Title

Airborne bacteria over the oceans shed light on global biogeodiversity patterns

Short title

Aerobiome Biogeography over Oceans

Authors

Naama Lang-Yona[1,6], J. Michel Flores[2], Rotem Haviv[1], Adriana Alberti[3,4], Julie Poulain[3,4], Caroline Belser[3,4], Miri Trainic[2], Daniella Gat[2], Hans-Joachim Ruscheweyh[5], Patrick Wincker[3,4], Shinichi Sunagawa[5], Yinon Rudich[2], Ilan Koren[2,*], Assaf Vardi[1,*]

Affiliations

[1]Department of Plant and Environmental Science, Weizmann Institute of Science, 7610001 Rehovot, Israel.

[2]Department of Earth and Planetary Sciences, Weizmann Institute of Science, 7610001 Rehovot, Israel.

[3]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France.

[4]Research Federation for the study of Global Ocean Systems Ecology and Evolution, FR2022/Tara Oceans GOSEE, 3 rue Michel-Ange, 75016 Paris, France.

[5]Department of Biology and Swiss Institute of Bioinformatics, ETH Zürich, Vladimir-Prelog-Weg 4, 8093 Zürich, Switzerland.

[6]Kinneret Limnological Laboratory, Israel Oceanographic and Limnological Research, 14950 Migdal, Israel.

[*]**Corresponding author**: Ilan Koren (Ilan.Koren@weizmann.ac.il) and Assaf Vardi (assaf.vardi@weizmann.ac.il)

Abstract

Microbes play essential roles in biogeochemical processes in the oceans and atmosphere. Studying the interplay between these two ecosystems can provide important insights into microbial biogeography and diversity. We simultaneously mapped the microbial diversity of airborne and marine bacterial communities across 15,000 kilometers in the Atlantic and Pacific oceans. Higher variability in microbial community composition was observed in the atmosphere than in the surface waters. In addition, a greater similarity was observed between oceans as compared to their overlaying atmosphere, and between atmospheric samples than with the ocean beneath. We detected higher coverage and relative abundance of marine bacteria in the Pacific atmosphere as compared to the Atlantic, while the dominant fraction in the Atlantic atmosphere was annotated as soil-associated bacteria. This study advances our understanding of microbial dispersion over oceans, and of their potential impact on ecology, and biogeochemistry.

Introduction

Microbes are ubiquitous in oceanic and atmospheric environments. Within the oceanic environment, they account for approximately 70% of the total marine biomass (*1*), playing a crucial role in biogeochemical cycles (such as the carbon, nitrogen and sulfur cycles) (*2*). In the atmosphere, the microbial community comprises a major part of atmospheric bioaerosols (*3, 4*), but little is known about the factors that affect their diversity, abundance (*5-7*), and role in this environment (*5, 8, 9*).

Understanding role of the oceans as a source and sink of microorganisms and the atmospheric transport of airborne bacteria can provide important insights into microbial biogeography and diversity as well as information on the interplay between terrestrial communities and their transmission over oceanic regions. Over land, airborne bacteria are emitted from a wide range of sources, from anthropogenic to natural ecosystems. Over the oceans, bacteria can be locally emitted as sea spray aerosols (SSAs), generated at the ocean's surface by wind-driven processes, or transferred into the marine atmosphere through long-range transport from terrestrial sources (*6*). Due to their aerodynamic sizes, it has been hypothesized that all airborne bacteria can disperse globally and may proliferate in any habitat with suitable environmental conditions (*10, 11*). Thus, it can be expected that the airborne bacterial community aloft a given marine environment would exhibit a large fraction of the local marine microbiome.

Indeed, biodiversity studies indicated a link between microbial community composition and geographic locations (*11*). Nevertheless, the geo-distribution of microorganisms in the interface between the ocean and the atmosphere is still underexplored, as the type, amount, and efficiency of particle emissions from the ocean (especially from SSA formation processes) or deposition from the atmosphere hold large uncertainties. Exploring such distributions and microbial fluxes can improve our understanding of the metabolic capabilities introduced into the ocean and the enrichment of local marine diversity and associated functional traits (*12*). Characterizing such biogeographical patterns is of great importance, as they may affect the biochemical and biogeochemical processes upon deposition of bacteria into the ocean or their release to the atmosphere.

In this study, we present the spatially-resolved composition of bacterial communities in the atmospheric marine boundary layer (AMBL) and the ocean surface along two basin-scale oceanic transects: The North Atlantic Ocean and the western Pacific Ocean, covering approximately 15,000 kilometers. We first explore microbial community compositions in the atmospheric and oceanic environments, revealing larger similarities between different oceans, in contrast to their overlaying atmospheres. By focusing on specific genera, we track terrestrial -associated bacteria in the airborne community and show distinct patterns of marine bacterial emission into the atmosphere. We thus suggest that bacterial geo-distribution theories and models should include constraints between the hydrosphere and atmosphere.

Results

Regional distribution of airborne and surface water bacterial phyla in the Pacific and Atlantic oceans

The two open ocean sailing transects examined in this study were as follows: western Pacific, starting from Keelung, Taiwan, towards Fiji (Fig. 1A), and the Atlantic crossing from Lorient, France, to Miami, U.S.A. (Fig. 1B). In the surface water, we found a relatively homogeneous phyla distribution within each transect, with Proteobacteria dominating both oceans (58 ± 3% in the Pacific and 66 ± 4% in the Atlantic; Fig. 1C and D, respectively). Cyanobacteria (29 ± 2% and 9 ± 4%) and Bacteroidetes (9 ± 1% and 17 ± 2%) were the next two most abundant phyla. When comparing to other marine microbiome studies, the phyla

distribution of the near-surface water environment is similar (*2*, *13*). For example, the Cyanobacteria to Proteobacteria ratios in our study are 0.49 ± 0.06 and 0.14 ± 0.07 in the Pacific and Atlantic surface water, respectively. Similarly-calculated ratios characterized in these oceanic regions are 0.43 and 0.16, for the Pacific (*13*) and Atlantic (*2*) regions, respectively.

In the AMBL, Proteobacteria was also the most dominant airborne bacterial phylum in both the Pacific and Atlantic oceans, with 69 ± 12% and 64 ± 8% average percentile abundance, respectively (Fig. 1E and F). However, we found a more heterogeneous distribution of bacterial phyla than that in ocean surface water samples (Fig. S1), even when considering an air mass presence of at least 120 hours over an oceanic path prior to sampling (Fig. 1A and 1B, colored lines). Other abundant phyla in the Pacific AMBL included Cyanobacteria (11 ± 9%), Bacteroidetes (8 ± 3%), Firmicutes (8 ± 3%), Actinobacteria (4 ± 3%), and Planctomycetes (2 ± 1%). In the Atlantic AMBL, the abundant phyla included Actinobacteria (11 ± 5%), Firmicutes (10 ± 6%), and Bacteroidetes (6 ± 3%; Fig. 1F), while Cyanobacteria was observed with an average abundance less than 0.5%.

Firmicutes, Actinobacteria, Fusobacteria, Deinococcus-Thermos and Acidobacteria were higher in abundance in the atmospheric samples of both oceans compared to the surface waters. Firmicutes were predominantly observed in the airborne samples (8 ± 3% and 10 ± 6% in the Pacific and Atlantic AMBL, respectively), with low (<1% in average) to non-significant abundance in the surface water samples. Firmicutes were previously detected in the oceanic environment, specifically the Bacillus genus (*14*), and Sul *et al.* found in low relative abundance (< 6% on average) across latitudes with little latitudinal dependence (*15*). Actinobacteria abundance was also significantly higher in the Atlantic AMBL compared to the surface water (*t*-test, *p*-value < 0.0001). Airborne Actinobacteria and Firmicutes likely represent a terrestrial source, as both phyla have been previously connected to desert dust samples in the Eastern Mediterranean (*16*-*19*), and detected in numerous studies of airborne bacteria (*20*-*23*). In addition, Firmicutes are usually more abundant in soils than in marine surface water (*2*, *14*).

Members of the airborne-detected Fusobacteria, Deinococcus-Thermos and Acidobacteria phyla exhibit high physiological diversity in cell shapes and sizes (*14*, *24*). Their presence in remote locations, such as Antarctica (*25*), airborne dust (*16*), and precipitation over the alpine (*26*), together with their detection above the remote western Pacific Ocean, after at least 120 h transport above the ocean before sampling, suggests that they are ubiquitous in the atmospheric environment. Furthermore, Deinococcus-Thermus are known to be highly resistant to extreme conditions such as high temperatures and dry environments. These traits can also enable the survival of these bacteria under extreme atmospheric conditions (*27*, *28*), and therefore, they may remain viable in the marine lower atmosphere for a long time (*29*).

The local primary production imprint in the AMBL was estimated by calculating the ratio of autotrophic to heterotrophic bacteria in the atmospheric and oceanic samples (Fig. S2A and b). The ratios in the Pacific atmosphere were more than an order of magnitude higher than those measured in the Atlantic atmosphere (mean values: 0.113 ± 0.109 and 0.006 ± 0.009, respectively, two sample *t*-test *p*-value <0.001; Fig. S2A and B). Additionally, the average ratio in the Pacific surface water is approximately four times higher than that in the Atlantic (0.392 ± 0.083, compared to 0.103 ± 0.052, respectively; two sample *t*-test *p*-value < 0.001). Similarly, a significantly higher relative abundance of cyanobacterial 16S rRNA gene was observed in the Pacific compared to the Atlantic atmosphere (Fig. S2C and D, Table S3) based on qPCR analysis of DNA extracted from the same air samples (*30*). The observed difference in oceanic cyanobacterial abundance is consistent with results reported

by Flombaum *et al.*, showing a higher abundance of Prochlorococcus and Synechococcus in the Pacific compared to the Atlantic. Flombaum *et al.* attributed this difference to the impact of temperature and photosynthetically active radiation (*31*). The depletion of cyanobacteria in the Atlantic atmosphere can also result from the dilution of the local emissions with dust-transported bacteria.

In general, the atmospheric phyla composition showed high variations between samples, whereas in the marine microbiome, a more homogeneous and stable community structure was observed. This can be attributed to the differences in the characteristic advection and mixing scales in the two media. While advection in the atmosphere is in the order of 10 m s^{-1}, typical advection and mixing speeds in the ocean are two orders of magnitude smaller (*32, 33*). Furthermore, while the wind-driven surface water currents show connectivity between oceans with time scales of years, the atmospheric circulation time scales are in the range of days to weeks (*34, 35*).

Similarities and differences in the atmospheric and oceanic microbiomes

To further explore the differences between the Pacific and Atlantic atmospheric microbiomes, we analyzed the marine and atmospheric microbiomes at the amplicon sequence variant (ASV) level. We found the variability in the airborne bacterial diversity and composition to be significantly higher compared to the surface water samples (Fig. 2A and B; Two-sample *t*-test, p-value < 0.001, and ANOSIM, p-value = 0.043, respectively), with a distinct difference between the air and water ASV composition (Fig. 2C; AMOVA, p-value < 0.01 between the atmosphere and surface water samples of both Atlantic and Pacific environments).

The ocean surface water biomes shared 166 taxa (comprising 28% of all ocean taxa; Fig. 2D), and the AMBL biomes shared 341 taxa (comprising 25% of all airborne taxa). Only 78 taxa in the Atlantic (7% of all Atlantic taxa) and 134 in the Pacific (15% of all Pacific taxa) were shared between an ocean and its corresponding atmosphere. We found different oceans to have a greater resemblance to one another than to their overlaying AMBL, and atmospheric samples from distinct locations (at least 13,000 km apart from each other) share more common taxa than with the ocean beneath. This suggests that the proximity of the sampled biomes is second in significance to the type of sampled environment. The high portions of shared airborne taxa suggest a potentially higher pool of air-resident bacteria with efficient long-range transport in the atmosphere.

A phylogenetic tree based on the bacterial 16S amplicon sequences provides an overview of the bacterial community and genetic distances between them in the observed marine environment (Fig. S3). Among the shared groups in the Atlantic and Pacific atmospheric biomes, the main phyla occurrences included 44% Proteobacteria, 19% Actinobacteria, 19% Firmicutes, and 10% Bacteroidetes. The high diversity in the common airborne bacterial taxa in both the Atlantic and Pacific is most likely a consequence of a short turnover time of the air mass, leading to continuously changing and dynamic community composition in the AMBL.

Spatial distributions of bacteria across the Pacific and Atlantic environments

The spatial distribution of specific bacteria in the AMBL allows the identification of transported bacteria and indicates a nonrandom contribution from the marine environment. The bacterial ASVs were clustered into taxonomic groups, and >5% taxon occurrences were ranked based on the averaged abundance of the different phyla (Fig. 3 and listed in Tables S4-S6). The most abundant Proteobacteria in the AMBL, which was not detected in the surface water samples, was *Paracoccus* (Table S4). *Paracoccus* strains have been isolated from soil (*36*), marine sediments (*37*), sewage (*38*), and biofilters (*39*), and have been

detected in other atmospheric bacterial studies (*19*, *40*, *41*). Some Proteobacteria, abundant in the air samples, were continuously detected in the surface water (e.g., *Halomonas*, *Pseudomonas*, *Idiomarina*, and *Methylobacterium*; Table S5), and thus assumed to be emitted from the local marine environment.

The most abundant genus in the phylum Actinobacteria was *Actinomarina*, which appeared in the surface waters of both oceans and in the Pacific air samples (Table S5).

The main Firmicutes included Bacillus, known for their endospores that can remain dormant for years. Bacillus is a common bacterium found in transported desert dust (*16*, *42*), and the deep marine environment (*43*). Some Bacillus species are known for their unique metabolites and antagonistic activity against pathogens (*44*, *45*). We further explored the differences between the surface waters and the atmospheric Firmicutes population, which was dramatically under-represented in the water samples, with a focused phylogenetic tree targeting only Firmicutes ASVs (Fig. S4).

Firmicutes ASVs were rarely detected in the surface water samples and differed phylogenetically from the atmospheric ASVs. The rare abundance of marine bacilli in the surface water may result from the preferential growth environment of the deep sea, coral and sediments (*43*, *46*), and their copiotrophic property (i.e., flourishing in environments with high nutrient availability) (*47*). Notably, *Tumebacillus*, and *Clostridium saccharobutylicum*, spore-forming Firmicutes, were detected in the Atlantic AMBL. Endospores can survive harsh and dry conditions and thus might be transported through the air at higher survival rates than others. *Tumebacillus* have been isolated from permafrost (*48*), soil and algal scum (*49*, *50*). We also found ASVs assigned to genera that are known to include human-associated microbes (i.e., *Micrococcus*, *Actinomyces*; Table S4). Although not detected in the blank filters, we cannot exclude the possibility that those taxa may originate from the human activity onboard Tara. Nevertheless, the abundance of these genera is low, ranging on average between ~0.1 – ~2.7 % of the ASVs per filter.

The most abundant genus in the phylum Bacteroidetes found in the atmospheric samples was *Sediminibacterium* (Table S5). This genus was previously found to contribute to the coral microbiome (*51*) and detected in air samples over the Great Barrier Reef (*20*) and the Mediterranean Sea (*41*). Nevertheless, in this study, it was detected in the surface water of the Pacific Ocean solely, with low relative abundance (< 0.01%) and spatial coverage.

We continuously detected water-originated species in the air samples, with higher relative abundance in the Pacific AMBL. One such genus is the Cyanobacteria *Prochlorococcus*, considered a key and most abundant autotroph (*52*), found mainly in oligotrophic oceans (*53*). A notable difference in the relative abundance and appearance of marine bacteria in the Pacific compared to the Atlantic AMBL is also seen for different phyla, including Proteobacteria, Bacteroidetes, Verrucomicrobia, Planctomycetes, etc. (Fig. 3, and Table S5). While the increased fraction of dust-borne bacteria could partially explain the reduction in relative abundance of the local marine bacteria, marine-associated taxa were absent from a significantly high fraction of the sampled Atlantic atmosphere, suggesting other factors may also play a role in the observed difference between the two AMBLs. A clear case of such difference is seen for the Pelagibacterales (SAR-11 clade), representing approximately one-third of the oceanic surface water microbial community (*54*), and highly abundant in both oceans' surface water samples. Their appearance in the atmospheric samples is significantly lower (Two-sample *t*-test, *p*-value < 0.0001 in both tests; Fig. S4B). However, while in the Atlantic AMBL, the spatial coverage is minimal, the Pacific AMBL show almost a full spatial coverage of SAR-11 ASVs in these samples, with significantly higher relative abundance than the Atlantic (two-sample *t*-test, *p*-value <

0.001; Tables S5). A reduced aerosolized fraction of SAR-11 compared to the seawater was also observed in the Arctic Sea by Fahlgren *et al.* in a marine bacterial aerosolization efficiency study (*55*). The difference in abundance between the Atlantic and Pacific AMBLs could be related to properties of the sea-surface microlayer (SML), including thickness, concentration, and chemical composition, which was shown to differ according to changes in heat exchanges, microbial composition, oceanic waves, pollution and dust storms (*56*). However, to fundamentally understand the causing factors for the reduced detection of marine bacteria in the Atlantic AMBL, further investigation is required.

Relationships between bacterial taxa in the marine boundary layer and other ecosystems

To predict the extent of the biogeography of the measured microbiome, from both the surface water and the AMBL, we performed an environmental ontology (ENVO) analysis on all bacterial ASVs (*57*). The ENVO analysis allows a comparison of our dataset to published microbiomes found in other environments, by grouping all environments where the ASVs were previously detected (Fig. 4 and Table S7 detailing the annotations grouped into the five environments presented).

The surface water samples showed 7% nonoceanic-annotated bacterial ASVs in the Pacific compared to 16% in the Atlantic (Fig. S5). Of these, 0.5%, in both environments, were from anthropogenic-annotated bacteria. In the Atlantic surface water samples, 9% were annotated as terrestrial bacteria, and might originate from sedimentation of dust particles or high maritime transportation activity in this environment (*58*).

The microbiome in the Pacific AMBL exhibited a higher fraction of marine-annotated bacteria, 47 ± 20% compared to 12 ± 5% in the Atlantic (Fig. 4A), while the Atlantic aerobiome was dominantly annotated as terrestrial (59 ± 16% compared to 35 ± 18% in the Pacific; Fig. 4B). A previous study by Mayol *et al.* (*59*) determined that overall, 25% (median) of the airborne bacteria over the ocean originate from the marine environment, and 42% originate from terrestrial sources based on parameterizations of sea spray and deposition flux calculations, assuming steady-state conditions. Higher soil-borne ENVO annotation was also observed in Firmicutes-targeted analysis (Fig. S6). This suggests that Firmicutes over the ocean may originate from terrestrial long-range transport, but the extent in which their sedimentation and proliferation in the ocean is yet to be determined. Anthropogenic-annotated ASVs were higher in the Atlantic AMBL, with 19 ± 6%, compared to 9 ± 6% in the Pacific).

Both environments presented freshwater annotations (6 ± 3% and 5 ± 3% in the Pacific and Atlantic AMBL, respectively). The contribution of bacteria to the formation of water precipitation is of high interest, and studies revealed bacterial proteins can promote droplet freezing (*60*) and even detected such bacterial activity in clouds (*22*). The freshwater annotation of the AMBL bacteria may indicate the presence of such bacteria in this environment.

Flores *et al.* (*61*) measured the air sampled on the Tara Pacific Expedition for the same transects investigated in the current study and found that Atlantic airborne particulates were composed of higher concentrations of larger particles related to the deposition of mineral dust compared to the Pacific. Other studies report massive dust quantities crossing over the Atlantic Ocean (*58, 62*), which may introduce bacteria into this environment. The ENVO annotation retrieved from genomic databases corroborate these findings and emphasize the vast contribution of terrestrial dust-borne bacteria into the Atlantic Ocean. Additionally, we detected significantly higher DNA biomass in the Atlantic air filters compared to the Pacific (average of 639.6 ± 468.2 and 128.4 ± 54.4 pg m^{-3}, respectively, *t*-test, *p*-value < 0.0001),

implying a higher concentration of microbial cells per air volume in this region. It has been reported that the imprint of dust transport and human activity on the Pacific Ocean is relatively small (*63*), and that this area is considered a pristine and remote environment, while the Atlantic Ocean experiences high loads of dust, and anthropogenic impact, with a fast pace of change (*64*). Although with lower rates, the terrestrial annotations observed in the Pacific environment emphasize the long-range transport and dispersion of bacteria in this remote environment, implying that it cannot be considered a pristine environment, as observed in the Southern Ocean of Antarctica, where marine contribution was reported to span between 33 – 91% of ASVs, with negligible levels of terrestrial contribution (*65*).

To obtain a more comprehensive overview of the differences between the Atlantic and Pacific AMBLs, we performed environmental ontology analysis for the shared and unique ASVs of each environment (Fig. 4C-E). The unique ASVs detected in the Pacific samples were mainly annotated as marine (37%) and terrestrial (34%) (1,703 ASVs in total; Fig. 4C). The airborne Pacific-Atlantic shared bacterial ASVs were also composed mostly of marine (47%) and terrestrial (32%) annotated bacteria (220 ASVs in total; Fig. 4D). A large percentage of unique Atlantic airborne bacteria (62%) were annotated as terrestrial bacteria, whereas anthropogenic- (16%) and marine (15%)-annotated bacteria were less abundant (1,511 ASVs in total; Fig. 4E).

Higher wind speed values increase aerosol emissions from the ocean (*66*). Indeed, when correlating wind speed with bacterial taxa diversity in the air samples, significant, positive correlations were observed for marine bacteria (Pearson correlation, Table S8) in both the Pacific and Atlantic environments.

The ubiquity of shared species found only in the atmospheric samples of the Atlantic and Pacific oceans suggest that some bacteria may occupy this environment as their growth niche, at least for transient phase (*67*). Specific conditions of humidity, and their attachment to airborne particles (such as dust) may allow their long-range transportation and survival. Nevertheless, to prove this hypothesis, additional analysis of RNA and bacterial activity is needed.

It seems that boundaries might be drawn between the atmosphere and hydrosphere, allowing a nonrandom distribution of species between them. One case is the underrepresentation of marine bacteria in the Atlantic air, and others are airborne taxa (e.g., Firmicutes species) not detected in the surface water samples. Thus, we propose another constraint to the hypothesis, referring to the geo-distribution of bacteria depending not only on distance and time but also on the chemical and physical differences between these ecosystems, dictating a selective transport of different bacteria.

Discussion

We mapped the microbial biodiversity in the oceans and the marine atmosphere across thousands of kilometers in the North Atlantic and western Pacific oceans. Our main findings are summarized schematically in Figure 5.

We found the atmospheric microbial composition to strongly vary even with air masses spending more than 120 hours over the open ocean (Fig. 5A), whereas the marine surface water showed a more stable and homogeneous microbial composition, even across latitudes and different oceans (Fig. 5B). This contrast can be attributed to the orders of magnitude differences in the characteristic advection and mixing scales in the two media, with higher stability and longer mixing cycles in the ocean and rapid changes of days to weeks in the atmosphere.

Furthermore, the Atlantic and Pacific surface water samples showed greater resemblance to one another than to the atmosphere atop, and thousands of kilometer distant atmospheric samples shared more common taxa than with the ocean beneath (Fig. 5C). This suggests that the proximity of the sampled biomes is less influential compared to the type of sampled environment.

We additionally detected differences in the relative abundance of marine bacteria in the Pacific compared to the Atlantic atmosphere, with a significantly reduced spatial coverage in the Atlantic's (Fig. 5D). This observation may be linked to the dilution effect in the Atlantic AMBL due to the relatively high terrestrial-associated bacteria in this environment. Nevertheless, the consistent low appearance of these taxa in the Atlantic air, together with the high sequencing coverage (Fig. S13) of the air samples, suggests an additional mechanism. Differences in the properties of these environments and of their SMLs could impact the aerosolization efficiency of local marine bacteria, and explain the biases in their atmospheric abundance, but this needs further validation. Although reduced in the Atlantic air, the airborne autotrophic cyanobacteria maintained a similar ratio vs. heterotrophs as in the surface waters in both environments (Fig. 5E).

While we found bacterial taxa associated with terrestrial environment (e.g., Firmicutes) to be significantly present in both oceans' AMBL, we did not detect them in the oceanic surface waters, suggesting no preferential proliferation there (Fig. 5F). Since prokaryote concentrations in the atmosphere are orders of magnitude smaller compared to the ocean (~ 10^3-10^4 m^{-3} in the atmosphere (*6*, *59*) vs. ~ 10^8-10^{12} l^{-1} in the surface waters (*14*, *59*)), sedimentation of terrestrial-originated bacteria from the atmosphere to the ocean are not expected to induce significant change in the water composition, unless proliferating. Nevertheless, it cannot be excluded that difference in atmospheric microbial composition could impact surface water composition and function.

In addition, we show that the local marine bacterial community in the AMBL is enriched by long-range transported bacteria, associated with particle transport from terrestrial and anthropogenic environments (Fig. 5G). Thus, even remote locations such as the atmospheric environment of the western Pacific Ocean cannot be defined as pristine.

Finally, we showed that the microbial composition of Atlantic aerosols is associated with dust sources, making these bacteria suitable biomarkers to trace dust transport (Fig. 5H).

This study depicts a new and high-resolution mapping of the spatial biodiversity and the transport mechanisms between the ocean and the atmosphere, at both local and global scales. This interplay between the ocean surface and atmospheric feedback provides new opportunities for future studies to further explore the selective properties of marine microbes within both spheres and how they in turn may affect biogeochemical cycles.

Materials and Methods

Sampling

Details of the expedition and sampling system used are described extensively in Flores *et al.* (*61*). In short, marine aerosols were collected aboard the *R/V* Tara during the first year of the Tara Pacific Expedition (*68, 69*). Airborne particles were collected at ~15 m above sea level (ASL) during the Atlantic crossing from Lorient, France, to Miami, U.S.A. After Miami, the inlet was relocated to ~27 m ASL. The inlet was constructed out of conductive tubing of 1.9 cm inner diameter and a funnel (allowing the collection of all diameters) and mounted on the rear backstay of Tara.

A custom-made aerosol filter system consisting of four 47 mm filter holders and one vacuum pump (Diaphragm pump ME 16 NT, VACUUBRAND BmbH & Co KG, Wertheim, Germany) was installed at the end of the inlet and used to collect the marine particles.

The flow through the filter system was 80 lpm (20 lpm for each filter) during the Atlantic crossing and 120 lpm (30 lpm for each filter) for the Pacific crossing. Three of the four filter holders were loaded with PVDF filters (47 mm diameter, 0.45 µm pore size, PAL, Port Washington, New York), and were used for the current study. The flow rates of each filter holder were monitored continuously and recorded at the beginning and the end of each sampling event. Aerosols were collected for periods between approximately 12 – 24 hours (see Table S1 for the exact times). The filters were folded into a 2 ml cryotube and immediately dropped into liquid nitrogen, and these conditions were maintained while on board. Blank filters were collected by placing filters on the filter holders, closing the system for a few seconds, reopening the holders, folding the filters into cryotubes, and immediately dropping them into liquid nitrogen. These conditions were maintained while on board. Thorough validation tests were conducted in order to verify no boat-originated contamination masked the airborne bacterial population composition (See SI).

Surface water sample collection. The surface water sample collection is described in detail in Gorsky *et al.* (*69*). In short, a “Dolphin sampler”, collecting < 2000 µm size particles, and connected to a large volume peristaltic pump installed on the deck (max flow rate = 3 $m^3\ h^{-1}$), was used for surface water sampling. Each sample endured for ~ 120 min. Water serial filtration (<0.22, 0.22–3, and 3–20 µm) was performed using 142 mm diameter polycarbonate filters (Millipore, Burlington, Massachusetts), and the 0.22-3 µm fraction, where free bacteria are expected to be concentrated, was used for the current study. The filters were folded into a 5 ml cryotube and immediately dropped into liquid nitrogen, and these conditions were maintained while on board. The samples were shipped to the laboratory on dry ice and kept at -80°C until DNA extraction was carried in the laboratory.

Back-Trajectory Analysis

The air mass origins were tracked using the National Oceanic and Atmospheric Administration HYSPLIT trajectory model and Global Data Assimilation System meteorological data (*70*). Although the average residence time for 3 µm particles is of about 4.7 days,(*5*) as considerable terrestrial influence was observed, the model was run to obtain the 240-hour (10 days) back trajectories using the “Ensemble option” at an endpoint height of 250 m, which is the minimum height for an optimal configuration of the ensemble. The presented back trajectories in Fig. 1 are the average of the 27 back trajectories produced by the Ensemble option.

DNA extraction and sequencing

DNA extraction from 0.22-3 μm water filters was performed as described in Alberti *et al.* (*71*). Extraction of DNA from the air filters, with minute amounts of DNA, was carried out using the DNeasy PowerWater Kit (Qiagen, Hilden, Germany), following thorough optimization procedure for the extraction of low DNA concentrations from air filters (see SI text and Fig. S8). The processing of the water and air samples was performed separately, with random DNA extraction order for the different transects, thus preventing cross-contaminations between the different environments, and geographic location (*72*).

The DNA concentrations were evaluated with a Qubit® 3.0 Fluorometer (ThermoFischer, Massachusetts, U.S.A), using the DeNovix (Wilmington, Delaware) dsDNA Ultra High Sensitivity Assay. For DNA sequencing, the bacterial V4–V5 region of the 16S rRNA gene (515Fa: 5'–GTGYCAGCMGCCGCGGTAA–3', and 926R: 5'–CCGYCAATTYMTTTRAGTTT–3') (*73*), was amplified. A PCR mix of 25 μl was prepared in triplicate using 1X Mytaq mix (Bioline, London, UK), 0.2 μM primers, 4 μl DNA extract, and PCR-grade water (Sigma Aldrich). A no-template control was included in all runs, as well as DNA from a mock community (ZymoBIOMICS Microbial Community DNA Standard; Zymo). The PCR products were validated on 1% agarose gel, and triplicates were pooled and sent to Genoscope, the French National DNA Sequencing Facility. The PCR products of the surface water samples from the Pacific transect were sent to the DNA Sequencing Facility at the University of Illinois at Chicago.

DNA sequencing in both facilities was conducted using Illumina MiSeq sequencing technology (maximum read length of 2×300 base pairs). No batch effects were observed, as validated by mock community positive control sequencing.

Quantitative PCR

The concentrations of universal bacterial and cyanobacterial genes in the sampled oceanic air were determined by quantitative PCR (qPCR, QuantStudio 3 real-time PCR system, Applied Biosystems), using [926F: 5'-AAACTCAAAKGAATTGACGG-3' and 1062R: 5'-CTCACRRCACGAGCTGAC-3'] (*74*), and [16SCF: 5'-GGCAGCAGTGGGGAATTTTC-3' and CYAN 377R: 5'-GTMTTACCGCGGCTGCTGG-3'] (*30*) primers for universal and cyanobacterial 16S genes, respectively. Calibration curves of counted cells mL-1 were conducted using DNA extracted from Sulfitobacter D7 for total bacteria and Prochlorococcus MIT604 for cyanobacteria. Triplicates of 10 μl reaction mixtures consisted of 1X fast SYBR Green master mix (Applied Biosystems), 1 μl extracted DNA, 0.2 μM of each primer, and PCR-grade water (Sigma Aldrich). The thermal cycling conditions consisted of an initial 20 s denaturation and enzyme activation at 95 °C, followed by 40 cycles of denaturation at 95 °C and 20 s annealing and extension at 60 °C.

Given the length of our inlet tubes and the differences in the sampling height between the Atlantic and the Pacific, we corrected the airborne DNA biomass concentrations for each transect based on the particle loss calculation following Flores et al 2020 (*61*).

Sequence Processing

Raw amplicon reads provided from the Genoscope sequencing facility were processed using the Tara Pacific metabarcoding (16S) pipeline up to the read margin stage using usearch v11 (*75*), followed by the DADA2 pipeline (version 1.12) (*76*), using R (dada2 package). The merged reads (provided as fastq files) were trimmed and filtered by removing reads exceeding the maximum expected error of 2 bp or an ambiguous read. The reads were dereplicated to acquire unique sequences, which were used to infer sequence variants with the trained error model. After chimeric sequence removal, amplicon

sequence variants (ASVs) were used to assign taxa. ASVs were taxonomically assigned using DADA2, with two steps: classification of sequences against the SILVA nr version 132 training dataset, followed by an exact matching between ASVs and the SILVA species assignment dataset (version 132) providing species-level assignment. Raw reads provided from the sequencing facility of the University of Illinois were also processed using the DADA2 pipeline. Forward and reversed reads were merged after trimming and filtering steps. Dereplication and annotation processes were carried out as described above. A total of 11,577 bacterial ASVs were obtained after error correction, chimera, Archean, chloroplast, and mitochondrial DNA removal. Possible shifts in microbial composition due to differences in analyses were excluded by comparing mock community positive controls sequenced and analyzed using both methods.

The microbial population sequencing analysis was performed only after a thorough decontamination procedure, to mitigate putative contamination in sequence libraries from the air samples with low microbial biomass (*72*), as described in details in the SI.

Environmental ontology

Environmental descriptive terms were extracted from the closest matches (97% identity) using the SEQenv pipeline for Python (version 1.3.0) with default parameters (*77, 78*) and ENVO terms. The input data included FASTA files of sequences after quality control check and removal of blank contaminants per each sample, to be compared to highly similar sequences from public repositories (such as GenBank, using the NCBI nucleotide data base). Detailed description of the pipeline flow is given in the SI file. In the current study, the terms were further clustered into five main groups: marine, terrestrial, fresh water, anthropogenic, and unclassified, as detailed in Table S7.

Statistical analysis

Statistical significance was analyzed using R (3.5.2) and Origin 2019. Analysis of similarities (ANOSIM) was used to verify the significance of the nonmetric multidimensional scaling (NMDS) ordination for taxonomic grouping (using the Bray-Curtis dissimilarity score, vegan package), as well as differences in phyla composition between environments. Analysis of molecular variance (AMOVA) was used to verify significance of the Bray-Curtis dissimilarity differences of bacterial composition between biomes. Alpha diversity values were represented by the Shannon index and were tested for differences using analysis of variance (ANOVA) after normal distribution was verified.

References

1. Y. M. Bar-On, R. Phillips, R. Milo, The biomass distribution on Earth. *Proc Natl Acad Sci U S A* **115**, 6506-6511 (2018).
2. S. Sunagawa *et al.*, Structure and function of the global ocean microbiome. *Science* **348**, 1261359 (2015).
3. W. Elbert, P. E. Taylor, M. O. Andreae, U. Pöschl, Contribution of fungi to primary biogenic aerosols in the atmosphere: wet and dry discharged spores, carbohydrates, and inorganic ions. *Atmos. Chem. Phys.* **7**, 4569-4588 (2007).
4. R. Jaenicke, Abundance of cellular material and proteins in the atmosphere. *Science* **308**, 73 (2005).
5. S. M. Burrows, W. Elbert, M. G. Lawrence, U. Pöschl, Bacteria in the global atmosphere – Part 1: Review and synthesis of literature data for different ecosystems. *Atmos. Chem. Phys.* **9**, 9263-9280 (2009).
6. J. Fröhlich-Nowoisky *et al.*, Bioaerosols in the Earth system: Climate, health, and ecosystem interactions. *Atmos Res* **182**, 346-376 (2016).
7. V. Després *et al.*, Primary biological aerosol particles in the atmosphere: a review. *Tellus B Chem Phys Meteorol* **64**, 15598 (2012).

8. S. M. Burrows *et al.*, Bacteria in the global atmosphere – Part 2: Modeling of emissions and transport between different ecosystems. *Atmos. Chem. Phys.* **9**, 9281-9297 (2009).
9. A. Barberán, J. Henley, N. Fierer, E. O. Casamayor, Structure, inter-annual recurrence, and global-scale connectivity of airborne microbial communities. *Sci Total Environ* **487**, 187-195 (2014).
10. L. G. M. Baas Becking, Geobiologie of inleiding tot de milieukunde. (1934).
11. J. B. H. Martiny *et al.*, Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**, 102-112 (2006).
12. E. Rahav *et al.*, Airborne Microbes Contribute to N2 Fixation in Surface Water of the Northern Red Sea. *Geophys Res Lett* **45**, 6186-6194 (2018).
13. E. F. DeLong *et al.*, Community Genomics Among Stratified Microbial Assemblages in the Ocean's Interior. *Science* **311**, 496 (2006).
14. T. B. Rodrigues, A. E. T. Silva, *Molecular Diversity of Environmental Prokaryotes*. (CRC Press, 2016).
15. W. J. Sul, T. A. Oliver, H. W. Ducklow, L. A. Amaral-Zettler, M. L. Sogin, Marine bacteria exhibit a bipolar distribution. *Proc Natl Acad Sci U S A* **110**, 2342-2347 (2013).
16. D. Gat, Y. Mazar, E. Cytryn, Y. Rudich, Origin-Dependent Variations in the Atmospheric Microbiome Community in Eastern Mediterranean Dust Storms. *Environ Sci Technol* **51**, 6709-6718 (2017).
17. E. Rahav, N. Belkin, A. Paytan, B. Herut, The Relationship between Air-Mass Trajectories and the Abundance of Dust-Borne Prokaryotes at the SE Mediterranean Sea. *Atmosphere* **10**, 280 (2019).
18. N. Lang-Yona *et al.*, Links between airborne microbiome, meteorology, and chemical composition in northwestern Turkey. *Sci Total Environ* **725**, 138227 (2020).
19. Y. Mazar, E. Cytryn, Y. Erel, Y. Rudich, Effect of Dust Storms on the Atmospheric Microbiome in the Eastern Mediterranean. *Environ Sci Technol* **50**, 4194-4202 (2016).
20. S. D. J. Archer *et al.*, Air mass source determines airborne microbial diversity at the ocean–atmosphere interface of the Great Barrier Reef marine ecosystem. *Isme j* **14**, 871–876 (2019).
21. T. Maki *et al.*, Aeolian Dispersal of Bacteria Associated With Desert Dust and Anthropogenic Particles Over Continental and Oceanic Surfaces. *J. Geophys. Res- Atmos* **124**, 5579-5588 (2019).
22. P. Amato *et al.*, Metatranscriptomic exploration of microbial functioning in clouds. *Sci Rep* **9**, 4383 (2019).
23. N. Els *et al.*, Microbial composition in seasonal time series of free tropospheric air and precipitation reveals community separation. *Aerobiologia* **35**, 671-701 (2019).
24. A. Quaiser *et al.*, Acidobacteria form a coherent but highly diverse group within the bacterial domain: evidence from environmental genomics. *Mol Microbiol* **50**, 563-575 (2003).
25. S. D. J. Archer *et al.*, Airborne microbial transport limitation to isolated Antarctic soil habitats. *Nature Microbiol* **4**, 925-932 (2019).
26. J. Caliz, X. Triado-Margarit, L. Camarero, E. O. Casamayor, A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proc Natl Acad Sci U S A* **115**, 12229-12234 (2018).
27. Z. Fang, W. Guo, J. Zhang, X. Lou, Influence of Heat Events on the Composition of Airborne Bacterial Communities in Urban Ecosystems. *Int J Environ Res Public Health* **15**, (2018).
28. J.-F. Gao, X.-Y. Fan, H.-Y. Li, K.-L. Pan, Airborne Bacterial Communities of PM2.5 in Beijing-Tianjin-Hebei Megalopolis, China as Revealed By Illumina MiSeq Sequencing: A Case Study. *Aerosol Air Qual Res* **17**, 788-798 (2017).

29. J. Ho, M. Adeolu, B. Khadka, R. S. Gupta, Identification of distinctive molecular traits that are characteristic of the phylum "Deinococcus-Thermus" and distinguish its main constituent groups. *Syst Appl Microbiol* **39**, 453-463 (2016).
30. K. H. Oh, D. H. Jeong, S. H. Shin, Y. C. Cho, Simultaneous quantification of cyanobacteria and Microcystis spp. using real-time PCR. *J Microbiol Biotechnol* **22**, 248-255 (2012).
31. P. Flombaum *et al.*, Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc Natl Acad Sci U S A* **110**, 9824-9829 (2013).
32. L. C. Laurindo, A. J. Mariano, R. Lumpkin, An improved near-surface velocity climatology for the global ocean from drifter observations. *Deep Sea Res Part I Oceanogr Res Pap* **124**, 73-92 (2017).
33. Y. Lehahn, F. d'Ovidio, M. Lévy, Y. Amitai, E. Heifetz, Long range transport of a quasi isolated chlorophyll patch by an Agulhas ring. *Geophys Res Lett* **38**, (2011).
34. T. DeVries, F. Primeau, Dynamically and Observationally Constrained Estimates of Water-Mass Distributions and Ages in the Global Ocean. *J Phys Oceanogr* **41**, 2381-2401 (2011).
35. J. H. Seinfeld, S. N. Pandis, *Atmospheric chemistry and physics : from air pollution to climate change*. (2016).
36. H. Siller, F. A. Rainey, E. Stackebrandt, J. Winter, Isolation and characterization of a new gram-negative, acetone-degrading, nitrate-reducing bacterium from soil, Paracoccus solventivorans sp. nov. *Int J Syst Bacteriol* **46**, 1125-1130 (1996).
37. J. H. Lee, Y. S. Kim, T. J. Choi, W. J. Lee, Y. T. Kim, Paracoccus haeundaensis sp. nov., a Gram-negative, halophilic, astaxanthin-producing bacterium. *Int J Syst Evol Microbiol* **54**, 1699-1702 (2004).
38. X. Y. Liu, B. J. Wang, C. Y. Jiang, S. J. Liu, Paracoccus sulfuroxidans sp. nov., a sulfur oxidizer from activated sludge. *Int J Syst Evol Microbiol* **56**, 2693-2695 (2006).
39. A. Lipski, K. Reichert, B. Reuter, C. Spröer, K. Altendorf, Identification of bacterial isolates from biofilters as Paracoccus alkenifer sp. nov. and Paracoccus solventivorans with emended description of Paracoccus solventivorans. *Int J Syst Bacteriol* **48 Pt 2**, 529-536 (1998).
40. N. DeLeon-Rodriguez *et al.*, Microbiome of the upper troposphere: Species composition and prevalence, effects of tropical storms, and atmospheric implications. *Proc Natl Acad Sci U S A* **110**, 2575-2580 (2013).
41. E. Mescioglu *et al.*, Aerosol Microbiome over the Mediterranean Sea Diversity and Abundance. *Atmosphere* **10**, 440 (2019).
42. K. Yoo *et al.*, Bacillus-Dominant Airborne Bacterial Communities Identified During Asian Dust Events. *Microbial ecology* **78**, 677-687 (2019).
43. E. P. Ivanova *et al.*, Characterization of Bacillus strains of marine origin. *Int Microbiol* **2**, 267-271 (1999).
44. M. A. M. Mondol, H. J. Shin, M. T. Islam, Diversity of secondary metabolites from marine Bacillus species: chemistry and biological activity. *Mar Drugs* **11**, 2846-2872 (2013).
45. Y. Chen, J. Li, P. Xiao, W. Zhu, Z. Mo, The ability of marine Bacillus spp. isolated from fish gastrointestinal tract and culture pond sediment to inhibit growth of aquatic pathogenic bacteria. **15**, 701-714 (2016).
46. J. Overmann, C. Lepleux, in *The Marine Microbiome: An untapped source of biodiversity and biotechnological potential*, L. J. Stal, M. S. Cretoiu, Eds. (Springer International Publishing, Switzerland, 2016), chap. 2, pp. 21-55.
47. N. Fierer, M. A. Bradford, R. B. Jackson, Toward an ecological classification of soil bacteria. *Ecology* **88**, 1354-1364 (2007).

48. B. Steven, M. Q. Chen, C. W. Greer, L. G. Whyte, T. D. Niederberger, Tumebacillus permanentifrigoris gen. nov., sp. nov., an aerobic, spore-forming bacterium isolated from Canadian high Arctic permafrost. *Int. J Syst Evol* **58**, 1497-1501 (2008).
49. J. Her, S. Srinivasan, S.-S. Lee, Tumebacillus luteolus sp. nov., isolated from soil. *Int. J Syst Evol* **65**, 4107-4112 (2015).
50. Y.-F. Wu, B. Zhang, P. Xing, Q.-L. Wu, S.-J. Liu, Tumebacillus algifaecis sp. nov., isolated from decomposing algal scum. *Int. J Syst Evol* **65**, 2194-2198 (2015).
51. W. Pootakham *et al.*, High resolution profiling of coral-associated bacterial communities using full-length 16S rRNA sequence data from PacBio SMRT sequencing system. *Sci Rep* **7**, 2774 (2017).
52. S. W. Chisholm *et al.*, A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**, 340-343 (1988).
53. S. W. Chisholm *et al.*, Prochlorococcus marinus nov. gen. nov. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll a and b. *Arch Microbiol* **157**, 297-300 (1992).
54. R. M. Morris *et al.*, SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**, 806-810 (2002).
55. C. Fahlgren *et al.*, Seawater mesocosm experiments in the Arctic uncover differential transfer of marine bacteria to aerosols. *Environ Microbiol Rep* **7**, 460-470 (2015).
56. O. Wurl, W. Ekau, W. Landing, C. Zappa, Sea surface microlayer in a changing ocean – A perspective. *Elementa-Sci Anthrop* **5**, (2017).
57. P. L. Buttigieg *et al.*, The environment ontology: contextualising biological and biomedical entities. *J Biomed Semant* **4**, 43 (2013).
58. Y. J. Kaufman *et al.*, Dust transport and deposition observed from the Terra-Moderate Resolution Imaging Spectroradiometer (MODIS) spacecraft over the Atlantic Ocean. *J. Geophys. Res- Atmos* **110**, (2005).
59. E. Mayol *et al.*, Long-range transport of airborne microbes over the global tropical and subtropical ocean. *Nature Com* **8**, 201 (2017).
60. K. C. Failor, D. G. Schmale, B. A. Vinatzer, C. L. Monteil, Ice nucleation active bacteria in precipitation are genetically diverse and nucleate ice by employing different mechanisms. *Isme j* **11**, 2740-2753 (2017).
61. J. M. Flores *et al.*, Tara Pacific expedition's atmospheric measurements. Marine aerosols across the Atlantic and Pacific Oceans Overview and Preliminary results. *Bull Am Meteorol Soc* **101**, 536-554 (2020).
62. M. Vázquez-Rodríguez *et al.*, Anthropogenic carbon distributions in the Atlantic Ocean: data-based estimates from the Arctic to the Antarctic. *Biogeosciences* **6**, 439-451 (2009).
63. K. R. Jones *et al.*, The Location and Protection Status of Earth's Diminishing Marine Wilderness. *Curr Biol* **28**, 2506-2512.e2503 (2018).
64. B. S. Halpern *et al.*, Recent pace of change in human impact on the world's ocean. *Sci Rep* **9**, 11609 (2019).
65. J. Uetake *et al.*, Airborne bacteria confirm the pristine nature of the Southern Ocean boundary layer. *Proc Natl Acad Sci U S A*, 202000134 (2020).
66. G. de Leeuw *et al.*, in *Ocean-Atmosphere Interactions of Gases and Particles*. (Springer Earth System Sciences, Berlin, Heidelberg, 2014), chap. 4, pp. 171–246.
67. J. B. Martiny *et al.*, Microbial biogeography: putting microorganisms on the map. *Nature reviews. Microbiology* **4**, 102-112 (2006).
68. S. Planes *et al.*, The Tara Pacific expedition—A pan-ecosystemic approach of the "-omics" complexity of coral reef holobionts across the Pacific Ocean. *PLOS Biology* **17**, e3000483 (2019).

69. G. Gorsky *et al.*, Expanding Tara Oceans Protocols for Underway, Ecosystemic Sampling of the Ocean-Atmosphere Interface During Tara Pacific Expedition (2016–2018). *Front Mar Sci* **6**, (2019).
70. A. F. Stein *et al.*, NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bull Am Meteorol Soc* **96**, 2059-2077 (2016).
71. A. Alberti *et al.*, Viral to metazoan marine plankton nucleotide sequences from the Tara Oceans expedition. *Sci Dat* **4**, 170093 (2017).
72. R. Eisenhofer *et al.*, Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol* **27**, 105-117 (2019).
73. A. E. Parada, D. M. Needham, J. A. Fuhrman, Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* **18**, 1403-1414 (2016).
74. T. Bacchetti De Gregoris, N. Aldred, A. S. Clare, J. G. Burgess, Improvement of phylum- and class-specific primers for real-time PCR quantification of bacterial taxa. *J Microbiol Methods* **86**, 351-356 (2011).
75. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461 (2010).
76. B. J. Callahan *et al.*, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581-583 (2016).
77. A. Z. Ijaz, T. C. Jeffries, U. Z. Ijaz, K. Hamonts, B. K. Singh, Extending SEQenv: a taxa-centric approach to environmental annotations of 16S rDNA sequences. *PeerJ* **5**, e3827-e3827 (2017).
78. L. Sinclair *et al.*, Seqenv: linking sequences to environments through text mining. *PeerJ* **4**, e2690 (2016).

References cited only in the Supplementary Materials

1. D. Hospodsky, N. Yamamoto, J. Peccia, Accuracy, precision, and method detection limits of quantitative PCR for airborne bacteria and fungi. *Appl Environ Microbiol* **76**, 7004-7012 (2010).
2. N. Lang-Yona et al., Annual distribution of allergenic fungal spores in atmospheric particulate matter in the Eastern Mediterranean; a comparative study between ergosterol and quantitative PCR analysis. *Atmos Chem Phys* **12**, 2681-2690 (2012).
3. S. Sharoni et al., Infection of phytoplankton by aerosolized marine viruses. Proc Natl Acad Sci U S A 112, 6643-6647 (2015).
4. N. M. Davis, D. M. Proctor, S. P. Holmes, D. A. Relman, B. J. Callahan, Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).
5. R. Eisenhofer et al., Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol* **27**, 105-117 (2019).
6. S. J. Salter et al., Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* **12**, 87 (2014).
7. E. R. Lewis, S. E. Schwartz, in Sea Salt Aerosol Production: Mechanisms, Methods, Measurements and Models, E. R. Lewis, S. E. Schwartz, Eds. (American Geophysical Union, 2013), pp. 9-99.
8. S.-L. von der Weiden, F. Drewnick, S. Borrmann, Particle Loss Calculator - a new software tool for the assessment of the performance of aerosol inlet systems. *Atmos Meas Tech Discuss* **2**, 1099 (2009).
9. J. M. Flores et al., Tara Pacific expedition's atmospheric measurements. Marine aerosols across the Atlantic and Pacific Oceans Overview and Preliminary results. *Bull Am Meteorol Soc* **101**, 536-554 (2020).

10. M. E. Allentoft et al., The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc R Soc Lond* **279**, 4724-4733 (2012).
11. J. Peccia, M. Hernandez, Incorporating polymerase chain reaction-based identification, population characterization, and quantification of microorganisms into aerosol science: A review. *Atmos Environ* (1994) **40**, 3941-3961 (2006).
12. A. Z. Ijaz, T. C. Jeffries, U. Z. Ijaz, K. Hamonts, B. K. Singh, Extending SEQenv: a taxa-centric approach to environmental annotations of 16S rDNA sequences. *PeerJ* **5**, e3827-e3827 (2017).
13. L. Sinclair et al., Seqenv: linking sequences to environments through text mining. *PeerJ* **4**, e2690 (2016).
14. P. L. Buttigieg et al., The environment ontology: contextualising biological and biomedical entities. *J Biomed Semant* **4**, 43 (2013).

Acknowledgments

We are grateful to the following institutions for their financial and scientific support in Tara Pacific expedition: CNRS, PSL, CSM, EPHE, Genoscope/CEA, and France Génomique funding (ANR-10-INBS-09-08), Inserm, Université Côte d'Azur, ANR, agnès b., UNESCO-IOC, the Veolia Environment Foundation, Région Bretagne, Serge Ferrari, Billerudkorsnas, Amerisource Bergen Company, Lorient Agglomeration, Oceans by Disney, the Prince Albert II de Monaco Foundation, L'Oréal, Biotherm, France Collectivités, Kankyo Station, Fonds Français pour l'Environnement Mondial (FFEM), Etienne BOURGOIS, and the Tara Foundation teams and crew. The authors particularly thank Dr. Serge Planes, Prof. Denis Allemand, and the Tara Pacific consortium for their management and coordination. The authors gratefully acknowledging Prof. Daniel Sher for providing *Prochlorococcus* cultures for qPCR quantification, and Dr. Flora Vincent for her helpful, critical feedback on the manuscript. This is publication number # 15 of the Tara Pacific Consortium.

Funding:
Scott Jordan and Gina Valdez (NLY, MF, MT, IK, AV)
The De Botton for Marine Science (RH, AV)
The Yeda-Sela Center for Basic Research (NLY, JMF, MT)
The Women Bridging position, Weizmann Institute of Science (NLY)
The Sustainability and Energy Research Initiative (SAERI), Weizmann Institute of Science (NLY).
The Israel Science Foundation grant #236/16 (YR)
The ETH and Helmut Horten Foundation (SSu)

Author contributions:
Conceptualization: NLY, JMF, IK, AV; Methodology: NLY, RH, MT, AA, JP, CB, HJR; Field Sampling: JMF, JP; Investigation: NLY, JMF, IK, AV; Data curation: NLY, JMF, RH, AA, DG, JP, CB, HJR; Visualization: NLY; Supervision: IK, AV; Writing—original draft: NLY; Writing—review & editing: NLY, JMF, MT, DG, SSu, YR, IK, AV; Resources: SSu, PW, YR, IK, and AV

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: The raw 16S amplicon sequences were deposited in the European Nucleotide Archive at EMBL-EBI (accession numbers: PRJEB39048, and PRJEB38899).

Figures and Tables

Figure 1.

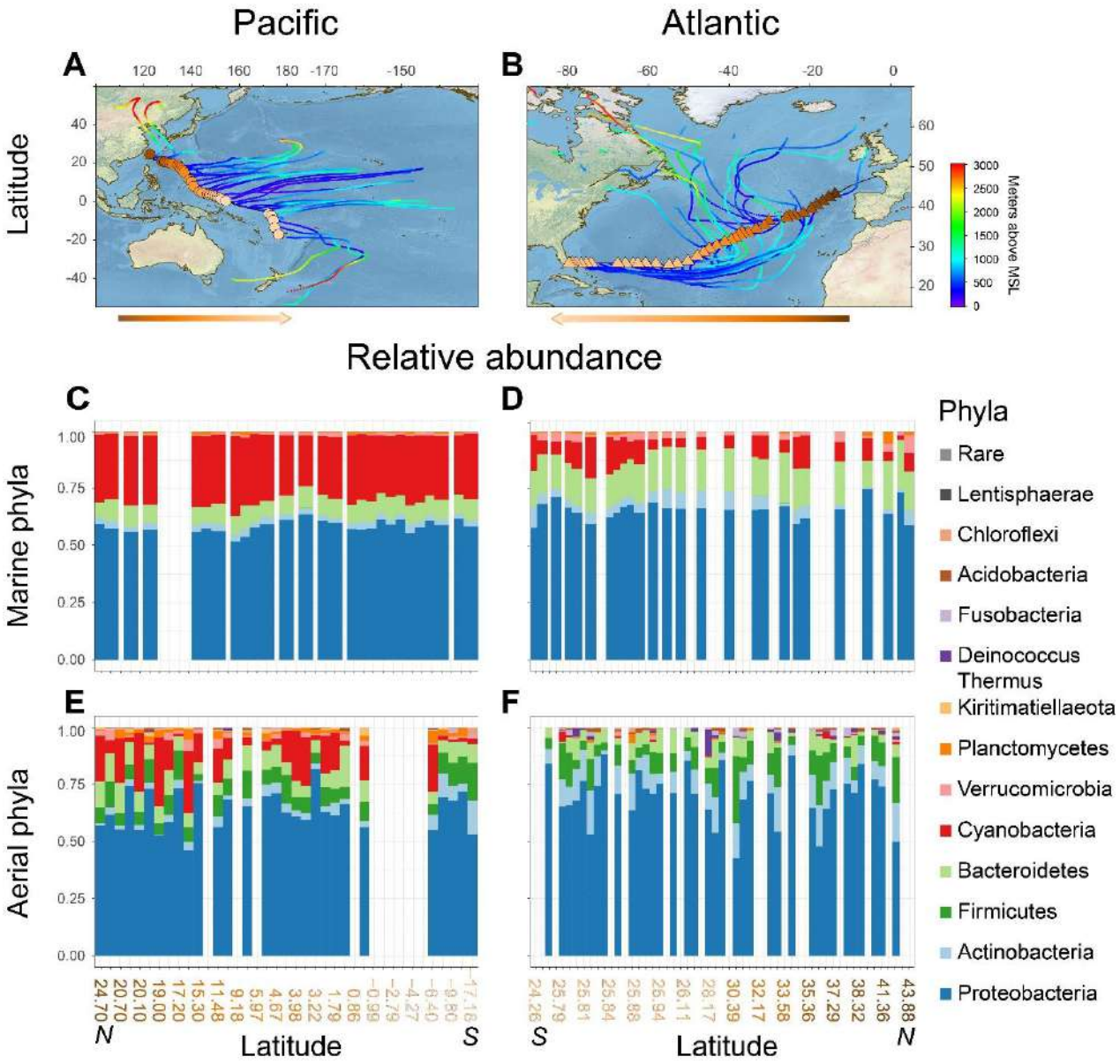


Fig. 1. Regional distribution of airborne bacterial phyla above the Pacific and Atlantic oceans. Air-sampling locations and back trajectories, with height above mean sea levels (MSL) indicated with color scale for the Pacific (A) and Atlantic (B) atmospheric samples. Phyla relative abundance in the water (C and D) and the air (E and F) for taxa observed in > 5% of the samples in the Atlantic and Pacific transects, respectively.

Figure 2.

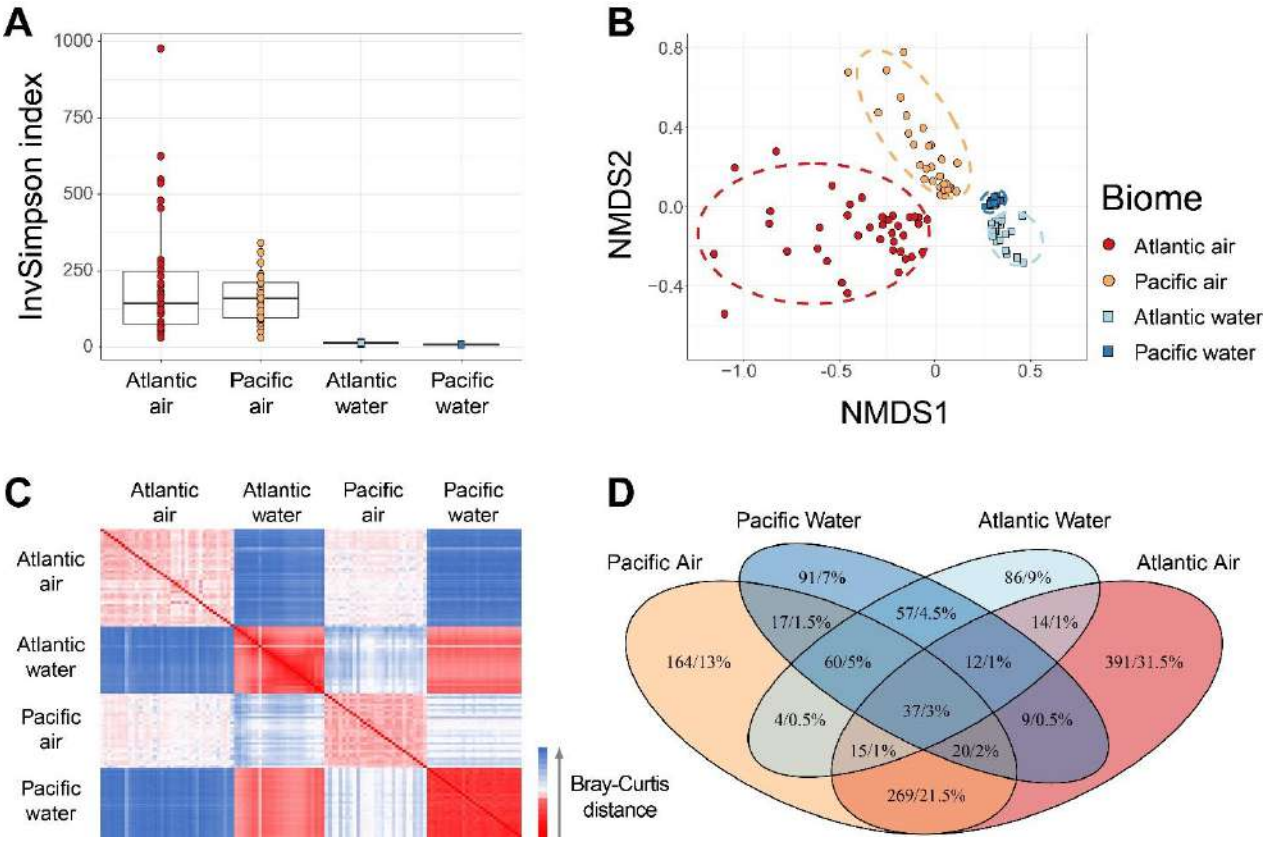
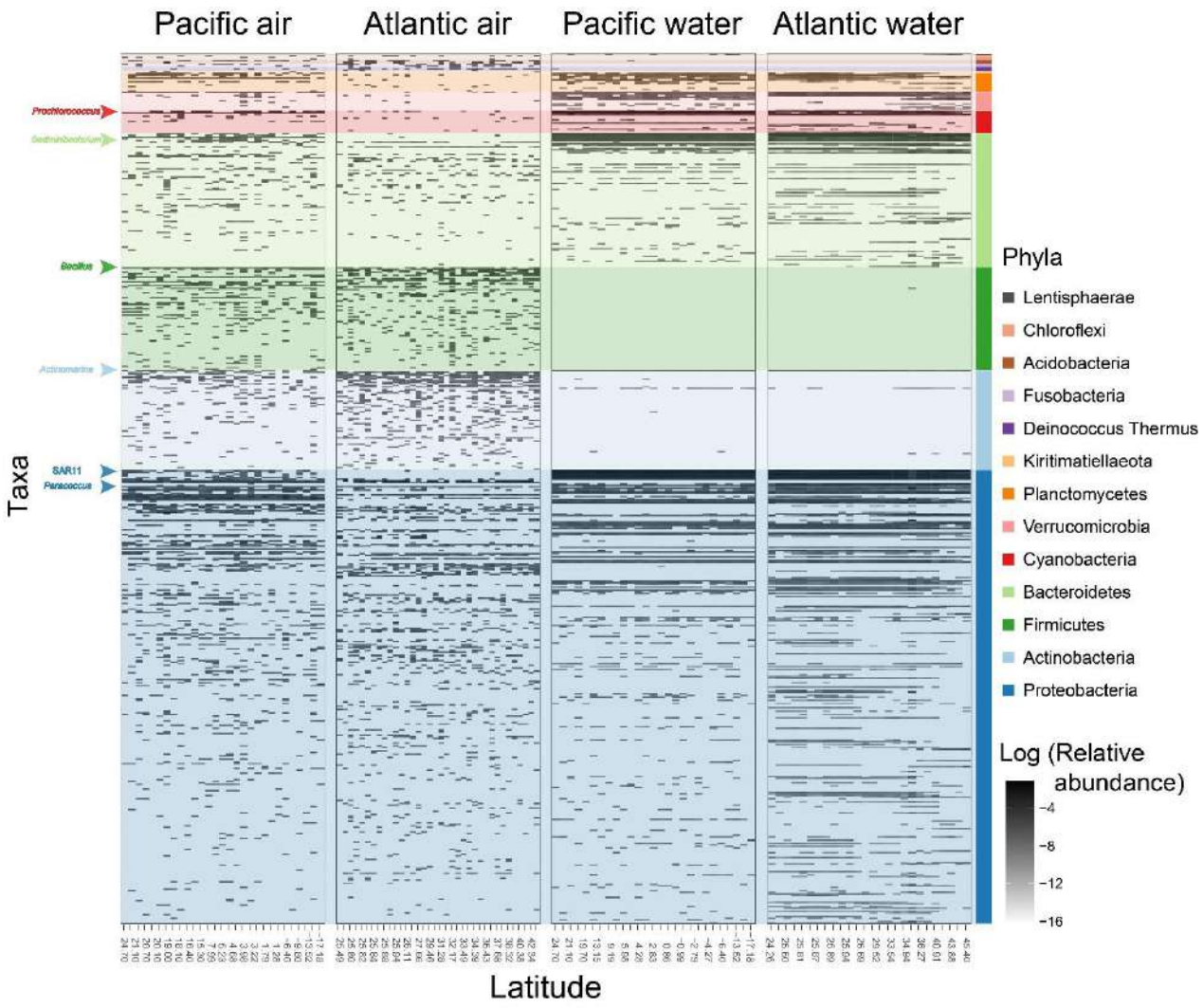


Fig. 2. Similarities and differences in bacterial microbiomes of different oceanic environments. Biome-based diversity (based on the Inverse Simpson index) of amplicon sequence variants (ASVs) detected in more than 5% of the samples (A) and clusters represented by nonmetric multidimensional scaling (NMDS) ordination, based on Bray-Curtis dissimilarity metrics (B), with 95% confidence ellipses. The distance in ASV composition is presented in a heatmap of the Bray-Curtis dissimilarity between all analyzed samples (Bray-Curtis index varies between 0, for an identical ASV composition, and 1, for most distant ASVs on the samples) (C). A Venn diagram (*VennDiagram* 1.6.20) summarizes the number and the percentage out of the total number of bacterial taxa that were observed in the different biomes (D).

772 **Figure 3.**



773

774 **Fig. 3. Spatial distribution of specific airborne microorganisms across the Atlantic and**
775 **Pacific transects.** The 16S amplicon sequence variants (ASVs) of Pacific air, Atlantic Air,
776 Pacific water, and Atlantic water samples are categorized based on abundance for each
777 phylum. ASVs were aggregated according to sequences and taxa annotation, and ASVs
778 covering > 5% of all four environments are presented. Phyla are distinguished based on
779 color code.

780

Figure 4.

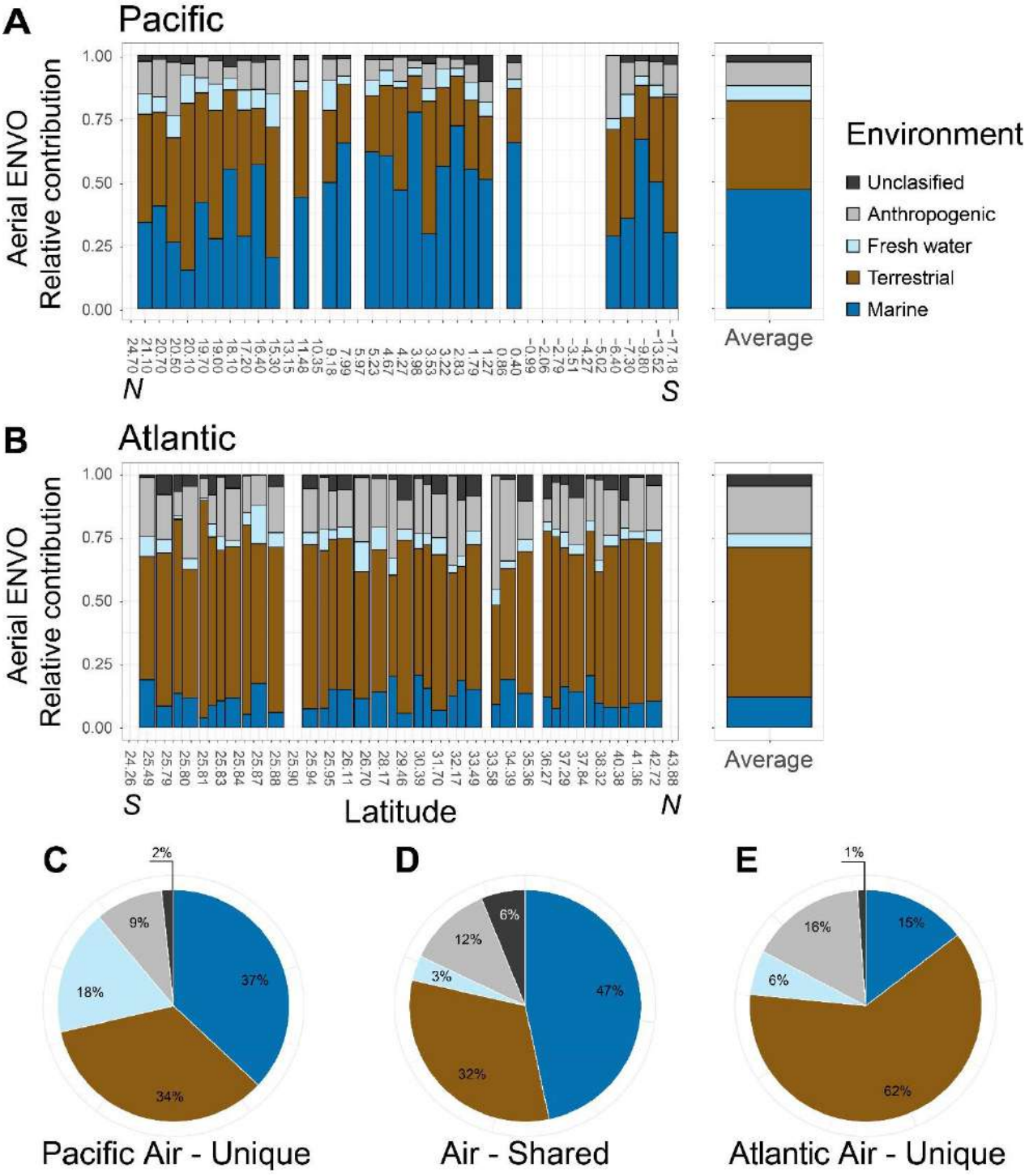


Fig. 4. Association of airborne bacterial communities with other environments. The relative environmental ontology (ENVO) annotation of all amplicon sequence variants (ASVs) identified in the air samples clustered into five main groups (detailed terms are listed in Table S6) for the Pacific (A) and Atlantic (B) transects, as well as for Pacific air-unique ASVs (C), Pacific and Atlantic air-shared ASVs (D), and Atlantic air-unique ASVs (E).

Figure 5.

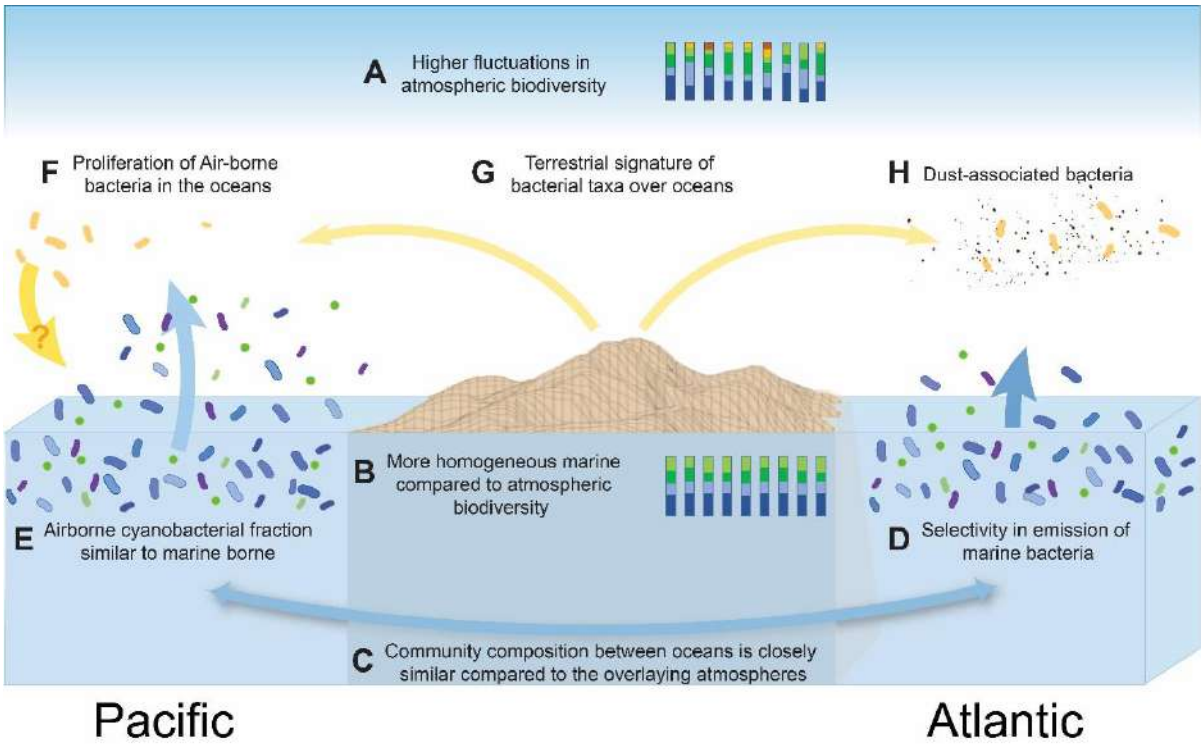


Fig. 5. Trends and transitions in bacterial abundance and richness between marine and atmospheric environments. The main findings of this study are presented as a conceptual model: Residence time induces fluctuations in microbial composition rather than physical mixing (A, and B). The microbial compositions of similar environments (e.g., Atlantic and Pacific Oceans) are more similar than geographical relations (e.g., Pacific atmosphere and surface water; C). The marine bacterial aerosolization efficiency is different for different oceans (D), and aerosolized autotrophs kept a higher ratio *vs*. heterotrophs in the Pacific compared to the Atlantic, similarly as in the water (E). The atmospheric contribution to marine ecology was not detected in the framework of this study (F), but it cannot be excluded that difference in atmospheric microbial composition could impact deposition and surface water composition and function. A terrestrial signature was detected in a remote oceanic environment (G), with a significant contribution of dust-associated bacteria in the Atlantic atmospheric marine boundary layer (H).

Pacific

Atlantic

A

B

Latitude

120 140 160 180 -170 -150

40 20 0 -20 -40

-80 -60 -40 -20 0

60 50 40 30 20

3000 2500 2000 1500 1000 500 0

Meters above MSL

Relative abundance

C

D

Marine phyla

1.00 0.75 0.50 0.25 0.00

Phyla

Rare

Lentisphaerae

Chloroflexi

Acidobacteria

Fusobacteria

Deinococcus Thermus

Kiritimatiellaeota

Planctomycetes

Verrucomicrobia

Cyanobacteria

Bacteroidetes

Firmicutes

Actinobacteria

Proteobacteria

E

F

Aerial phyla

1.00 0.75 0.50 0.25 0.00

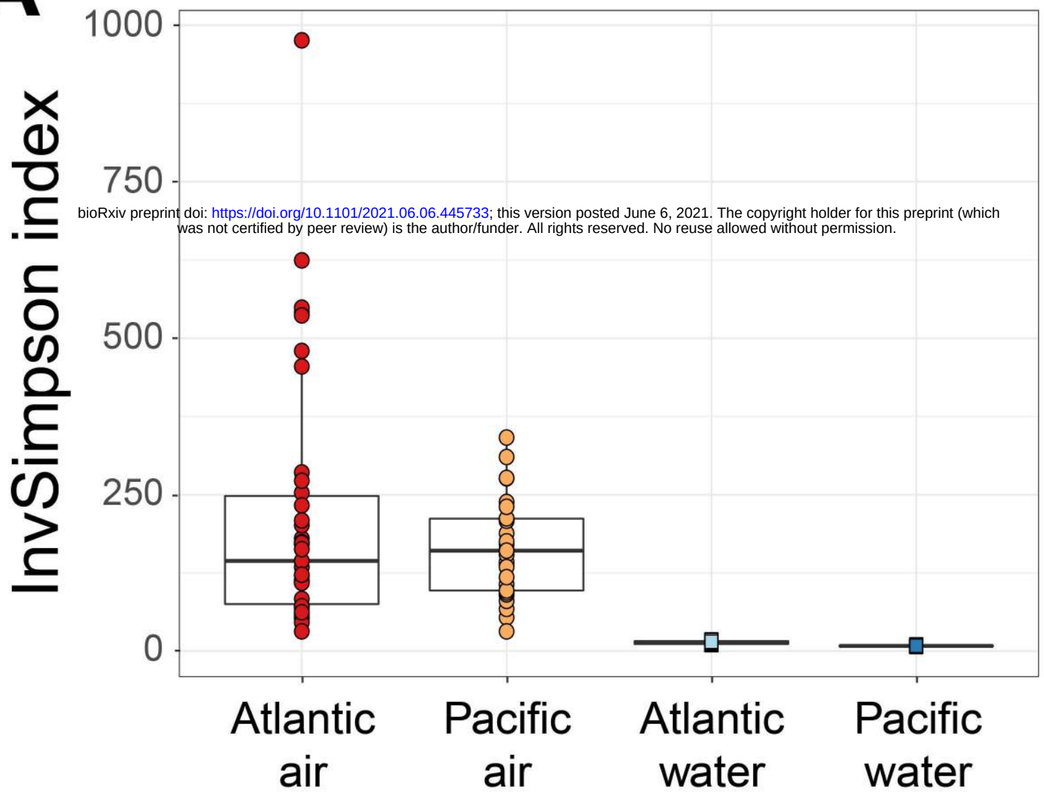
24.70 20.70 20.10 19.00 17.20 15.30 11.48 9.18 5.97 4.67 3.98 3.22 1.79 0.86 −0.99 −2.79 −4.27 −6.40 −9.80 −17.18

N Latitude S

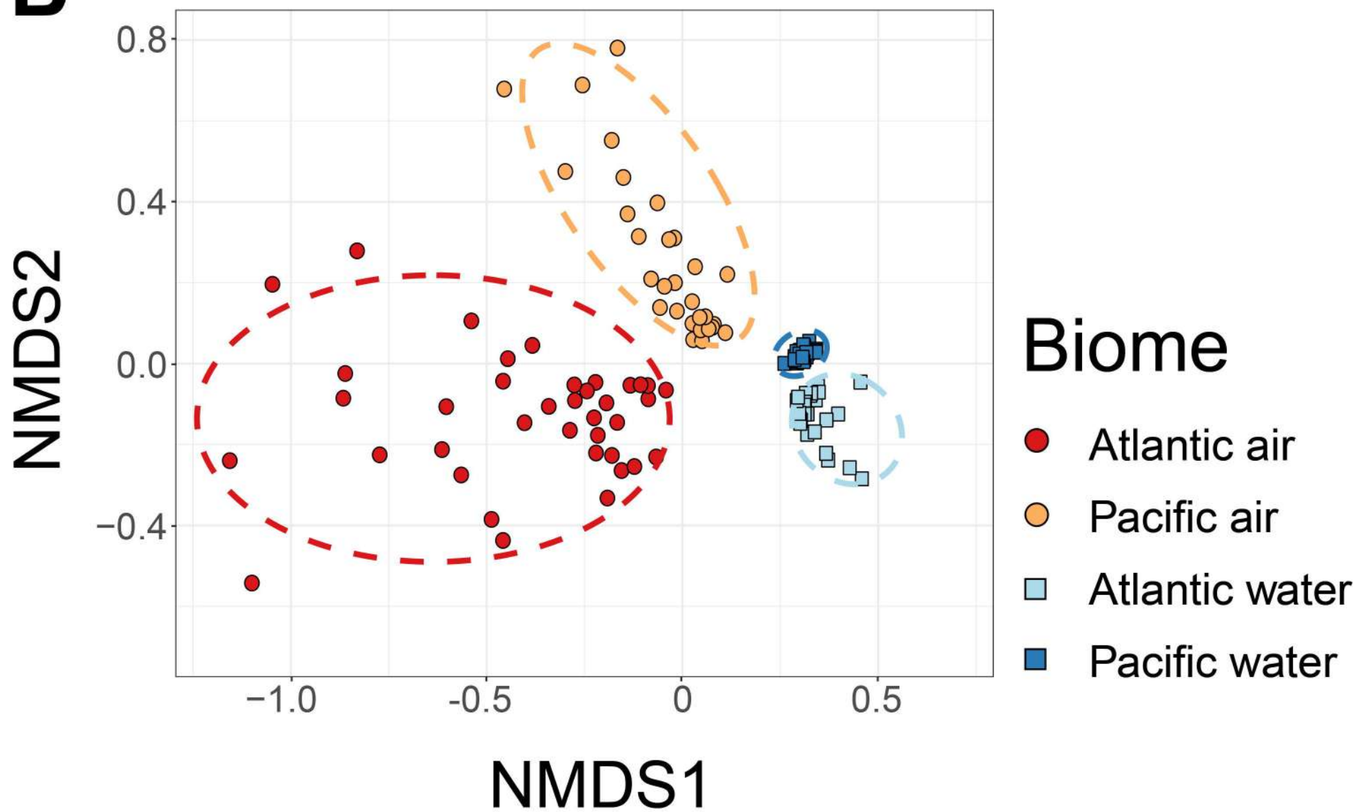
24.26 25.79 25.81 25.84 25.88 25.94 26.11 28.17 30.39 32.17 33.58 35.36 37.29 38.32 41.36 43.88

S Latitude N

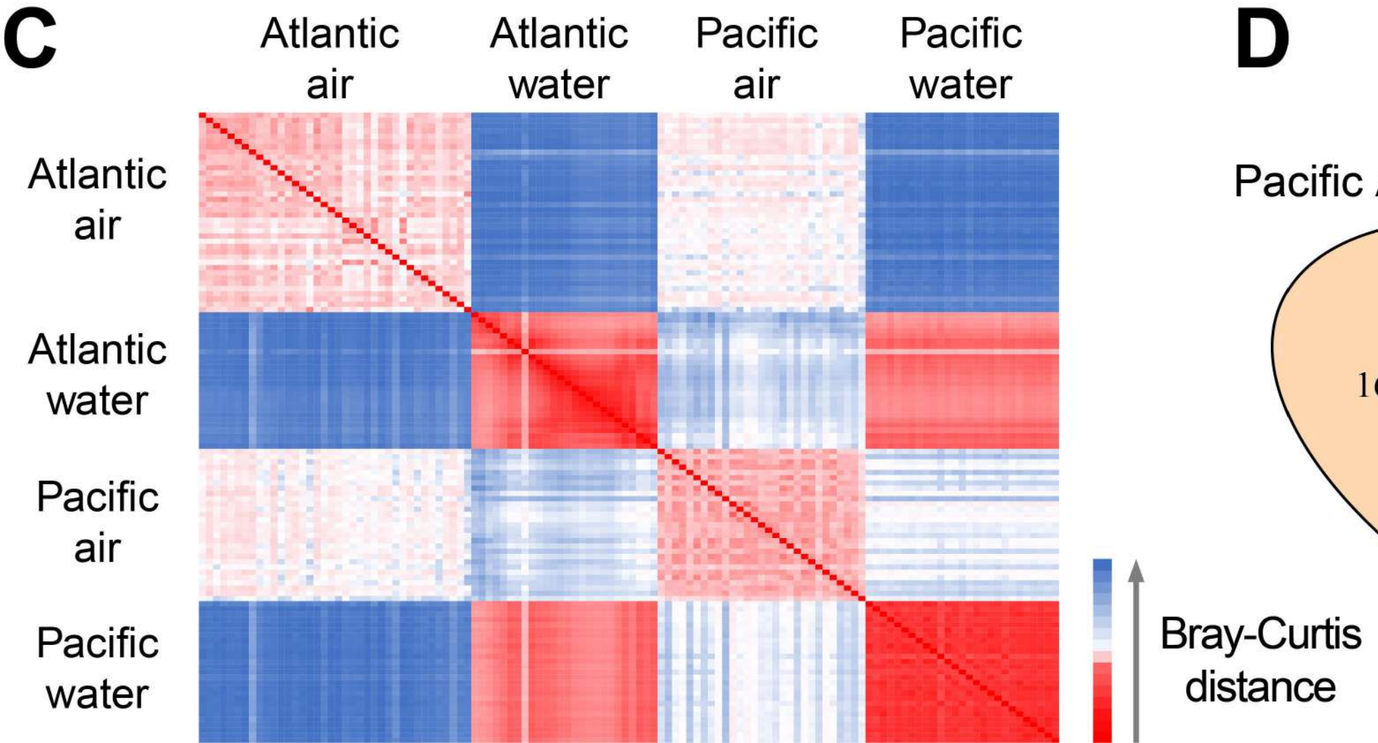
A



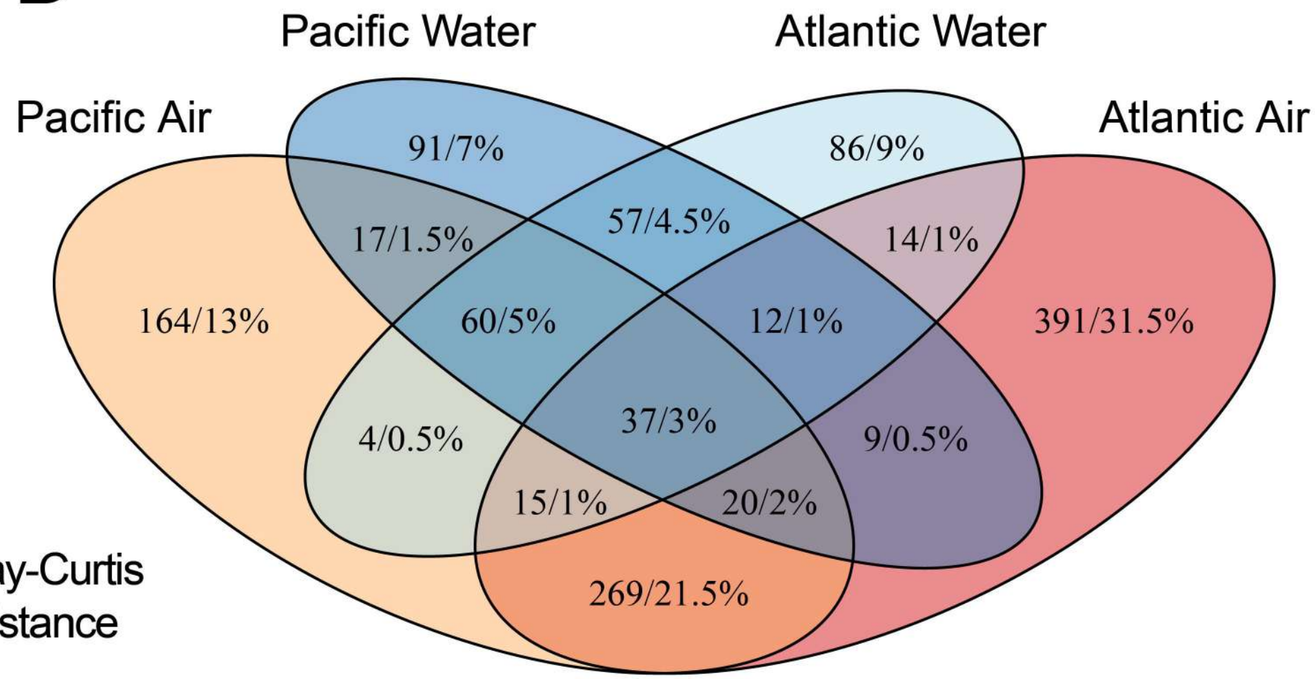
B



C



D



Pacific air

Atlantic air

Pacific water

Atlantic water

Taxa

Prochlorococcus

Sediminibacterium

Bacillus

Actinomarina

SAR11

Paracoccus

24.70 21.10 19.70 13.15 9.19 5.98 4.28 2.83 0.86 −0.99 −2.79 −4.27 −6.40 −13.52 −17.18

25.49 25.80 25.82 25.84 25.88 25.94 26.11 27.66 29.46 31.28 32.17 33.49 34.39 36.43 37.68 38.32 40.38 42.34

24.70 21.10 20.70 20.10 19.00 18.10 16.40 15.30 7.99 5.23 4.68 3.98 3.22 1.79 1.28 −6.40 −9.80 −13.52 −17.18

24.26 25.60 25.81 25.87 25.89 25.94 26.69 29.52 33.54 34.94 36.27 40.91 43.88 45.40

Latitude

Phyla

- Lentisphaerae
- Chloroflexi
- Acidobacteria
- Fusobacteria
- Deinococcus Thermus
- Kiritimatiellaeota
- Planctomycetes
- Verrucomicrobia
- Cyanobacteria
- Bacteroidetes
- Firmicutes
- Actinobacteria
- Proteobacteria

Log (Relative abundance)

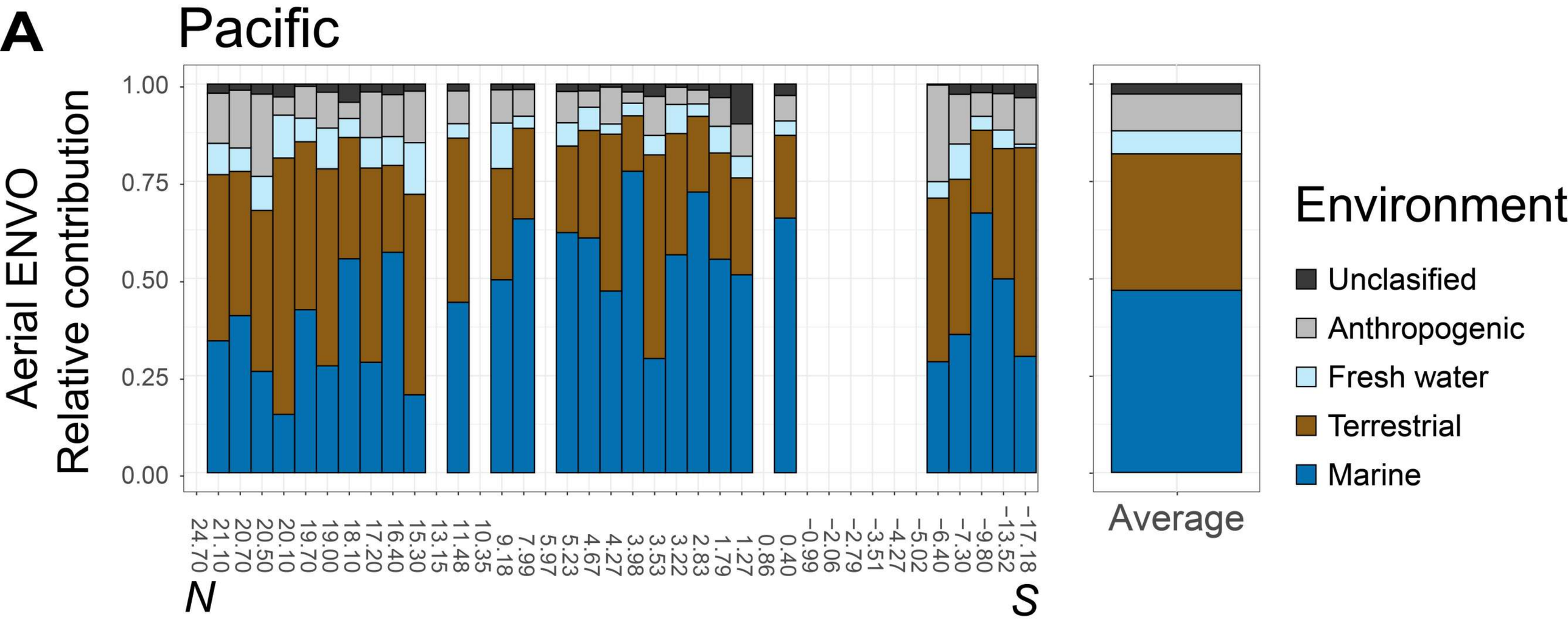
−4

−8

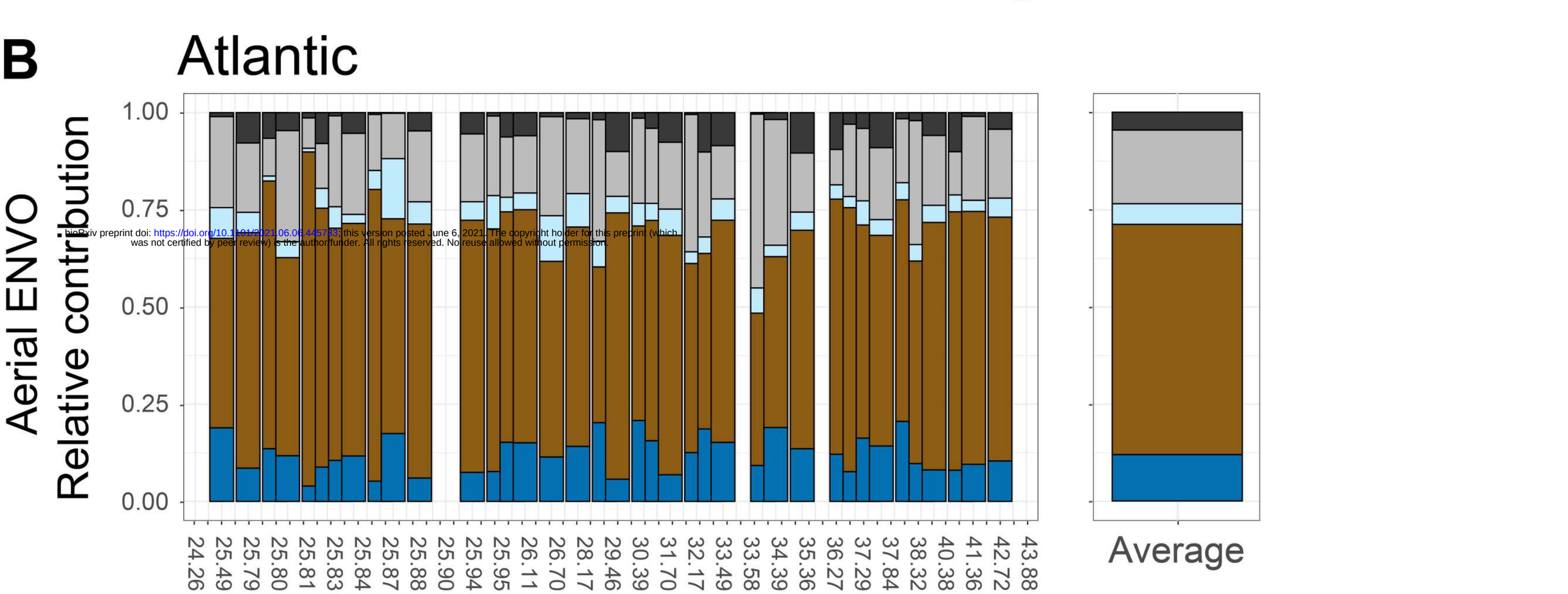
−12

−16

A

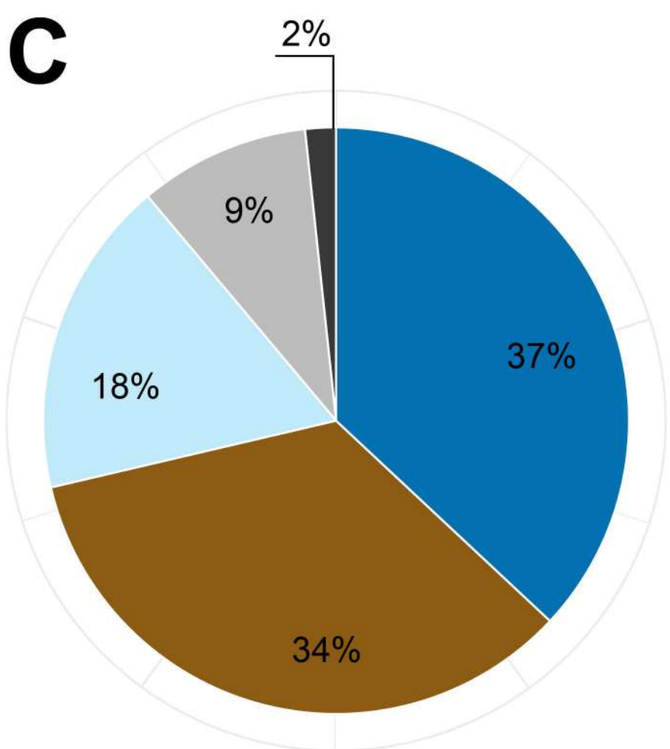


B



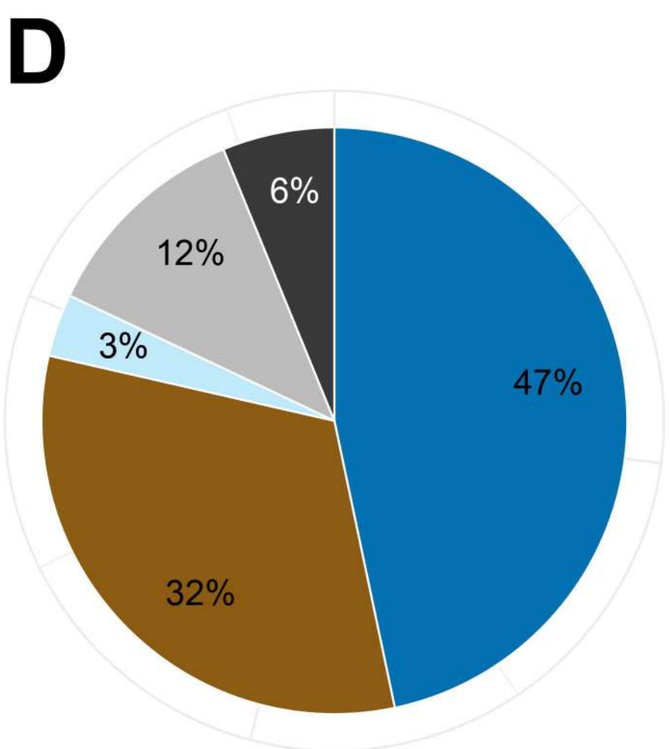
Latitude

C



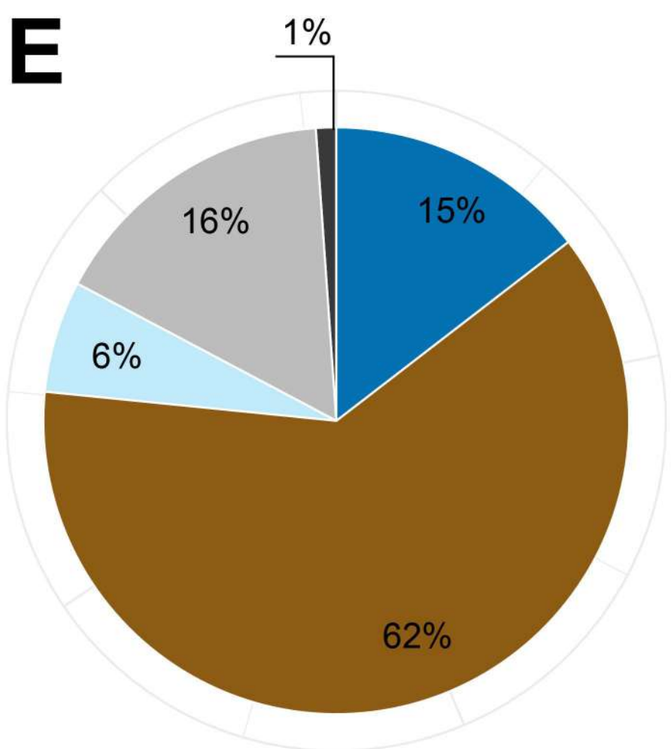
Pacific Air - Unique

D



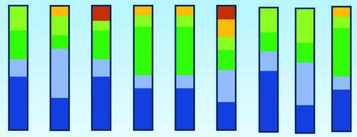
Air - Shared

E



Atlantic Air - Unique

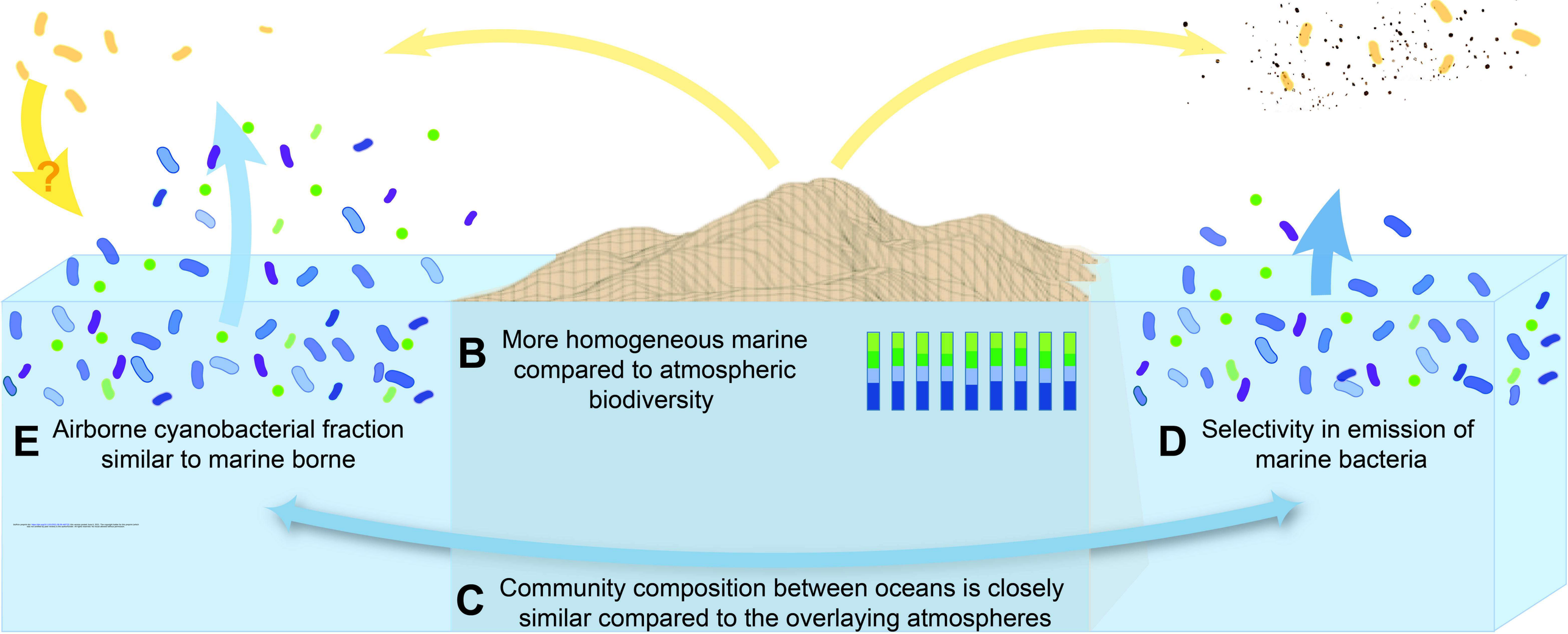
A Higher fluctuations in atmospheric biodiversity



F Proliferation of Air-borne bacteria in the oceans

G Terrestrial signature of bacterial taxa over oceans

H Dust-associated bacteria



B More homogeneous marine compared to atmospheric biodiversity

E Airborne cyanobacterial fraction similar to marine borne

D Selectivity in emission of marine bacteria

C Community composition between oceans is closely similar compared to the overlaying atmospheres

Pacific

Atlantic