jensen, L.J.; et al. eggNOG v2.0: Extending the Evolutionary Genealogy of Genes with Enhanced Non-Supervised Orthologous Groups, Species and Functional Annotations. *Nucleic Acids Res.* **2010**, *38*, D190–D195. [CrossRef]
28. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [CrossRef]
29. Xue, C.-X.; Lin, H.; Zhu, X.-Y.; Liu, J.; Zhang, Y.; Rowley, G.; Todd, J.D.; Li, M.; Zhang, X.-H. DiTing: A Pipeline to Infer and Compare Biogeochemical Pathways From Metagenomic and Metatranscriptomic Data. *Front. Microbiol.* **2021**, *12*, 698286. [CrossRef]
30. Hyatt, D.; Chen, G.-L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification. *BMC Bioinform.* **2010**, *11*, 119. [CrossRef]
31. Aramaki, T.; Blanc-Mathieu, R.; Endo, H.; Ohkubo, K.; Kanehisa, M.; Goto, S.; Ogata, H. KofamKOALA: KEGG Ortholog Assignment Based on Profile HMM and Adaptive Score Threshold. *Bioinformatics* **2020**, *36*, 2251–2252. [CrossRef]
32. Topping, J.N.; Heywood, J.L.; Ward, P.; Zubkov, M.V. Bacterioplankton Composition in the Scotia Sea, Antarctica, during the Austral Summer of 2003. *Aquat. Microb. Ecol.* **2006**, *45*, 229–235. [CrossRef]
33. Gupta, R.S. The Phylogeny of Proteobacteria: Relationships to Other Eubacterial Phyla and Eukaryotes. *FEMS Microbiol. Rev.* **2000**, *24*, 367–402. [CrossRef]
34. Fernández-Gómez, B.; Richter, M.; Schüler, M.; Pinhassi, J.; Acinas, S.G.; González, J.M.; Pedrós-Alió, C. Ecology of Marine Bacteroidetes: A Comparative Genomics Approach. *ISME J.* **2013**, *7*, 1026–1037. [CrossRef]
35. Enisoglu-Atalay, V.; Atasever-Arslan, B.; Yaman, B.; Cebecioglu, R.; Kul, A.; Ozilhan, S.; Ozen, F.; Catal, T. Chemical and Molecular Characterization of Metabolites from *Flavobacterium* sp. *PLoS ONE* **2018**, *13*, e0205817. [CrossRef]
36. Hahnke, R.L.; Harder, J. Phylogenetic Diversity of Flavobacteria Isolated from the North Sea on Solid Media. *Syst. Appl. Microbiol.* **2013**, *36*, 497–504. [CrossRef]
37. Małecka-Adamowicz, M.; Kubera, Ł. Patterns of Structural and Functional Bacterioplankton Metacommunity along a River under Anthropogenic Pressure. *Sustainability* **2021**, *13*, 11518. [CrossRef]
38. Sui, Q.; Huang, J.; Deng, S.; Chen, W.; Yu, G. Occurrence and Removal of Pharmaceuticals, Caffeine and DEET in Wastewater Treatment Plants of Beijing, China. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 417–426. [CrossRef]
39. Gilbert, J.A.; Steele, J.A.; Caporaso, J.G.; Steinbrück, L.; Reeder, J.; Temperton, B.; Huse, S.; McHardy, A.C.; Knight, R.; Joint, I.; et al. Defining Seasonal Marine Microbial Community Dynamics. *ISME J.* **2012**, *6*, 298–308. [CrossRef]
40. Yu, Z.; Yang, J.; Liu, L.; Zhang, W.; Amalfitano, S. Bacterioplankton Community Shifts Associated with Epipelagic and Mesopelagic Waters in the Southern Ocean. *Sci. Rep.* **2015**, *5*, 12897. [CrossRef]
41. Evans, F.F.; Egan, S.; Kjelleberg, S. Ecology of Type II Secretion in Marine Gammaproteobacteria. *Environ. Microbiol.* **2008**, *10*, 1101–1107. [CrossRef]

42. Ghai, R.; Rodríguez-Valera, F.; McMahon, K.D.; Toyama, D.; Rinke, R.; de Oliveira, T.C.S.; Garcia, J.W.; de Miranda, F.P.; Henrique-Silva, F. Metagenomics of the Water Column in the Pristine Upper Course of the Amazon River. *PLoS ONE* **2011**, *6*, e23785. [CrossRef] [PubMed]
43. Pohlner, M.; Dlugosch, L.; Wemheuer, B.; Mills, H.; Engelen, B.; Reese, B.K. The Majority of Active Rhodobacteraceae in Marine Sediments Belong to Uncultured Genera: A Molecular Approach to Link Their Distribution to Environmental Conditions. *Front. Microbiol.* **2019**, *10*, 659. [CrossRef] [PubMed]
44. Pujalte, M.J.; Lucena, T.; Ruvira, M.A.; Arahal, D.R.; Macián, M.C. The Family Rhodobacteraceae. In *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*; Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 439–512. [CrossRef]
45. Morris, R.M.; Rappé, M.S.; Connon, S.A.; Vergin, K.L.; Siebold, W.A.; Carlson, C.A.; Giovannoni, S.J. SAR11 Clade Dominates Ocean Surface Bacterioplankton Communities. *Nature* **2002**, *420*, 806–810. [CrossRef] [PubMed]
46. Smith, D.P.; Thrash, J.C.; Nicora, C.D.; Lipton, M.S.; Burnum-Johnson, K.E.; Carini, P.; Smith, R.D.; Giovannoni, S.J. Proteomic and Transcriptomic Analyses of "Candidatus Pelagibacter Ubique" Describe the First PII-Independent Response to Nitrogen Limitation in a Free-Living Alphaproteobacterium. *mBio* **2013**, *4*. [CrossRef]
47. Priest, S. Flood Risk Research for Improving Flood Risk Outcomes. *J. Flood Risk Manag.* **2023**, *16*, e12888. [CrossRef]
48. Wani, A.K.; Akhtar, N.; Sher, F.; Navarrete, A.A.; Américo-Pinheiro, J.H.P. Microbial Adaptation to Different Environmental Conditions: Molecular Perspective of Evolved Genetic and Cellular Systems. *Arch. Microbiol.* **2022**, *204*, 144. [CrossRef] [PubMed]
49. Casciotti, K.L.; Marshall, T.A.; Fawcett, S.E.; Knapp, A.N. Advances in Understanding the Marine Nitrogen Cycle in the GEOTRACES Era. *Oceanography* **2024**, *37*, 85–101. [CrossRef]
50. Siokou-Frangou, I.; Christaki, U.; Mazzocchi, M.G.; Montresor, M.; Ribera d'Alcalá, M.; Vaqué, D.; Zingone, A. Plankton in the Open Mediterranean Sea: A Review. *Biogeosciences* **2010**, *7*, 1543–1586. [CrossRef]
51. Pfister, C.A.; Cardini, U.; Mirasole, A.; Montilla, L.M.; Veseli, I.; Gattuso, J.-P.; Teixido, N. Microbial Associates of an Endemic Mediterranean Seagrass Enhance the Access of the Host and the Surrounding Seawater to Inorganic Nitrogen under Ocean Acidification. *Sci. Rep.* **2023**, *13*, 19996. [CrossRef]
52. Hu, X.; Liu, J.; Liu, H.; Zhuang, G.; Xun, L. Sulfur Metabolism by Marine Heterotrophic Bacteria Involved in Sulfur Cycling in the Ocean. *Sci. China Earth Sci.* **2018**, *61*, 1369–1378. [CrossRef]
53. Urvoy, M.; Labry, C.; L'Helguen, S.; Lami, R. Quorum Sensing Regulates Bacterial Processes That Play a Major Role in Marine Biogeochemical Cycles. *Front. Mar. Sci.* **2022**, *9*, 834337. [CrossRef]
54. Pajares, S.; Ramos, R. Processes and Microorganisms Involved in the Marine Nitrogen Cycle: Knowledge and Gaps. *Front. Mar. Sci.* **2019**, *6*, 739. [CrossRef]