Environmental Microbiology

ENVIRONMENTAL MICROBIOLOGY

RESEARCH ARTICLE

Nutrient Transfer Hypothesis for the Fluctuation and Regional Variation of Microbial Diversity in the Pacific Ocean

Xu Dong[1] | Xiangang Hu[1] | Kai Hu[1] | Zehou Li[1] | Weilei Wang[2]

[1]Key Laboratory of Pollution Processes and Environmental Criteria (Ministry of Education), Carbon Neutrality Interdisciplinary Science Centre, College of Environmental Science and Engineering, Nankai University, Tianjin, China | [2]State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, Fujian, China

Correspondence: Xiangang Hu (huxiangang@nankai.edu.cn)

Received: 18 February 2026 | **Revised:** 17 August 2026 | **Accepted:** 10 September 2026

Keywords: diversity | grazing | metagenomics | nutrient transfer | ocean | threshold

ABSTRACT

Microbial diversity affects marine biogeochemical cycles and sustainability. However, the processes and feedbacks that control the spatial and temporal trends in species diversity and functional gene diversity remain largely unknown at the global marine scale. We analysed 1582 marine metagenomic samples and 19 variables related to marine climate and microbial community dynamics. Over the past 12 years, marine microbial species diversity has significantly decreased across the surface ocean, especially in the Pacific. The greatest decrease reached 24.29% in the central region of the North Pacific Ocean (NPO). We propose a food chain-based nutrient transfer hypothesis in which a decrease in grazing intensity is associated with a decrease in microbial diversity. This hypothesized relationship was further examined using 27 years of independent time series data from a Pacific site. The Pacific Ocean presented high sensitivity to grazing intensity coupled with environmental fluctuations of < 5%, in which species diversity in the NPO and South Pacific Ocean (SPO) regions changed from −9.62% to 13.02% and from −8.38% to 15.03%, respectively. The incorporation of nutrient transfer concepts to reduce the unpredictable uncertainty in marine biogeochemical cycles and ecological stability related to microbial diversity is urgent, especially in the vulnerable Pacific Ocean.

1 | Introduction

As key components of Earth's ecosystems, marine microbial communities constitute 90% of global ocean biomass and constitute the core of global biogeochemical cycles and ecological processes (Arrigo 2005; Cavicchioli et al. 2019). Microbial species diversity and functional gene diversity contribute to ecosystem stability and functionality, driving matter cycling and energy flow through species interactions and functional gene complementarity, which are crucial for ecological sustainability (Zhivkoplias et al. 2024). Taxonomic and functional diversity reflects how microbial communities respond and adapt to changes in resources and stresses (Wang et al. 2023). Both dimensions exhibit extensive spatial and temporal variability. Global marine microbial communities are undergoing rapid changes, which are driven by various climate changes and anthropogenic activities (Salazar et al. 2019). Although some studies have investigated the spatial and temporal patterns of microbial diversity (Pommier et al. 2007; Giovannoni and Vergin 2012; Louca et al. 2016), the processes and feedbacks that control the spatial and temporal trends in species diversity and functional gene diversity remain unclear, and their sensitivity to fluctuations in the existing environment is not well explained.

Microbial community surveys based on field sampling exhibit significant geographic sparsity (Ladau et al. 2013).

Comprehensive frameworks and data to characterize the distribution of specific species or functional genes accurately at the global or local regional scale are lacking, resulting in geographic mysteries. Species distribution modelling (SDM) has been used for the prediction of specific species diversity, but it typically relies on species occurrence or absence data, tends to ignore community structure, and simplifies environmental situations (Norberg et al. 2019). Patterns of microbial community dynamics require systematic attention; the incorporation of multiple factors, such as pollution, community grazing relationships, and environmental variables, into intelligent models; and more attention to potential cascading effects and geographic differences.

The Kunming–Montreal Global Biodiversity Framework prioritizes the conservation of global biodiversity. With intensifying changes in microbial species composition and functional expression, the long-term dynamic patterns, hotspots of diversity loss and sensitivity to environmental changes still face considerable uncertainty (Ladau et al. 2013). Existing models (e.g., species distribution models and ecological network models) may underestimate the risk of loss and the protection potential of marine microbial diversity under climate change and anthropogenic pressure (Chen et al. 2024). It is difficult to systematically analyse the biotic and abiotic factors affecting marine microbial diversity (Ladau and Eloe-Fadrosh 2019). To elucidate the underlying mechanisms and complex feedback processes (Bissett et al. 2013), it is necessary to integrate the composition and function of marine microbial communities to explore the synergistic effects of environmental gradients, biotic interactions and evolutionary processes on the formation and maintenance of diversity.

Here, we subjected 1582 marine metagenomic samples to classification and functional annotation (Figure S1A–C) and constructed a dataset ($1 \times 1°$) spanning 12 years (2009–2020). The predicted spatial and temporal patterns of marine microbial diversity (including species and functional genes) quantitatively and accurately revealed the hotspots of biodiversity change. On this basis, we propose the nutrient transfer hypothesis, which posits that nutrient availability, phytoplankton biomass, and grazing together shape regional variation in upper-ocean microbial species and functional gene diversity through linked bottom-up resource supply, nutrient transfer, and trophic regulation (Figure S2 and Table S1). Sensitivity analysis revealed that minor environmental fluctuations could lead to a large loss and decreased protection potential for microbial diversity. The significant decreasing trends and high sensitivity of diversity indicated that critical marine areas require further investigation in terms of ecological status. Accurate estimations of the underestimated losses and variations in marine microbial diversity could help address the hidden instability of marine ecosystems.

2 | Materials and Methods

2.1 | Analysis of Ocean Metagenomic Data and Processing Protocols

Marine metagenomic sequences, including 1582 samples collected between 2009 and 2020, were retrieved from the Sequence Read Archive (SRA) (file size ≥ 3 Gb for each sample). These samples originated primarily from three major global ocean sampling programs—Tara Oceans, Bio-GO-SHIP, and BioGEOTRACES (Sunagawa et al. 2020; Larkin et al. 2021)—which are well-organized projects and can minimize sample bias (Table S2). The original sampling was accompanied by corresponding time, latitude, longitude, depth, and measured environmental data for preliminary analyses. All samples were collected within the upper 200 m of the water column with a balanced seasonal distribution, effectively preventing potential biases from environmental and seasonal characteristics (Supporting Information 1). According to the species annotation results, the metagenomic samples contained sequences of bacteria, archaea, viruses, and eukaryotes. Microbial species diversity and KEGG-annotated functional gene diversity were characterized separately. Species diversity was calculated from the relative abundances of microbial species, whereas functional gene diversity was calculated from the relative abundance of KEGG-annotated functional units. Therefore, functional gene diversity represents the diversity and evenness of annotated functional gene profiles rather than realized microbial activity or functional redundancy. The Shannon–Wiener index was calculated for both profiles using the "vegan" package in R 4.2.1. The detailed processing steps can be found in the Supplementary text.

2.2 | Marine Data Preparation

Seventeen marine physicochemical and biological properties were obtained from the Copernicus Marine Environment Monitoring Service (CMEMS) and the Coupled Model Intercomparison Project Phase 6 (CMIP6). These environmental factors include the eastward sea water velocity (LatV), northward sea water velocity (LonV), mixed layer thickness (Mix), salinity, temperature (Tem), precipitation, pH, total grazing of phytoplankton by zooplankton (grazing), and concentrations of chlorophyll *a* (Chla), dissolved iron (Fe), dissolved oxygen (O_2), nitrate (NO_3^-), phosphate (PO_4), silicate (Si), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), and particulate organic carbon (POC). Human activity data were calculated from cumulative human impacts (CHIs), accounting for 14 human stressors related to climate change, fishing, land-based pressures, and other commercial activities (Halpern et al. 2019), spanning from 2003 to 2013. The marine MP concentration (items/km^2) data between 37 degrees north and south latitudes were obtained from the Cyclone Global Navigation Satellite System L3 Ocean Microplastic Concentration V1.0, whereas other data were obtained from the Van Sebille model (van Sebille et al. 2015). Sensitivity analyses further indicated that the absence of some elemental forms (e.g., particulate organic phosphorus) did not materially affect model performance, with the validation R^2 changing by +0.002 for species diversity and −0.004 for functional gene diversity (Figure S3A–C). Gridded environmental variables were matched to each metagenomic sample according to its sampling time (consistent month), geographic coordinates (average difference = 0.06°), and depth (average difference = 1.32 m) to ensure spatial and temporal consistency. Global ocean abbreviations can be found in Table S3. Detailed information concerning the data sources is listed in Table S4 and the Supplementary text. All the variables were derived from well-organized reanalysis products and curated datasets, ensuring the reliability of the environmental information used in this study (Hawco et al. 2025). To verify the accuracy

of the matched product-based variables, we fitted them against the in situ measurements obtained from sensors or bottle samples collected during the cruises. The global products used in this study were highly consistent with the in situ observations (average $R^2 = 0.88$; Figure S4).

2.3 | Machine Learning Model Construction and Validation

Machine learning approaches were employed for regression modelling of marine microbial diversity (with species diversity and functional gene diversity as the output labels), with marine physicochemical properties, climate, and human activity exogenous pressures as the input variables. Random forest (RF), extreme gradient boosting (XGBoost), support vector machine (SVM), and multilayer perceptron (MLP) models were used for performance comparisons, and a grid search was used during the modelling process to find the optimal hyperparameters for the models. Tenfold cross-validation (by ShuffleSplit) was used during the training process to improve data availability while reducing the likelihood of overfitting. Details of the modelling process, parameter tuning, and overfitting tests can be found in the Supplementary text.

To examine whether the RF models showed spatial autocorrelation in the residuals, we applied a spatial eigenvector-based approach (Carle et al. 2025). Moran's eigenvector maps (MEMs) were constructed to represent multiscale spatial structures arising from the geographic distribution of the 1582 metagenomic sampling sites. Spatial eigenvectors were generated using a spatial weighting matrix implemented by the "adespatial" package in R 4.2.1, combining site connectivity (Gabriel graph) with distance-based weights (Bauman et al. 2018). Model residuals (observed minus predicted diversity values) were regressed against the spatial eigenvectors, and the presence of spatial structure in the residuals was assessed with permutation tests (9999 permutations). The RF models performed well ($p = 0.109$ and 0.532 for species diversity and functional gene diversity, respectively), indicating that residual spatial autocorrelation was negligible.

2.4 | Interannual Trend Analysis and Validation of Microbial Diversity

To investigate the spatial distribution patterns of microbial diversity, predictions were made by the trained RF models, which performed better than the other models (XGBoost, SVM, and MLP). On the basis of the constructed database of marine variables, we created a prediction raster with a 1°×1° spatial resolution and performed predictions of marine microbial species diversity and functional gene diversity from 2009 to 2020. We identified spatial clusters of high microbial diversity using the Getis–Ord Gi* statistic. For each grid cell, Gi* compares the diversity values within its local neighbourhood with the mean and variance across all grid cells:

$$G_i^* = \frac{\sum_j w_{ij}x_j - \bar{x}\sum_j w_{ij}}{S\sqrt{\frac{n\sum_j w_{ij}^2 - \left(\sum_j w_{ij}\right)^2}{n-1}}}$$

where x_j is the predicted multiyear mean diversity in grid cell j, w_{ij} defines the spatial neighbourhood, and n, $\bar{x}$, and S represent the number, mean, and standard deviation of grid cells, respectively. For the primary analysis, each grid cell was evaluated together with its eight queen-contiguous neighbours. Positive Gi* z scores indicate spatial clustering of high diversity values, whereas negative scores indicate clustering of low values. The false discovery rate was controlled using the Benjamini–Hochberg procedure, and grid cells were classified with $z > 0$ and $q < 0.05$ as hotspots and those with $q < 0.01$ as core hotspots. Pairwise phylum-level taxonomic beta diversity was partitioned into turnover and nestedness components using the "betapart" package in R 4.2.1.

The temporal trend of marine microbial diversity was determined by the Mann–Kendall (M-K) test, which is a nonparametric trend test (Yu et al. 2024). The M-K test outputs the slope and significance. The slope represents the magnitude of the trend, which can be displayed on the map for each grid cell, and the significance determines whether the calculated trend is statistically significant. The M-K test was implemented using the "pymannkendall" package for Python 3.8. Maps of each microbial diversity and slope in the global surface ocean were generated by the "Basemap" package for Python 3.8. The species abundance and functional gene pathway expression of specific microorganisms were also calculated. We calculated and ranked the relative abundances and defined the species with the top five relative abundances as the dominant taxa. Representative functional gene pathways were also selected (top ranked relative abundance). Recursive trend analysis was applied to ensure that the predicted diversity trends did not simply originate from the temporal structure of the environmental variables (You et al. 2025). We characterized the extrapolation uncertainty (application domain) by comparing the environmental predictor vector of each grid cell with the training data using multivariate environmental similarity surface (MESS) values and Mahalanobis distances. A grid cell with MESS < -20 or a Mahalanobis distance above the 99th percentile of the training distribution was flagged as having higher extrapolation risk (Liang et al. 2025). The detailed processing steps can be found in the Supplementary text.

Additionally, an independent metagenomic dataset (267 samples) from the MiCRO time series, collected at Newport Pier, Newport Beach, California, USA (33.608°N, 117.928°W) between September 2009 and 2022, was used to validate the accuracy of the RF models with the interannual trends (Larkin et al. 2025). Species diversity was calculated on the basis of the provided taxonomic annotations. The corresponding environmental variables were extracted according to the sampling time and geographic coordinates. These data were then input into the established RF models to predict microbial diversity and compared with observed diversity to assess predictive performance.

2.5 | Identification and Spatial Patterns of Major Predictors Associated With Microbial Diversity

The multiway feature importance analysis included seven metrics of three types: model error-based metrics (mean square error increase), tree model structure-based metrics (node

purity increase, mean minimal depth, times as root, number of nodes and p values), and game theory-based metrics (SHAP value). The above methods (except the SHAP value) were implemented via the "randomForestExplainer" package of R 4.2.1. These metrics measure and rank the importance of features by considering changes in model accuracy and the contributions of specific features to the tree structure, thus avoiding errors caused by the use of a single metric (Ishwaran et al. 2010). The SHAP value is based on the concept of Shapley values in game theory and is used to calculate the importance of features by measuring the contribution of each feature to the model prediction (Jia et al. 2024). SHAP analysis was implemented using the "SHAP" package in Python 3.8. To assess the sampling stability of predictor importance, we generated 1000 bootstrap resamples of the training data and calculated permutation importance. Mean importance values with 95% intervals were summarized across iterations. These values quantify correlations rather than causal effects. To understand the spatial distribution patterns of the major predictors associated with microbial diversity, the Shapley value was calculated for each grid cell in 2020. The variable with the largest Shapley absolute value was selected as the dominant model predictor for that grid.

2.6 | Assessment of the Nutrient Transfer Hypothesis

To assess whether the proposed nutrient transfer hypothesis was consistent with independent temporal observations, we used long-term primary production observations collected at the Hawaii Ocean Time-series (HOT) station in the North Pacific (Hynes et al. 2024). After removal of missing values, a total of 253 surface observations from 1995 to 2022 were obtained. The dataset includes measurements of chlorophyll a (mg/m^3), pigments (mg/m^3), light-12 ($mg\ C/m^3/day$), and heterotrophic bacteria (cells/mL). These four variables were selected as proxies for phytoplankton biomass, zooplankton grazing pressure, primary productivity, and heterotrophic microbes to test the nutrient transfer hypothesis (Collos et al. 2005). Convergent cross mapping (CCM) was applied to examine whether the temporal relationships among variables were consistent with the nutrient transfer hypothesis. In nonlinear complex systems, CCM reconstructs the state space attractor of each time series through delay embedding and assesses whether the historical values of one variable can reliably predict the concurrent values of another, thereby inferring causality (Sugihara et al. 2012). We constructed a fully connected network among the four indicators and systematically tested the bidirectional causal relationships for each pair of variables. To assess statistical significance, we conducted 100 surrogate tests by randomly permuting the effect variable to generate null distributions of predictive skill (rho). The results were considered reliable if the forecast skill converged and if the p value was <0.05. CCM was implemented by the "rEDM" package of R 4.2.1.

2.7 | Feature Interaction and Response Analysis

A generalized linear mixed model (GLMM) was used to quantify the effects of different variables on species diversity and functional gene diversity (Van't Veen et al. 2020). After screening predictors using variance inflation factors (VIFs) and standardizing the retained variables, Project_name was included as a random intercept to explain the unobserved heterogeneity among the sampling programs. To assess whether seasonal variation acted as a confounding factor, we additionally fitted two models in which season was included as a fixed effect or random intercept. For each response variable (species diversity and functional gene diversity), we specified a gamma distribution with a log-link function to accommodate the right-skewed distribution. Model performance was evaluated against corresponding null models using Akaike's information criterion (AIC) to ensure adequate model fit. Finally, we computed confidence intervals for fixed effect estimates and extracted associated p values. Variables with $p<0.05$ were considered statistically significant. All results were visualized using coefficient plots with 95% confidence intervals, and significant predictors were highlighted.

To understand how marine microbial species diversity and functional gene diversity respond to changes in environmental factors (climate, ocean physicochemical properties, exogenous pressures, etc.), trends and thresholds corresponding to the impacts were determined by fitting from low to high dimensions on the basis of the established RF model, which includes three parts. (1) Single-feature continuous responses (SHAP value fitting): For SHAP value fitting, the calculated SHAP values and the corresponding raw feature values were fitted by the generalized additive model (GAM). The fitted curves reflected the associations of the feature values to the output, where a y-axis value of 0 was set as the boundary to divide the threshold region; this analysis was performed by using the GAM from the "pygam" package for Python 3.8. (2) Two-feature continuous responses: The values of two features are changed at the same time to observe the different responses, which are used to determine the threshold transformation intervals of influence; this analysis was performed in the "pdp" package for R 4.2.1. (3) Feature interaction network: On the basis of the structural characteristics of the tree models, the strength of interactions between features can be reflected by the number of times two features appear in the forest, and the interaction coefficients are reformulated to reduce the influence of external factors (Yu et al. 2021). Interaction networks can be constructed according to the corresponding node and edge weights; these networks were visualized with Gephi-0.9.2 software, and the interactions were calculated with the "randomForestExplainer" package for R 4.2.1.

2.8 | Biogeochemical Cycling of Marine Microorganisms

To determine the relative abundance of metabolic and biogeochemical functional pathways in the metagenomic data, DiTing software (Xue et al. 2021) was used to perform calculations and draw the corresponding cycling pathways (carbon, nitrogen, and sulfur cycles). On the basis of the KEGG database and the created cycling gene database, the relative abundance of each pathway corresponding to the KO (KEGG Orthology) number associated with the functional cycles was calculated. KEGG pathways with significantly different functional abundances

between data with positive and negative correlations between species diversity and functional gene diversity were explored. In addition, Spearman's correlations between functional gene abundance and variables such as climate, physical and chemical properties, and human activities were determined by the "seaborn" package of Python 3.8.

2.9 | Construction of Constrained Scenarios for Microbial Diversity

To explore the potential decrease and increase in microbial diversity, a multiobjective optimization algorithm (nondominated sorting genetic algorithm, NSGA-II) was used to construct constrained exploratory scenarios (downward and upward). NSGA-II is a fast sorting and elite multiobjective genetic algorithm that searches in a preset solution space to achieve a specific optimization objective (adjusting both species diversity and functional gene diversity) (Deb et al. 2002). NSGA-II was implemented through the "geatpy" package in Python 3.8. The construction of the best- and worst-case scenarios was based on the global data for 2020 (containing both input features and prediction results). To assess sensitivity to the assumed optimization domain, we repeated the NSGA-II analysis using relative bounds of ±5%, ±10%, ±15%, ±20%, ±25% and ±30% around the 2020 baseline values. We also applied ecological and physical constraints (based on grid-specific historical data), defining the admissible range of each predictor from its 2009–2020 minimum and maximum annual values (e.g., nitrate at a given grid cell was restricted to its local 2009–2020 concentration range). All the solutions were interpreted as constrained exploratory scenarios rather than predictive outcomes. Details can be found in the Supplementary text. On the basis of the constructed upstream and downstream tasks of prediction and optimization, the best- and worst-case scenarios of marine microbial diversity were identified. The results were mapped using the "Basemap" package of Python 3.8.

2.10 | Statistical and Uncertainty Analysis

All the statistical analyses were conducted with Python 3.8 software. Data transformation, normalization, and spatial–temporal matching were performed as described in the corresponding sections. Linear regressions of microbial species diversity and functional gene diversity were performed through the "statsmodels" package for Python 3.8, and statistical tests were performed through two-sided *t*-tests, with *p* values reflecting whether the significance test was passed ($p < 0.05$). The fit of the GAM to the SHAP values was determined with the same *t*-tests as above. Structural equation models (SEMs) were used to understand the network responses of microbial species diversity and functional gene diversity to complex ecosystems. The SEMs were constructed in the "piecewiseSEM" package of R 4.2.1, which evaluated prespecified directional associations among multiple variables (Lefcheck 2016).

Monte Carlo simulation was used to estimate the uncertainty. The uncertainty from the data was simulated by adding random Gaussian noise (mean of 0 and standard deviation of 0.1) to the training data, and the uncertainty from the model was simulated by presetting the change intervals for the model hyperparameters. Model uncertainty was evaluated on the basis of the RF regressions, with the hyperparameters max _features and *n*_estimators. Simulations were obtained after the iterations ($n = 1000$), and the coefficients of variation (standard deviation/mean) were used to represent the uncertainty. The uncertainty results were mapped to $1° \times 1°$ grid cells to understand the spatial distribution of the uncertainty in 2020. In this study, model uncertainty refers to limitations in model estimates or inference, whereas ecological variability refers to observed spatial and temporal variation. These analyses quantify responses to prescribed noise and hyperparameter ranges without fully capturing measurement error, model structural uncertainty or sampling bias. Model-derived relationships were therefore interpreted as associations consistent with the nutrient transfer hypothesis rather than deterministic or causal responses.

3 | Results

3.1 | Hotspots and Significant Declining Trends in Microbial Diversity in the Pacific Ocean

Here, marine microbial species diversity and functional gene diversity were predicted at the global scale, and prediction maps spanning 12 years (2009–2020) were generated (Figure S5). The established random forest (RF) models performed well and did not suffer from spatial autocorrelation in the residuals ($p = 0.109$ and 0.532 for species diversity and functional gene diversity, respectively; Figure S6). Most grid cells fell within the environmental conditions represented by the training data. Only 6.38% had MESS values < -20, and 5.18% exceeded the 99th percentile threshold for Mahalanobis distance (Figure S7A–D). Environmentally dissimilar cells were concentrated mainly at high latitudes and in distinct coastal regions, where predictions of microbial diversity should be interpreted with caution. The additional 11 years of validation data also provided evidence of good model fit (the correlation coefficient was 0.713, Figure S8). Getis–Ord Gi* analysis revealed a prominent species diversity hotspot across the low-latitude Pacific near South America (Figure S9A,B), where diversity was 10.46% greater than that in other regions (Figure 1A; Figure S10). In contrast, the species diversity of the South Atlantic Ocean (SAO) was relatively low (Figure 1A). Microbial functional genes showed greater diversity in high-latitude areas (Figure 1B; Figure S9A,B). Across the modelled surface ocean (0–5 m) without considering a deeper water column and high latitude, marine microbial species diversity significantly decreased by 0.16% per year ($p < 0.05$). The turnover component (0.050) exceeded the nestedness component (0.035), indicating broadly similar phylum-level community composition among samples and a greater role of taxon replacement in the observed differences (Figure S11). However, the functional gene diversity decreased slightly (Figure 1C). As hotspots of species diversity, both the North Pacific Ocean (NPO) and South Pacific Ocean (SPO) showed significant declining trends from 2009 to 2020, with decreases in species diversity of 1.87% and 0.70%, respectively (Figure S12). Compared with the SPO, the NPO had greater loss risks, with the greatest decrease (24.29%) in species diversity in the

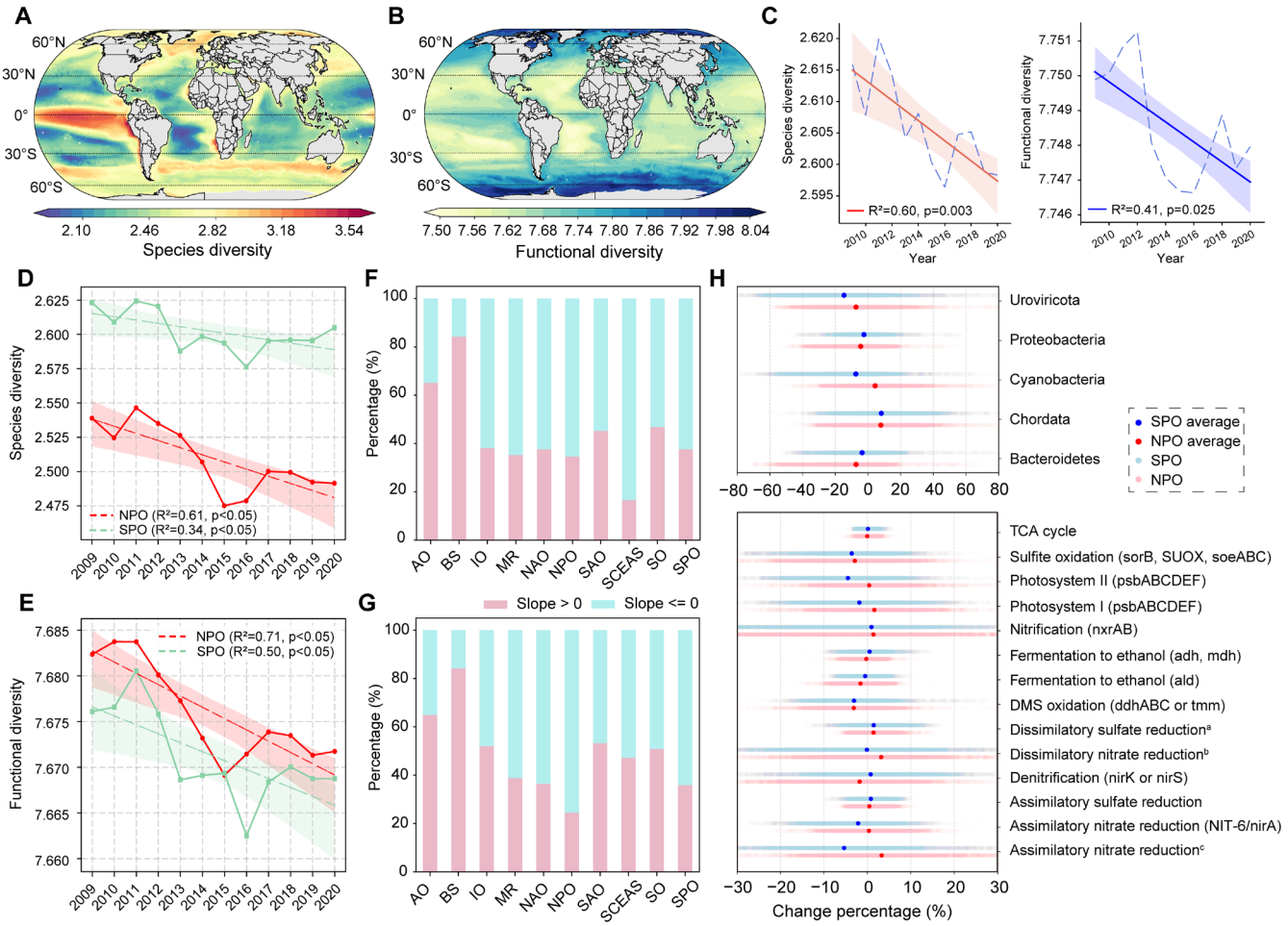

FIGURE 1 | Distribution patterns and change trends of marine microbial species diversity and functional gene diversity from 2009 to 2020. (A, B) Distribution of average microbial species diversity and functional gene diversity from 2009 to 2020. (C) Multiyear trends in the average values ($N = 12$) of microbial diversity. (D, E) Change trends of marine microbial diversity in the North Pacific Ocean (red) and South Pacific Ocean (green) from 2009 to 2020. (F, G) Proportion of increasing or decreasing diversity change trends in different marine areas from 2009 to 2020 (Mann–Kendall test). Species diversity (F) and functional gene diversity (G). (H) Percentages of change in the relative abundances of dominant taxa and functional gene pathways in the North Pacific Ocean and South Pacific Ocean microbial communities in 2020 compared with those in 2009. Dissimilatory sulfate reduction (sat, aprAB)[a], dissimilatory nitrate reduction (nirBD/nrfAH)[b], and assimilatory nitrate reduction (narB/NR/nasAB)[c].

central region (from 17.0°N to 18.0°N and from −148.0°E to −147.0°E) (Figure 1D,E; Figure S13).

Over the past 12 years, microbial species diversity and functional diversity have shown decreasing trends across ~56.9% (with 19.5% of the regions showing $p < 0.05$) and 55.5% (with 19.3% of the regions showing $p < 0.05$) of the global surface ocean, respectively, with species diversity showing a more prominent decreasing trend in the NPO (0°N–30°N) than in other regions (Figure 1F,G; Figure S14). The SCEAS exhibited the greatest percentage decrease (Figure 1F,G) in species diversity. With respect to functional gene diversity, the NPO was the region with the greatest decrease in its area (76%). Changes in diversity reflected variations in the evenness of individuals across taxa. The dominant taxa, such as cyanobacteria, proteobacteria, uroviricota, and bacteroidetes, which are related to the global carbon and nitrogen cycles, all showed more pronounced changes in the Pacific Ocean and polar regions (Figure S15). The variations in cyanobacterial abundance differed significantly between the NPO and the SPO, reflecting the risk heterogeneity of the Pacific region and potential perturbations to biogeochemical cycling (Figure 1H). The abundance of genes related to nitrogen cycles changed more significantly, and the abundance of genes related to assimilatory nitrate reduction (nitrite to ammonia) decreased significantly in the high-latitude regions of the SPO (Figure 1H; Figure S16). Significant declines in microbial species and functional diversity occurred mainly in the Pacific Ocean, with differences between the NPO and the SPO.

3.2 | Nutrient Transfer Hypothesis Dominated by Grazing, Nitrate and Chlorophyll Levels

The long-term trends in diversity were not mechanical replication of the temporal trend in the predictors but rather changes driven by the cascading effects of the complex environment (Figure 2A; Figure S17). After recursive trend analysis of all the variables with significant temporal trends, the predicted diversity still showed declining trends, although the slope was reduced compared with the original predictions (Figure S18). The integration of multiple metrics ensured that potential bias was controlled and

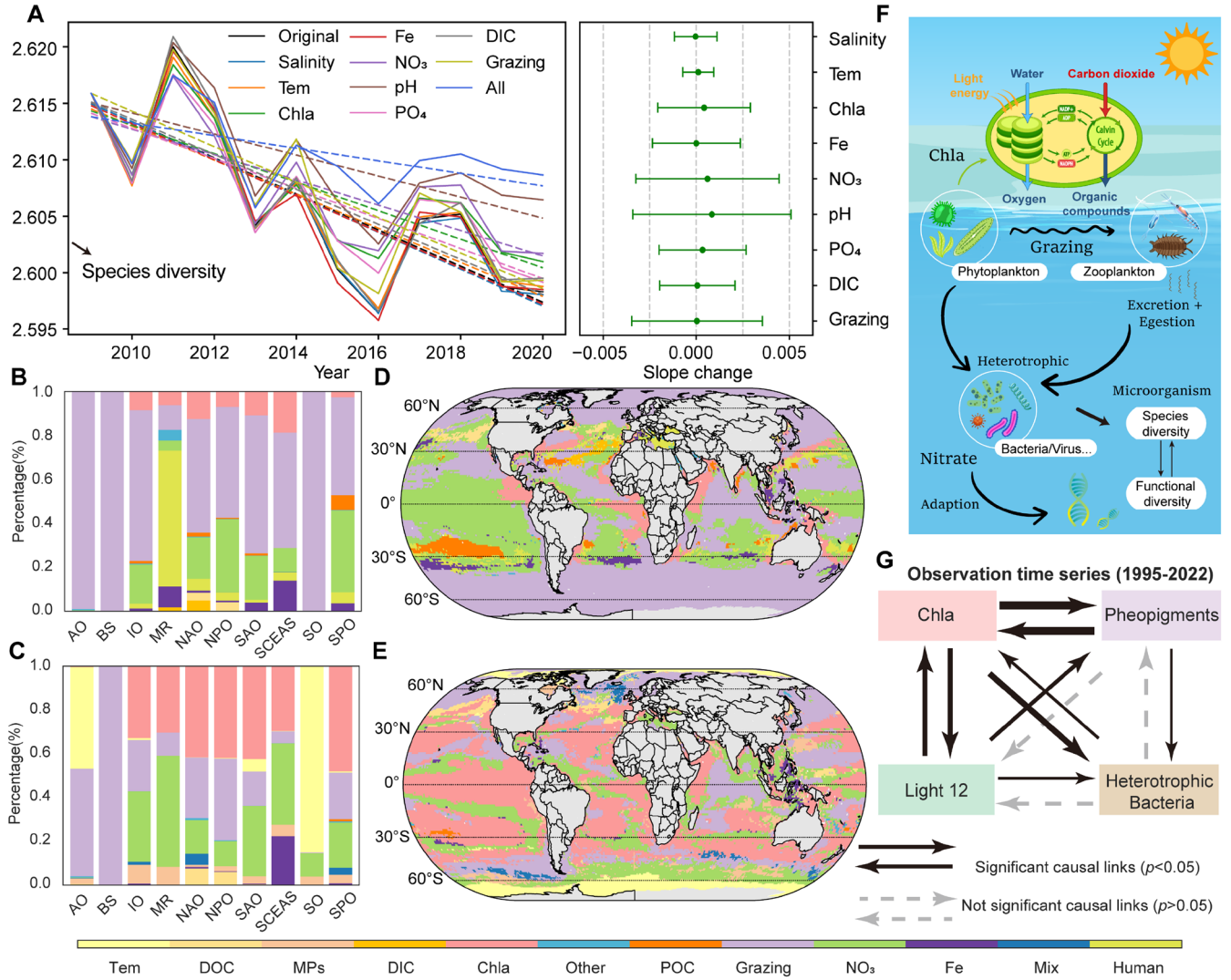

FIGURE 2 | Microbial diversity is regulated predominantly by grazing coupled with nitrate and chlorophyll levels. (A) Annual trends in species diversity generated using recursive trend analysis. The solid lines indicate predictions with individual features held constant (2009), and the dashed lines show their linear fits. The change in slope reflects the contribution of each feature. (B, C) Percentages of the different associated variables in the surface ocean measured by SHAP values for species diversity (B) and functional gene diversity (C). (D, E) Spatial distribution patterns of major factors of species diversity (D) and functional gene diversity (E) in the surface ocean. (F) The proposed nutrient transfer hypothesis, by which key factors influence microbial species diversity and functional gene diversity. (G) Assessment of the nutrient transfer hypothesis using convergent cross mapping (CCM) at a Pacific observational time series site (1995–2022). Line thickness represents the magnitude of the cross-map skill for each proposed relationship.

revealed that grazing was the variable most strongly associated with species diversity. Notably, functional gene diversity was critically regulated by multiple factors, including grazing, nitrate (NO_3^-), and chlorophyll (Chla) levels (Figure 2A; Figure S19A,B). Moreover, there were strong interactive effects of grazing, NO_3^-, and Chla on diversity (Figure S20A,B). With respect to species diversity, grazing was identified as the dominant predictor in most regions, especially in the BS, Arctic Ocean (AO), and Southern Ocean (SO), where trophic level limitations were associated with microbial species diversity and structure (Figure 2B). With respect to functional gene diversity, the spatial patterns of the dominant predictors are more diverse, with Chla having an important role in most areas. NO_3^- was also identified as an important predictor of functional gene diversity (Figure 2C). As emerging contaminants, microplastics (MPs) were also identified as notable model predictors in parts of the Mediterranean region (MR) and Indian Ocean (IO). Grazing, NO_3^-, and Chla were identified as the dominant predictors associated with microbial diversity across most regions of the Pacific Ocean (Figure 2D,E). Bootstrap resampling revealed that grazing, NO_3^- and Chla were the strongest predictors of species diversity and functional gene diversity (Figure S21).

Here, we propose the nutrient transfer hypothesis, according to which changes in the food chain may lead to changes in microbial species and functional diversity (Figure 2F; Figure S2 and Table S1). The nutrient transfer hypothesis supports the dynamic relationships between phytoplankton, zooplankton, and microbes (heterotrophic bacteria and viruses). At the long-term Pacific time series station (1995–2022), we tested this hypothesis using convergent cross mapping (CCM). Among the proxies for the nutrient transfer hypothesis, significant causal relationships ($p < 0.05$) were detected between chlorophyll a concentration, primary productivity (Light-12), grazing intensity

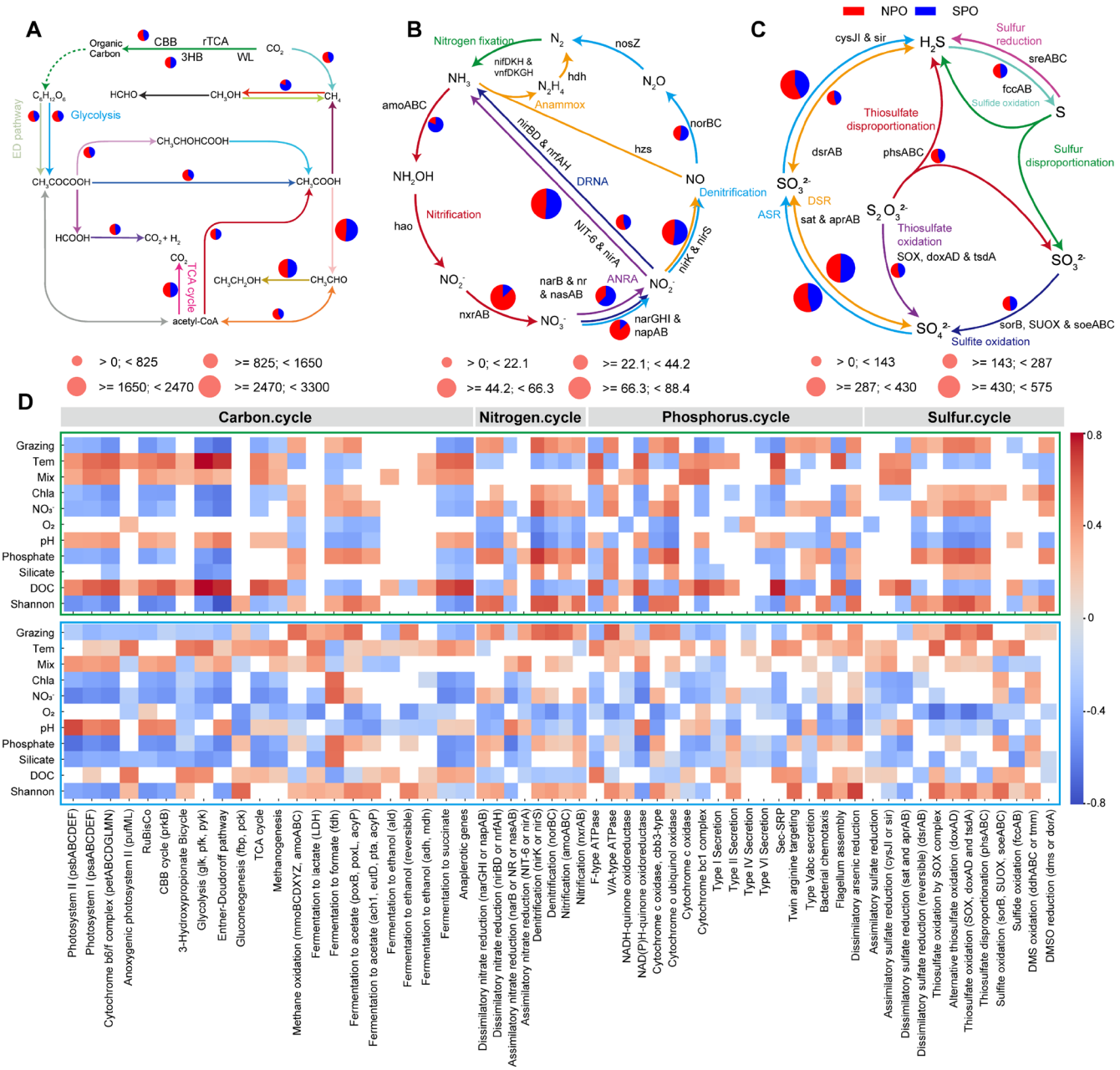

FIGURE 3 | Elemental cycling and metabolic pathways associated with functional gene diversity in the North Pacific Ocean and South Pacific Ocean. (A–C) Biogeochemical pathways represented by the carbon, nitrogen, and sulfur cycles, with pie charts representing relative gene abundances in the North Pacific Ocean and the South Pacific Ocean. CBB, Calvin–Benson–Bassham cycle; rTCA, reductive citric acid cycle; WL, Wood–Ljung–Dahl pathway; 3HB, 3-hydroxypropionate bicycle; DHC, dicarboxylate–hydroxybutyrate cycle. ANRA, assimilatory nitrate reduction to ammonium; DNRA, dissimilatory nitrate reduction to ammonium; Anammox, anaerobic ammonium oxidation. ASR, assimilatory sulfate reduction; DSR, dissimilatory sulfate reduction. (D) Correlation heatmaps (Pearson correlation) of gene pathways with environmental variables and species diversity in the North Pacific Ocean (upper) and South Pacific Ocean (lower).

(pheopigments), and heterotrophic bacteria. These results support the hypothesized trophic microbial linkage (Figure 2G and Table S5). In the surface ocean, when grazing significantly increased in 2011, both microbial species diversity and functional gene diversity increased (Figure S22A,B). Among all the oceans, the Pacific region experienced the most pronounced changes in grazing and nitrate (Figure S23). The nutrient transfer hypothesis offers an integrative, testable framework for interpreting region-dependent declines in microbial diversity through linked nutrient availability, phytoplankton, and grazing associations.

3.3 | Microbial Functional Pathways Mediated by Nutrient Transfer Hypothesis-Related Environmental Factors

The differences in functional pathways and genes involved in the carbon, nitrogen, and sulfur cycles between the North and South Pacific Ocean are provided in Figure 3A–C. With respect to the carbon cycle, most of the functional pathway abundances in the NPO were greater than those in the SPO; for example, the fermentation of organic matter to formate

and ethanol increased by 13.07% and 4.71%, respectively (Figure 3A). However, functional genes involved in methane oxidation were significantly enriched in the SPO, with a more than 18-fold increase in relative abundance compared with that in the NPO. The North and South Pacific regions showed more pronounced differences in the nitrogen cycle. The expression levels of functional genes in the NPO involved in the transformation of NO_2^- to NO_3^- during nitrification increased by more than 5-fold, but the transformation of NH_3 to NH_2OH and the oxidation of NO_2^- to NO_3^- in SPO increased by 333.22% and 65.41%, respectively (Figure 3B). These differences may be influenced by specific environmental background conditions, such as eutrophication and high-intensity grazing (Figure S24). The functional gene abundances in the transformation pathways of $S_2O_3^{2-}$(thiosulfate) to H_2S and SO_3^{2-}(sulfide and sulfite) and SO_4^{2-} (sulfate) in the NPO increased by 27.41% and 18.35%, respectively (Figure 3C). Similar correlations were observed between environmental variables (grazing and NO_3^-) and most of the gene pathways (e.g., photosystem I and denitrification) in the North and South Pacific Ocean. However, temperature and DOC showed significantly stronger relationships with gene pathways in the North Pacific than in the South Pacific, especially for the glycolysis and Entner–Doudoroff pathways (Figure 3D). Inconsistent changes in environmental variables, such as grazing and NO_3^-, shaped microbial functions (Figures S22–S24) and led to the risk of diversity loss.

3.4 | Decline in Microbial Diversity in the Pacific Ocean Associated With Grazing and Nitrate

Using generalized linear mixed models (GLMMs), we accounted for potential biases associated with project type. Grazing, NO_3^-, and chlorophyll a had the strongest effects on species diversity and functional gene diversity. The estimated effects of these variables were consistent between models that included season as a random effect or fixed effect. This consistency indicated that seasonal heterogeneity did not confound our inference on the major variables (Figure 4A). Furthermore, on the basis of the above nutrient transfer hypothesis, thresholds were identified as the values of the dominant predictors at which microbial diversity showed rapid and continuous changes (Figure 4B–E). The negative effects of grazing on species diversity were <0.0075 and $0.01\,mol{\cdot}m^{-3}{\cdot}s^{-1}$ in the North and South Pacific, respectively (Figure 4B). In contrast, NO_3^- concentrations above $0.6\,mmol{\cdot}m^{-3}$ had positive effects on species diversity, but a NO_3^- concentration above $18.07\,mmol{\cdot}m^{-3}$ also led to a negative effect in the South Pacific Ocean (Figure 4C). The lower thresholds and greater variability for the NPO suggest that the microbial diversity in this region was more sensitive to environmental change than that in the SPO, which may be associated with the contrasting risks of diversity loss between the two regions. Approximately 56.13% of the Pacific Ocean faced negative effects of grazing (51.42% for NO_3^-), which resulted in an ~11.77% decrease in species diversity (16.15% for NO_3^-), in contrast to positive effects (Figure S25A–D). Species diversity led to rapid changes at a grazing intensity of $\sim 0.0075\,mol{\cdot}m^{-3}{\cdot}s^{-1}$ and an NO_3^- concentration of $0.6\,mmol{\cdot}m^{-3}$ in the NPO (Figure 4D). Changes in SPO were more continuous (Figure 4E), and similar trends were observed for changes in functional diversity (Figure S26A–D).

Differences in the estimated pathways between the North and South Pacific were further discussed through the nutrient transfer hypothesis (Figure 4F,G). The path coefficient between species diversity and functional gene diversity was greater in the North Pacific Ocean (0.409, $p<0.05$) than in the SPO (0.339, $p<0.05$), suggesting a greater risk of synergistic declines in both types of diversity. In addition, the positive effects of grazing on species diversity and functional gene diversity were significantly greater in the NPO region (0.761 and 0.331 in NPO, 0.241 and 0.242 in SPO, $p<0.05$), whereas the effects of NO_3^- on microbial diversity in the NPO region were mediated by grazing (0.785, $p<0.05$). In the SPO, the indirect effects on species diversity exhibited by both grazing and NO_3^- through Chla were significantly positive. The NO_3^- pathway for gene diversity in the SPO had a large effect value (0.718, $p<0.05$).

3.5 | Minor Environmental Fluctuations Lead to Large Changes in Microbial Diversity

Through adaptive adjustment of key predictors within the ±10% constraint, the downward exploratory scenario indicated potential decreases of 7.3% and 0.28% for species diversity and functional gene diversity in the whole tested region, respectively (Figure 5A,B; Figure S21). The greatest concurrent decreases were identified in the North Pacific Ocean (Figure 5C,D), with mean reductions of 9.62% and 0.42% and maximum reductions reaching 35.18% and 1.61%, respectively. In the South Pacific Ocean, species diversity and functional gene diversity declined by −8.38% and −0.25%, respectively. Although the decreases in diversity may have been underestimated, marine microbial diversity has significant improvement potential. Species diversity and functional gene diversity have the potential to increase in the surface ocean by 14.63% and 0.45%, respectively. Under the upward exploratory scenario, the greatest potential increases were identified in the SAO and SPO (16.94% and 0.57% for the SAO and 15.03% and 0.51% for the SPO, respectively) for both types of diversity. The increases were lower in the NPO near the equator and in the Baltic Sea, with the potential for improvement being less than 6% and 0.3% for species diversity and functional gene diversity, respectively (Figure 5B; Figure S27A–D). The range of variation in diversity can also indicate sensitivity or vulnerability (Figure 5E). The NPO and SPO were more sensitive to environmental fluctuations, with the NPO having the highest functional diversity sensitivity among all the oceans (0.84%) and the NPO and SPO having higher species diversity sensitivities than most of the oceans (22.63% and 23.41%, respectively).

Across the optimization bounds of ±5% to ±30%, both diversity measures changed progressively as the admissible range widened, while the ±10% scenario showed similar spatial patterns to adjacent bounds and consistent directions of change under grid-specific 2009–2020 historical constraints. Spatial rankings were also generally stable, with median rank SDs below the 12th percentile across all four scenarios and greater stability for

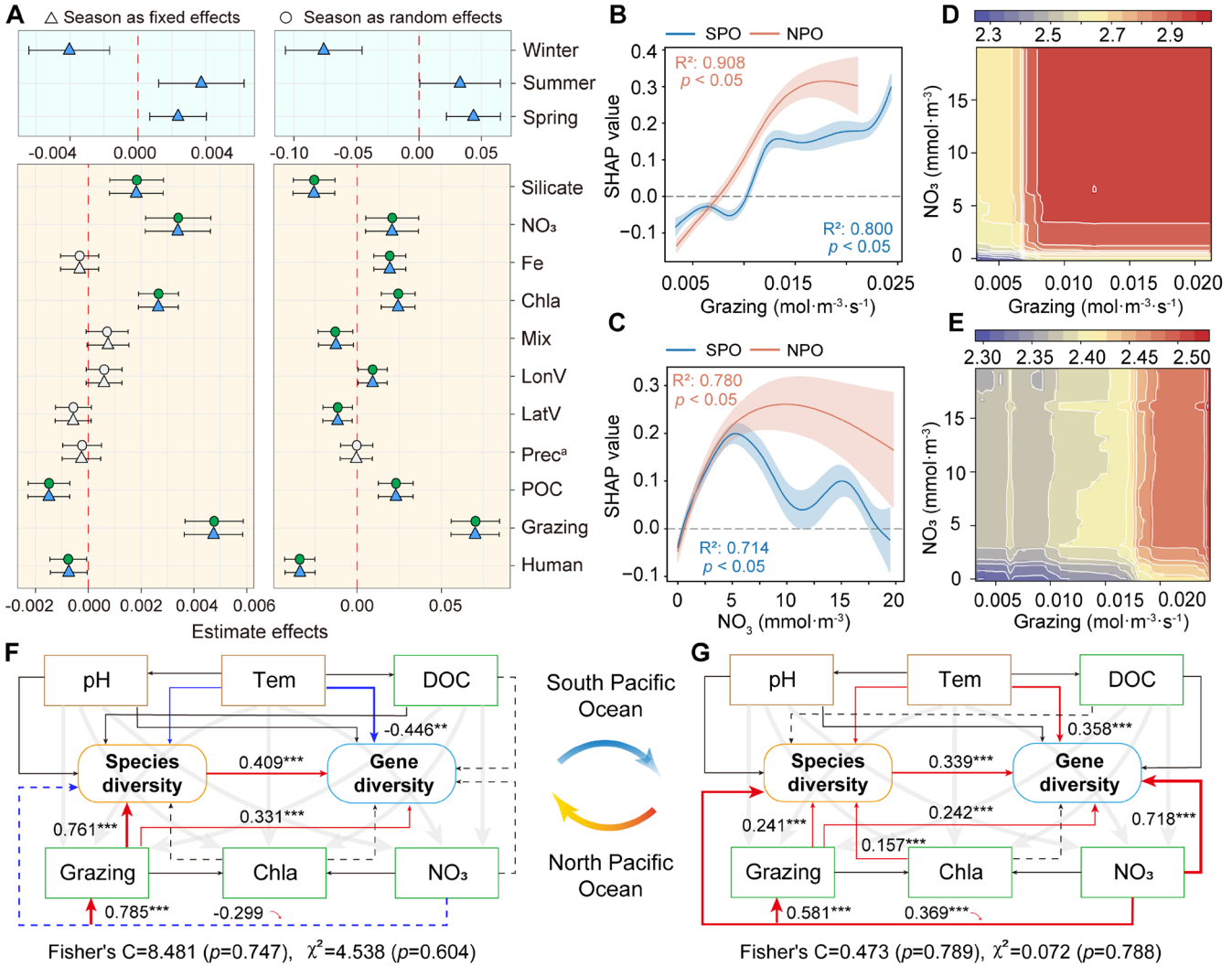

FIGURE 4 | Decreases in microbial diversity associated with grazing and NO_3^-. (A) Generalized linear mixed model (GLMM) for microbial species diversity (right) and functional gene diversity (left). Seasons were considered as random effects and fixed effects for separate modelling. The solid shapes represent significant effects ($p < 0.05$). Prec[a], precipitation. (B, C) Single-feature continuous responses of microbial species diversity. The grey shading represents the 95% confidence interval, with a *y*-axis value of 0 as the baseline for distinguishing positive and negative effects. (D, E) Two-feature continuous responses of microbial species diversity to grazing and NO_3^- in the North Pacific Ocean (D) and South Pacific Ocean (E). (F, G) Structural equation modelling of the hypothesized pathways linking key predictors to microbial diversity in the North ($N = 63$) and South Pacific Oceans ($N = 288$). Asterisks indicate significance: *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$.

species diversity than for functional gene diversity. These results indicate consistent qualitative patterns but range-dependent response magnitudes, supporting ±10% as a conservative reference (Figures S28A–F and S29A–D). The change directions for the key predictors (e.g., grazing and NO_3^-) corresponding to the highest (diversity protection) and lowest (diversity loss) scenarios were determined (Figure 5F,G; Figure S30A–F). Overall, decreases of 4.02%, 1.59% and 2.25% in grazing, NO_3^- and Chla levels, respectively, could result in significant losses in diversity (−7.3% and −0.28% for species diversity and functional gene diversity, respectively). In contrast, increases of 2.7%, 0.68% and 1.46% in the grazing, NO_3^- and Chla levels, respectively, could provide a significant increase of diversity (14.63% and 0.45% for species diversity and functional gene diversity, respectively). We predicted microbial diversity in the Pacific Ocean with and without consideration of nutrient transfer mechanisms, and the results revealed that species diversity and functional gene diversity decreased by 12.60% and 1.12%, respectively, after the relevant variables were removed (Figure S31).

4 | Discussion

Recent studies have increasingly resolved the biogeographic and regional macroecological patterns of ocean microbial diversity (Yáñez et al. 2025). However, globally coordinated observations that jointly track taxonomic and functional diversity over time remain limited. Here, we quantified diversity trends in the global surface ocean over the past 12 years and identified hotspots of diversity loss in the NPO and SPO. Changes in functional diversity are slighter than those in species diversity, which may be related to the adaptability of microbial functional expression to changes in external factors and functional redundancy. Nevertheless, KEGG-annotated

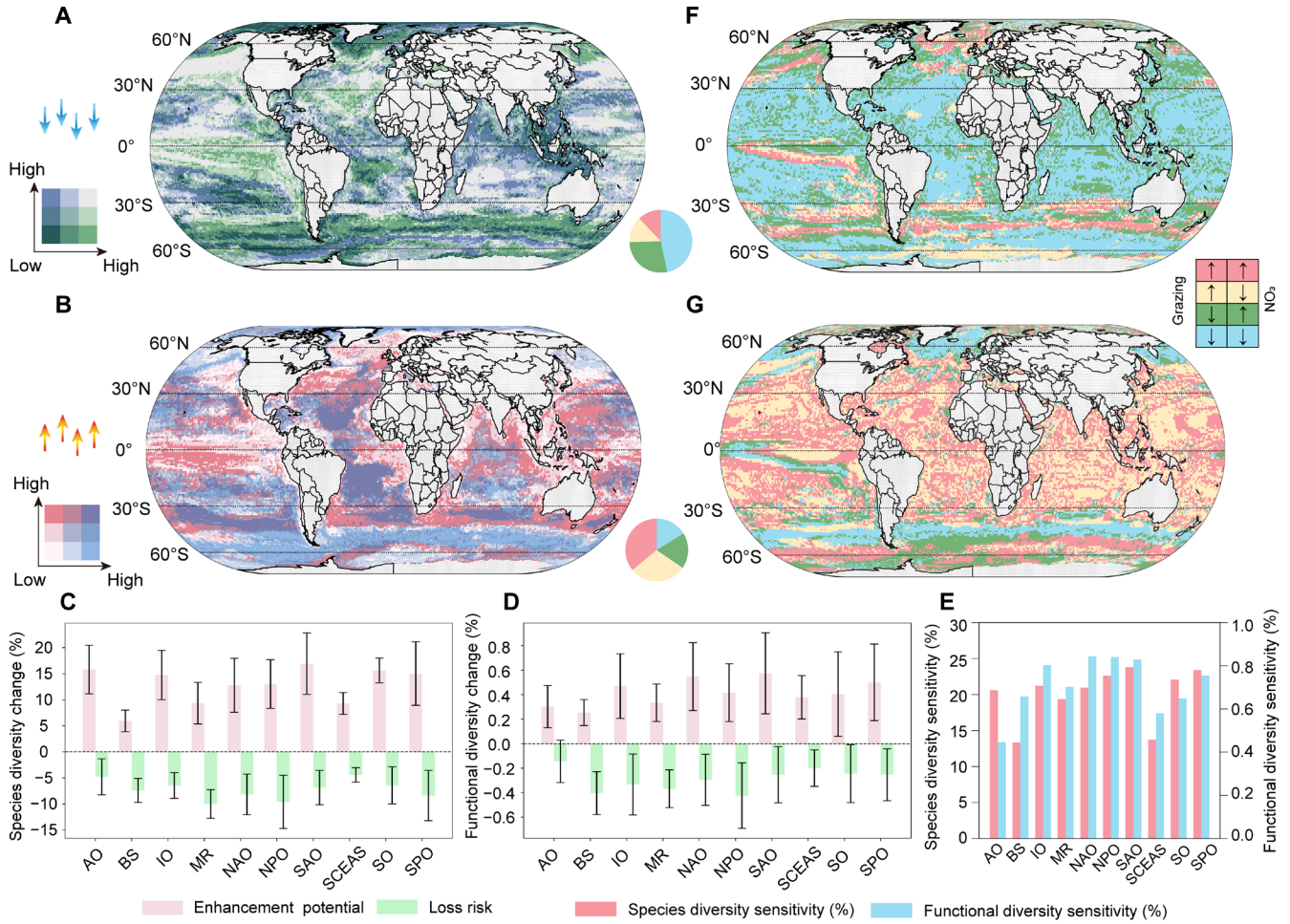

FIGURE 5 | Large changes in diversity and geographical heterogeneity with minor environmental fluctuations in the Pacific Ocean. (A, B) Maximizing diversity loss (A) and diversity protection (B). (C, D) Percentages of change in species diversity (C) and functional gene diversity (D) under different scenarios. (E) Sensitivity of microbial species diversity and functional diversity. (F, G) Change directions for grazing and NO_3^- levels to maximize diversity loss risks (F) and diversity protection (G).

functional gene diversity reflects encoded functional profiles rather than realized activity and remains subject to annotation bias and unmeasured functional redundancy. We interpret these predictions as region-specific surface ocean patterns rather than evidence of diversity trends and patterns. Uneven high-latitude sampling and the lack of observations below 200 m limit broader generalization and highlight the need for expanded polar and deep ocean sampling and sustained time series. Furthermore, spatial differences in the factors associated with marine microbial species diversity and functional gene diversity remain incompletely understood (Chen et al. 2024; Zakem et al. 2025). Grazing, nitrate (NO_3^-), and chlorophyll (Chla) levels have been identified as important predictors of microbial diversity in the Pacific region, whereas atmospheric deposition and subtropical gyre dynamics may also characterize regional trophic conditions (Dai et al. 2023). Interestingly, microplastics (MPs) also have been shown to have driving effects in parts of the Mediterranean region (MR) and Indian Ocean (IO), which may be related to the combined effects of MPs on microbial habitats and regional characteristics (Li et al. 2020; Wang, Liu, et al. 2024). Functional gene diversity responses depend on environmental changes related to gene expression, amplification, or transfer (Hoffmann and Willi 2008). However, in addition to environmental changes, species diversity involves ecological feedback between communities (e.g., grazing). Among the associated variables, Chla represents phytoplankton communities, provides major nutrients and changes heterotrophic microbial species diversity and the selective expression of genes on the basis of responses to environmental characteristics (resource stress or interspecific competition) (Li et al. 2024). The role of grazing in microbial diversity has been overlooked in previous studies (Ratnarajah et al. 2023). The size of the upper predator population may indirectly influence the composition of the microbial community (Ferguson et al. 2021; Zhu et al. 2021).

Although the bootstrap and application domain analyses support the stability and spatial applicability of the models, the identified predictors should not be interpreted causally. Their apparent importance may partly reflect collinearity (Figure S32), spatial structure, or indirect ecological pathways. Our results concern the association of nitrate concentration with microbial diversity within a trophic context rather than a general effect of nutrient type. We therefore interpret grazing, NO_3^- and Chla as ecological predictors consistent with the nutrient transfer hypothesis and regard predictions outside the application domain as exploratory.

The models performed well without evidence of overfitting (Figures S33A,B and S34). In addition, we screened the key factors associated with microbial diversity in the Pacific by proposing a nutrient transfer hypothesis that integrates the potential dynamic relationships. Each pathway was linked to a predefined testable prediction and assessed using complementary analyses (e.g., machine learning response patterns, threshold identification, structural equation modelling and convergent cross mapping). The nutrient transfer hypothesis is an integrative, region-specific framework that builds on established ecological theory. It extends the microbial loop perspective (Azam et al. 1983) by treating microbial species diversity and functional gene diversity as explicit responses to nutrient availability, phytoplankton biomass and grazing intensity and integrates bottom-up nutrient supply with the grazer-associated pathways emphasized by trophic cascade theory (Glibert and Mitra 2022). Rather than modelling resource-ratio theory (Held and Manhart 2024), it links these established processes through testable pathways to microbial taxonomic and functional diversity. The nutrient transfer hypothesis offers an integrative, testable framework for interpreting region-dependent declines in microbial diversity through linked nutrient availability, phytoplankton, and grazing associations.

There are also differences in functional gene pathway expression in the NPO and SPO; for example, anaerobic methane oxidation can be enhanced by alternative electron acceptors (e.g., nitrite, nitrate, and trivalent iron) (Shen et al. 2019). Microorganisms may also utilize organic matter and its metabolites (e.g., methane) produced through grazing for growth and reproduction, which in turn may increase the abundance of genes in related pathways (Middelburg 2019). NO_3^- and SO_4^{2-} may compete as electron acceptors, which in turn affects sulfur morphology and cycling, while grazing intensity affects the sulfur cycle by altering microbial communities and organic matter decomposition rates. These regional differences may be associated with variations in microbially mediated biogeochemical cycling. Although the core functions of marine microorganisms (e.g., photosynthesis and denitrification) are similar, the regulation of energy metabolic pathways (e.g., glycolysis) may be determined by the regional environment, resulting in different microbial metabolic phenotypes between the NPO (which is sensitive to environmental changes and interspecific relationships) and the SPO. In addition, the different correlations between metabolic pathways and species diversity may be associated with the limitations on species survival linked to metabolic specialization. The function of a community is influenced by the trade-offs between metabolic specialization and the versatility of individuals (Wang, Chen, et al. 2024). It may also be related to resource use efficiency, ecological niche competition, and environmental adaptation (Stocker 2012; Ghaderiardakani et al. 2020; Lambert et al. 2021). Even in the vulnerable Pacific, internal variations still occur.

In response to environmental variables, changes in the SPO were more continuous, reflecting the ability of the microbial community in the SPO to adapt to the environment (Cao et al. 2020; Wani et al. 2022). The microbial communities in the NPO were more strongly influenced by interspecific competition, and the external environment indirectly affected diversity by shaping community dynamics. The different response relationships were influenced by cascade effects among marine ecosystems and taxa. Under eutrophic conditions, warming could reduce phytoplankton biomass through increased zooplankton grazing (Lewandowska et al. 2014; Karakuş et al. 2022). Under oligotrophic conditions, the grazing intensity of zooplankton is influenced by the growth of phytoplankton; in other words, to some extent, the grazing intensity also reflects the nutrient conditions in the region, which has a more direct effect on microbial diversity at the trophic level. The above results highlight the dynamic relationships and feedback mechanisms of nutrient transfer, which could be an early warning index of diversity loss. In the SPO, the indirect effects on species diversity exhibited by both grazing and NO_3^- through Chla were significantly positive; these factors support primary productivity by phytoplankton and consequently drive the increase in heterotrophic microbial diversity (Mühlenbruch et al. 2018). The NO_3^- pathway for gene diversity in the SPO had a large effect value (0.718, $p < 0.05$), which may be related to the expression of microbial differential nitrogen metabolic pathways (e.g., assimilatory reduction or denitrification) in oligotrophic environments (Wang et al. 2022; Deng et al. 2024). Differential impact pathways involving similar predictors may be associated with the variations in diversity determined by regional environmental and microbial community characteristics. The relative monotony of impact pathways in the North Pacific may lead to vulnerability and deserves greater attention under increased environmental change.

Before practical approaches for conserving diversity are established, understanding the best and worst-case scenarios and the sensitivity of diversity to environmental fluctuations is critical (Ramírez et al. 2017). Owing to the environmental characteristics of the oligotrophic North Pacific Subtropical Gyre, the oxygen-depleted eastern tropical North Pacific, and nutrient-rich equatorial upwelling (Saunders et al. 2022), microorganisms in the Pacific Ocean are sensitive to nutrient transport. The high sensitivity of the NPO and SPO regions often involves the global transport of nutrients accompanying circulation (Saunders et al. 2022). With the westerly circulation as well as the monsoon, rapid nutrient changes due to sand and dust significantly impact marine ecology (represented by grazing, NO_3^- and Chla) (Rae et al. 2020). Indicators involved in nutrient transfer mechanisms are critical to the global estimation of microbial diversity. The proposed nutrient transfer concept explains the obvious risk of diversity loss in the NPO region after minor environmental fluctuations. With the significant decreasing trends and high sensitivity of microbial diversity, both risk assessment and appropriate conservation practices are urgently needed for the Pacific Ocean, further considering potential differences between the NPO and SPO. The constrained exploratory scenarios represent conditional model sensitivities within specific bounds rather than predictive outcomes or attainable ecological optima. Although ±10% provided a conservative reference, the estimated response magnitudes increased as the optimization bounds widened. Previous models that do not consider the nutrient transfer concept may incorrectly underestimate microbial diversity (Norberg et al. 2019). Large changes in diversity loss and protection potential may challenge the stabilization of biogeochemical cycles, and sufficient interventions are urgently needed to mitigate dramatic changes in the future. Pacific diversity hotspots should be prioritized for monitoring microbial diversity, grazing, nutrient levels and phytoplankton biomass. These observations can further guide surveillance and field validation.

5 | Conclusions

Microbial communities participate in marine biogeochemical cycles and ecological processes, but most related studies are limited by geographic sparsity and lack an understanding of microbial community dynamics. Here, we use a wide range of metagenome samples and marine databases to analyse the spatial and temporal patterns of microbial species diversity and functional gene diversity (including the migration of specific species and changes in gene expression) at a global scale over 12 years and highlight the risk of diversity loss in the Pacific Ocean. Dynamic nutrient transfer as a hypothesized conceptual feedback mechanism is proposed to explain the geographic differences between the NPO and SPO, where the NPO presents lower thresholds for grazing intensity and the NO_3^- level, along with fewer impact pathways of microbial diversity. Causal links in the hypotheses are also verified by time series data. The risk and uncertainty for the loss of microbial diversity in the Pacific Ocean are considerable. Even with minor environmental fluctuations (< 5%), there is an underestimated risk of decline of species diversity of −7.3% in the global surface ocean, especially for the Pacific Ocean (−9.62% and −8.38%). In contrast, in terms of protection, the enhancement potential for species diversity is 14.63% at most. Given the remarkable loss risks and sensitivities concerning microbial diversity in the Pacific Ocean, a complex cascade effect on marine biogeochemical cycles and the diversity of large organisms could result. It is urgent to incorporate this dynamic nutrient transfer concept to support the recognition of risk areas and reduce the uncertainty and variation of microbial biodiversity.

Author Contributions

Xu Dong: methodology, software, data curation, validation, visualization. **Xiangang Hu:** conceptualization, funding acquisition. **Kai Hu:** validation. **Zehou Li:** validation. **Weilei Wang:** validation.

Acknowledgements

The authors have nothing to report.

Funding

This work was supported by Fujian Ocean Innovation Center, 25FV0CWZ02. National Natural Science Foundation of China, 42577318. Fundamental Research Funds for the Central University and 111 Program, T2017002.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Github at https://github.com/NKUHuLab/Ocean-microorganism.

References

Arrigo, K. R. 2005. "Marine Microorganisms and Global Nutrient Cycles." *Nature* 437: 349–355. https://doi.org/10.1038/nature04159.

Azam, F., T. Fenchel, J. G. Field, J. S. Gray, L.-A. Meyer-Reil, and F. Thingstad. 1983. "The Ecological Role of Water-Column Microbes in the Sea." *Marine Ecology Progress Series* 10: 257–263. https://doi.org/10.7208/chicago/9780226125534-024.

Bauman, D., T. Drouet, M.-J. Fortin, and S. Dray. 2018. "Optimizing the Choice of a Spatial Weighting Matrix in Eigenvector-Based Methods." *Ecology* 99: 2159–2166. https://doi.org/10.1002/ecy.2469.

Bissett, A., M. V. Brown, S. D. Siciliano, and P. H. Thrall. 2013. "Microbial Community Responses to Anthropogenically Induced Environmental Change: Towards a Systems Approach." *Ecology Letters* 16: 128–139. https://doi.org/10.1111/ele.12109.

Cao, S., W. Zhang, W. Ding, et al. 2020. "Structure and Function of the Arctic and Antarctic Marine Microbiota as Revealed by Metagenomics." *Microbiome* 8: 47. https://doi.org/10.1186/s40168-020-00826-9.

Carle, H., D. Bauman, M. N. Evans, et al. 2025. "Aboveground Biomass in Australian Tropical Forests Now a Net Carbon Source." *Nature* 646: 611–618. https://doi.org/10.1038/s41586-025-09497-8.

Cavicchioli, R., W. J. Ripple, K. N. Timmis, et al. 2019. "Scientists' Warning to Humanity: Microorganisms and Climate Change." *Nature Reviews Microbiology* 17: 569–586. https://doi.org/10.1038/s41579-019-0222-5.

Chen, J., Y. Jia, Y. Sun, et al. 2024. "Global Marine Microbial Diversity and Its Potential in Bioprospecting." *Nature* 633: 371–379. https://doi.org/10.1038/s41586-024-07891-2.

Collos, Y., J. Husseini-Ratrema, B. Bec, et al. 2005. "Pheopigment Dynamics, Zooplankton Grazing Rates and the Autumnal Ammonium Peak in a Mediterranean Lagoon." *Hydrobiologia* 550: 83–93. https://doi.org/10.1007/s10750-005-4365-1.

Dai, M., Y.-W. Luo, E. P. Achterberg, et al. 2023. "Upper Ocean Biogeochemistry of the Oligotrophic North Pacific Subtropical Gyre: From Nutrient Sources to Carbon Export." *Reviews of Geophysics* 61: e2022RG000800. https://doi.org/10.1029/2022RG000800.

Deb, K., A. Pratap, S. Agarwal, and T. J. I. T. O. E. C. Meyarivan. 2002. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II." *IEEE Transactions on Evolutionary Computation* 6: 182–197. https://doi.org/10.1109/4235.996017.

Deng, D., Z. Yang, Y. Yang, W. Wan, W. Liu, and X. Xiong. 2024. "Metagenomic Insights Into Nitrogen-Cycling Microbial Communities and Their Relationships With Nitrogen Removal Potential in the Yangtze River." *Water Research* 265: 122229. https://doi.org/10.1016/j.watres.2024.122229.

Ferguson, R. M. W., E. J. O'Gorman, D. J. McElroy, et al. 2021. "The Ecological Impacts of Multiple Environmental Stressors on Coastal Biofilm Bacteria." *Global Change Biology* 27: 3166–3178. https://doi.org/10.1111/gcb.15626.

Ghaderiardakani, F., M. L. Quartino, and T. J. F. I. M. S. Wichard. 2020. "Microbiome-Dependent Adaptation of Seaweeds Under Environmental Stresses: A Perspective." *Frontiers in Marine Science* 7: 575228. https://doi.org/10.3389/fmars.2020.575228.

Giovannoni, S. J., and K. L. Vergin. 2012. "Seasonality in Ocean Microbial Communities." *Science* 335: 671–676. https://doi.org/10.1126/science.1198078.

Glibert, P. M., and A. Mitra. 2022. "From Webs, Loops, Shunts, and Pumps to Microbial Multitasking: Evolving Concepts of Marine Microbial Ecology, the Mixoplankton Paradigm, and Implications for a Future Ocean." *Limnology and Oceanography* 67: 585–597. https://doi.org/10.1002/lno.12018.

Halpern, B. S., M. Frazier, J. Afflerbach, et al. 2019. "Recent Pace of Change in Human Impact on the World's Ocean." *Scientific Reports* 9: 11609. https://doi.org/10.1038/s41598-019-47201-9.

Hawco, N. J., T. M. Conway, S. N. Coesel, et al. 2025. "Anthropogenic Iron Alters the Spring Phytoplankton Bloom in the North Pacific Transition Zone." *Proceedings of the National Academy of Sciences of the United States of America* 122: e2418201122. https://doi.org/10.1073/pnas.2418201122.

Held, N. A., and M. Manhart. 2024. "Are Microbes Colimited by Multiple Resources?" *Current Opinion in Microbiology* 80: 102509. https://doi.org/10.1016/j.mib.2024.102509.

Hoffmann, A. A., and Y. Willi. 2008. "Detecting Genetic Responses to Environmental Change." *Nature Reviews Genetics* 9: 421–432. https://doi.org/10.1038/nrg2339.

Hynes, A. M., J. Winter, C. T. Berthiaume, et al. 2024. "High-Frequency Sampling Captures Variability in Phytoplankton Population-Specific Periodicity, Growth, and Productivity." *Limnology and Oceanography* 69: 2516–2531. https://doi.org/10.1002/lno.12683.

Ishwaran, H., U. B. Kogalur, E. Z. Gorodeski, A. J. Minn, and M. S. Lauer. 2010. "High-Dimensional Variable Selection for Survival Data." *Journal of the American Statistical Association* 105: 205–217. https://doi.org/10.1198/jasa.2009.tm08622.

Jia, Y. Y., X. G. Hu, W. L. Kang, and X. Dong. 2024. "Unveiling Microbial Nitrogen Metabolism in Rivers Using a Machine Learning Approach." *Environmental Science & Technology* 58: 6605–6615. https://doi.org/10.1021/acs.est.3c09653.

Karakuş, O., C. Völker, M. Iversen, W. Hagen, and J. Hauck. 2022. "The Role of Zooplankton Grazing and Nutrient Recycling for Global Ocean Biogeochemistry and Phytoplankton Phenology." *Journal of Geophysical Research: Biogeosciences* 127: e2022JG006798. https://doi.org/10.1029/2022JG006798.

Ladau, J., and E. A. Eloe-Fadrosh. 2019. "Spatial, Temporal, and Phylogenetic Scales of Microbial Ecology." *Trends in Microbiology* 27: 662–669. https://doi.org/10.1016/j.tim.2019.03.003.

Ladau, J., T. J. Sharpton, M. M. Finucane, et al. 2013. "Global Marine Bacterial Diversity Peaks at High Latitudes in Winter." *ISME Journal* 7: 1669–1677. https://doi.org/10.1038/ismej.2013.37.

Lambert, S., J.-C. Lozano, F.-Y. Bouget, and P. E. Galand. 2021. "Seasonal Marine Microorganisms Change Neighbours Under Contrasting Environmental Conditions." *Environmental Microbiology* 23: 2592–2604. https://doi.org/10.1111/1462-2920.15482.

Larkin, A. A., M. L. Brock, A. J. Fagan, et al. 2025. "Climate-Driven Succession in Marine Microbiome Biodiversity and Biogeochemical Function." *Nature Communications* 16: 3926. https://doi.org/10.1038/s41467-025-59382-1.

Larkin, A. A., C. A. Garcia, N. Garcia, et al. 2021. "High Spatial Resolution Global Ocean Metagenomes From Bio-GO-SHIP Repeat Hydrography Transects." *Scientific Data* 8: 107. https://doi.org/10.1038/s41597-021-00889-9.

Lefcheck, J. S. 2016. "piecewiseSEM: Piecewise Structural Equation Modelling in R for Ecology, Evolution, and Systematics." *Methods in Ecology and Evolution* 7: 573–579. https://doi.org/10.1111/2041-210X.12512.

Lewandowska, A. M., D. G. Boyce, M. Hofmann, B. Matthiessen, U. Sommer, and B. Worm. 2014. "Effects of Sea Surface Warming on Marine Plankton." *Ecology Letters* 17: 614–623. https://doi.org/10.1111/ele.12265.

Li, J., W. Huang, R. Jiang, X. Han, D. Zhang, and C. Zhang. 2020. "Are Bacterial Communities Associated With Microplastics Influenced by Marine Habitats?" *Science of the Total Environment* 733: 139400. https://doi.org/10.1016/j.scitotenv.2020.139400.

Li, Z., D. Sun, S. Wang, et al. 2024. "Ocean-Scale Patterns of Environment and Climate Changes Driving Global Marine Phytoplankton Biomass Dynamics." *Science Advances* 10: eadm7556. https://doi.org/10.1126/sciadv.adm7556.

Liang, W., J. Li, X. Wang, J. P. Giesy, and X. Zhao. 2025. "Machine Learning-Driven Cross-Species Toxicity Prediction for Advancing Ecologically Relevant PFAS Water Quality Criteria." *Environmental Science & Technology* 59: 25688–25702. https://doi.org/10.1021/acs.est.5c12013.

Louca, S., L. W. Parfrey, and M. Doebeli. 2016. "Decoupling Function and Taxonomy in the Global Ocean Microbiome." *Science* 353: 1272–1277. https://doi.org/10.1126/science.aaf4507.

Middelburg, J. J. 2019. "Carbon Processing at the Seafloor." In *Marine Carbon Biogeochemistry: A Primer for Earth System Scientists*, edited by J. J. Middelburg, 57–75. Springer International Publishing.

Mühlenbruch, M., H.-P. Grossart, F. Eigemann, and M. Voss. 2018. "Mini-Review: Phytoplankton-Derived Polysaccharides in the Marine Environment and Their Interactions With Heterotrophic Bacteria." *Environmental Microbiology* 20: 2671–2685. https://doi.org/10.1111/1462-2920.14302.

Norberg, A., N. Abrego, F. G. Blanchet, et al. 2019. "A Comprehensive Evaluation of Predictive Performance of 33 Species Distribution Models at Species and Community Levels." *Ecological Monographs* 89: e01370. https://doi.org/10.1002/ecm.1370.

Pommier, T., B. CanbÄCk, L. Riemann, et al. 2007. "Global Patterns of Diversity and Community Structure in Marine Bacterioplankton." *Molecular Ecology* 16: 867–880. https://doi.org/10.1111/j.1365-294X.2006.03189.x.

Rae, J. W. B., W. R. Gray, R. C. J. Wills, et al. 2020. "Overturning Circulation, Nutrient Limitation, and Warming in the Glacial North Pacific." *Science Advances* 6: eabd1654. https://doi.org/10.1126/sciadv.abd1654.

Ramírez, F., I. Afán, L. S. Davis, and A. Chiaradia. 2017. "Climate Impacts on Global Hot Spots of Marine Biodiversity." *Science Advances* 3: e1601198. https://doi.org/10.1126/sciadv.1601198.

Ratnarajah, L., R. Abu-Alhaija, A. Atkinson, et al. 2023. "Monitoring and Modelling Marine Zooplankton in a Changing Climate." *Nature Communications* 14: 564. https://doi.org/10.1038/s41467-023-36241-5.

Salazar, G., L. Paoli, A. Alberti, et al. 2019. "Gene Expression Changes and Community Turnover Differentially Shape the Global Ocean Metatranscriptome." *Cell* 179: 1068–1083.e1021. https://doi.org/10.1016/j.cell.2019.10.014.

Saunders, J. K., M. R. McIlvin, C. L. Dupont, et al. 2022. "Microbial Functional Diversity Across Biogeochemical Provinces in the Central Pacific Ocean." *Proceedings of the National Academy of Sciences of the United States of America* 119: e2200014119. https://doi.org/10.1073/pnas.2200014119.

Shen, L.-d., L. Ouyang, Y. Zhu, and M. Trimmer. 2019. "Active Pathways of Anaerobic Methane Oxidation Across Contrasting Riverbeds." *ISME Journal* 13: 752–766. https://doi.org/10.1038/s41396-018-0302-y.

Stocker, R. 2012. "Marine Microbes See a Sea of Gradients." *Science* 338: 628–633. https://doi.org/10.1126/science.1208929.

Sugihara, G., R. May, H. Ye, et al. 2012. "Detecting Causality in Complex Ecosystems." *Science* 338: 496–500. https://doi.org/10.1126/science.1227079.

Sunagawa, S., S. G. Acinas, P. Bork, et al. 2020. "Tara Oceans: Towards Global Ocean Ecosystems Biology." *Nature Reviews Microbiology* 18: 428–445. https://doi.org/10.1038/s41579-020-0364-5.

van Sebille, E., C. Wilcox, L. Lebreton, et al. 2015. "A Global Inventory of Small Floating Plastic Debris." *Environmental Research Letters* 10: 124006. https://doi.org/10.1088/1748-9326/10/12/124006.

Van't Veen, H., L. Chalmandrier, N. Sandau, et al. 2020. "A Landscape-Scale Assessment of the Relationship Between Grassland Functioning, Community Diversity, and Functional Traits." *Ecology and Evolution* 10: 9906–9919. https://doi.org/10.1002/ece3.6650.

Wang, C., Q.-Y. Yu, N.-N. Ji, et al. 2023. "Bacterial Genome Size and Gene Functional Diversity Negatively Correlate With Taxonomic Diversity Along a pH Gradient." *Nature Communications* 14: 7437. https://doi.org/10.1038/s41467-023-43297-w.

Wang, M., X. Chen, Y. Fang, et al. 2024. "The Trade-Off Between Individual Metabolic Specialization and Versatility Determines the Metabolic Efficiency of Microbial Communities." *Cell Systems* 15: 63–74.e65. https://doi.org/10.1016/j.cels.2023.12.004.

Wang, Y., M. Guo, X. Li, et al. 2022. "Shifts in Microbial Communities in Shallow Lakes Depending on Trophic States: Feasibility as an Evaluation Index for Eutrophication." *Ecological Indicators* 136: 108691. https://doi.org/10.1016/j.ecolind.2022.108691.

Wang, Y.-F., Y.-J. Liu, Y.-M. Fu, et al. 2024. "Microplastic Diversity Increases the Abundance of Antibiotic Resistance Genes in Soil." *Nature Communications* 15: 9788. https://doi.org/10.1038/s41467-024-54237-7.

Wani, A. K., N. Akhtar, F. Sher, A. A. Navarrete, and J. H. P. Américo-Pinheiro. 2022. "Microbial Adaptation to Different Environmental Conditions: Molecular Perspective of Evolved Genetic and Cellular Systems." *Archives of Microbiology* 204: 144. https://doi.org/10.1007/s00203-022-02757-5.

Xue, C.-X., H. H. Lin, X. X.-Y. Zhu, et al. 2021. "DiTing: A Pipeline to Infer and Compare Biogeochemical Pathways From Metagenomic and Metatranscriptomic Data." *Frontiers in Microbiology* 12: 698286. https://doi.org/10.3389/fmicb.2021.698286.

Yáñez, S., L. Tavera, G. Guzmán-Rendón, et al. 2025. "Macroecological Patterns of Planktonic Unicellular Eukaryotes Richness in the Southeast Pacific Ocean." *Scientific Reports* 15: 18833. https://doi.org/10.1038/s41598-025-03220-3.

You, C., S. Chen, Z. Tu, et al. 2025. "Decadal Trends in Global Grassland Growth Peaks and Their Drivers Since the 1980s." *Nature Communications* 16: 9501. https://doi.org/10.1038/s41467-025-64565-x.

Yu, F. B., C. H. Wei, P. Deng, T. Peng, and X. G. Hu. 2021. "Deep Exploration of Random Forest Model Boosts the Interpretability of Machine Learning Studies of Complicated Immune Responses and Lung Burden of Nanoparticles." *Science Advances* 7: eabf4130. https://doi.org/10.1126/sciadv.abf4130.

Yu, H., M. Yang, L. Wang, and Y. Chen. 2024. "A Non-Parametric Method to Investigate Internal Trends in Time Sequence: A Case Study of Temperature and Precipitation." *Ecological Indicators* 158: 111373. https://doi.org/10.1016/j.ecolind.2023.111373.

Zakem, E. J., J. McNichol, J. L. Weissman, et al. 2025. "Functional Biogeography of Marine Microbial Heterotrophs." *Science* 388: eado5323. https://doi.org/10.1126/science.ado5323.

Zhivkoplias, E., J.-B. Jouffray, P. Dunshirn, A. Pranindita, and R. Blasiak. 2024. "Growing Prominence of Deep-Sea Life in Marine Bioprospecting." *Nature Sustainability* 7: 1027–1037. https://doi.org/10.1038/s41893-024-01392-w.

Zhu, D., M. Delgado-Baquerizo, J. Ding, M. R. Gillings, and Y.-G. Zhu. 2021. "Trophic Level Drives the Host Microbiome of Soil Invertebrates at a Continental Scale." *Microbiome* 9: 189. https://doi.org/10.1186/s40168-021-01144-4.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Metagenomic sample sampling locations (A) and information on the species and functional gene diversity of microbial communities (B, C). Assimilatory nitrate reduction (1), nitrate > nitrite; assimilatory nitrate reduction (2), nitrite > ammonia. **Figure S2:** Conceptual model and testable pathways of the nutrient transfer hypothesis. The solid arrows show testable pathways. The dashed arrows denote proposed transfer processes that were not directly measured. P1–P4 corresponds to Table S1. **Figure S3:** Sample-level sensitivity and redundancy analysis of particulate organic phosphorus (POP). (A, B) R^2 for random forest models with and without POP. Lines connect results from the same validation group, and horizontal bars show mean values. (C) Spearman correlation coefficients between POP and the 10 most strongly associated original predictors. The analysis included 768 of the 1582 metagenomic samples matched to GOPOPCORN V2 by distance (≤1 km), date (≤1 day), and depth (≤10 m). **Figure S4:** Validation of product-based environmental variables against in situ cruise measurements. Points represent matched observations with 1:1 dashed lines. **Figure S5:** Global marine microbial species diversity and functional gene diversity (from 2009 to 2020). **Figure S6:** Performance of machine learning models. RF, random forest; SVM, support vector machine; MLP, multilayer perceptron; XGBoost, extreme gradient boosting. **Figure S7:** Application domain of the diversity models. (A) Global distribution of multivariate environmental similarity surface (MESS) values. (B) Density distribution of MESS values, with the dashed line indicating the threshold of −20; lower values represent stronger environmental novelty relative to the training data. (C) Global distribution of Mahalanobis distances. (D) Density distribution of Mahalanobis distances, with the dashed line indicating the 99th-percentile threshold derived from the training data. Grid cells exceeding this threshold were considered outside the model application domain. **Figure S8:** Additional metagenomic samples to validate the fit of the random forest model for species diversity (267 samples from 2009 to 2022). Average values of the observed and predicted diversity and the corresponding error bars were calculated for each year. **Figure S9:** Global hotspot and coldspot patterns of marine microbial diversity. (A) Spatial distribution of species diversity hotspots and coldspots. (B) Spatial distribution of functional diversity hotspots and coldspots. Hotspots and coldspots were identified using the Getis–Ord Gi* statistic with queen contiguity spatial weights and classified at the 90%, 95%, and 99% confidence levels; nonsignificant grid cells are shown separately. **Figure S10:** Comparison of 12-year average values of microbial species diversity in hotspots and other regions. Hotspots were located in the North and South Pacific Ocean regions near South America (−15°–5° latitude in the Western Hemisphere). **Figure S11:** Distribution of turnover and nestedness components of beta diversity. Boxplots show total Sørensen dissimilarity (βSOR), the turnover component (βSIM), and the nestedness component (βSNE) across all pairwise comparisons among the included samples. **Figure S12:** Global trends in marine microbial species diversity in different oceans (from 2009 to 2020). **Figure S13:** Global marine microbial change trends in functional diversity in different oceans (from 2009 to 2020). **Figure S14:** Latitudinal patterns of multiyear trends in microbial diversity (from 2009 to 2020, based on the Mann–Kendall test). Species diversity (lower) and functional gene diversity (upper). **Figure S15:** Changes in the relative abundances of dominant taxa in ocean microbial communities in 2020 compared with those in 2009. **Figure S16:** Changes in the relative abundances of functional gene pathway expression in ocean microbial communities in 2020 compared with those in 2009. **Figure S17:** Environmental variables with significant time trends between 2009 and 2020. **Figure S18:** Annual functional gene diversity trends generated using recursive trend analysis. The solid lines indicate predictions with individual features held constant (2009), and the dashed lines show their linear fits. The change in slope reflects the contribution of each feature. **Figure S19:** Summary of multiple importance metrics for species diversity (A) and functional gene diversity (B). **Figure S20:** Feature interaction networks for species diversity (A) and functional gene diversity (B). The five strongest interactions are shown as red lines (the strength of the interaction between features is reflected by the number of times the two features appear in the RF model). **Figure S21:** Bootstrap stability of predictor importance for species and functional diversity. Points show the mean out-of-bootstrap permutation importance for species diversity and functional diversity, and horizontal bars indicate the 95% bootstrap intervals from 1000 iterations. Predictor importance reflects contributions to random forest predictions and should not be interpreted as causal effects. **Figure S22:** Change trends for grazing (A) and NO_3^- (B) from 2009 to 2020 compared with the baseline values for 2009. **Figure S23:** Percentage changes in grazing (upper) and NO_3^- (lower) in each grid cell compared with the values in the previous year (from 2009 to 2020). **Figure S24:** Average values of grazing (upper) and NO_3^- (lower) from 2009 to 2020 in the North Pacific Ocean and South Pacific Ocean. **Figure S25:** Spatial areas and species diversity based on the division of grazing (A, B) and NO_3^- (C, D) thresholds in the North Pacific Ocean and South Pacific Ocean. Thresholds were defined according to

Figure 4B–E, distinguishing between positive and negative effects. **Figure S26:** Single- and two-feature continuous responses of microbial functional diversity to grazing and NO_3^-. (A, B) Single-feature continuous responses of microbial functional diversity. The grey shading represents the 95% confidence interval, with a y-axis value of 0 as the baseline for distinguishing positive and negative effects. (C, D) Two-feature continuous responses of microbial functional diversity to grazing and NO_3^- in the North Pacific Ocean (C) and South Pacific Ocean (D). **Figure S27:** Global spatial distributions of the percentage changes in species diversity and functional gene diversity under different scenarios. (A, B) Enhancement potential for species diversity (A) and functional gene diversity (B). (C, D) Risks of loss for species diversity (C) and functional gene diversity (D). **Figure S28:** Sensitivity of constrained exploratory scenarios to the optimization parameter ranges. (A, B) Mean responses of species diversity (A) and functional gene diversity (B) across relative bounds of ±5% to ±30%. (C, D) Grid-level Pearson correlation coefficient between each relative bound result and the ±10% baseline for species diversity (C) and functional gene diversity (D). (E, F) Mean responses under the baseline ±10% optimization bounds (E) and grid-specific bounds derived from the 2009–2020 historical range of each grid cell (F). **Figure S29:** Spatial robustness across the different ranges of optimization bounds. (A, B) Rank standard deviation (SD) for species diversity under upward (A) and downward (B) constrained exploratory scenarios. (C, D) Corresponding rank SDs for functional gene diversity under upward (C) and downward (D) constrained exploratory scenarios. The rank SD was calculated from percentile ranks across ±5%, ±10%, ±15%, ±20%, ±25% and ±30% bounds for 43,254 ocean grid cells; lower values indicate more stable spatial ranking. **Figure S30:** Global spatial distributions of the percentage changes for grazing, NO_3^- and Chla levels under different scenarios. (A–C) Change percentage for grazing (A), NO_3^- (B), and Chla (C) under scenarios of maximize the enhancement potential of the diversity. (D–F) Change percentage for grazing (D), NO_3^- (E), and Chla (F) under scenarios of maximize the loss risk of the diversity. **Figure S31:** Comparison of microbial diversity in 2020 predicted by models with and without nutrient transfer-related indices (grazing, NO_3^- concentration and Chla concentration, set to zero) in the Pacific Ocean. **Figure S32:** Pearson correlation coefficient matrix for the variables in the datasets. **Figure S33:** Permutation test for overfitting. Species diversity (A) and functional gene diversity (B). The *x*-axes represent the correlation coefficients of the raw label values and permutated label values, and the *y*-axes represent the Q^2 values. The permutation test results revealed that none of the models were overfitted (intercept <0.05). **Figure S34:** Uncertainty analyses of species diversity (upper) and functional gene diversity (lower) in 2020. **Table S1:** Testable pathways and their evaluation in this work. **Table S2:** Project-level summary of the 1582 marine metagenomic samples. IO, Indian Ocean; MS, Mediterranean Sea; NAO, North Atlantic Ocean; SAO, South Atlantic Ocean; NPO, North Pacific Ocean; SPO, South Pacific Ocean; SO, Southern Ocean. **Table S3:** Global ocean zoning and abbreviations. **Table S4:** Sources of variables in the model. **Table S5:** Validation of the nutrient transfer hypothesis by convergent cross mapping (CCM) at Pacific observational time series site (1995–2022). **Data S1:** Suppplementary data 1.