Exploring planktonic microbiome diversity in the coastal ecosystems of the North of Portugal

Simão Matos Conceição Horta

Extended summary

November 2022

1. Introduction

1.1 The relevance of marine microorganism

In the marine ecosystem, microorganisms are present everywhere[1]. Their diversity allowed them to be essential for biogeochemical cycling processes and therefore, crucial for the marine ecosystems and Earth´s climate[1]. These microorganisms are deeply involved in the carbon [1], nitrogen [1,2], phosphorus [3] and silica [4] cycles. Additionally, they are also involved in the marine food chains. The microbial loop is the most striking example[5]. The importance of marine microbial communities is notorious since microbes are key players in the marine ecosystems. Before the industrial revolution, these microbial communities adapted to the natural climate changes that happened in the marine environment. With the advent of the industrial revolution, the environmental conditions are now changing at higher rates than before, turning the global climate even more unpredictable and unstable. These new stresses and challenges are specially noticed by the marine microorganisms since they can rapidly adapt to the new conditions[6]. Thus, become important to understand if and how the microbial communities are actually changing. In this perspective, monitoring programs will allow us to be aware of the whole environmental picture, evaluating the marine microbial biodiversity, over time and space, and understand the possible

consequences of a certain factor. To keep track of the alterations in the microbial composition, constant and significant data must be available[8]. With such meaningful information available, researchers, alongside with politicians, can establish a solid maritime spatial planning, to successfully meet ecosystem preservation goals[7].

1.2 Marine Microbial Distribution

Marine microbial communities can change over space, time and over several environmental conditions. The spatial distribution of the microbial communities is shaped by the interplay of multiple factors. However, factors like dispersal, distance, and environmental gradient conditions, seem to be the ones that will influence more significantly the microbial composition across space [8,9]. The water column has also a gradient of environmental conditions that change according with depth. Salinity is also an important agent in the microbial spatial distribution. It is one of the biggest barriers to the dispersal of microorganism from freshwater to the oceans, and vice-versa[10], due to the selective halotolerance of each microbial species. The temporal distribution of the marine microorganisms is essential shaped by seasons, followed by the intersessional and interannual variability [11]. Additionally, the environmental parameters that are the main drivers of intersessional variability are the light, temperature, and nutrients [12]. However, the interaction of multiple parameters will be ultimately, the main factor that drives communities' assemblages. This interaction is often very complex since many variables are at play.

There are several monitoring campaigns around the world that are trying to characterize the marine microbial communities. In Portugal, the monitoring programs that tried to describe the distribution and composition of the marine microbial communities are rare. In order to fill this gap, the present thesis aims to evaluate and catalog the biodiversity of marine microorganisms on the north coast of Portugal. The present dissertation aims to meet three main objectives. The first objective proposed is to evaluate the spatial biogeography of the microplankton communities in the northern coast of Portugal. The second goal is to determine the microbiome distribution across an estuary gradient. In this topic the Douro estuarine samples are going to be analyzed spatially, across a transect and at depth, and temporally, between the Autumn and Winter season. The third objective proposed was to evaluate the impact of the tides in the estuarine microplankton communities in the Douro River. The samples taken in 2016, will be analyzed and a posterior assessment will be done to understand the tide effect on communities' assemblages.

2. Materials and methods

The research took part in sampling of the coast NW of Portugal (integrated in the Atlântida campaign), plus the sampling of the Douro estuary. The samples from the NW Portuguese coastal area, plus the ones from the river will allow a much deeper and integrated understanding of the

communities' dynamics in the northern region of Portugal. Additionally, it was also integrated a third dataset from 2016 where it will be possible to understand the tides influence in the microbial assemblages. The surface water samples were taken from predefined stations that stretch across the coastal and estuarine zones. Additionally, in each station, the multiparametric probe was submerged in order to save the CTD profile for posterior analyses. The probe measured a wide range of parameters like the depth of the station (meters), the temperature (°C), the salinity (PSU), and the pH. After sample collection and storage, the water was filtrated to discard the liquid phase and recover the biomass present in the water sample. At this step, the particulate environmental DNA (e-DNA) retrieved in the Sterivex cartridges was isolated. To do that, the DNeasy® PowerWater Sterivex Kit (QIAGEN) was selected and the manufacturer instructions followed. The quantification step followed, enabling to understand the quantity of DNA extracted from the previous step and realize if it was successful or not. The kit used to do the quantification was Invitrogen Qubit™ 4 Fluorometer (Thermo Fisher Scientific Inc.). The Polymerase chain reaction (PCR) was a molecular technique applied to amplify specific target regions of the 16S rRNA gene and 18S rRNA gene able to describe microplankton diversity harbored in the water samples collected. In the sequencing step, the samples were sent to the Integrated Microbiome Resource (IMR) group, however the samples from 2016 were sent to the LGC group. In any case, the library preparation was accomplished by PCR amplification of specific regions of the ribosomal RNA genes. The amplification of the 16S ribosomal RNA gene was performed by using the primer pair 515F (5′-GTGYCAGCMGCCGCGGTAA-3′) and the recently revised 926R (5'-CCGYCAATTYMTTTRAGTTT-3') [13], specially created to target the V4-V5 hypervariable regions in prokaryotes [14]. The Region V4 of the 18S ribosomal RNA gene was amplified using the forward primer TAReuk454FWD1 (5′-CCAGCASCYGCGGTAATTCC-3′) and TAReukREV3_modified (5′-ACTTTCGTTCTTGATYRATGA-3′) [15] in unicellular eukaryotes (protists).

In the meantime, the nutrient analyses were carried out. To do that, the water from each station was collected and filtered through a cellulose acetate membrane with 0.45 µm pore´s wide. After that, to retrieve the concentration of ammonia, nitrites, nitrates, phosphate, chlorophyll *a*, and silica, in each station, a specific protocol was followed. After the arrival of the sequencing datasets, the bioinformatic workflow can take place. The filtering of the raw Illumina fastq files, into high-quality data must be assured to allow the proper analyses. The paired-end reads were treated using the DADA2 pipeline version 1.16.0, applying the standard parameters. The downstream analyses used the data generated previously by the DADA2, to investigate the distribution and diversity of the microplankton community in the different datasets. These analyses rely mostly on the phyloseq package (v. 1.22.3), that was loaded on the Rstudio, in order to execute their functions and tools.

3. Results

3.1 Microbiome in the Douro River Estuary

Biogeochemical Gradients across the Douro River Estuary

The seasonal biogeochemical gradients in the Douro estuary demonstrate higher concentrations of temperature, chlorophyll-a, pH, nitrates and phosphor in the Autumn than in the Winter. On the contrary, the concentrations of silica and nitrites showed to be higher in the Winter than in the Autumn. The nitrogen forms like nitrates and nitrites were mainly present in the upstream regions. The downstream regions displayed a higher salinity, and a higher concentration of ammonia. Regarding depth, no major differences were noticed.

Alpha and beta diversity

For alpha diversity analyses, the Shannon index showed that in the Autumn, the gradient of prokaryotic diversity increases from the estuary mouth to the upstream region of the estuary. In the Winter, the prokaryotic diversity was generally higher in the most downstream part of the transept, decreasing towards the upstream region. The prokaryotic beta diversity analyses allowed to visualize a clearly prokaryotic community's separation. Results showed that samples from the Winter campaign clustered together, although the samples from the Autumn campaign did not cluster as tightly (Figure 1a). For the 18S rRNA gene dataset it was also possible to notice four clusters, (Figure 1b). A close analysis showed that the samples from Winter and Autumn were more disperse in the plot, showing that the unicellular eukaryotic communities were not as well separated by season, as the prokaryotic communities (Figure 1).

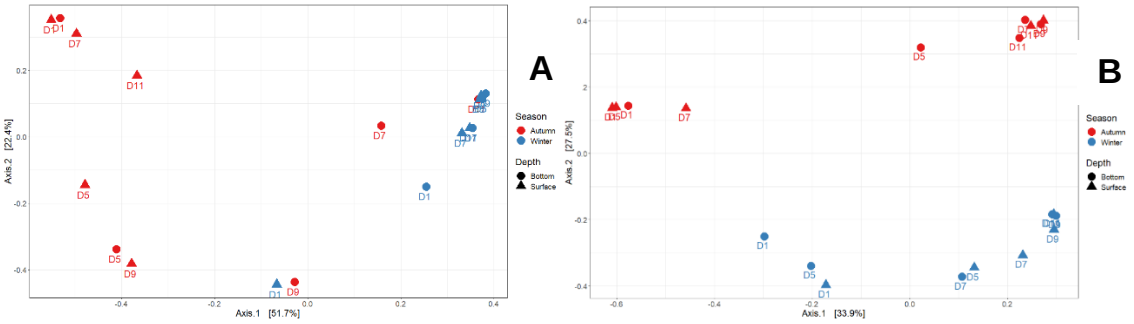

Figure 1: Beta diversity analyses of the 16S rRNA gene (A) and 18S rRNA gene Douro dataset (B), collected during 2021/2022; In each figure was presented the PCoA ordination plot with the different depths (surface and bottom) distinguished by different shapes (circle and triangle), and the different seasons (Autumn and Winter) were notable with a different color (red and clue).

Environmental drivers of microbiome distribution

Additionally, a PERMANOVA was carried out to understand which were the environmental parameters that shaped more significantly the microbial community's distribution. From this analysis it was possible to notice that the most significant conditions that shaped the prokaryotic communities were the temperature (R2=0.2290, p-value=0.001) and nitrate concentration

(R2=0.133, p-value=0.043). For the unicellular eukaryotic communities, the most relevant variables were pH (R2=0.127, p-value=0.010), temperature (R2=0.237, p-value=0.001), salinity (R2=0.124, p-value=0.019), ammonia (R2=0.1424, p-value=0.014), nitrite (R2=0.132, p-value=0.015) and nitrate (R2=0.159, p-value=0.002) concentrations.

3.2 Tides impact in the Douro River Estuary

Biogeochemical Gradients across the Douro River Estuary

In most of the environmental parameters, no patterns were noticed regarding the effect of tides. However, salinity was always higher in the stations that were near the estuary mouth in both tides. Regarding temperature, was it possible to notice a difference of 0.8 ºC between tides, in the stations R1H and R1L. Ammonia increases substantially in the station R1 at high tide and phosphate was generally higher when low tide was happening.

Alpha and beta diversity

In the prokaryotic dataset the samples R1H, R1L and R2L presented higher diversity levels Regarding tide variation, the only remarkable station that displayed a higher Alpha diversity metrics fluctuations was the station R2. In its turn, the unicellular eukaryotic communities showed similar patterns, since R1H and R2L showed to be a higher Shannon diversity. Regarding the tides variability in alpha diversity were especially noticed in the R1 and R2 stations. For the beta diversity, the results were similar for both prokaryotic and unicellular eukaryotic communities, showing that the stations R1H, R1L and R2L strongly differ in terms of community structure (Figure 2). Also, the samples closer to the mouth of the estuary show a higher variation between tides since R1H, R1L, R2H and R2L cluster separately. On the contrary, the samples more upstream (R3, R4, R5) cluster together showing the increase similarity in the microbial composition on these parts of the transect.

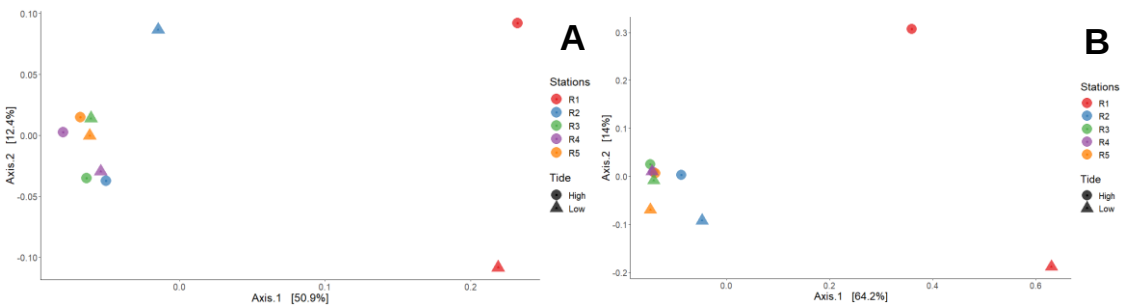

Figure 2: Beta diversity analyses of the Douro samples collected during 2016, for 16S rRNA gene (figure 2a) and 18S rRNA gene (figure 2b) datasets. The analyses were graphically represented on the PCoA ordination plot, where the different stations have different colors, and the tides (High and low) were distinguished by different shapes (circle and triangle).

Environmental drivers of microbiome distribution

The PERMANOVA analyses showed that across all the transect, the tides did not display any statistically significant influence in the prokaryotic communities (R2=0.063, p-value= 0.764) and in the unicellular eukaryotic communities (R2=0.061, p-value= 0.690). However, temperature (R2=0.4365, p-value= 0.011), salinity (R2=0.4643, p-value= 0.004) and pH (R2=0.2730, p-value= 0.006) were a significant condition that shaped the prokaryotic communities (Annex-11 a.). By its turn, the unicellular eukaryotic communities were mostly influenced by the temperature (R2=0.4010, p-value= 0.007), salinity (R2=0.4573, p-value= 0.001) and pH (R2=0.2968, p-value= 0.031) concentrations.

3.3 Microbiome distribution along the NW coast of Portugal

Biogeochemical Gradients across the NW coast of Portugal

The spatial biogeochemical gradients in the Douro estuary demonstrate higher concentrations of temperature, chlorophyll *a* and salinity in the regions twelve miles from the coast. On the contrary, environmental parameters like ammonia, nitrite, nitrate, silica and phosphate demonstrate higher concentrations in two miles from the coast. Regarding the longitudinal gradient, it was possible to notice an increase concentration in the nitrites, chlorophyll *a,* and in temperature in north on the transept. In the south, the higher concentrations were from phosphate, ammonia, and pH.

Alpha and beta diversity

An overall increase in the diversity metrics was registered, where the lowest Shannon values were registered south. From there, the values gradually increase towards the north region. The analyzes of the unicellular eukaryotic communities demonstrate no important patterns in the diversity metrics. For beta diversity it was possible to notice that the prokaryotic communities were separated in two clusters according to their different longitudinal amplitudes. Within each of these clusters was possible to notice that were separated based on the distance to the coast (Figure 3). The unicellular eukaryotic communities (18S rRNA gene dataset) samples were distributed within two clusters. With these arrangement was possible to notice that there was not a clear longitudinal pattern, like the one seen for the prokaryotic communities.

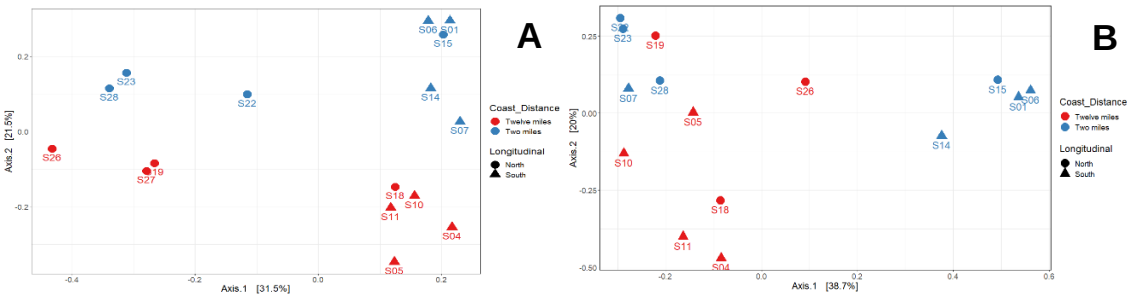

Figure 3: Beta diversity analyses of the NW Portuguese Coast dataset; PCoA ordination plot with the 16S rRNA gene (Figure 3a) and 18S rRNA gene (Figure 3b) samples distinguished by the distance from the coast (two miles and twelve miles from the coast) with different colors (red and blue) and distinguished by the longitudinal gradient (north and south) with different shapes (circle and triangle).

Environmental drivers of microbiome distribution

The PERMANOVA results, the prokaryotic community's distribution along the NW coast of Portugal were also strongly influenced by environmental parameters like pH (R2=0.2290, p-value=0.001), ammonia (R2=0.1218, p-value=0.039), nitrite (R2=0.1830, p-value=0.003), nitrate (R2=0.1308, p-value=0.043), and phosphate (R2=0.1545, p-value=0.011). Additionally, for the unicellular eukaryotic communities, showed that the environmental parameters pH (R2=0.1474, p-value=0.020), chlorophyl a (R2=0,1351, p-value=0.028), silica (R2=0.1479, p-value=0.026) and phosphate (R2=0.1519, p-value=0.023), were statistical significantly in shaping these group of microbial communities.

4. Discussion

4.1 Estuarine Microbiome Horizontal distribution

Although not supported statistically, the alpha and beta diversity analyses showed a particular spatial pattern for the prokaryotic microbiomes. In the beta diversity it was possible to notice that the upstream and downstream regions were clustered separately. By its turn, the unicellular eukaryotic microbial communities showed statistically significant differences across the different stations. Additionally, the beta diversity showed a clear distinction between the upper and lower areas of the estuarine transect. These results lead us to robustly conclude that the similarity of the unicellular eukaryotic communities decreases more strongly than prokaryotic communities, as geographic distance increases, even between small distances. The same conclusions were obtained by Liu et al. (2013)[16], arguing that these findings could be explained by the better dispersal ability that bacteria possess, since they were smaller than microbial eukaryotes. Additionally, since eukaryotes were more complex than prokaryotes, they tend to get their appendages and feeding apparatuses damaged more frequently when carried by the watercourses[17,18]. This will result in lower rates of colonization and therefore the of dispersal of

these populations were constrained to a small geographic area. The results showed that depth was not statistically significant for both prokaryotic and unicellular eukaryotic communities. Additionally, there was no pattern concerning depth in the alpha and beta analysis. Similar results were found by Seguro et al. (2015)[19] in the Gulf of Nicoya. However, the influence of depth was noticed by Ohore et al. (2022)[20] in the estuary of the Rongjiang River. Such results were expected where depth can have more influence where it produces enough environmental variation in nutrients availability, pressure, and temperature, to shape the microbial communities.

4.2 Influence of Estuarine Tides

The influence of tides in the surface microbial communities of the Douro estuary showed to be not significant when a PERMANOVA analysis was performed, for both prokaryotic and unicellular eukaryotic communities. Nevertheless, there were some trends that can be discussed. Mainly, it was found that the Douro estuary tides did not significantly influence the microbial communities located in the area of the estuary more influenced by the river. However, the microbial communities from the downstream station were highly influenced by tides showing dissimilar communities at high and low tide.

4.3 Estuarine Microbiome Distribution with Season

The prokaryotic communities were also shaped significantly by seasonality, supported by either statistically and alpha and beta diversity analyses. In agreement, the unicellular eukaryotic communities also showed that season was a statistically significant. Nevertheless, the alpha and beta diversity analyses, performed for the unicellular eukaryotic dataset, showed small differences and less clear clusters differentiation between seasons, when compared with the prokaryotic communities. This means that prokaryotic communities were more influenced by seasons than unicellular eukaryotic communities. In this study, it was possible to notice that the majority of the environmental parameters registered have a degree of seasonality. Conditions like temperature, pH, nitrates, phosphates were at higher levels in Autumn while nitrites and silica were at higher levels in Winter. The higher concentration of nutrients in Autumn can be due to the low rainfall that was registered in this period, which led to decrease in the river flow and consequently the decrease of the dilution of the nutrients in the estuary[16,21].

4.4 Spatial Microbiome Distribution Across the NW Coast of Portugal

The statistical and the diversity metrics support the fact that the longitudinal and latitudinal gradients of the NW coast of Portugal influence greatly the prokaryotic community's distribution. In its turn, the unicellular eukaryotic communities did not show the same robust patterns. This means that the prokaryotic communities were much more influenced by distance-decay

relationships than the unicellular eukaryotic communities, in the Portuguese NW coast. These strong spatial patterns were commonly explained by the hydrological processes like the upwelling, the magnitude of mixing by river flow, the winds regimes, the fluctuation of tides, etc[22]. The physical processes were also tightly related with the environmental gradients across the coastal areas.

5. Conclusion

Spatially, the unicellular eukaryotic communities, inhabiting the surface waters of the Douro estuary, were increasingly dissimilar as the distance between stations increases. In its turn, the prokaryotic communities, inhabiting the surface waters of the Douro estuary, showed a less pronounced dissimilarity as the distance between stations increases. The microbial communities from the downstream station were highly influenced by tides showing dissimilar communities at high and low tide. Depth was not significantly influencing the microbial communities in the Douro estuary at high tide. Seasonally, Autumn and Winter significantly shaped the prokaryotic communities and the unicellular eukaryotic communities of the Douro estuary, although prokaryotic communities were the ones that showed to be more influenced. Spatially, it was found that the prokaryotic communities were increasingly dissimilar as the distance between stations increases. In its turn, the unicellular eukaryotic communities, inhabiting the surface waters of the NW coast of Portugal, showed a less pronounced dissimilarity as the distance between stations increases.

The present thesis delivers for the first time new molecular data on the diversity of prokaryotes and unicellular eukaryotes for the Douro estuary and coastal zone associated with it. This research will help to enhance our understanding of such important ecosystems, providing a theoretical foundation for the marine ecological health management.

6. References

1. Stal, L. J. & Cretoiu, M. S. *The marine microbiome: An untapped source of biodiversity and biotechnological potential*. *The Marine Microbiome: An Untapped Source of Biodiversity and Biotechnological Potential* (Springer International Publishing, 2016). doi:10.1007/978-3-319-33000-6.

2. Zehr, J. P. & Kudela, R. M. Nitrogen cycle of the open ocean: From genes to ecosystems. *Ann Rev Mar Sci* **3**, 197–225 (2011).

3. Paytan, A. & McLaughlin, K. The oceanic phosphorus cycle. *Chem Rev* **107**, 563–576 (2007).

4. Tréguer, P. J. & de La Rocha, C. L. The world ocean silica cycle. *Ann Rev Mar Sci* **5**, 477–501 (2013).

5. Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A., & Thingstad, F. (1983). The ecological role of water-column microbes in the sea. Marine ecology progress series, 257-263.

6. Jeffries, T. C. *et al.* Bacterioplankton dynamics within a large anthropogenically impacted urban estuary. *Front Microbiol* **6**, (2016).

7. Douvere, F. & Ehler, C. N. The importance of monitoring and evaluation in adaptive maritime spatial planning. *J Coast Conserv* **15**, 305–311 (2011).

8. Martiny, J. B. H. *et al.* Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology* vol. 4 102–112 Preprint at https://doi.org/10.1038/nrmicro1341 (2006).

9. Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C. & Martiny, J. B. H. Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology* vol. 10 497–506 Preprint at https://doi.org/10.1038/nrmicro2795 (2012).

10. Fortunato, C. S. & Crump, B. C. Microbial gene abundance and expression patterns across a river to ocean salinity gradient. *PLoS One* **10**, (2015).

11. Fuhrman, J. A., Cram, J. A. & Needham, D. M. Marine microbial community dynamics and their ecological interpretation. *Nature Reviews Microbiology* vol. 13 133–146 Preprint at https://doi.org/10.1038/nrmicro3417 (2015).

12. Giovannoni, S. J. & Vergin, K. L. Seasonality in ocean microbial communities. *Science* vol. 335 671–676 Preprint at https://doi.org/10.1126/science.1198078 (2012).

13. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME Journal* **6**, 1621–1624 (2012).

14. Apprill, A., Mcnally, S., Parsons, R. & Weber, L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology* **75**, 129–137 (2015).

15. Piredda, R. *et al.* Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. *FEMS Microbiol Ecol* **93**, (2017).

16. Liu, L., Yang, J., Yu, X., Chen, G. & Yu, Z. Patterns in the composition of microbial communities from a subtropical river: Effects of environmental, spatial and temporal factors. *PLoS One* **8**, (2013).

17. External control of bacterial community structure in lakes.

18. Bergström, A. K. & Jansson, M. Bacterioplankton production in humic Lake Ortrasket in relation to input of bacterial cells and input of allochthonous organic carbon. *Microb Ecol* **39**, 101–115 (2000).

19. Seguro, I. *et al.* Seasonal changes of the microplankton community along a tropical estuary. *Reg Stud Mar Sci* **2**, 189–202 (2015).

20. Ohore, O. E. *et al.* Vertical characterisation of phylogenetic divergence of microbial community structures, interaction, and sustainability in estuary and marine ecosystems. *Science of the Total Environment* **851**, (2022).

21. Zhou, L. *et al.* Environmental filtering dominates bacterioplankton community assembly in a highly urbanized estuarine ecosystem. *Environ Res* **196**, (2021).

22. Ghiglione, J. F. , L. M. , & L. P. Spatial and temporal scales of variation in bacterioplankton community structure in theNW Mediterranean Sea. *Aquatic Microbial Ecology* **40**, 229–240 (2005).