# Metagenomics-Based Microbial Ecological Community Threshold and Indicators of Anthropogenic Disturbances in Estuarine Sediments

Bao, Yingyu; Ruan, Yuefei; Wu, Jiaxue; Wang, Wen-Xiong; Leung, Kenneth M. Y.; Lee, Patrick K. H.

**Published in:**
Environmental Science & Technology

**Published: 09/01/2024**

**Document Version:**
Post-print, also known as Accepted Author Manuscript, Peer-reviewed or Author Final version



**Published version (DOI):**
10.1021/acs.est.3c08076

**Publication details:**
Bao, Y., Ruan, Y., Wu, J., Wang, W.-X., Leung, K. M. Y., & Lee, P. K. H. (2024). Metagenomics-Based Microbial Ecological Community Threshold and Indicators of Anthropogenic Disturbances in Estuarine Sediments. *Environmental Science & Technology*, *58*(1), 780–794. https://doi.org/10.1021/acs.est.3c08076

# Metagenomics-based microbial ecological community threshold and indicators of anthropogenic disturbances in estuarine sediments

Yingyu Bao,[1] Yuefei Ruan,[2,3] Jiaxue Wu,[3] Wen-Xiong Wang,[1] Kenneth M. Y. Leung,[1,2,3] and Patrick K. H. Lee[1]*

[1]School of Energy and Environment and State Key Laboratory of Marine Pollution, City University of Hong Kong, Hong Kong SAR, China

[2]Department of Chemistry, City University of Hong Kong, Hong Kong SAR, China

[3]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519000, China

**Correspondence:** *B5423, Yeung Kin Man Academic Building, School of Energy and Environment, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong SAR, China; E-mail: patrick.kh.lee@cityu.edu.hk; Tel: (852) 3442-4625; Fax: (852) 3442-0688.

**Abstract**

Assessing the impacts of cumulative anthropogenic disturbances on estuary ecosystem health is challenging. Using spatially distributed sediments from the Pearl River Estuary (PRE) in southern China, which is significantly influenced by anthropogenic activities, we demonstrated that metagenomics-based surveillance of benthic microbial communities is a robust approach to assess anthropogenic impacts on estuarine benthic ecosystems. Correlational and threshold analyses between microbial compositions and environmental conditions indicated that anthropogenic disturbances in the PRE sediments drove the taxonomic and functional variations in benthic microbial communities. An ecological community threshold of anthropogenic disturbances was identified, which delineated the PRE sediments into two groups (H and L) with distinct taxa and functional traits. Group H, located nearshore and subjected to a higher level of anthropogenic disturbances, was enriched in pollutant degraders, putative human pathogens, fecal pollution indicators, and functional traits related to stress tolerance. In contrast, Group L, located offshore and subjected to a lower level of anthropogenic disturbances, was enriched in halotolerant and oligotrophic taxa and functional traits related to growth and resource acquisition. The machine learning random forest model identified a number of taxonomic and functional indicators that could differentiate PRE sediments between Groups H and L. The identified ecological community threshold and microbial indicators highlight the utility of metagenomics-based microbial surveillance in assessing the adverse impacts of anthropogenic disturbances in estuarine sediments, which can assist environmental management to better protect ecosystem health.

**Synopsis:** Metagenomics-based surveillance of benthic microbial communities is an effective tool to assess cumulative anthropogenic impacts on estuarine benthic ecosystems.

**Keywords:** anthropogenic impacts, estuarine sediments, ecological community threshold, microbial community, microbial indicator

## Introduction

Estuaries are vital ecosystems that sustain human and marine life and provide diverse ecological services[1]. Given the rapid population growth and development in coastal regions worldwide, anthropogenic impacts on estuaries are a growing concern[2]. The uncontrolled release of environmental pollutants (e.g., nutrients, metals, pathogens, and other micropollutants) has resulted in their accumulation in estuarine sediments, leading to the formation of a transition zone with a steep gradient of anthropogenic disturbances between land and sea (hereinafter referred to as the anthropogenic gradient)[3-5]. The adverse impacts on ecosystem health along such an anthropogenic gradient are challenging to assess and monitor[6,7].

Benthic organisms' adaptive or passive responses to anthropogenic disturbances may increase their stress[8,9]. Consequently, abrupt alteration of species distribution may occur at a critical point in an anthropogenic gradient, which is referred to as an ecological community threshold[10-12]. Anthropogenic disturbances that exceed this threshold could decrease biodiversity due to loss of intolerant species[13], resulting in changes in ecosystem goods and services[6,14]. Studies applying this concept of ecological community threshold have used vegetation[15] and animals[16] as indicators to assess anthropogenic impacts, but the use of microbial responses for such purposes has been limited[13]. Environmental microbes and their functional genes have emerged as robust indicators of various anthropogenic disturbances such as metal pollution[17], eutrophication[4], and urbanization[13]. This is because microbes respond rapidly and sensitively to trace-level environmental stressors[18,19], with such responses manifesting across the biological hierarchy of gene, organism, and community[20], offering a comprehensive perspective for field-based assessments of anthropogenic disturbances[21,22]. An ecological community threshold based on microbial responses can thus be useful for developing environmental criteria for the conservation of biodiversity and ecosystem functions[6,9].

Metagenomic sequencing is an invaluable tool to assess the taxonomic and functional profiles of indigenous microbial communities[23]. The phylogenetic markers and genes in metagenomes can provide a comprehensive view of the *in situ* taxonomic diversity and functions of microbes, while the key functional traits in metagenomes can reflect the optimized microbial life-history strategies applied by microbes to adapt to different environmental conditions[24,25]. Life-history strategies represent a crucial dimension of a microbial species' ecological niche because these strategies are considered to have evolved via constraints among functional traits that have had consequences for reproduction and cell fitness under different environments[26]. Furthermore, life-history strategies can explicitly reflect the relationships between microbes and environmental conditions[24]. However, whether and how anthropogenic disturbances are recorded in the microbial life-history strategies of microbial communities in estuarine sediments remain largely unknown.

The Pearl River Delta metropolitan region in southern China, which is centered around the Pearl River Estuary (PRE), is among the most populated and industrialized areas worldwide. Substantial evidence has shown that PRE sediments encounter intensive anthropogenic disturbances, such as eutrophication[27], metals[28], and organic pollutants[29]. Hence, the PRE is an ideal model ecosystem in which to test the hypothesis that microbial taxa and functions strongly correspond to anthropogenic disturbances along a gradient in estuarine sediments. In this study, metagenomic sequencing was used to determine variations in the taxonomic and functional compositions of microbial communities along an anthropogenic gradient in the PRE sediments. Using the microbial responses to anthropogenic disturbances in a gradient, we showed that an ecological community threshold could be established and that the life-history strategies of the microbial communities delineated by the threshold varied. Furthermore, machine learning (ML) identified taxonomic and functional indicators of the anthropogenic disturbances in these sediments. This study demonstrates and supports the use of metagenomics-based microbial

surveillance to assess anthropogenic impacts on ecosystem health in estuarine sediments.

## Materials and Methods

### Sampling sites and sample collection

Two large-scale sampling events were conducted at 18 and 16 PRE sites in July 2020 and 2021 (eight sites were re-sampled between the 2 years; these were geographically close but non-identical between the years), respectively (**Fig. 1a**). The GPS coordinates of the sampling sites were displayed on a map by Ocean Data View (v.5.6.3)[30] and are shown in **Table S1**. The sampling sites were chosen to provide a wide spatial coverage of PRE sediments from nearshore to offshore (~5,116 $km^2$). To investigate aerobic microbial communities and relatively recent pollution, surface sediments (depth: 0–10 cm) were collected in triplicate per sampling site and per timepoint using a Van Veen stainless steel grab sampler and homogenized into a composite sample. The composite samples were stored in sterile bags in the dark at 4°C during transport.

### Physicochemical analysis and the anthropogenic index

The environmental conditions of sediments at each site were evaluated based on 11 anthropogenic factors and 5 indigenous estuarine factors. The anthropogenic factors, namely total nitrogen, total phosphorus, total sulfur, total organic carbon, and seven metal species (iron, manganese, arsenic, cobalt, copper, lead, and cadmium), were measured using ~100 mg of composite surface sediments to indicate nutrient, organic, and metal pollution. The indigenous estuarine factors other than water depth, namely temperature, dissolved oxygen, salinity, and turbidity, were measured *in situ* at a depth of 1 m above the seabed. The measurement methods and results are shown in **Table S1**.

The anthropogenic disturbance level at each site was represented by an anthropogenic

index obtained using a model for calculating the cumulative anthropogenic impacts on multiple types of marine ecosystems[31,32]. However, because only estuarine sediments were considered as the target ecosystem in this study, the model was simplified as shown in **Text S1**. The average of the unweighted z-scores of the 11 anthropogenic factors was used to calculate the anthropogenic index at each site. The z-scores the 11 anthropogenic factors were calculated using the "scale" function in R (v.4.1.1)[33]. All anthropogenic factors were equally weighted because their impacts on microbial community compositions were shown to be similar (for details of the comparison of the individual effects of each anthropogenic factor, see **Text S1**).

**DNA extraction, metagenomic sequencing, and bioinformatic analysis**

Genomic DNA was extracted from ~300 mg of composite surface sediments using the FastDNA™ SPIN Kit for Soil (MP Biomedical, OH, USA) according to the manufacturer's instructions. Three mixtures of only extraction reagents were processed in parallel as negative controls. Sequencing libraries were generated using the Rapid Plus DNA Lib Prep Kit for Illumina (Abclonal, MA, USA) according to the manufacturer's protocols, and the library quality was assessed on an Agilent Bioanalyzer Agilent 5400 system (Agilent Technologies, CA, USA). Paired-end 150-bp sequencing of the libraries was performed on an Illumina NovaSeq 6000 platform (Novogene Co., Ltd., Beijing, China) according to the manufacturer's protocol. The raw sequences have been deposited in the NCBI Sequence Read Archive (BioProject accession number: PRJNA958616). High-quality reads were obtained after quality control and contaminant removal; **Table S2** shows the number of reads retained per sample. The high-quality reads were used for taxonomic and functional profiling and pathogen identification (**Table S3**). Bioinformatic analysis was performed as described in **Text S2**.

**Determination of taxonomic responders and the ecological community threshold**

The algorithm Threshold Indicator Taxa ANalysis (TITAN; v.2.4.1)[9] was executed by the function "titan" in the R package "TITAN2" (v.2.4.1) to identify the most probable change points along an environmental gradient where synchronous changes in the relative abundances of most of the taxa in the community would occur. Briefly, TITAN detects the changes in the relative abundance and occurrence frequency of individual species along a gradient by calculating a scaled indicator value score[34], and the change points of individual species are identified where changes in their relative abundances are maximized along the gradient. Species whose relative abundance increased and decreased with increasing anthropogenic index (collectively, taxonomic responders) are deemed increasers and decreasers, respectively. TITAN uses bootstrapping resampling ($n$ = 500) and permutation ($n$ = 250) procedures to estimate the reliability, purity, and uncertainty of responders and change points to ensure robustness. A responder is deemed statistically reliable if (i) 95% of 500 bootstrap replicates are significantly different from a random distribution ($P < 0.05$, high reliability) and (ii) the changes in its occurrence frequency and abundance are in the same direction for 95% of the 500 bootstrap replicates ($P < 0.05$, high purity). In our analysis, species present in at least 25% of the samples were included in TITAN. A maximum likelihood phylogenetic tree of the identified taxonomic responders was constructed using NCBI Taxonomy[35] and visualized using iTOL (v.6.7.3)[36]. Based on the change points of individual species, TITAN calculates the community-level change points for increasers [i.e., CP(+)] and decreasers [i.e., CP(-)] by considering where the most synchrony in the change points of increasers and decreasers is observed. In this study, CP(-) was adopted as the ecological community threshold (see results below).

**Differential function analysis and inference of life-history strategies**

Differences in the functional profiles of communities delineated by the ecological community threshold were assessed using permutational multivariate analysis of variance (PERMANOVA), permutational analysis of multivariate dispersion (PERMDISP) and non-metric multidimensional scaling (NMDS) based on Bray–Curtis dissimilarity by the "metaMDS" function in the R package "vegan" (v.2.5-7). The similarity percentage (SIMPER) analysis with 1,000 permutations was executed by the "simper" function in the R package "vegan" to identify functions that significantly contributed to the dissimilarity ($P < 0.05$), with those that varied > two-fold in relative abundances designated as the key functional traits that drove the functional differences between microbial communities delineated by the ecological community threshold[37]. The identified key functional traits were further analyzed to determine the microbial life-history strategies. The key functional traits were first grouped into their ancestral classes, which were referred to as the functional sub-levels (levels 2 and 3) according to hierarchical classifications from public databases (e.g., Gene Ontology, Kyoto Encyclopedia of Genes and Genomes, and UniProt) or the literature. Subsequently, the functional sub-levels were manually classified into three main strategies (level 1)—growth yield (Y), resource acquisition (A), and stress tolerance (S)—according to the trait-based classification scheme of the Y-A-S framework, as defined previously[24]. The Y-A-S framework reflects the microbial life-history strategies in different scenarios with different levels of resources and stresses[24, 25]. The Y-strategy refers to the maximization of microbial growth yield, which includes the functional sub-levels for enhancing central metabolism and biosynthesis and predominates in unstressed and resource-abundant habitats. The A-strategy refers to cells' enhanced investment in gaining resources, which includes the functional sub-levels for improving motility, competition, substrate transportation, and biodegradation and prevails in unstressed but resource-scarce habitats. The S-strategy refers to the capability to enhance cell tolerance under

environmental stresses, which includes the functional sub-levels such as sporulation, biomolecular damage repair, and osmotic protection. Functions that were poorly classified were grouped under the unclassified (U) strategy. The classification results were confirmed by searching the key functional traits against the microbial trait database proposed by Karaoz et al.[38], which was used for classifying genomics markers into Y-A-S strategies.

**ML-based identification of microbial indicators**

ML models were developed based on the relative abundances of the identified taxonomic responders and key functional traits to distinguish sediments with anthropogenic indexes above and below the ecological community threshold. Six models that have been commonly used to identify indicators[39,40]—decision tree, gradient-boosting decision tree, linear discriminant analysis, Naïve Bayes, random forest (RF), and support vector machine—were applied and compared to identify the model with the best performance. Training and evaluation of all models were performed using the Python package scikit-learn (v.1.1.3)[41] (for details, see **Text S2**). The top 20 taxonomic or functional predictors with the highest mean decrease accuracies in the best performing model were designated as indicators. The potential correlations between the taxonomic indicators and the top 500 most abundant species, and between the taxonomic and functional indicators were investigated using network analysis (for the methods to construct and visualize the networks, see **Text S2**).

**Data analysis and statistics**

Data analyses and statistical tests—pairwise geographical distance between sampling sites, linear and non-linear relationships between variables, principal component analysis (PCA) of

213 the environmental parameters, Spearman's correlations, Wilcoxon test, and partial Mantel
214 test—were performed using R packages (for details, see **Text S2**). Quantitative data are
215 presented as means ± standard deviations.
216
217 **Results**
218 **Anthropogenic gradient in the PRE sediments**
219 PCA of the 16 environmental parameters measured in the samples collected in 2020 and
220 2021 revealed nonsignificant differences in the environmental conditions under which the PRE
221 sediments were collected between the 2 years (PERMANOVA, $R^2$ = 0.037, $P$ = 0.237;
222 PERMDISP, $P$ = 0.533). The 11 anthropogenic factors (red vectors, **Fig. 1b**) exhibited
223 collinearity (**Fig. S1**), and their respective loadings on the first principal component (PC1) were
224 high and comparable to one another (0.287 ± 0.049) (**Fig. S2**). The anthropogenic index,
225 calculated by integrating the 11 anthropogenic factors, formed an anthropogenic gradient that
226 strongly correlated with PC1 (linear regression, $R^2$ = 0.970, $P$ < 0.001, **Fig. 1c**). In contrast, the
227 5 indigenous estuarine factors (blue vectors, **Fig. 1b**) were distinct from the anthropogenic
228 gradient, indicating that anthropogenic and natural processes played different roles in driving
229 the environmental conditions of the PRE sediments. The anthropogenic index decreased
230 spatially from nearshore to offshore sediments (**Fig. 1d and 1e**), consistent with the fact that
231 the offshore environment is less impacted by anthropogenic activities[42].
232
233 **Environmental conditions affected the microbial community structures**
234 Microbial community compositions in the PRE sediments were analyzed to investigate the
235 effects of environmental conditions on community assemblages. Although only 14.1 ± 2.2% of
236 the high-quality reads could be classified at the species level, 6,812 bacterial and archaeal
237 species from 45 phyla were identified. From the classified high-quality reads, Proteobacteria

(55.4 ± 4.4%), Actinobacteria (24.7 ± 2.8%) and Firmicutes (9.2 ± 3.6%) were the predominant bacterial phyla, while Euryarchaeota (0.8 ± 0.2%) and Thaumarchaeota (0.7 ± 0.6%) were the predominant archaeal phyla (**Fig. 2a**) in the PRE sediments in both 2020 and 2021, consistent with previous studies in the same region[4,43]. Moreover, aerobic genera dominated the microbial communities, accounting for 97.5 ± 0.8% of the total relative abundance of the classified high-quality reads in all of the samples. Neither the relative abundances of the top eight most abundant phyla (Wilcoxon test, $P > 0.05$) nor the community structure (PERMANOVA, $R^2 = 0.058$, $P = 0.059$; PERMDISP, $P = 0.054$) differed significantly between sediment samples collected in 2020 and 2021; hence, our subsequent analyses focused on spatial rather than temporal differences driven by environmental conditions.

At the species level, the α- and β-diversity of microbial communities varied with the environmental conditions. Regarding α-diversity, species richness increased with increasing anthropogenic index (Spearman's $r = 0.505$, $P = 0.002$) and decreased with increasing salinity (Spearman's $r = -0.428$, $P = 0.012$) and water depth (Spearman's $r = -0.491$, $P = 0.003$), while Simpson's diversity index decreased with increasing anthropogenic index (Spearman's $r = -0.365$, $P = 0.034$) (**Fig. 2b**), indicating a relatively uneven distribution of species in sediments with stronger anthropogenic influences[44]. For β-diversity, the anthropogenic index could explain the most variance (multiple regression, $R^2 = 0.584$, $P = 0.001$) in the microbial community structure, followed by salinity (multiple regression, $R^2 = 0.185$, $P = 0.048$) (**Fig. 2c**). Furthermore, the similarity between microbial communities was high and increased with increasing similarity between the anthropogenic factors (partial Mantel test, $r = 0.254$, $P = 0.003$, **Fig. 2d**), as well as with differences in the anthropogenic index (partial Mantel test, $r = -0.227$, $P = 0.007$, **Fig. S3a**), after controlling for the effects of indigenous estuarine factors and pairwise geographical distances. However, the microbial community structure did not vary significantly with the indigenous estuarine factors (partial Mantel test, $r = 0.049$, $P = 0.332$,

**Fig. S3b**) or pairwise geographical distances (partial Mantel test, $r$ = -0.124, $P$ = 0.920, **Fig. 2e**). Altogether, environmental conditions, especially the anthropogenic factors, had pivotal effects on community assemblages in the PRE sediments.

**Diverse taxonomic responders were identified along the anthropogenic gradient**

To further understand the impacts of anthropogenic disturbances on microbial composition in the PRE sediments, responses of individual species to the anthropogenic gradient were investigated using TITAN[9]. In total, 1,123 increasers and 1,077 decreasers were identified, accounting for 11.7 ± 4.2% and 27.7 ± 3.9% of the relative abundances of the classified species in all of the samples, respectively (**Fig. 3a**). The taxonomic information of all responders is shown in **Table S4**.

The responses of responders were consistent among phylogenetically related taxa (**Fig. 3b**). At the phylum level, the increasers predominantly belonged to Actinobacteria ($n$ = 236; 6.8 ± 2.0% of the relative abundance of the classified species in all of the samples), Firmicutes ($n$ = 540; 3.1 ± 2.1%), and Euryarchaeota ($n$ = 94; 0.3 ± 0.2%) (**Fig. 3a and 3b**). At the genus level under Actinobacteria, the increasers predominantly belonged to *Streptomyces*, *Mycolicibacterium*, and *Mycobacterium* (**Fig. S4**)—genera that have been reported to be able to transform pesticides and heavy metals[45,46]. Among the increasers belonging to Firmicutes, species closely related to human health, including putative human pathogens and fecal pollution indicators (e.g., *Enterococcus faecalis*, *E. durans*, and *E. hirae*[47]), were predominant (**Table S4**). The increasers belonging to Euryarchaeota were mainly methanogens and sulfate-reducing archaea, accounting for 86.7 ± 4.3% and 9.3 ± 3.2%, respectively, of the relative abundance of this phylum in all of the samples.

At the phylum level, the largest proportion of decreasers belonged to Proteobacteria ($n$ = 962; 22.5 ± 2.8% of the relative abundance of the classified species in all of the samples),

followed by Planctomycetes ($n$ = 56; 3.0 ± 0.6%) and Thaumarchaeota ($n$ = 16; 0.8 ± 0.6%) (**Fig. 3a and 3b**). At the genus level under Proteobacteria, the decreasers predominantly belonged to *Pseudomonas*, followed by many halotolerant genera (e.g., *Halomonas*[48] and *Marinobacter*[49]) and oligotrophic genera (e.g., *Caulobacter*[50] and *Marinomonas*[51]) (**Fig. S4**). Under Planctomycetes, the decreasers were mostly from the Planctomycetaceae and Pirellulaceae families, which are typically found in brackish or marine habitats[52-54]. The decreasers belonging to Thaumarchaeota were all ammonia-oxidizing archaea.

**Determination of the ecological community threshold of anthropogenic disturbances**

As anthropogenic disturbances had pivotal effects on community assemblages and the relative abundances of species in the PRE sediments, the anthropogenic disturbance level that would trigger synchronous changes in the relative abundances of most of the responders [i.e., the community-level change point (CP)] was investigated. Based on the anthropogenic index, the community-level change points of increasers [i.e., CP(+)] and decreasers [i.e., CP(-)] were 0.033 (90% confidence interval [CI]: [-0.475, 0.160]) and 0.202 (90% CI: [0.075, 0.428]), respectively.

Considering that the change points were distinct for the increasers and decreasers, we sought to determine a specific change point to represent the ecological community threshold, which would reflect the potential adverse impacts of anthropogenic disturbances, such as loss of indigenous intolerant species or a decrease in community diversity[6, 9]. Therefore, the ecological significance of CP(+) and CP(-) were compared to evaluate their suitability as the ecological community threshold in the PRE sediments. The individual responders that underwent significant changes ($P < 0.05$) ≥ CP(-) accounted for 20.7 ± 1.8%, whereas those that underwent significant changes ≥ CP(+) only accounted for 13.6 ± 1.1% of the relative abundances in all of the samples (**Table S5**). Furthermore, Simpson's diversity index decreased

significantly above CP(-) (Wilcoxon test, $P = 0.004$) but showed no change above or below CP(+) (Wilcoxon test, $P = 0.081$) (**Fig. S5**). Altogether, CP(-) represented a threshold above which widespread variations in the relative abundances of individual species and decreases in community diversity occurred; hence, it was deemed suitable as the ecological community threshold of anthropogenic disturbances in the PRE sediments.

By applying this threshold to PRE sediments, the 34 sampling sites from the 2 years were classified into two groups (H and L) presenting distinct levels of anthropogenic disturbances. Group H included sites ($n = 16$) with a relatively higher level of anthropogenic disturbances (i.e., anthropogenic index above the threshold), and all were located nearshore (**Fig. 3c**), while Group L included sites ($n = 18$) with a relatively lower level of anthropogenic disturbances (i.e., anthropogenic index below the threshold), and all but one were located offshore (**Fig. 3c**).

**Differential functional compositions between Groups H and L**

Functional traits in a microbial community reflect the optimized life-history strategies adopted in response to environmental variations[24]. To understand how anthropogenic disturbances influence the life-history strategies of microbial communities in Groups H and L, the functional compositions of members in these groups were compared. The anthropogenic index was the only significant factor in explaining the variance in the functional compositions (multiple regression, $R^2 = 0.674$, $P = 0.001$, **Fig. 4a**). Furthermore, the functional compositions differed significantly between the two groups (PERMANOVA, $R^2 = 0.164$, $P = 0.001$; PERMDISP, $P = 0.197$; **Fig. 4a**), suggesting that the ecological community threshold can be used to delineate microbial functions.

To determine the differences in microbial life-history strategies between Groups H and L, the key functional traits classified into Y-A-S strategies and the more detailed functional sub-levels 2 and 3 were compared (**Table S6**). Based on the 561 identified key functional traits in

Groups H and L, the microbial communities in the two groups had distinct life-history strategies (**Fig. 4b**). A large proportion of the functional traits enriched in Group H were closely related to the S-strategy (**Fig. 4c**). Specifically, cell communication was the predominant S-strategy subcategory in level 2 (45.5 ± 15.2% of the enriched functional traits, based on the relative abundance in Group H), and this mainly comprised the phosphorelay signal transduction function[55] in level 3. Cell homeostasis maintenance was also abundant among the S-strategy subcategories in level 2 (18.3 ± 9.3%), in which the functional traits could be further categorized in level 3 as biomolecular damage repair, sporulation, oxidative stress tolerance, antibiotic and metal tolerance, and other environmental stressor tolerance, suggesting that defense against multiple environmental stresses was important in Group H. Most of the key functional traits enriched in Group L were related to Y- and A-strategies (**Fig. 4d**). Specifically, central metabolism was the dominant Y-strategy subcategory in level 2 (49.4 ± 17.6% of the enriched functional traits, based on the relative abundance in Group L), which was further categorized in level 3 as functions for precursor metabolite and energy generation, translation, transcription, and replication, indicating a greater potential for cell growth in this group. Substrate transportation and utilization were predominant functions among the A-strategy subcategories in level 2 in both Groups H and L, with the functions in level 3 for transportation and utilization of complex substrates from phosphorus sources (e.g., phosphonates) and carbohydrates (e.g., polysaccharides) being predominant in Group L (14.5 ± 5.5%), while transportation of metal and other substrates was predominant in Group H (12.0 ± 5.4%). Interestingly, mobility-related functions in level 3 were only found in Group L, suggesting that microbes in this group may move in search of resources[56].

To investigate the connections between the key functional traits and taxonomic compositions, the functions were stratified according to species. The key functional traits enriched in Groups H and L were associated with distinct taxa, including taxonomic responders

(**Fig. S6 and S7**). For the functional traits enriched in Group H, members of Firmicutes and Actinobacteria, some of which were increasers, were the key contributors, contributing 23.1 ± 11.8% to 63.7 ± 24.1%, respectively, of the relative abundance of cell communication (GO:0000155 and GO:0009372) and cell homeostasis maintenance (GO:0010127, GO:0030435, GO:00030612, GO:0008941, and GO:0000724) (**Fig. S6**). For the functional traits enriched in Group L, decreasers and non-responders belonging to Proteobacteria and Thaumarchaeota were the key contributors, contributing 30.2 ± 12.7% to 48.8 ± 19.7%, respectively, of the relative abundance of central metabolism (GO:0022904, GO:0009061, GO:0003954, and GO:0002181), substrate utilization (GO:005980 and GO:0050421), and motility (GO:1902208 and GO:0009297) (**Fig. S7**).

**Microbial indicators differentiated sediments between Groups H and L**

As the ecological community threshold separated the PRE benthic microbial communities into two groups with distinct taxonomic and functional profiles, the species and functional traits that could function as indicators to differentiate between sediments with high and low levels of anthropogenic disturbances were identified. Six ML models, with taxonomic responders or functional traits as input predictors, were evaluated for their ability to classify the PRE sediments into Groups H and L. Among the different models, the RF model consistently yielded the highest mean prediction accuracy (86.1–99.9%) and area under the receiver operating characteristic curve (0.94–0.99) for both the taxonomic responders and functional traits as predictors (**Fig. S8**) and was thus selected for the subsequent indicator analysis.

The top 20 species functioning as taxonomic indicators accounted for 27.4% of the total importance in the RF model and could distinguish the sediments between Groups H and L (PERMANOVA, $R^2 = 0.456$, $P = 0.001$; PERMDISP, $P = 0.061$; **Fig. S9a and c**). Most of the taxonomic indicators belonged to Firmicutes and Proteobacteria (**Fig. 5a**), and similar changes

388 in relative abundances along the anthropogenic gradient were observed for the taxonomic
389 indicators in the same phyla (**Fig. 5c**). Moreover, co-occurrence network analysis showed that
390 the taxonomic indicators were strongly connected with the top 500 most abundant species in
391 all of the samples (**Fig. S10a**); some of the taxonomic indicators acted as module hubs (i.e.,
392 nodes highly connected to other members in a module) and connectors (i.e., nodes linking
393 different modules), indicating their important roles in community interactions, assemblages,
394 and functioning (**Fig. S10b**). The top 20 functions functioning as functional indicators
395 accounted for 56.8% of the total importance in the RF model and could differentiate the
396 sediments between Groups H and L (PERMANOVA, $R^2 = 0.896$, $P = 0.001$; PERMDISP, $P =$
397 0.077; **Fig. S9b and d**). The functional indicators were mainly related to functions for precursor
398 metabolite and energy generation, substrate transportation, and environmental stressor
399 tolerance (**Fig. 5b**). The relative abundances of most of the functional indicators related to the
400 S-strategy increased along the anthropogenic gradient, while a majority of those related to the
401 Y-strategy decreased (**Fig. 5d**).
402 Association network analysis was used to investigate the potential associations between
403 the taxonomic and functional indicators. Covariations between these indicators were observed
404 after controlling for the covariate effects of the anthropogenic index (**Fig. S11**). In the
405 association network, taxonomic indicators in the same module were closely related
406 phylogenetically, and many of the functional indicators were positively related to the
407 taxonomic indicators that had a similar response along the anthropogenic gradient (**Fig. S11**).
408 Furthermore, some of the taxonomic indicators contributed to the relative abundances of their
409 positively correlated functional indicators (**Fig. S12-S14**). For example, the taxonomic
410 indicators from Firmicutes and Actinobacteria (e.g., *Bacillus pseudomycoides*, *Bacillus*
411 *dafuensis*, and *Mycobacterium noviomagense*) were positively correlated with the S-strategy-
412 related functional indicators (GO:0000155, GO:0006298, GO:0032300, and GO:0006969) in

Module 1 of the network and contributed 21.2 ± 12.9% to 31.0 ± 25.6% of their relative abundances in all of the samples (**Fig. S12**). Similarly, the taxonomic indicators from Proteobacteria (e.g., *Pseudomonas mosselii*, *P. mucidolens*, and *P. rhodesiae*) were positively correlated with the Y-strategy-related functional indicators (GO:0045278, GO:0006426, and GO:0022900) in Module 2 of the network and contributed 1.1 ± 0.8% to 30.5 ± 24.0% of their relative abundances in all of the samples (**Fig. S13**).

**Discussion**

Living at the interface between human inhabitation and marine environment, estuarine benthic microbial communities face constant long-term anthropogenic disturbances[57]. An ecological community threshold based on microbial responses to anthropogenic disturbances may indicate adverse impacts on ecosystem health and inform appropriate control measures to minimize anthropogenic impacts[9]. However, no study has comprehensively investigated microbial taxonomic and functional responses to cumulative anthropogenic disturbances in estuaries. In this study, metagenomic sequencing was used to assess the responses of microbial communities along an anthropogenic gradient in PRE sediments. By integrating environmental conditions with metagenomic data, an ecological community threshold that distinguished two groups of microbial populations with distinct taxa and adaptation strategies in response to anthropogenic disturbances was identified in the PRE sediments.

Natural ecosystems are constantly exposed to diverse anthropogenic stressors[32,58]; it is thus important to understand the cumulative effects these stressors exert on benthic microbial communities. By representing 11 anthropogenic factors as an anthropogenic index, a decreasing anthropogenic gradient from nearshore to offshore was identified, consistent with the fact that anthropogenic releases in the PRE are originated from land-based coastal sources[59,60]. This result also agrees with source apportionment analysis results that 69.1% of

438 nutrients[61] and 67.0% of metals[28] in the PRE are due to livestock, agricultural, industrial, and
439 domestic releases. Anthropogenic releases in the PRE were found to significantly influence
440 benthic microbial communities, with the anthropogenic index explaining most of the variations
441 in taxonomic and functional community compositions. The cumulative effects of multiple
442 anthropogenic stressors can be additive or synergistic[62]. For example, the presence of multiple
443 metal species can significantly inhibit soil microbial enzyme activity and change soil microbial
444 community compositions in manners that individual metals cannot[63,64], while the presence of
445 mixed carbohydrates can alter the relative abundances of soil microbes in an additive manner[65].
446 Therefore, the cumulative effects of multiple stressors in the PRE sediments are probably a
447 pivotal factor driving microbial community assemblages, similar to other pollution-impacted
448 ecosystems[66,67]. By assessing changes in the relative abundances of individual taxa along the
449 anthropogenic gradient, an ecological community threshold that reflects anthropogenic
450 disturbances in the benthic microbial communities of PRE sediments was identified. When the
451 anthropogenic index exceeded this threshold in the sediments, sharp changes in diversity and
452 relative abundances were observed. This observation is consistent with growing evidence that
453 many members of a biological community (e.g., zooplankton, fish, and vegetation), in addition
454 to microbes, undergo significant changes in occurrence frequency and abundance when
455 subjected to anthropogenic disturbances exceeding a critical threshold[8,13,68,69]. A biological
456 community can typically only tolerate a narrow range of fluctuations in environmental
457 conditions (e.g., physicochemical, hydrological, and thermal); hence, when anthropogenic
458 disturbances are long-term and intense, irreversible shifts in its composition are
459 inevitable[10,70,71].

460 The ecological community threshold in the benthic ecosystem of PRE sediments
461 distinguished two groups of taxa and functions, consistent with the environmental conditions.
462 Group H was mainly located nearshore where nutrient and metal concentrations were higher.

Consistent with the intensified anthropogenic disturbances, many of the taxonomic responders (i.e., increasers) enriched in Group H could cope with anthropogenic stresses. For example, the predominant genera of increasers belonging to Actinobacteria and Firmicutes can degrade and tolerate multiple pollutants (e.g., pesticides, heavy metals, and antibiotics)[46,72] and thus are often enriched in polluted sites[73,74]. In contrast, Group L was mainly located offshore where salinity was higher and nutrient concentrations were lower. Accordingly, characteristics such as halotolerance (coping with osmotic stress) and oligotrophy (coping with resource-limited conditions) were widespread among the taxonomic responders (i.e., decreasers) enriched in this group. Furthermore, the identified taxonomic indicators for differentiating Groups H and L were strongly connected with the most abundant species in the microbial communities, suggesting that these species interacted with other community members and co-varied with many other taxa according to the ecological community threshold. Notably, many taxa (60.6% of the total relative abundance) in the microbial communities were identified as non-responders to anthropogenic disturbances, suggesting that a large proportion of community members were broadly tolerant to environmental variations[75]. By properly allocating resources and energy and utilizing widely available nutrients, these non-responders may have been able to survive under both high and low levels of anthropogenic disturbances and perform their essential functions[76]. In contrast, the responders may have possessed special adaptive traits that were not essential for generalism but were important for maintaining particular services within a community under a specific environmental condition (i.e., specialism). Both generalism and specialism collectively contribute to the stability and ecosystem functions of a microbial community[77].

The distinct adaptive responses to high and low levels of anthropogenic disturbances were further supported by the optimized microbial life-history strategies. In Group H, the S-strategy was enriched, consistent with the need to counter a high level of anthropogenic disturbances[24]. The enriched functional traits related to cell communication reflect an investment in

488 recognizing and responding to environmental stressors[55], while those related to cell
489 homeostasis maintenance reflect an investment in defense against the adverse impacts of
490 environmental stressors[24]. In Group L, Y- and A-strategies played more important roles,
491 consistent with an environment with a lower level of anthropogenic disturbances. Various
492 functional traits related to central metabolism and biosynthesis were enriched in Group L,
493 indicating that microbes tend to allocate more resources toward energy and precursor
494 generation for synthesizing cellular components for growth[78]. A-strategy enrichment was also
495 observed in Group L, most likely to facilitate adaptation to a nutrient-limited environment
496 where simple substrates are less available and complex substrates need to be utilized as
497 resources and energy[24,78,79]. As many taxonomic responders contributed the functional traits
498 that were enriched in the same group, they probably shared these functions to cope with the
499 corresponding environmental conditions[80], resulting in the covariations of taxa and functions
500 according to the ecological community threshold.

501 Despite showing the feasibility and utility of determining an ecological community
502 threshold of anthropogenic disturbances based on indigenous microbial communities, this
503 study has limitations. First, the effects of the 11 anthropogenic factors on benthic microbial
504 communities were weighted equally, and their cumulative effects were investigated using only
505 a correlational analysis; therefore, the individual and cumulative importance of the
506 anthropogenic factors must be verified by laboratory and/or field-based manipulative
507 experiments. Furthermore, other anthropogenic stressors that were excluded may also influence
508 the microbial community composition[81]. The application of time-of-flight mass spectrometry
509 for high-throughput screening of diverse chemical stressors, together with laboratory-based
510 microcosm experiments, is recommended for a comprehensive evaluation of anthropogenic
511 disturbances[32]. Second, the established ecological community threshold and microbial
512 indicators are mostly specific to the PRE sediments analyzed, and it remains unknown whether

513 they vary temporally or spatially; therefore, the applicability of the derived threshold and
514 microbial indicators to other ecosystems warrants verification using a larger set of samples
515 collected across time and space. Third, metagenomic sequencing data can only inform the
516 presence of taxa and potential functions; thus, cellular metabolic activity and gene expression
517 changes remain unknown. Furthermore, the life-history strategies are inferred from putative
518 functions based on metagenomic data; functional traits that reflect phenotypic and metabolic
519 plasticity under specific environmental conditions also remain unclear. Therefore, direct
520 measurement of the functional traits that indicate life-history strategies (e.g., growth rate,
521 biomass concentration, and respiration rate) is required for confirming metagenomics-based
522 functional traits[24], and in future studies, metatranscriptomic sequencing and metaproteomic
523 mass spectrometry data could be further integrated to enhance the robustness of life-history
524 strategies and the ecological community threshold[82,83].
525
526 **Environmental implications**
527 Understanding the impacts of cumulative anthropogenic disturbances in the environment
528 provides essential information for environmental management to maintaining ecosystem
529 health[31]. This study demonstrated the utility of metagenomics-based microbial surveillance in
530 delineating the adverse impacts of anthropogenic disturbances in estuarine sediments. The
531 observed taxonomic and functional shifts highlight the sensitivity and specificity of
532 microorganisms' responses to anthropogenic disturbances in estuarine sediments. The
533 conventional approach to assessing anthropogenic impacts in estuarine sediments relies on
534 evaluating morphological changes in indicator benthic macro-organisms such as
535 Chironomidae[84]. Compared with morphology-based approaches, metagenomics-based
536 microbial surveillance can provide information across the biological hierarchy of gene,
537 organism, and community, offering a higher resolution for more accurate diagnosis[20,85].

Additionally, metagenomics-based microbial surveillance is less time-consuming and labor intensive than morphology-based approaches[6,86]. Compared with analytical measurements of a limited set of anthropogenic chemicals for environmental monitoring and impact assessment[87], metagenomics-based microbial surveillance offers the advantage of assessing the effects of multiple stressors[88], making the evaluation more direct[89]. Although the cost of metagenomic sequencing is higher than that of targeted molecular techniques such as quantitative PCR and amplicon 16S rRNA gene sequencing, metagenomic sequencing provides a higher taxonomic resolution (i.e., species or strains) and more information about metabolic functions without amplification bias[90].

In the metagenomics-based microbial surveillance framework of an anthropogenic-impacted area, the identified ecological community threshold and microbial indicators across time and space can provide references for routine environmental monitoring. For example, the impacts of anthropogenic disturbances at a site may be assessed by comparing the calculated anthropogenic index with the ecological community threshold, and the findings can inform appropriate control and mitigation actions to reduce the disturbances. Moreover, the microbial indicators can be used to determine the extent of anthropogenic disturbances and sources across space and time. In the case of PRE sediments, the enrichment of microbial indicators of putative human pathogens (e.g., *Mycoplasma hyosynoviae*[91] and *Mycobacterium noviomagense*[92]) and xenobiotic degraders (e.g., *Mycolicibacterium vanbaalenii*[93] and *Bacillus pseudomycoides*[94]) in the nearshore samples suggests a high level of anthropogenic disturbances and the potential release of fecal matter and hazardous chemicals. Local environmental managers can accordingly trace the sources and implement control strategies to minimize their releases, and routine monitoring of the microbial indicators over time (e.g., by quantitative PCR) can inform the effectiveness of those strategies. In summary, metagenomics-based microbial surveillance is an effective tool for environmental management in complex ecosystems to aid the control of

563 anthropogenic disturbances and hence, better protect ecosystem health.

564

565 **Supporting Information**

566 Spearman's correlations between the anthropogenic index, anthropogenic factors, and
567 indigenous estuarine factors; principal component analysis (PCA) loadings of the
568 environmental parameters on the first principle component (PC1); relationships between the
569 microbial community similarity and the difference in the magnitude of anthropogenic index
570 and the environmental similarity between the 5 estuarine indigenous factors; average relative
571 compositions of genera in the top three most abundant increasers and decreasers at the phylum
572 level; boxplots of the Simpson's diversity index of the microbial communities delineated using
573 the community-level change points of decreasers and increasers; stratified relative abundances
574 of the key functional traits enriched in Groups H and L according to species; performance of
575 six machine learning (ML) models in identifying taxonomic and functional indicators;
576 performance of the RF model and the microbial indicators in differentiating sediments between
577 Groups H and L; co-occurrence network between the taxonomic indicators and the top 500
578 most abundant species in the microbial communities; association network of Spearman's
579 correlations between the taxonomic and functional indicators; stratified relative abundances of
580 the functional indicators in Modules 1, 2, and 3 of the network according to species;
581 relationships between the anthropogenic index that includes all 11 anthropogenic factors, the
582 indexes that include any 10 of these factors, and the taxonomic and functional compositions of
583 the microbial communities; individual and shared effects of each environmental parameter on
584 the taxonomic and functional compositions of the microbial communities; Spearman's
585 correlations of the pollution load indexes for nutrient factors, metal factors, and all of the
586 anthropogenic factors with the anthropogenic index; calculation and applicability of the
587 anthropogenic index; detailed materials and methods (PDF).

Measurement results of 16 environmental parameters; number of paired-end reads in each sample after quality control and contaminant removal; putative human pathogens identified in the samples; taxonomic responders identified using TITAN; community-level change points determined using the increasers and decreasers; life-history strategies inferred from key functional traits (XLSX).

**Acknowledgments**

This research was supported by the Research Grants Council of Hong Kong (11205923), the State Key Laboratory of Marine Pollution (SKLMP) Seed Collaborative Research Fund (SKLMP/IRF/0026), the City University of Hong Kong (7005918 and 9380128), and the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (311020003 and 311020004). SKLMP receives regular research funding from Innovation and Technology Commission (ITC) of the HKSAR Government. However, any opinions, findings, conclusions or recommendations expressed in this publication do not reflect the views of the HKSAR Government or the ITC. This work is part of the United Nations endorsed Global Estuarine Monitoring (GEM) Program led by KMYL under the UN Decade of Ocean Science for Sustainable Development (2021-2030). We thank Linjie Jin, Rongben Wu, and Jing Li for assisting with the field work and Lan Ma for the analysis of metals.

606 **References**

607 (1) Kennish, M. J., Environmental threats and environmental future of estuaries. *Environ.*
608 *Conserv.* **2002,** *29*, (1), 78-107.
609 (2) Cai, W.; Feely, R. A.; Testa, J. M.; Li, M.; Evans, W.; Alin, S. R.; Xu, Y.; Pelletier, G.;
610 Ahmed, A.; Greeley, D. J.; Newton, J. A.; Bednarsek, N., Natural and anthropogenic drivers of
611 acidification in large estuaries. *Ann. Rev. Mar. Sci.* **2021,** *13*, 23-55.
612 (3) Dürr, H. H.; Laruelle, G. G.; van Kempen, C. M.; Slomp, C. P.; Meybeck, M.; Middelkoop,
613 H., Worldwide typology of nearshore coastal systems: defining the estuarine filter of river
614 inputs to the oceans. *Estuaries Coast* **2011,** *34*, (3), 441-458.
615 (4) Chen, J.; McIlroy, S. E.; Archana, A.; Baker, D. M.; Panagiotou, G., A pollution gradient
616 contributes to the taxonomic, functional, and resistome diversity of microbial communities in
617 marine sediments. *Microbiome* **2019,** *7*, (1), 104.
618 (5) Broman, E.; Izabel-Shen, D.; Rodriguez-Gijon, A.; Bonaglia, S.; Garcia, S. L.; Nascimento,
619 F. J. A. Microbial functional genes are driven by gradients in sediment stoichiometry, oxygen,
620 and salinity across the Baltic benthic ecosystem. *Microbiome* **2022**, 10 (1), 126.
621 (6) Zhang, X., Environmental DNA shaping a new era of ecotoxicological research. *Environ.*
622 *Sci. Technol.* **2019,** *53*, (10), 5605-5612.
623 (7) Goodwin, K. D.; Thompson, L. R.; Duarte, B.; Kahlke, T.; Thompson, A. R.; Marques, J.
624 C.; Caçador, I., DNA sequencing as a tool to monitor marine ecological status. *Front. Mar. Sci.*
625 **2017,** *4,* 107.
626 (8) Yang, J.; Zhang, X.; Xie, Y.; Song, C.; Sun, J.; Zhang, Y.; Giesy, J. P.; Yu, H., Ecogenomics
627 of zooplankton community reveals ecological threshold of ammonia nitrogen. *Environ. Sci.*
628 *Technol.* **2017,** *51*, (5), 3057-3064.
629 (9) Baker, M. E.; King, R. S., A new method for detecting and interpreting biodiversity and
630 ecological community thresholds. *Methods Ecol. Evol.* **2010,** *1*, (1), 25-37.
631 (10) Hughes, T. P.; Bellwood, D. R.; Folke, C.; Steneck, R. S.; Wilson, J., New paradigms for
632 supporting the resilience of marine ecosystems. *Trends Ecol. Evol.* **2005,** *20*, (7), 380-386.
633 (11) Baker, M. E.; King, R. S., Of TITAN and straw men: an appeal for greater understanding

634 of community data. *Freshw. Sci.* **2013,** *32*, (2), 489-506.
635 (12) Brenden, T. O.; Wang, L.; Su, Z., Quantitative identification of disturbance thresholds in
636 support of aquatic resource management. *Environ. Manage.* **2008,** *42*, (5), 821-832.
637 (13) Simonin, M.; Voss, K. A.; Hassett, B. A.; Rocca, J. D.; Wang, S. Y.; Bier, R. L.; Violin, C.
638 R.; Wright, J. P.; Bernhardt, E. S., In search of microbial indicator taxa: shifts in stream bacterial
639 communities along an urbanization gradient. *Environ. Microbiol.* **2019,** *21*, (10), 3653-3668.
640 (14) Delgado-Baquerizo, M.; Maestre, F. T.; Reich, P. B.; Jeffries, T. C.; Gaitan, J. J.; Encinar,
641 D.; Berdugo, M.; Campbell, C. D.; Singh, B. K., Microbial diversity drives multifunctionality
642 in terrestrial ecosystems. *Nat. Commun.* **2016,** *7*, 10541.
643 (15) Wilkins, K.; Clark, C.; Aherne, J., Ecological thresholds under atmospheric nitrogen
644 deposition for 1200 herbaceous species and 24 communities across the United States. *Glob.*
645 *Chang. Biol.* **2022,** *28*, (7), 2381-2395.
646 (16) Brejão, G. L.; Hoeinghaus, D. J.; Pérez-Mayorga, M. A.; Ferraz, S. F. B.; Casatti, L.,
647 Threshold responses of Amazonian stream fishes to timing and extent of deforestation. *Conserv.*
648 *Biol.* **2018,** *32*, (4), 860-871.
649 (17) Yang, J.; Jeppe, K.; Pettigrove, V.; Zhang, X., Environmental DNA metabarcoding
650 supporting community assessment of environmental stressors in a field-based sediment
651 microcosm study. *Environ. Sci. Technol.* **2018,** *52*, (24), 14469-14479.
652 (18) Hu, A.; Ju, F.; Hou, L.; Li, J.; Yang, X.; Wang, H.; Mulla, S. I.; Sun, Q.; Burgmann, H.;
653 Yu, C. P., Strong impact of anthropogenic contamination on the co-occurrence patterns of a
654 riverine microbial community. *Environ. Microbiol.* **2017,** *19*, (12), 4993-5009.
655 (19) Ma, X.; Zhang, Q.; Zheng, M.; Gao, Y.; Yuan, T.; Hale, L.; Van Nostrand, J. D.; Zhou, J.;
656 Wan, S.; Yang, Y., Microbial functional traits are sensitive indicators of mild disturbance by
657 lamb grazing. *ISME J.* **2019,** *13*, (5), 1370-1373.
658 (20) Yang, Y., Emerging patterns of microbial functional traits. *Trends Microbiol.* **2021,** *29*,
659 (10), 874-882.
660 (21) Pierre Louis, A.M.; Yu, H.; Shumlas, S. L.; Van Aken, B.; Schoonen, M. A.; Strongin, D.
661 R., Effect of phospholipid on pyrite oxidation and microbial communities under simulated acid

662 mine drainage (AMD) conditions. *Environ. Sci. Technol.* **2015,** *49*, (13), 7701-7708.
663 (22) Dai, T.; Zhang, Y.; Ning, D.; Su, Z.; Tang, Y.; Huang, B.; Mu, Q.; Wen, D., Dynamics of
664 sediment microbial functional capacity and community interaction networks in an urbanized
665 coastal estuary. *Front. Microbiol.* **2018,** *9*, 2731.
666 (23) Simon, C.; Daniel, R., Achievements and new knowledge unraveled by metagenomic
667 approaches. *Appl. Microbiol. Biotechnol.* **2009,** *85*, (2), 265-276.
668 (24) Malik, A. A.; Martiny, J. B. H.; Brodie, E. L.; Martiny, A. C.; Treseder, K. K.; Allison, S.
669 D., Defining trait-based microbial strategies with consequences for soil carbon cycling under
670 climate change. *ISME J.* **2020,** *14*, (1), 1-9.
671 (25) Li, C.; Liao, H.; Xu, L.; Wang, C.; He, N.; Wang, J.; Li, X., The adjustment of life history
672 strategies drives the ecological adaptations of soil microbiota to aridity. *Mol. Ecol.* **2022,** *31*,
673 (10), 2920-2934.
674 (26) Darling, E. S.; Alvarez-Filip, L.; Oliver, T. A.; McClanahan, T. R.; Côté, I. M., Evaluating
675 life-history strategies of reef corals from species traits. *Ecol. Lett.* **2012,** *15*, (12), 1378-1386.
676 (27) Duprey, N. N.; Wang, T. X.; Kim, T.; Cybulski, J. D.; Vonhof, H. B.; Crutzen, P. J.; Haug,
677 G. H.; Sigman, D. M.; Martínez-García, A.; Baker, D. M., Megacity development and the
678 demise of coastal coral communities: evidence from coral skeleton $\delta^{15}N$ records in the Pearl
679 River estuary. *Glob. Chang. Biol.* **2020,** *26*, (3), 1338-1353.
680 (28) Niu, L.; Cai, H.; Jia, L.; Luo, X.; Tao, W.; Dong, Y.; Yang, Q., Metal pollution in the Pearl
681 River Estuary and implications for estuary management: The influence of hydrological
682 connectivity associated with estuarine mixing. *Ecotoxicol. Environ. Saf.* **2021,** *225*, 112747.
683 (29) Wu, R.; Ruan, Y.; Huang, G.; Li, J.; Lao, J. Y.; Lin, H.; Liu, Y.; Cui, Y.; Zhang, K.; Wang,
684 Q.; Yan, M.; Wu, J.; Huang, B.; Lam, P. K. S., Source apportionment, hydrodynamic influence,
685 and environmental stress of pharmaceuticals in a microtidal estuary with multiple outlets in
686 South China. *Environ. Sci. Technol*. **2022**, 56 (16), 11374-11386.
687 (30) Schlitzer, R., Ocean Data View. **2023**. Available from: https://odv.awi.de.
688 (31) Halpern, B. S.; Walbridge, S.; Selkoe, K. A.; Kappel, C. V.; Micheli, F.; D'Agrosa, C.;
689 Bruno, J. F.; Casey, K. S.; Ebert, C.; Fox, H. E.; Fujita, R.; Heinemann, D.; Lenihan, H. S.;

690 Madin, E. M. P.; Perry, M. T.; Selig, E. R.; Spalding, M.; Steneck, R.; Watson, R., A global map
691 of human impact on marine ecosystems. *Science* **2008,** *319*, (5865), 948-952.
692 (32) Halpern, B. S.; Frazier, M.; Potapenko, J.; Casey, K. S.; Koenig, K.; Longo, C.; Lowndes,
693 J. S.; Rockwood, R. C.; Selig, E. R.; Selkoe, K. A.; Walbridge, S., Spatial and temporal changes
694 in cumulative human impacts on the world's ocean. *Nat. Commun.* **2015,** *6*, 7615.
695 (33) R Core Team, R: A language and environment for statistical computing. R Foundation for
696 Statistical Computing, Vienna, Austria. **2023**. Available from: https://www.R-project.org/.
697 (34) Dufrêne, M.; Legendre, P., Species assemblages and indicator species: the need for a
698 flexible asymmetrical approach. *Ecol. Monogr.* **1997,** *67*, (3), 345-366.
699 (35) Schoch, C. L.; Ciufo, S.; Domrachev, M.; Hotton, C. L.; Kannan, S.; Khovanskaya, R.;
700 Leipe, D.; McVeigh, R.; O'Neill, K.; Robbertse, B.; Sharma, S.; Soussov, V.; Sullivan, J. P.;
701 Sun, L.; Turner, S.; Karsch-Mizrachi, I., NCBI Taxonomy: a comprehensive update on curation,
702 resources and tools. *Database* 2020, *2020*, baaa062.
703 (36) Letunic, I.; Bork, P., Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic
704 tree display and annotation. *Nucleic Acids Res.* **2021,** *49*, (W1), 293-296.
705 (37) Glasl, B.; Robbins, S.; Frade, P. R.; Marangon, E.; Laffy, P. W.; Bourne, D. G.; Webster,
706 N. S., Comparative genome-centric analysis reveals seasonal variation in the function of coral
707 reef microbiomes. *ISME J.* **2020,** *14*, (6), 1435-1450.
708 (38) Karaoz, U.; Brodie, E. L., MicroTrait: A Toolset for a Trait-Based Representation of
709 Microbial Genomes. *Front. Bioinform.* **2022,** *2*, 918853.
710 (39) Yu, J.; Tang, S. N.; Lee, P. K., Universal Dynamics of Microbial Communities in Full-
711 Scale Textile Wastewater Treatment Plants and System Prediction by Machine Learning.
712 *Environ. Sci. Technol.* **2023,** *57*, (8), 3345-3356.
713 (40) Fang, W.; Chang, X.; Su, X.; Xu, J.; Zhang, D.; Ning, K., A machine learning framework
714 of functional biomarker discovery for different microbial communities based on metagenomic
715 data. *2012 IEEE 6th International Conference on Systems Biology (ISB)*, **2012**, 106-112.
716 (41) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel,
717 M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J., Scikit-learn: machine learning in

718 Python. *J. Mach. Learn. Res.* **2011,** *12*, 2825-2830.
719 (42) Jennerjahn, T. C.; Mitchell, S. B., Pressures, stresses, shocks and trends in estuarine
720 ecosystems – An introduction and synthesis. *Estuar. Coast. Shelf Sci.* **2013,** *130*, 1-8.
721 (43) Zhou, L.; Wang, P.; Huang, S.; Li, Z.; Gong, H.; Huang, W.; Zhao, Z.; Yu, Z.,
722 Environmental filtering dominates bacterioplankton community assembly in a highly
723 urbanized estuarine ecosystem. *Environ. Res.* **2021,** *196*, 110934.
724 (44) DeJong, T. M., A comparison of three diversity indices based on their components of
725 richness and evenness. *Oikos* 1975, *26*, 2, 222-227.
726 (45) Mawang, C. I.; Azman, A. S.; Fuad, A. M.; Ahamad, M., Actinobacteria: an eco-friendly
727 and promising technology for the bioaugmentation of contaminants. *Biotechnol Rep (Amst).*
728 **2021,** *32*, e00679.
729 (46) Mohammadipanah, F.; Wink, J., Actinobacteria from Arid and Desert Habitats: Diversity
730 and Biological Activity. *Front. Microbiol.* **2015,** *6*, 1541.
731 (47) Pinto, B.; Pierotti, R.; Canale, G.; Reali, D., Characterization of 'faecal streptococci' as
732 indicators of faecal pollution and distribution in the environment. *Lett. Appl. Microbiol.* **1999,**
733 *29*, (4), 258-263.
734 (48) Xu, M.; Wang, P.; Wang, F.; Xiao, X., Microbial diversity at a deep-sea station of the
735 Pacific nodule province. *Biodivers. Conserv.* **2005,** *14*, (14), 3363-3380.
736 (49) Oueriaghli, N.; Castro, D. J.; Llamas, I.; Bejar, V.; Martinez-Checa, F., Study of bacterial
737 community composition and correlation of environmental variables in Rambla Salada, a
738 hypersaline environment in south-eastern Spain. *Front. Microbiol.* **2018,** *9*, 1377.
739 (50) Wilhelm, R. C., Following the terrestrial tracks of Caulobacter-redefining the ecology of
740 a reputed aquatic oligotroph. *ISME J.* **2018,** *12*, (12), 3025-3037.
741 (51) Xue, J.; Zhang, B.; Zhang, F.; Liu, Y..; Wu, W.; Wu, Z.; Si, Y.; Yang, P.; Xing, X.; Zhao,
742 L., Comparative genomic analysis of the genus *Marinomonas* and taxonomic study of
743 *Marinomonas algarum* sp. nov., isolated from red algae *Gelidium amansii*. *Arch. Microbiol.*
744 **2022,** *204*, (9), 586.
745 (52) Fuerst, J. A.; Sagulenko, E., Beyond the bacterium: Planctomycetes challenge our

746 concepts of microbial structure and function. *Nat. Rev. Microbiol.* **2011,** *9*, (6), 403-13.
747 (53) Guo, M.; Zhou, Q.; Zhou, Y.; Yang, L.; Liu, T.; Yang, J.; Chen, Y.; Su, L.; Xu, J.; Chen, J.;
748 Liu, F.; Chen, J.; Dai, W.; Ni, P.; Fang, C.; Yang, R., Genomic evolution of 11 type strains
749 within family Planctomycetaceae. *PLoS One* **2014,** *9*, (1), e86752.
750 (54) Shu, Q.; Xiong, W.; Peng, S.; Huang, P., Molecular progresses of marine Planctomycetes:
751 a review. *Afr. J. Microbiol. Res.* **2011,** *5*, (33), 6018-6023.
752 (55) Hoch, J. A., Two-component and phosphorelay signal transduction. *Curr. Opin. Microbiol.*
753 **2000,** *3*, (2), 165-170.
754 (56) Fernandez, V. I.; Yawata, Y.; Stocker, R., A foraging mandala for aquatic microorganisms.
755 *ISME J.* **2019,** *13*, (3), 563-575.
756 (57) Meziti, A.; Tsementzi, D.; Ar Kormas, K.; Karayanni, H.; Konstantinidis, K. T.,
757 Anthropogenic effects on bacterial diversity and function along a river-to-estuary gradient in
758 Northwest Greece revealed by metagenomics. *Environ. Microbiol.* **2016,** *18*, (12), 4640-4652.
759 (58) Halpern, B. S.; McLeod, K. L.; Rosenberg, A. A.; Crowder, L. B., Managing for
760 cumulative impacts in ecosystem-based management through ocean zoning. *Ocean Coast*
761 *Manag.* **2008,** *51*, (3), 203-211.
762 (59) Zhao, G.; Ye, S.; Yuan, H.; Ding, X.; Wang, J., Surface sediment properties and heavy
763 metal pollution assessment in the Pearl River Estuary, China. *Environ. Sci. Pollut. Res. Int.*
764 **2017,** *24*, (3), 2966-2979.
765 (60) Hu, J.; Li, S., Modeling the mass fluxes and transformations of nutrients in the Pearl River
766 Delta, China. *J. Mar. Syst.* **2009,** *78*, (1), 146-167.
767 (61) Mao, Y.; Zhang, H.; Tang, W.; Zhao, J.; Wang, Z.; Fan, A., Net anthropogenic nitrogen
768 and phosphorus inputs in Pearl River Delta region (2008-2016). *J. Environ. Manage.* **2021,** *282*,
769 111952.
770 (62) Crain, C. M.; Kroeker, K.; Halpern, B. S., Interactive and cumulative effects of multiple
771 human stressors in marine systems. *Ecol. Lett.* **2008,** *11*, (12), 1304-15.
772 (63) Gong, P.; Sun, T.; Beudert, G.; Hahn, H. H., Ecological effects of combined organic or
773 inorganic pollution on soil microbial activities. *Water, Air, Soil Pollut.* **1997,** *96*, (1), 133-143.

774 (64) Khan, S.; Hesham Ael, L.; Qiao, M.; Rehman, S.; He, J. Z., Effects of Cd and Pb on soil
775 microbial community structure and activities. *Environ. Sci. Pollut. Res. Int.* **2010,** *17*, (2), 288-
776 96.
777 (65) Estrela, S.; Sanchez-Gorostiaga, A.; Vila, J. C.; Sanchez, A., Nutrient dominance governs
778 the assembly of microbial communities in mixed nutrient environments. *Elife* **2021,** *10*, e65948.
779 (66) Gao, H.; LaVergne, J. M.; Carpenter, C. M. G.; Desai, R.; Zhang, X.; Gray, K.; Helbling,
780 D. E.; Wells, G. F., Exploring co-occurrence patterns between organic micropollutants and
781 bacterial community structure in a mixed-use watershed. *Environ. Sci. Process. Impacts.* **2019,**
782 *21*, (5), 867-880.
783 (67) Philippot, L.; Griffiths, B. S.; Langenheder, S., Microbial community resilience across
784 ecosystems and multiple disturbances. *Microbiol. Mol. Biol. Rev.* **2021,** *85*, (2), e00026-20.
785 (68) Rompré, G.; Boucher, Y.; Bélanger, L.; Côté, S.; Robinson, W. D., Conserving biodiversity
786 in managed forest landscapes: the use of critical thresholds for habitat. *For. Chron.* **2010,** *86*,
787 (5), 589-596.
788 (69) Wang, L.; Seelbach, P. W.; Lyons, J., Effects of levels of human disturbance on the
789 influence of catchment, riparian, and reach-scale factors on fish assemblages. *Am. Fish. Soc.*
790 *Symp.* **2006,** *2006*, 199-219.
791 (70) Song, H. S.; Renslow, R. S.; Fredrickson, J. K.; Lindemann, S. R., Integrating ecological
792 and engineering concepts of resilience in microbial communities. *Front. Microbiol.* **2015,** *6*,
793 1298.
794 (71) Bissett, A.; Brown, M. V.; Siciliano, S. D.; Thrall, P. H., Microbial community responses
795 to anthropogenically induced environmental change: towards a systems approach. *Ecol. Lett.*
796 **2013,** *16 Suppl 1*, 128-39.
797 (72) Arora, P. K., Bacilli-mediated degradation of xenobiotic compounds and heavy metals.
798 *Front. Bioeng. Biotechnol.* **2020,** *8*, 570307.
799 (73) Fatima, K.; Afzal, M.; Imran, A.; Khan, Q. M., Bacterial rhizosphere and endosphere
800 populations associated with grasses and trees to be used for phytoremediation of crude oil
801 contaminated soil. *Bull. Environ. Contam. Toxicol.* **2015,** *94*, (3), 314-320.

(74) Wu, T.; Li, X.; Xu, J.; Liu, L.; Ren, L.; Dong, B.; Li, W.; Xie, W.; Yao, Z.; Chen, Q.; Xia, J., Diversity and functional characteristics of endophytic bacteria from two grass species growing on an oil-contaminated site in the Yellow River Delta, China. *Sci. Total Environ.* **2021,** *767*, 144340.

(75) Bell, T. H.; Bell, T., Many roads to bacterial generalism. *FEMS Microbiol. Ecol.* **2021,** *97*, (1), fiaa240.

(76) Sriswasdi, S.; Yang, C.; Iwasaki, W., Generalist species drive microbial dispersion and evolution. *Nat. Commun.* **2017,** *8*, (1), 1162.

(77) Xu, Q.; Vandenkoornhuyse, P.; Li, L.; Guo, J.; Zhu, C.; Guo, S.; Ling, N.; Shen, Q., Microbial generalists and specialists differently contribute to the community diversity in farmland soils. *J. Adv. Res.* **2022,** *40*, 17-27.

(78) Noack, S.; Voges, R.; Gätgens, J.; Wiechert, W., The linkage between nutrient supply, intracellular enzyme abundances and bacterial growth: new evidences from the central carbon metabolism of Corynebacterium glutamicum. *J. Biotechnol.* **2017,** *258*, 13-24.

(79) Litchman, E.; Edwards, K. F.; Klausmeier, C. A., Microbial resource utilization traits and trade-offs: implications for community structure, functioning, and biogeochemical impacts at present and in the future. *Front. Microbiol.* **2015,** *6*, 254.

(80) Martiny, J. B.; Jones, S. E.; Lennon, J. T.; Martiny, A. C., Microbiomes in light of traits: a phylogenetic perspective. *Science* **2015,** *350*, (6261), aac9323.

(81) Andersen, J. H.; Al-Hamdani, Z.; Harvey, E. T.; Kallenbach, E.; Murray, C.; Stock, A., Relative impacts of multiple human stressors in estuaries and coastal waters in the North Sea-Baltic Sea transition zone. *Sci. Total Environ.* **2020,** *704*, 135316.

(82) Van Den Bossche, T.; Arntzen, M. Ø.; Becher, D.; Benndorf, D.; Eijsink, V. G. H.; Henry, C.; Jagtap, P. D.; Jehmlich, N.; Juste, C.; Kunath, B. J.; Mesuere, B.; Muth, T.; Pope, P. B.; Seifert, J.; Tanca, A.; Uzzau, S.; Wilmes, P.; Hettich, R. L.; Armengaud, J., The Metaproteomics Initiative: a coordinated approach for propelling the functional characterization of microbiomes. *Microbiome* **2021,** *9*, (1), 243.

(83) Lambert, B. S.; Groussman, R. D.; Schatz, M. J.; Coesel, S. N.; Durham, B. P.; Alverson,

A. J.; White, A. E.; Armbrust, E. V., The dynamic trophic architecture of open-ocean protist communities revealed through machine-guided metatranscriptomics. *Proc. Natl. Acad. Sci. U.S.A.* **2022,** *119*, (7), e2100916119.

(84) Gomes, W. I. A.; da Silva Jovem-Azevêdo, D.; Paiva, F. F.; Milesi, S. V.; Molozzi, J., Functional attributes of Chironomidae for detecting anthropogenic impacts on reservoirs: a biomonitoring approach. *Ecol. Indic.* **2018,** *93*, 404-410.

(85) Ko, K. K. K.; Chng, K. R.; Nagarajan, N. Metagenomics-enabled microbial surveillance. *Nat. Microbiol.* **2022**, 7 (4), 486-496.

(86) Cordier, T.; Esling, P.; Lejzerowicz, F.; Visco, J.; Ouadahi, A.; Martins, C.; Cedhagen, T.; Pawlowski, J., Predicting the ecological quality status of marine environments from eDNA metabarcoding data using supervised machine learning. *Environ. Sci. Technol.* **2017,** *51*, (16), 9118-9126.

(87) Zhou, S.; Di Paolo, C.; Wu, X.; Shao, Y.; Seiler, T.; Hollert, H., Optimization of screening-level risk assessment and priority selection of emerging pollutants – The case of pharmaceuticals in European surface waters. *Environ. Int.* **2019,** *128*, 1-10.

(88) Altenburger, R.; Ait-Aissa, S.; Antczak, P.; Backhaus, T.; Barceló, D.; Seiler, T.; Brion, F.; Busch, W.; Chipman, K.; de Alda, M. L.; de Aragão Umbuzeiro, G.; Escher, B. I.; Falciani, F.; Faust, M.; Focks, A.; Hilscherova, K.; Hollender, J.; Hollert, H.; Jäger, F.; Jahnke, A.; Kortenkamp, A.; Krauss, M.; Lemkine, G. F.; Munthe, J.; Neumann, S.; Schymanski, E. L.; Scrimshaw, M.; Segner, H.; Slobodnik, J.; Smedes, F.; Kughathas, S.; Teodorovic, I.; Tindall, A. J.; Tollefsen, K. E.; Walz, K.-H.; Williams, T. D.; Van den Brink, P. J.; van Gils, J.; Vrana, B.; Zhang, X.; Brack, W., Future water quality monitoring — Adapting tools to deal with mixtures of pollutants in water resource management. *Sci. Total Environ.* **2015,** *512-513*, 540-551.

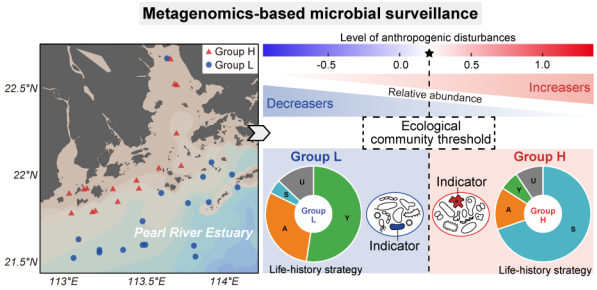
(89) de Baat, M. L.; Wieringa, N.; Droge, S. T. J.; van Hall, B. G.; van der Meer, F.; Kraak, M. H. S., Smarter sediment screening: effect-based quality assessment, chemical profiling, and risk identification. *Environ. Sci. Technol.* **2019,** *53*, (24), 14479-14488.

(90) Scholz, M. B.; Lo, C. C.; Chain, P. S., Next generation sequencing and bioinformatic

858 bottlenecks: the current state of metagenomic data analysis. *Curr. Opin. Biotechnol.* **2012,** *23*,
859 (1), 9-15.
860 (91) Roachford, O.; Nelson, K. E.; Mohapatra, B. R., Virulence and molecular adaptation of
861 human urogenital mycoplasmas: a review. *Biotechnol. Biotechnol. Equip.* **2019,** *33*, (1), 689-
862 698.
863 (92) van Ingen, J.; Boeree, M. J.; de Lange, W. C.; de Haas, P. E.; van der Zanden, A. G.; Mijs,
864 W.; Rigouts, L.; Dekhuijzen, P. R.; van Soolingen, D., *Mycobacterium noviomagense* sp. nov.;
865 clinical relevance evaluated in 17 patients. *Int. J. Syst. Evol. Microbiol.* **2009,** *59*, (4), 845-849.
866 (93) Stingley, R. L.; Khan, A. A.; Cerniglia, C. E., Molecular characterization of a phenanthrene
867 degradation pathway in *Mycobacterium vanbaalenii* PYR-1. *Biochem. Biophys. Res. Commun.*
868 **2004,** *322*, (1), 133-146.
869 (94) Hossain, M. S.; Paul, G. K.; Mahmud, S.; Saleh, M. A.; Uddin, M. S.; Dutta, A. K.; Roy,
870 A. K.; Saha, A. K.; Sheam, M. M.; Ahmed, S.; Rahman, M. M.; Paul, D. K.; Biswas, S. K.,
871 Mixed dye degradation by *Bacillus pseudomycoides* and *Acinetobacter haemolyticus* isolated
872 from industrial effluents: a combined affirmation with wetlab and in silico studies. *Arab. J.*
873 *Chem.* **2022,** *15*, (9), 104078.

874

# For Table of Contents Only




875

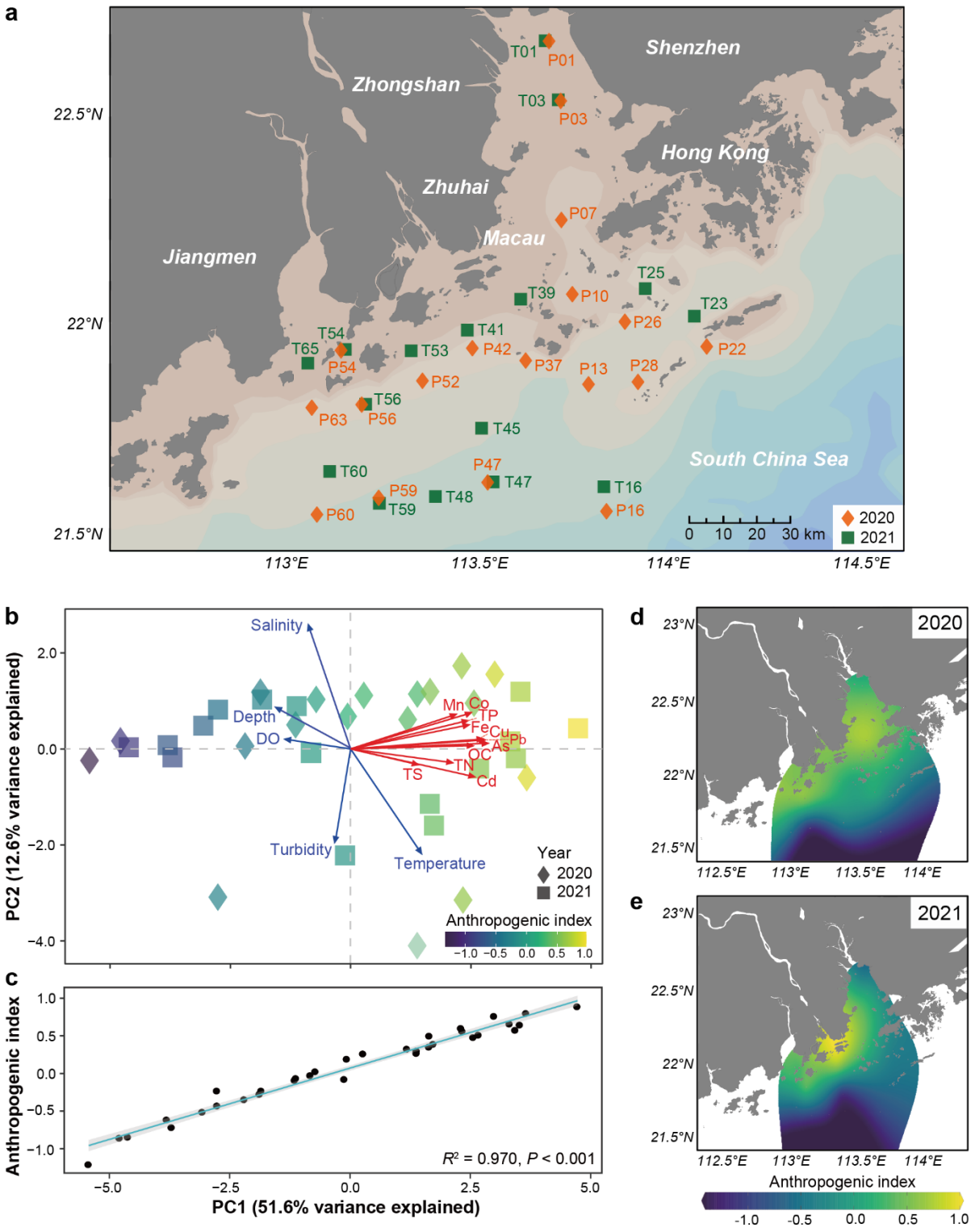



**Figure 1. Variations in environmental conditions along the anthropogenic gradient. (a)** Sampling sites in the Pearl River Estuary (PRE). **(b)** Principal component analysis (PCA) plot of the environmental parameters of the PRE sediment samples collected in 2020 and 2021. The samples are colored according to the magnitude of the anthropogenic index, the sampling year is indicated by the symbol shape, and the loadings of the anthropogenic and indigenous estuarine factors are shown as red and blue vectors, respectively. TN: total nitrogen; TP: total phosphorus; TS: total sulfur; OC: total organic carbon; Fe: iron; Mn: manganese; As: arsenic; Co: cobalt; Cu: copper; Pb: lead; Cd: cadmium; DO: dissolved oxygen. **(c)** Relationships between the first principal component (PC1) and anthropogenic index. The shaded area represents the 95% confidence interval of the linear regression model. **(d, e)** Gradient of the magnitude of anthropogenic index of the PRE sediments based on samples collected in (d) 2020 and (e) 2021.

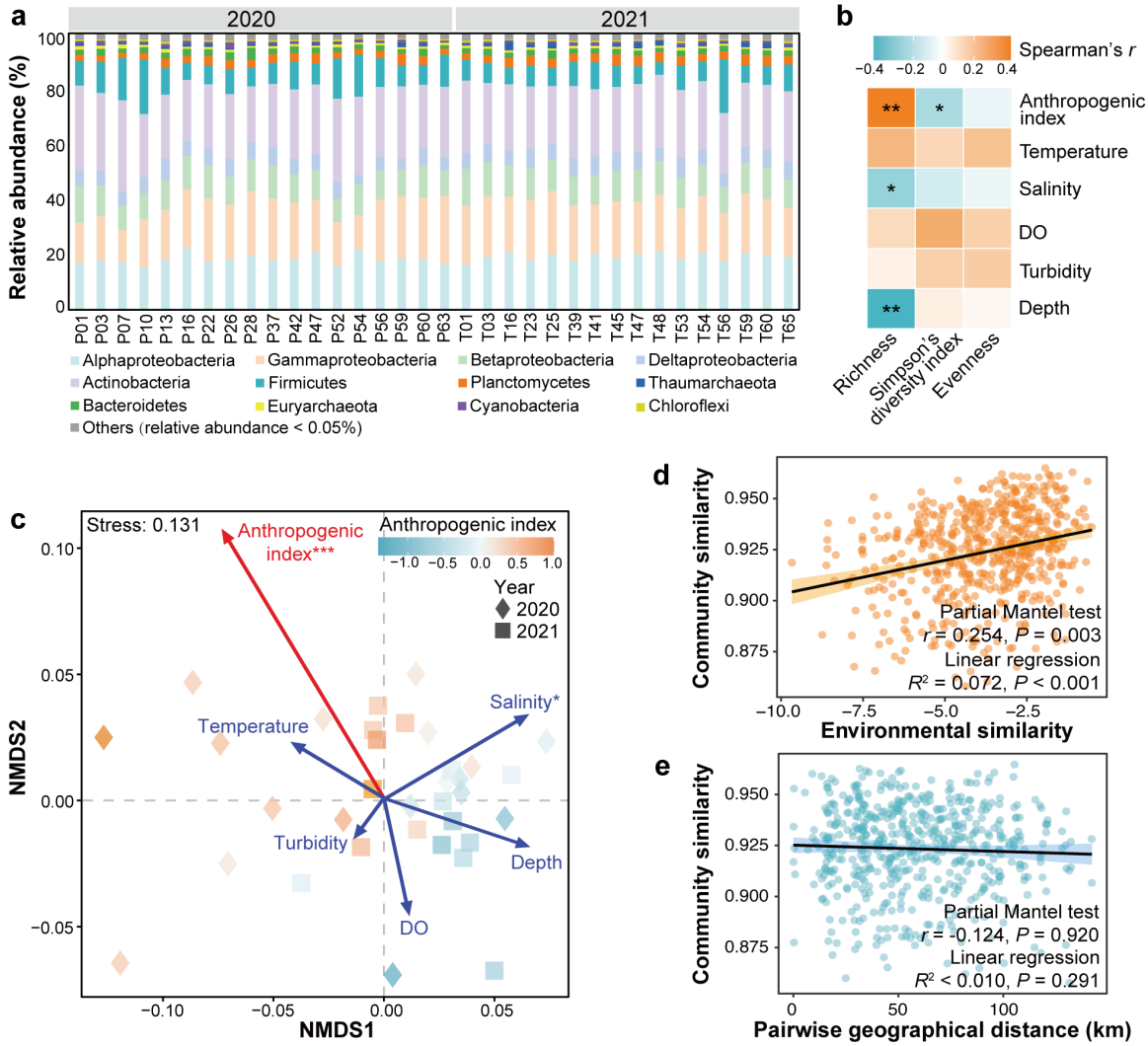



**Figure 2. Variations in microbial community compositions along the anthropogenic gradient. (a)** Relative abundances of the classified high-quality reads at the phylum level in each sample (class level is shown for Proteobacteria). **(b)** Heatmap of correlations between α-diversity, the anthropogenic index, and indigenous estuarine factors (** $P < 0.01$; * $P < 0.05$). **(c)** Non-metric multidimensional scaling (NMDS) ordination of the microbial community structure based on Bray–Curtis dissimilarity. The samples are colored according to the magnitude of the anthropogenic index, and the sampling year is indicated by the symbol shape. Vectors representing the anthropogenic index (red) and indigenous estuarine factors (blue) are fitted onto the NMDS ordination and scaled according to the strength of their correlation with the microbial community structure (*** $P < 0.001$; * $P < 0.05$). **(d, e)** The relationships of the microbial community similarity (1 − Bray–Curtis dissimilarity) with the (d) similarity between the 11 anthropogenic factors (1 − Euclidean distance) and (e) pairwise geographic distance. The shaded area represents the 95% confidence interval of the linear regression model.

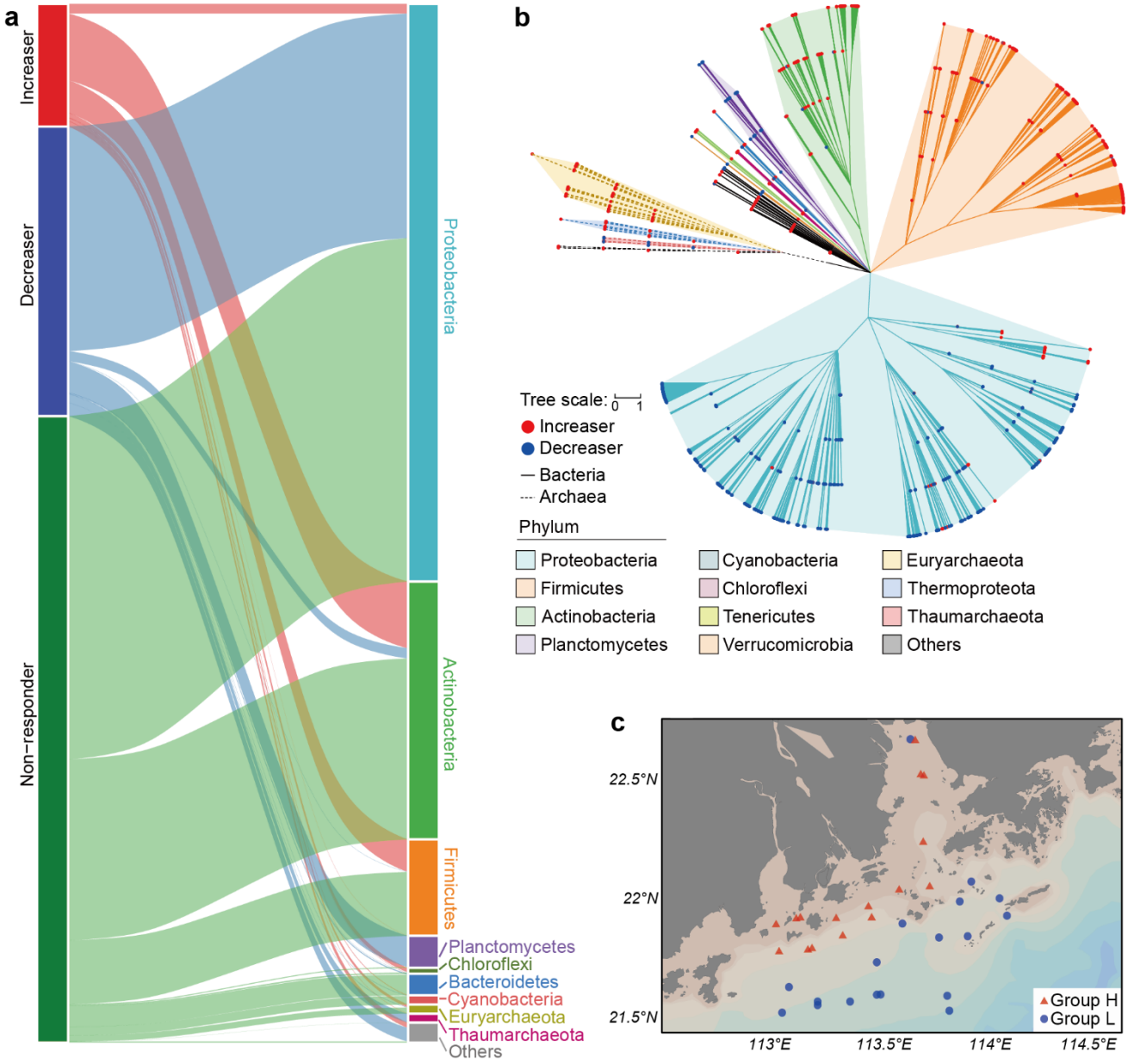



**Figure 3. Taxonomy and relative compositions of the taxonomic responders to the anthropogenic gradient. (a)** The average relative abundances of the taxonomic responders and non-responders at the phylum level in all of the samples. **(b)** A phylum-level maximum likelihood phylogenic tree of the increasers and decreasers identified using TITAN. **(c)** Classification of the sampling sites into Groups H and L based on the ecological community threshold.


**Figure 4. Variations in functional compositions between Groups H and L. (a)** Non-metric multidimensional scaling (NMDS) ordination of the functional compositions based on Bray–Curtis dissimilarity. Samples are colored according to the magnitude of the anthropogenic index, and Groups H and L are indicated by the symbol shape. Vectors representing the anthropogenic index (red) and indigenous estuarine factors (blue) are fitted onto the NMDS ordination and scaled according to the strength of their correlation with the microbial community structure (*** $P < 0.001$). The ellipses represent the 90% confidence ellipse based on a multivariate t-distribution. **(b)** Differences in Y-A-S strategies between Groups H and L based on the relative abundances in the respective groups. **(c, d)** The life-history strategies that were enriched in Groups H and L, respectively. The inner to outer rings show the proportions of Y-A-S strategies (level 1) and their subcategories (levels 2 and 3) based on the relative abundances in the respective groups.

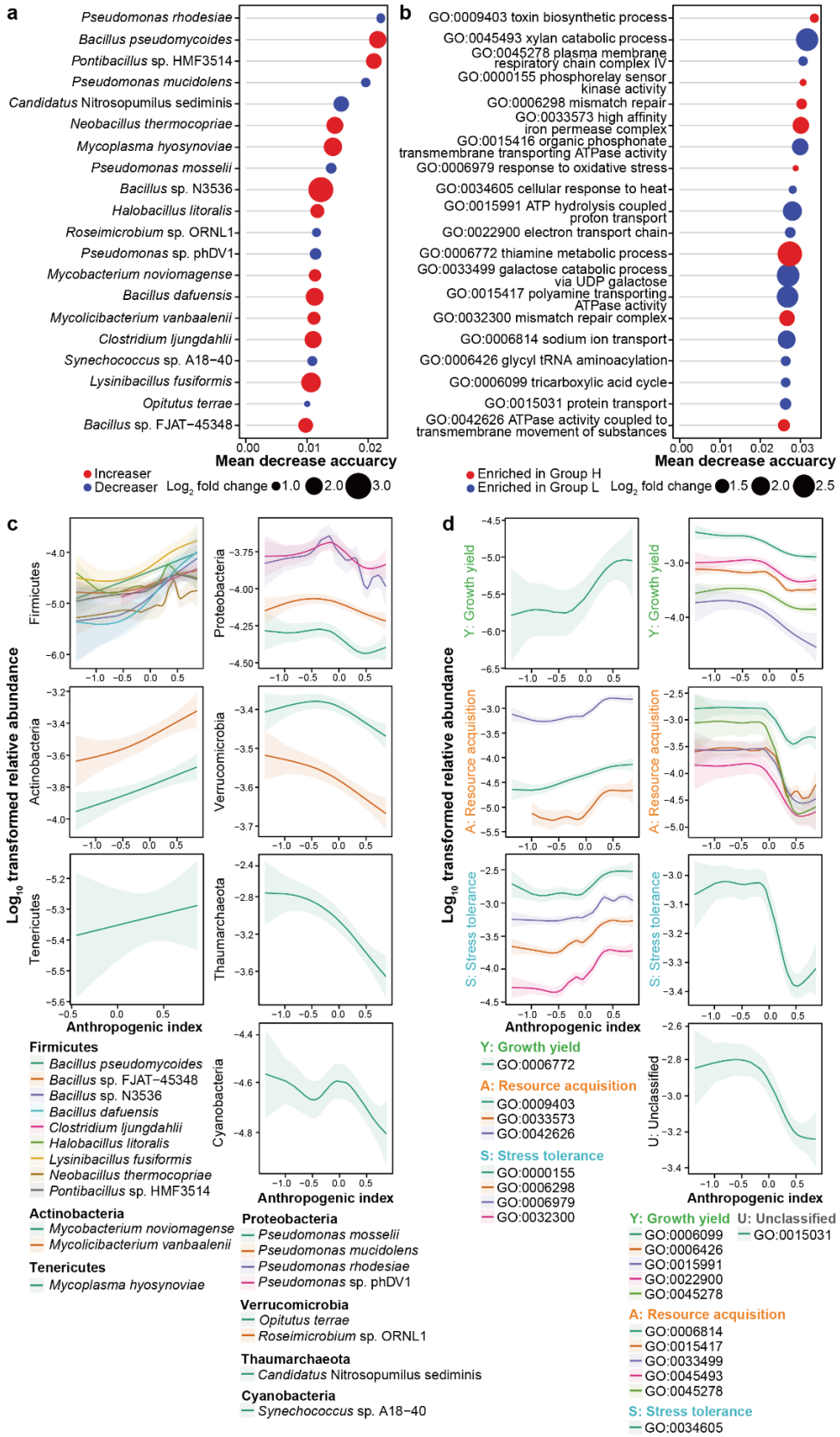



**Figure 5. Microbial indicators identified by the random forest (RF) model for differentiating sediments between Groups H and L. (a, b)** Importance of the top 20 (a) taxonomic and (b) functional indicators for reducing uncertainty in the prediction of sample groups based on the mean decrease accuracies. The size of the circle represents the $\log_2$ fold change in the relative abundance of each indicator in the group in which it was enriched relative to the other group. **(c, d)** Changes in the relative abundances of the top 20 (c) taxonomic and (d) functional indicators along the anthropogenic gradient. The shaded area represents the 95% confidence interval of the generalized additive model.