scientific reports

Macroecological patterns of planktonic unicellular eukaryotes richness in the Southeast Pacific Ocean

Sonia Yáñez[1,2], Laura Tavera[2,8]✉, Garen Guzmán-Rendón[2,4], Leonardo D. Fernández[5], Nicolás Espinoza-Arevena[2,7], Margarita Marchant[3] & Cristián E. Hernández[2,6]✉

In recent years, studies focusing on microbial biogeography have been developed, but macroecological processes in marine microorganisms remain unclear, especially in seemingly continuous environments such as the Southeast Pacific Ocean (SPO), where information on microbial distribution patterns is limited, and they may vary depending on the habitat and lifestyle. We used unicellular planktonic eukaryotes as model organisms to determine their biogeographic patterns in the SPO, identify the underlying ecological and historical-evolutionary processes and compare with other microorganism groups. Our analyses were based on the Niche Theory to model species diversity distribution using large open-access ecological and physical-biogeochemical databases based on Bayesian approaches, an integrated nested Laplace approximation (INLA), and Generalized Additive Models (GAM). As a result, two richness hotspots were observed, which are associated with coastal and offshore regions in the central southern areas of SPO. The richness hotspots were associated mainly with nutrients (N/Si ratio) and Mixed Layer Depth (MLD), which could be explained by highly productive upwelling events in the SPO. In contrast, the negative correlation of predicted richness with low pH is strongly related to the effect of calcareous shells (tests), as lower pH levels hinder the formation and stability of calcium carbonate shells in protists like foraminifera and radiolaria, thereby affecting overall unicellular planktonic eukaryote diversity. Our results support the role of ecological processes related to productivity, energy dynamics, and ecological limits in shaping broad-scale diversity patterns of unicellular planktonic eukaryotes in the SPO. The results show colonization and extinction dynamics through species replacement (i.e. High Turnover) along the Chilean and Equatorial coasts associated mainly with the Hotspots of their biodiversity, but also a gradual species loss (i.e. High Nestedness) along the Peruvian Coast associated mainly with the Coldspots of their biodiversity; highlighting how local environmental fluctuations can shape these planktonic microorganisms' behavior, ecology and distribution. The distribution patterns of planktonic unicellular eukaryotes show little evidence of the effects of historical and evolutionary processes. This is because the high dispersal capacity of planktonic microbes probably dilutes the influence of these processes in environments lacking clear barriers to species dispersal. Additionally, the effect of historical events could be highlighted in specific taxonomic groups at the kingdom, phylum level or habitat type and addressing gaps about latitudinal richness in the SPO. This provides insight into the spatial distribution of marine microbes and contributes to conservation efforts, as these organisms are an essential foundation of the upper levels of the food web.

Keywords Biogeographic patterns, Humboldt Current System, Latitudinal gradient, Planktonic microorganisms, Protists, Species diversity distribution

[1]Departamento de Zoología, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Concepción, Chile. [2]Facultad de Medicina Veterinaria, Universidad San Sebastián, Concepción, Chile. [3]Departamento de Zoología, Universidad de Concepción, Casilla 160-C, Concepción, Chile. [4]Facultad de Ciencias, Universidad Católica de la Santísima Concepción, Concepción, Chile. [5]Núcleo de Investigación en Sustentabilidad Agroambiental (NISUA), Facultad de Medicina Veterinaria y Agronomía, Universidad de Las Américas, Manuel Montt 948, Providencia, Santiago, Chile. [6]Universidad Católica de Santa María, Arequipa, Perú. [7]Programa de Doctorado en sistemática y biodiversidad, Facultad de Ciencias Naturales y Oceanográficas, Universidad de

Concepción, Concepción, Chile. [8] Museo de Zoología, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Casilla 160- C, Concepción, Chile, Concepción, Chile. ✉email: ltavera@udec.cl; cristian.hernandez@uss.cl

The question of spatial distribution and the processes that determine diversity in microorganisms has been considered one of the most relevant topics in ecology. However, it is still widely debated and has not reached a consensus to clarify wide-scale distribution patterns[1,2]. Furthermore, the uncertainty and scarce knowledge of the global diversity of microorganisms[3–5] have restricted advances in the biogeography and macroecology of this group[6–8], despite its relevance in understanding the general rules of biodiversity[9].

Planktonic eukaryotic microorganisms have traits that enable them to disperse passively over long distances. Key traits that facilitate passive dispersal are their typically large population sizes, which increase the likelihood of dispersal events; small size, which allows them to be carried by wind or water currents; and the ability to form resistant cysts or spores, which provide protection during adverse environmental conditions[6,10,11]. They also exhibit rapid diversification due to their short generation times[12–15], compared to macroorganisms.

Generally, the biogeographic processes determining diversity distribution can be ecological and/or historical-evolutionary (e.g. Martiny et al.[14]; Fine[16]). These processes act at different temporal scales, either through generational times (ecological processes)[17] or geological (historical-evolutionary processes)[18]. In the case of ecological processes affecting the species' dispersal and distribution, the co-existence and survival of species are modulated by abiotic parameters. The range of environmental tolerance plays a fundamental role, reflecting the limits (upper and lower) of an environmental parameter at which a species thrives[19]. This is related to the environmental requirements of each species (e.g. generalists or specialists), where specialization increases the number of available niches, influencing speciation and extinction rates. This will depend on the carrying capacity for species within localities or regions, determined by resource availability[16,20]. Therefore, the extent of ranges is modulated by the persistence of populations, which respond to changing conditions at temporal and spatial scales[21].

In the ocean, ecological processes are focused on hypotheses such as energy availability (measured through environmental variables such as Mixed Layer Depth, and Eddy kinetic Energy), ecological limits (e.g. salinity, dissolved oxygen concentration and pH) and productivity (e.g. nutrients, chlorophyll-a and upwelling index), which modulate marine diversity[14,22–29].

Meanwhile, historical-evolutionary processes operate on geological time scales and depend on the time and rate of diversification, involving the effect of historical geographic barriers on biota's speciation and extinction rates. This affects species' distribution ranges, either expanding or fragmenting them[12,26,30]. These abrupt changes in past climatic conditions (e.g. transgressions and regressions) lead to local extinctions and changes in the dispersal of the most tolerant species, as evidenced during the Pleistocene in southwestern South America[31]. Therefore, these processes can be assessed by the niche conservatism hypothesis[32,33], which indicates the persistence of species in environmental conditions similar to those in which their ancestors evolved[32–34]. Additionally, dispersal-extinction dynamics[35,36] reflect variation in species composition within a community, exhibiting a turnover-nestedness pattern (beta diversity)[37], suggesting that variability in species richness results from a non-random mechanism.

The effect of these biogeographical processes on the spatial distribution of microorganisms may differ depending on the taxonomic group studied or their lifestyle (planktonic or benthic), which has been widely debated based on the principle "everything is everywhere, but the environment selects"[38,39]. Among the ecological processes that have determined the distribution of microorganisms are: (a) productivity measured through chlorophyll and nutrients in marine protists of phytoplankton and microzooplankton[40–43], (b) energy availability reflected by temperature gradients according to the thermal tolerance limits of terrestrial and marine protists species[24,40,41,77] and (c) ecological limits due to constraints in resource availability (e.g. salinity, oxygen) in marine protists, bacterioplankton and microeukaryotes[20,45,46]. On the other hand, historical-evolutionary processes have been found to have influenced the distribution of microorganisms, where the distribution pattern is determined by regional and local historical conditions that have generated and contributed to the formation of geographic barriers. These barriers restrict dispersal and produce physical isolation, leading to colonization limited to purely local scales in free-living bacteria and cyanobacteria[47,48].

Therefore, it has been suggested that new microbial biogeography studies be developed, where diverse data sets from various habitats and environmental conditions are analyzed[3,14,49] to address the mechanisms driving microbial biogeographic patterns. For this reason, it is essential to study the biogeography of microorganisms at a macroecological level, using unicellular planktonic eukaryotes from the Southeastern Pacific Ocean as a study model. Additionally, more information is needed on their latitudinal distribution and the biogeographic processes that shape diversity in this region, as it is the largest and one of the most productive oceans globally[50,51]. This will allow for comparisons with the biogeographic processes observed in other microorganisms with different lifestyles and ecosystems.

Material and methods

Study area: the coastal regions in the South Pacific Ocean (SPO) as an ecosystem model

The Humboldt Current Large Marine Ecosystem (HCLME) in the Southeast Pacific Ocean is widely recognized as the most productive marine ecosystem for the open ocean and closed areas[52,53] (Fig. 1), where the coastal upwelling process plays a fundamental role driven by the wind-induced offshore movement of surface waters mainly in central-south and northern area[54,55]. It occurs along certain coastlines where the prevailing wind blows parallel to the shore, causing surface waters to move away from the coast[56,57]. This movement brings up nutrient-rich, cold, and poorly oxygenated deep waters[58]. Southward, the Patagonian Fjords ecosystem is one of the most extensive fjord regions of the world[59], and some of the most biogeochemically active areas of

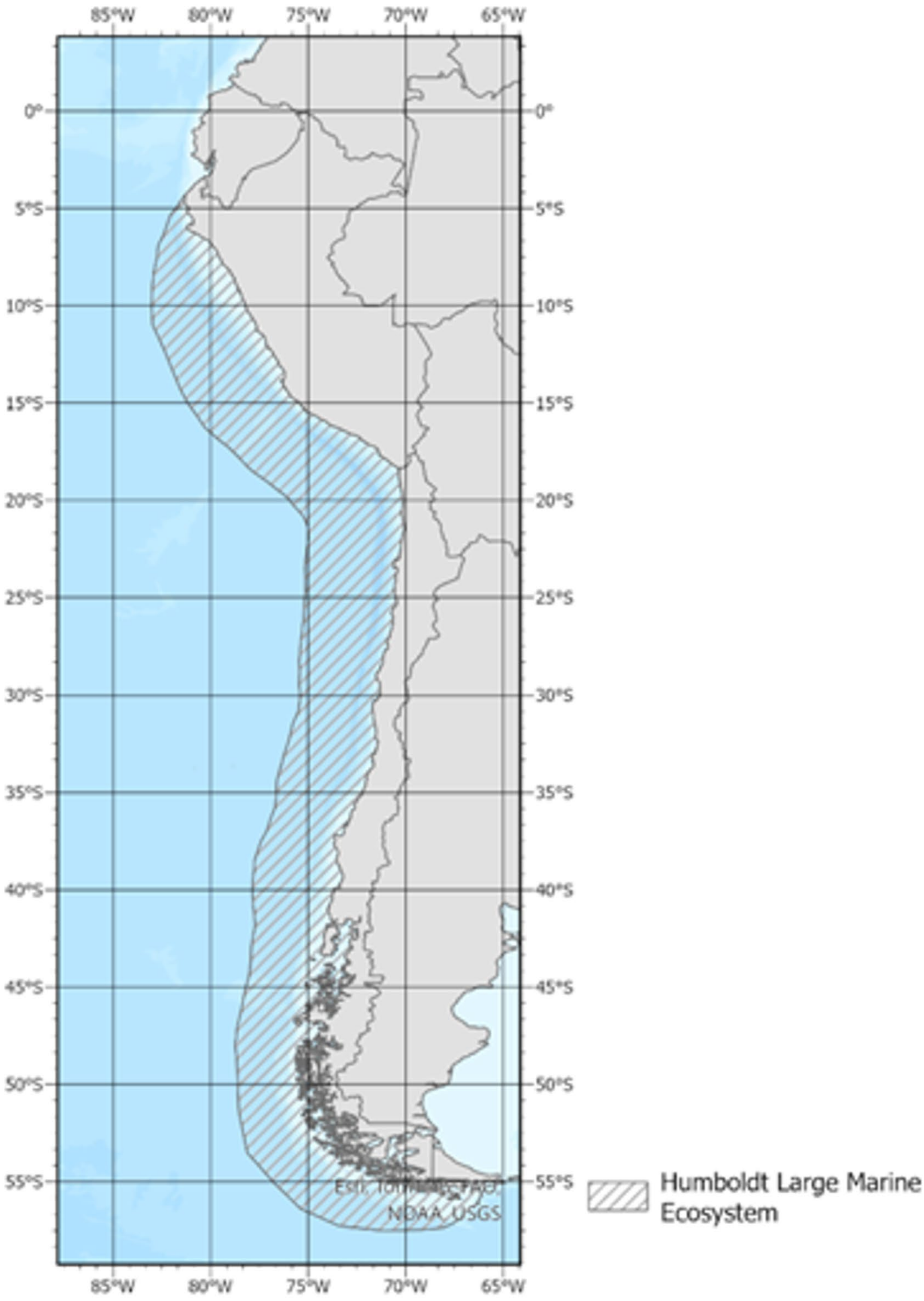

Fig. 1. A map of the study area with the Humboldt Large Marine Ecosystem is shown as a hatched fill area.

the biosphere with invaluable ecosystem goods and services primarily through aquaculture and tourism. The Equatorial divergence, on the other hand, has been described as an ecosystem with high N/Si nutrient ratios and elevated chlorophyll concentrations at the pelagic layers (100 m depth). In comparison, the subtropical band exhibits a high N/Si nutrient ratio at the surface and 100 m depth. Nutrient availability enhances productivity in the upper layer, supporting phytoplankton (diatoms) patches in these ecosystems[60].

Environmental database and processing

Data from the Marine Environment Monitoring Service was used to compile and build a database of environmental variables to a 0.25 × 0.25 degrees resolution. The Copernicus (http://marine.copernicus.eu/services-portfolio/access-to-products) open-access databases were explored and used for selected variables: Salinity (Sal), Mixed Layer Depth (MLD), Phosphate (PO_4^{-3}), Silica (Si), Potential temperature (Tpot), pH and Dissolved Oxygen Concentration (DO) and from 0 to 200 m in discrete vertical levels; Chlorophyll-a (Chl-a),Eddy Kinetic Energy (EKE), Sea Level Anomaly (SLA) at surface. All data were compiled from 1992 to 2022.

Pearson's rank correlation index and the variance inflation factor[61] were calculated before running the models to avoid the correlation and collinearity among explanatory variables (Supplementary Fig. 2). Specifically, correlations among the variables were assessed by performing a Pearson's correlation test and VIF with the R packages corrplot and usdm, respectively[62,63]. Pairs of variables with high correlation values (Pearson correlation $r > 0.6$) or low variance inflation ($VIF > 2$) were identified, and only one was included in the modelling process. We adopted this threshold based on what Jeng[64] proposes for $r = 0.6$, $k = 8$. We found that the main environmental variables drivers of richness were N/Si ratio, MLD and pH.

Data sources-occurrences database and richness estimation

The ecological/biological data were obtained from the Ocean Biogeographic Information System (OBIS) and Global Biodiversity Information Facility (GBIF.org [11th November 2020; GBIF Occurrence Download http://doi.org/https://doi.org/10.15468/dl.wnjjkc]). We matched the latitude and longitude coordinates for OBIS and GBIF species records to seas after cleaning the data for taxonomy. Considering the posterior analyses based on species distribution models (SDMs), we only included taxa with at least ten occurrences in the database[65]. The data were obtained for all occurrences between 0 and 200 m of the study area from 1992 to 2022. Following the data retrieval, we eliminated records without information on the geographic coordinates, coordinates equal to zero, or records located inside continents. We selected only records at the level of genus/species and excluded duplicate records. The different sources of information were compiled in a spatial database using QGIS 3.12.1 (QGIS Development Team, 2020), a free and open-source Geographic Information System and ArcGIS Pro (Esri, 2023).

The raw data was organized to evaluate geographic areas with high levels of species richness within the Humboldt Current Large Marine Ecosystem (HCLME), based on the proportion of sites occupied. This metric helps detect clusters of cells containing higher-than-expected richness values, indicating spatial clustering[66]. To assess this observed species richness (alpha diversity), we constructed a presence-absence matrix (PAM) using species occurrence data within the study area. We defined the study area using a shapefile and later imported a second shapefile containing species occurrence records with geographic coordinates (longitude and lA grid-based PAM was created using the lets.presab.points() function from the letsR package. The study area was divided into 1° × 1° grid cells, where species presence (1) or absence (0) was recorded. Cells without species records and species absent from the dataset were removed to optimize the matrix. Species richness per grid cell was extracted from the richness raster (PAM$Richness_Raster) and visualized. The richness raster was exported in ASCII format for GIS applications, while the PAM was saved as a CSV file for further analysis. Beta diversity and its components (i.e. species turnover and nestedness) were analyzed using a presence-absence matrix generated with the "fuzzySim", "betapart"[67] and CommEcol[68], packages in R software (R Core Team, 2021).

Species richness (alpha diversity) is the absolute number of species living in each area, given equal weight to all resident species[69]. Beta diversity measures the differences in species composition between two ecological communities[37]. In this study, we calculated alpha diversity as species richness. Also, we assessed two components of Beta diversity in the Southeast Pacific Ocean space using the R package `betapart` following the methodology proposed by Baselga (2010, 2012)[67] that allows us to decompose beta diversity into two distinct components: species turnover and nestedness-resultant dissimilarity, providing a clearer understanding of the processes driving community composition across sites (see equations in Supplementary Table S4). High turnover reflects increased environmental variability, while nestedness represents the effects of variability on species colonization and local extinction. The identification of high species richness levels was done using the Gi* statistic[70], which detects spatial concentrations of high values of an entity (i.e., species richness). A significant area is characterized by high values surrounded by other high-value cells. Using the Z-score statistic, the local sum for an entity and its neighbourhood is compared to the total sum of all entities to determine if the local sum deviates significantly from random expectations. Significant Z-scores above 0 indicate hotspots (high diversity), while values below 0 suggest coldspots (low diversity). Additionally, Alpha and Beta diversity and hotspots were visualized using empirical Bayesian Kriging interpolation. Hotspots and Kriging models were determined using ArcGIS Pro software (Esri, 2023).

Because species richness in a sample depends on sample size and sampling effort[71], in oceanic regions, spatial biases in the distribution of sampling locations[72] can lead to under-sampled areas, potentially distorting observed patterns of marine species richness. To address this, redundancy statistics were used to estimate the total number of species, and sample completeness was assessed using the R package iNEXT, which can be applied to both abundance and incidence data[71]. This package provides functions to compute and visualise seamless rarefaction and extrapolation sampling curves. The rarefaction-extrapolation curve (Supplementary Fig. S3), which shows the estimated species richness as a function of sampling effort, exhibits an asymptotic trend, suggesting that while most species in the area have likely been sampled, further sampling could reveal additional diversity. This supports the adequacy of the sampling to describe the general macroecological patterns, but continued sampling is necessary to ensure a comprehensive assessment of marine biodiversity in the study area.

We created a biological-environmental spatial database, where each cell contained biogeographic information on eukaryotic cell richness alongside environmental data from selected variables. This database was generated using the 'raster', 'purr', and 'sf' packages in R software. This approach enabled the extraction of environmental data values corresponding to the latitudinal/longitudinal richness data, ensuring that both datasets shared the same spatial resolution. We conducted a Pearson correlation analysis to identify and remove highly correlated variables (Supplementary Fig. 2). As seen in the matrix, there is a high negative correlation (highlighted in dark red) between pH_mean and sal_mean (− 0.80) and between DO_mean and sal_mean (− 0.71). While strong positive correlations (highlighted in dark blue) between DO_mean and pH_mean (0.85). Other variables exhibit moderate to weak correlations, with richness showing relatively weak correlations with the other parameters**.** Therefore, three predicted variables were selected as explanatory variables in the predicted richness models: MLD, pH and N/Si ratio.

Validation of species records

Once the table of occurrence data from OBIS and GBIF was available, the information was filtered: (1) it was corroborated that each species recorded was indeed a marine eukaryote and planktonic species. The most updated marine database, the World Register of Marine Species (WoRMS; https://www.marinespecies.org/), was consulted for this. In the case of species not found in this database, we reviewed scientific articles that recorded the species of interest; (2) The taxonomy of each eukaryotic species was updated using the "Match Taxa" tool of the WoRMS (https://www.marinespecies.org/aphia.php?p=match) to ensure the most accurate and current information possible. Due to the frequent changes in the classification and taxonomic categories of microorganisms (and the fact that some still have an uncertain taxonomic status), this step was essential to ensure data consistency and reliability (Table S1).

Species spatial distribution models

The Integrated Nested Laplace Approximation (INLA) framework[73] was implemented using the R package inla to implement a hierarchical Bayesian spatial GAM for modelling eukaryotic unicells richness. INLA uses the Stochastic Partial Differential Equations approach[74] to account for spatial effects. Under this framework, the smoothness of the spatial field is typically denoted by the Kappa statistics[75]. We incorporated a stochastic spatial effect using the Stochastic Partial Differential Equation (SPDE) approach[76]; this numeric vector links each observation to a spatial location, capturing region-specific noise that the available covariates cannot be explained[77]. Following Lindgren and Rue[75], the spatial component was modelled using multivariate Gaussian distributions with zero means and a spatially structured covariance matrix; this allows us to model spatial autocorrelation in species distribution over a triangulated mesh. The mesh was constructed to balance computational efficiency and spatial resolution, ensuring that areas with higher sampling density had finer resolution while maintaining adequate coverage in under-sampled regions.

The species richness was modelled using the common Poisson family and logit link function. Except for the set type, selected explanatory variables (pH, N/Si ratio, MLD) were modelled using a second-order random walk (RW2) latent model, accommodating potential non-linear relationships[76]. Posterior distributions were obtained for all the parameters, with the 0.025 and 0.975 quantiles defining the 95% credible intervals, indicating that each unknown parameter has a 95% probability of falling within this range of values[78].

Multiple models were generated to obtain the best-fitting model. First, variables with linear relationships were included using a GAM. Second, the influence of the spatial effect was evaluated by removing it from the model. Final model selection was performed using a forward stepwise procedure. The performance of these models was assessed based on Watanabe-Akaike Information Criterion (WAIC), where a lower value indicates a better model fit, meaning it is more likely to accurately predict unseen data, as it balances model complexity with how well it explains the observed data within a Bayesian framework; essentially, a lower WAIC value represents a better trade-off between model fit and overfitting, with smaller values being preferred when comparing different models. We use WAIC mainly because, according to Gelman et al.[79], it is a fast and computationally convenient alternative, is fully Bayesian (using the posterior distribution rather than a point estimate), gives reasonable results, and has a more-or-less explicit connection to cross-validation.

A cross-validation approach was applied with a k-fold partitioning method (with $k=5$) to assess model performance[80,81]. The relationship between occurrence data and the environmental variables was modelled using a training dataset comprising 80% of the data. At the same time, the quality of predictions was assessed using a test dataset for validation with the remaining 20%[82–84]. Validation was repeated five times for the best-performing model, and results were averaged across the different random subsets. Models were evaluated to formally assess their overall spatial effect and mean of the posterior distribution of predictive performance by calculating the Spearman rank correlation. Finally, posteriori GAMs models were run to test the effects of pH, N/Si ratio associated with Upwelling index and Mixed Layer Depth (MLD) on Predicted richness.

Ecological processes

We tested three ecological hypotheses to investigate contemporary processes' role in shaping planktonic protists' latitudinal diversity gradient (i.e. Productivity, Ecological Limits, and Energy; Table 1). Each hypothesis was represented by a set of explanatory variables frequently used as surrogates for these hypotheses. The model parameters used to test the productivity hypothesis were associated with upwelling events, such as chlorophyll, nutrients (e.g. phosphates, silica, nitrogen) and their export fluxes through N/Si ratios, which affect the growth of many planktonic eukaryotes species and food web structure[85,86]. Additionally, the upwelling index was used, which estimates upwelling intensity over a time series, reflecting the distribution and spatial dynamics of plankton and marine resources[87,88]. On the other hand, the ecological limits hypothesis was explained by resource availability constraints[20] (salinity, pH, oxygen), determined by the environmental tolerance ranges for each species. This can lead to adaptation to strictly local conditions, as seen in specialist organisms[89–91]. Finally, the variables used to estimate the energy hypothesis were Eddy Kinetic Energy (EKE), Temperature potential (Tpot), Sea Level Anomaly (SLA) and Mixed Layer Depth (MLD), which correspond to physical processes driving

Ecological hypothesis	Model parameters	Prediction	References
Productivity	Phosphate (PO4-3), Nitrogen, Silica (Si), Chlorophyll-a (Chl-a), upwelling index, N/Si ratio	Positive relationship between population sizes and available energy. Higher probability of speciation and lower probability of extinction in areas with higher productivity. Low temperatures and oxygen associated with upwelling	Bode et al.[87], Jetz and Fine[96], Hurlbert and Stegen[27]; Coogan et al.[97], Goldenberg et al.[85]
Ecologal limits	Salinity, Dissolved oxygen concentration (DO), pH	The carrying capacity of species within an area is determined by the availability of resources, so species diversity is balanced by immigration, speciation or extinction. Predominance of competition between species with similar ecological requirements	Balzano et al.[90], Rabosky and Hurlbert[20], Rabosky et al.[28]; Fine[16]
Energy	Eddy Kinetic Energy (EKE), Potential temperature (Tpot), Mixed Layer Deep (MLD) and Sea Level Anomaly (SLA)	Species richness increases according to the energy input to the ecosystem	Currie[98]; Evans et al.[25], Brown[34], Freilich et al.[93], Moreles et al.[94]

Table 1. Ecological hypotheses tested to explain the latitudinal diversity gradient of planktonic unicellular eukaryotes richness in the Southeast Pacific Ocean.

the vertical movement of water masses and the export of phytoplankton from de surface[92,93]. This determines the available thermal energy and the energy and its exchange between the atmosphere and the ocean[94,95].

The model parameters for each hypothesis were preselected based on Deviance Information Criterion (DIC), AUC, and Logarithmic Conditional Predictive Ordinate criteria (LCPO). All parameters correspond to environmental variables frequently used as surrogates for productivity, ecological limits, and energy (Table 1).

Historical-evolutionary processes

Given the absence of adequate and well-resolved phylogenies for unicellular planktonic eukaryotes of our dataset, the Index of Taxonomic Distinction ($\Delta+$) was conducted in PRIMER 6. This index was estimated using different taxonomic levels (species, genus, families) based on how occurrences were recorded in the database extracted from GBIF. However, according to Clarke and Warwick[99], this analysis accounts for taxonomic precision by incorporating the taxonomic hierarchy level recorded for each sample (e.g. species, genus, family, order, class). With this approach, the index calculates the average taxonomic distance between all pairs of taxa in a community, using a taxonomic classification tree to approximate evolutionary relationships[29,100]. It assigns a weight to the taxonomic distance (ω) according to the taxonomic hierarchy and weights taxa richness according to each of these hierarchies[99].

The observed $\Delta+$values were compared to a null model generated through 10,000 iterations (repeatedly subsampling), resampling species richness for each latitude[99]. This is intended to compare the observed values with the expected values and detect deviations in diversity, indicating communities significantly different than expected by random chance[101]. Communities with $\Delta+$values similar to those predicted by the null model suggest that historical-evolutionary processes are weak and contemporary ecological factors primarily shape diversity. In contrast, $\Delta+$values significantly higher than expected indicate taxonomically diverse communities, where evolutionary novelties may have arisen as adaptations to environmental challenges[44]. On the other hand, $\Delta+$values significantly lower than expected suggest communities dominated by closely related taxa, likely due to niche conservatism and environmental filtering[32,33,44].

Results

Diversity of planktonic unicellular eukaryotes in the SPO

We found occurrences data from OBIS and GBIF of 110 genera distributed in two kingdoms: Chromista and Protozoa, with 24 phyla, 34 classes, 56 orders and 70 families (Table S2).The most representative phylum was Foraminifera, found in 31% of the total latitudes evaluated. Followed by Euglenozoa and Amoebozoa, which had 19% and 16%, respectively (Fig. 2). The most widely distributed protozoan genus was the flagellate kinetoplastid *Rhynchobodo* sp., recorded at 23 latitudes, primarily between 5°S and 19°S, and the species *Bodo saltans,* recorded at 20 latitudes, mainly between 8°S and 18°S. These taxa were followed by the euglenids *Eutreptia* sp. and *Eutreptiella gymnastica*, found at 15°S and 17°S, respectively (mostly between 8°S and 15°S) (Fig. 3, Table S1).

Chromista was mainly represented by foraminiferal species. Among these, *Globigerina bulloides, Globigerinita glutinata*, and *Neogloboquadrina* species (Fig. 4) were the most widely distributed, occurring across 17°S to 19°S, predominantly between 28°S and 40°S (Fig. 3, Supplementary Table S1).

Spatial gradients of unicellular planktonic unicellular eukaryotes richness

Two major hotspots of species richness were identified (Fig. 5a,b): (1) along the equatorial coast (0° to 5°S), and (2) in the Chilean central/south coastal regions (28°S to 44°S). This latter region was more extensive, with higher hotspot intensity detected in offshore areas (32° to 35°S) and coastal zones (38°to 43°S). In contrast, coldspots were located in the Peruvian coastal region (8°S to 17°S), the northern coast (18°S to 22°S), and the Patagonian channels of southern Chile (Fig. 5a,b).

On another hand, the species composition across the SPO exhibited high turnover rates associated with the Chilean coast, except between 25° and 30ºS, and in equatorial regions (Fig. 6a). In contrast, the Peruvian coast exhibited low turnover but high nestedness values (Fig. 6b). These spatial variations in turnover displayed a negative correlation from high latitude to equator (Supplementary Fig. S1a, $\text{Rho} = -0.55$; $P < 0.05$), while nestedness showed a positive correlation from high latitude to equator (Supplementary Fig. S1b, $\text{Rho} = 0.53$; $P < 0.05$).

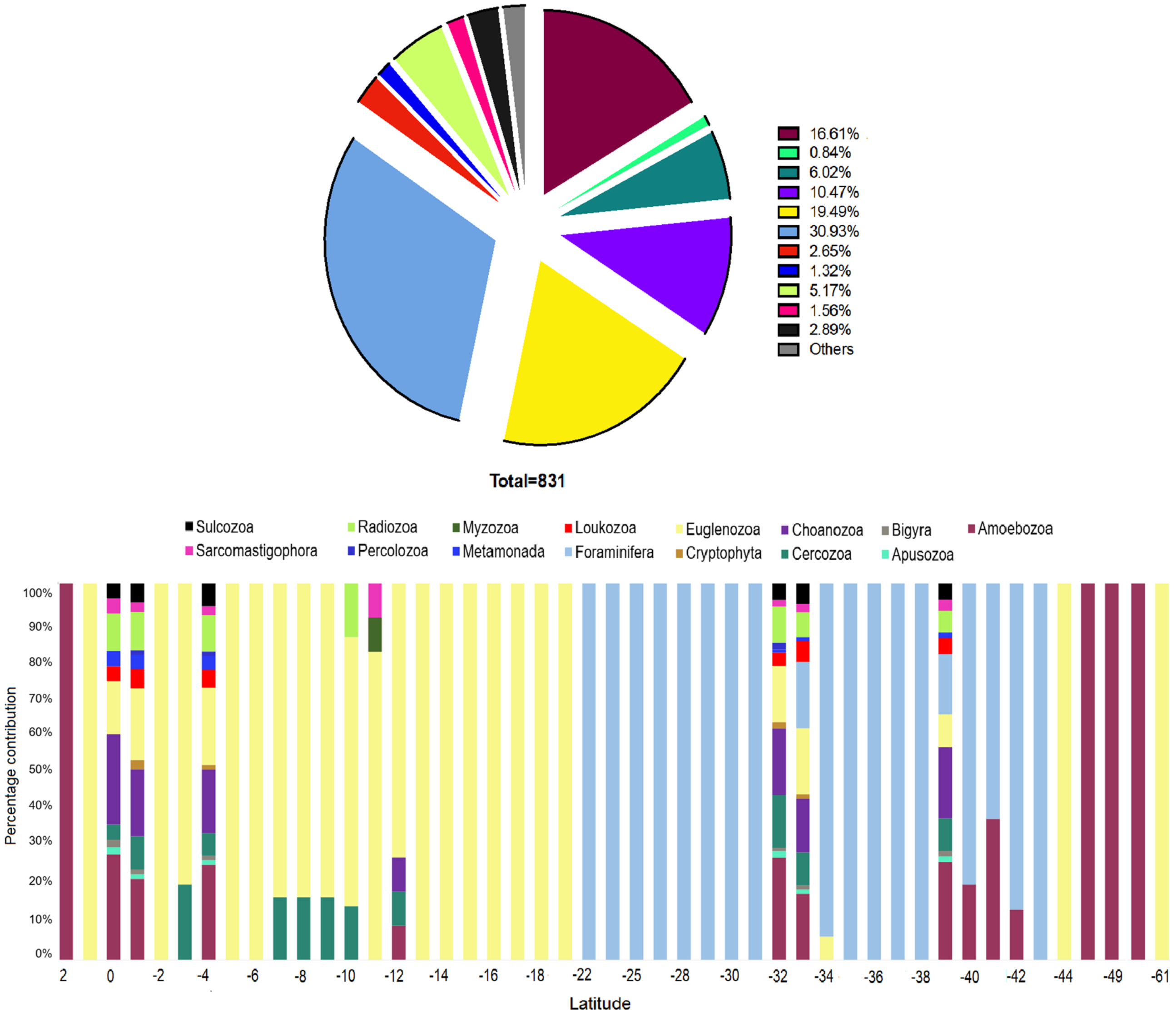

Fig. 2. Taxonomic composition of phyla of planktonic unicellular eukaryotes of Southeast Pacific Ocean.

Species spatial distribution models and Ecological processes

All possible models derived from the predictors were evaluated, and the results of the model comparisons based on WAIC criteria identified the best-performing model as one that combined the N/Si ratio and spatial effect. This model was followed closely by models that included Mixed Layer Depth (MLD) and the N/Si ratio plus spatial effect (Table 2 and Fig. 7). Thus, data analysis revealed that the primary environmental drivers of richness were the N/Si ratio, MLD, and spatial effect.

The richness patterns predicted by the selected model were consistent with the observed richness, validating the model with a high positive Pearson correlation (Fig. 8, $r = 0.8$; $p > 0.001$). Also, the spatial outputs from this model effectively account for habitat variability mainly from N/Si ratio (Table 2), producing a more natural representation of unicellular eukaryotic richness patterns, including the spatial effect (Fig. 8a) and the mean of the posterior predictive probability distribution (Fig. 8b) coherent with the observed richness distribution, and validated with also high positive Spearman rank correlation $r = 0.81$, $p < 0.001$ (see Fig. 8c).

Finally, to better understand the ecological processes underlying the spatial distribution of unicellular eukaryotes, we run a posteriori GAMs models to test the effects of pH, N/Si ratio associated with Upwelling index and Mixed Layer Depth (MLD) on predicted richness; those results showed that all of them are significative positive related with variability of predicted richness ($p < 0.001$).

Historical-evolutionary processes

The taxonomic distinctiveness ($\Delta +$) does not show communities for each latitude with values higher than the null model (Fig. 9), suggesting that historical-evolutionary processes are weak and contemporary ecological factors primarily shape diversity. These results show that the species are not relatively distantly related to each other, mainly northern 20° and southern 39° S ($\Delta +$: 1.53 to 2, Supplementary Table S3). On the other hand, the

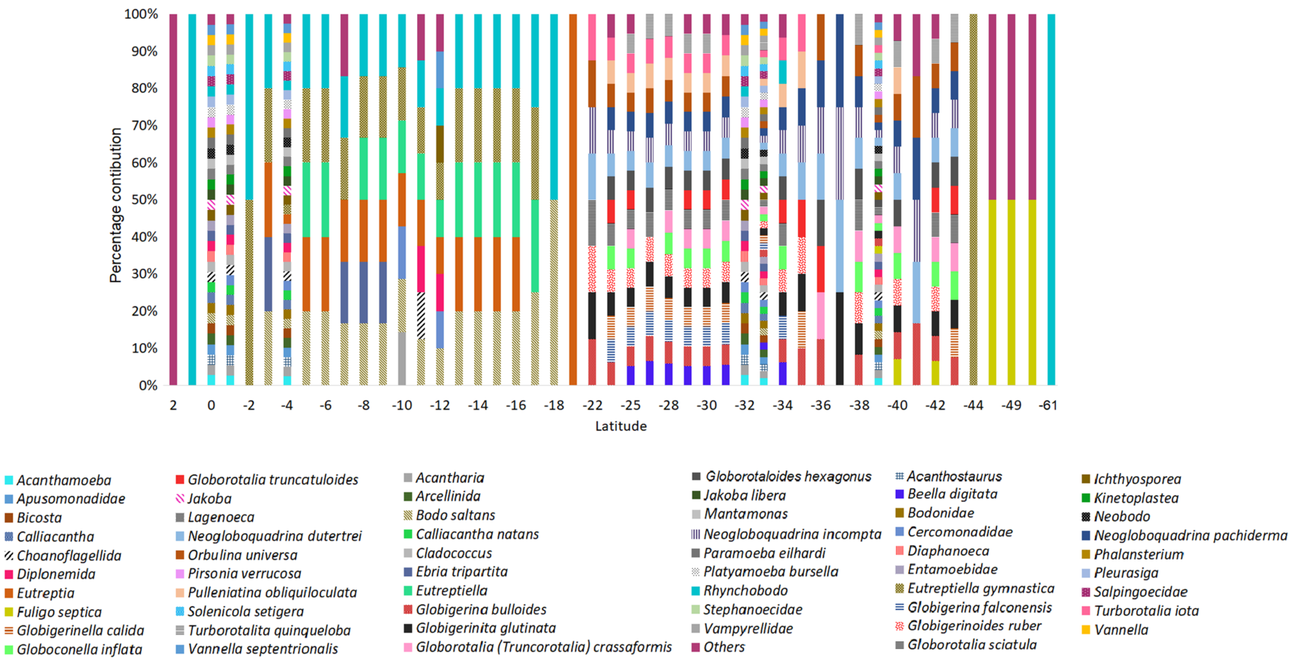

Fig. 3. Percentage contribution of species or genera of planktonic unicellular eukaryotes from the Southeast Pacific Ocean.

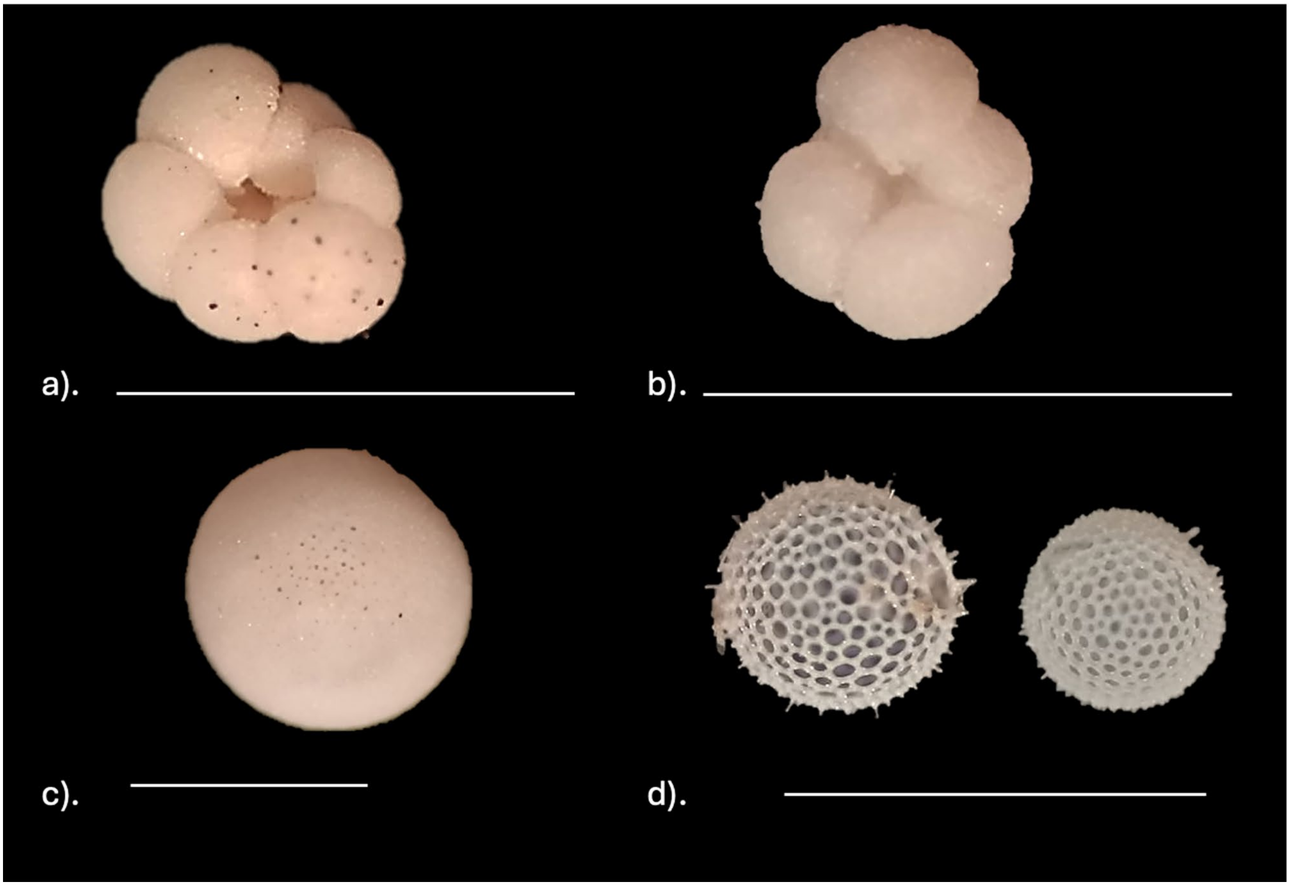

Fig. 4. Examples of widely distributed planktonic unicellular eukaryotes in the Southeast Pacific Ocean. Phylum Foraminifera: (**a**) *Neogloboquadrina dutertrei*, (**b**) *Neogloboquadrina incompta*, (**c**) *Orbulina universa* and (**d**). Phylum Radiozoa: *Carposphaera* sp. *The white bar indicates a size scale of 1 mm.

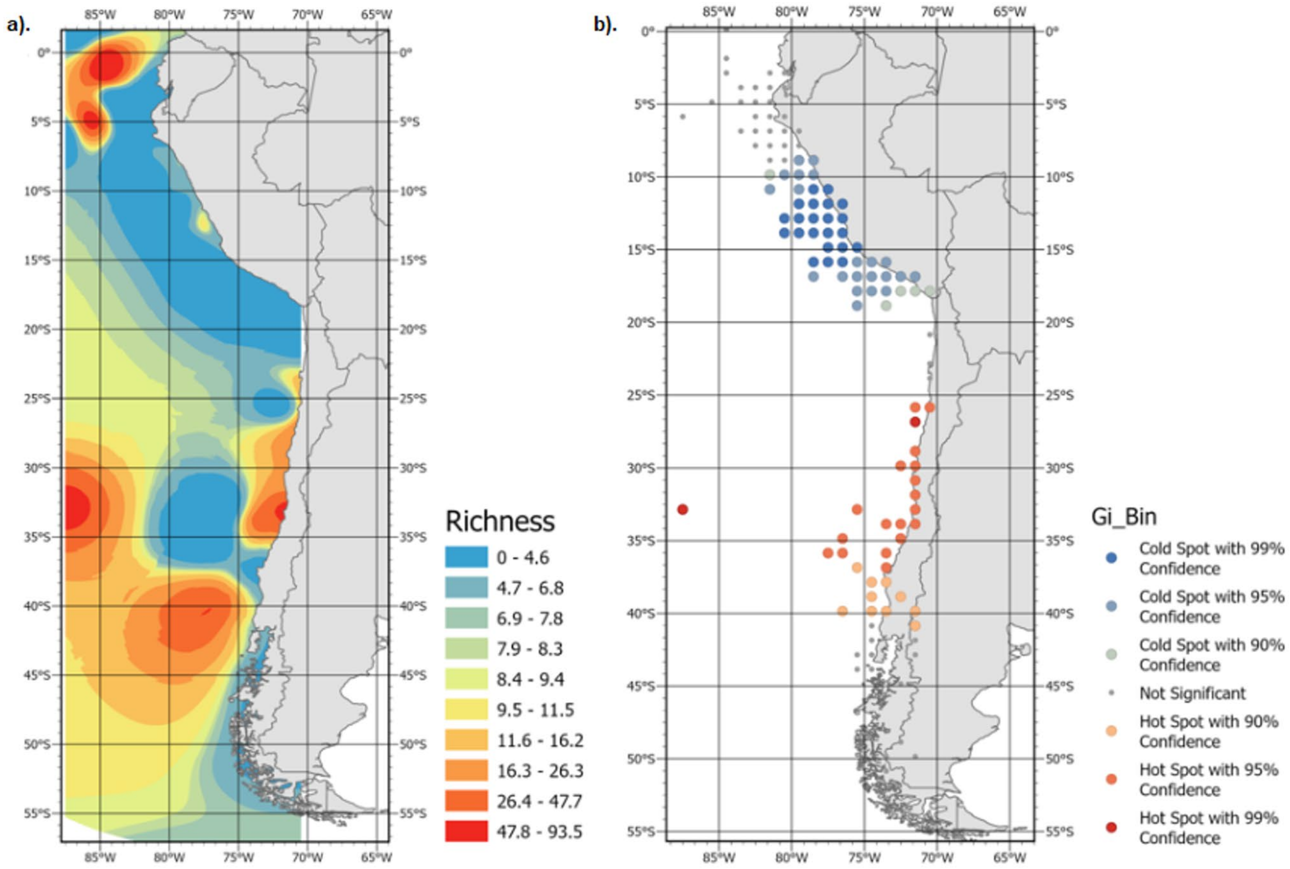

Fig. 5. Survey of planktonic unicellular eukaryotes richness of the Southeast Pacific Ocean. (**a**). Observed Richness, and (**b**) Getis-ord analysis.

results indicated a trend of lower values than expected by the null model between 20°S and 44°S (Δ+: 1—1.24, Supplementary Table S3), indicating communities with species that are more closely related to each other, suggesting a narrower range of taxonomic groups likely due to niche conservatism and environmental filtering.

Discussion

Our results, based on Bayesian species distribution models with key environmental variables and spatial effects to investigate the spatial distribution patterns and drivers of unicellular eukaryote richness along the Southeast Pacific Ocean (SPO), support the idea that ecological processes (i.e. productivity, energy dynamics, and ecological limits) are the fundamental drivers modulating the spatial distribution of these microorganisms along the SPO. This is because the high dispersal capacity of planktonic microbes probably dilutes the influence of historical-evolutionary processes in environments lacking clear barriers to species dispersal. The general patterns of its biodiversity show two hotspots (i.e. coastal and offshore regions in the central southern areas of SPO) associated mainly with nutrients (N/Si ratio) and Mixed Layer Depth (MLD), which could be explained by highly productive upwelling events in the SPO. These hotspots were associated with colonization and local extinction dynamics through species replacement (i.e. High Turnover); in contrast, the coldspots were associated with a gradual species loss (i.e. High Nestedness), highlighting how local environmental fluctuations can shape these planktonic microorganisms' behaviour, ecology and distribution.

Richness hotspots

The first hotspot, located along the equatorial coast (0° to 5°S), was mainly composed of amoeboids of the class Discosea, Tubulinea and Variosea, recorded in three latitudinal bands. These results may reflect the thermal preferences of amoeboid communities, reflecting both homothermal and heterothermal systems[102,103]. The lobose testate amoebae species (Tubulinea) have a K-strategy by homothermal conditions, by tropical currents (e.g. Gunther Current and Peru–Chile Countercurrent), without presenting a cyst stage in conditions with lower temperatures[102]. Conversely, filose testate amoebae species (Discosea and Variosea) present an r-strategy, indicating a heterothermal system[102] by mixing between tropical currents and the Humboldt Current in the Ecuador and Perú latitudes (2°S to 5°S).

Yet, the second hotspot (28°S to 44°S) was formed by many taxonomic groups of foraminifera, Choanozoa, Cercozoa, Amoebozoa, Euglenozoa and Radiozoa, forming the highest probability of hotspot in offshore regions (32°S–35°S, Fig. 5). In coastal regions (38°S–44°S), the hotspots were represented mainly by foraminifera species

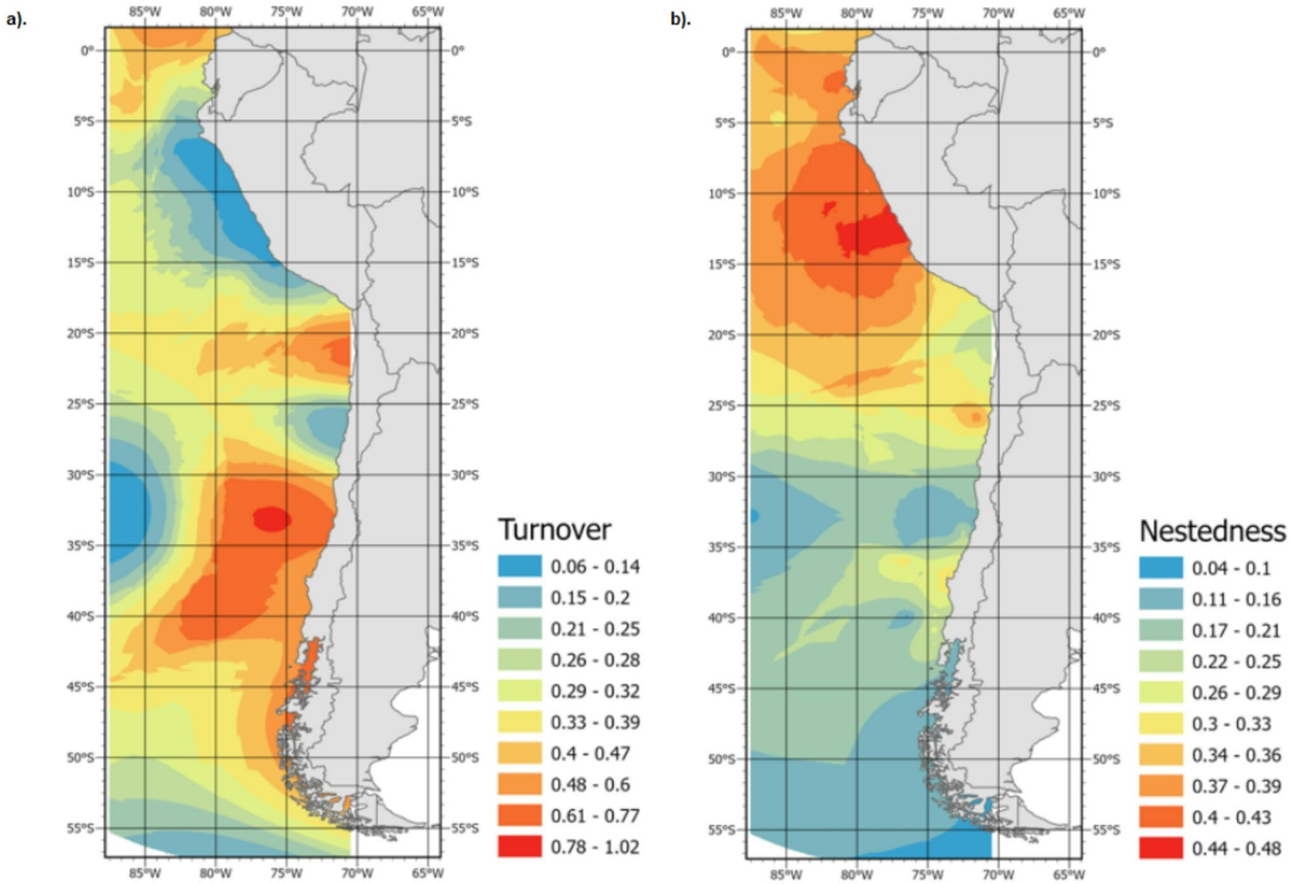

Fig. 6. Spatial patterns of Beta diversity of the planktonic unicellular eukaryote's richness in the Southeast Pacific Ocean. Where: (**a**) Turnover, showing spatial variation in species replacement across latitudinal gradients; and (**b**) Nestedness, highlighting the contribution of species loss to patterns of richness.

belonging to Globigerinidae and Globorotaliidae families. The greatest representation of the diversity hotspot in the central-southern Chile regions coincided with the three most important upwelling areas in Chile (23°S, 30°S and 33°S)[104–107]. The increase in nutrients and primary productivity leads to a rise in tropical and subtropical species, which are characteristic of subsurface water upwelling and well-adapted to strong mixing and eutrophic conditions[107–109]. Similar patterns have also been observed in regions like the Benguela Upwelling System[109]. South of 39°S, low temperatures driven by the increased influence of the coastal branch of the Humboldt Current System and the Cape Horn Current[110,111] limit the distribution of warm-water and subtropical foraminifera species. Instead, these areas show a higher occurrence of typical cold environment species, polar waters with broad temperature tolerances, such as *N. incompta* and *G. bulloides*.

Bimodal richness gradient

Several studies have evaluated the latitudinal diversity gradient of organisms, highlighting geographic bimodal patterns[112–114]. These studies emphasize areas of high diversity associated with high productivity conditions, such as upwelling zones, which align with the biodiversity hotspots observed along the south-central coast of Chile and Ecuador. In our study, a distinctive bimodal pattern is also evident in unicellular eukaryotes cells, which could indicate that these organisms respond more sensitively to local factors at a smaller scale. This contrasts with other planktonic organisms, whose diversity follows a continuous and monotonic latitudinal diversity gradient and is less influenced by local productivity, likely due to their greater mobility and ability to adapt to broad environmental gradients[115,116].

In the case of SPO, the Humboldt Current and the coastal upwelling in the central-southern region of Chile are well-known for sustaining communities rich in zooplankton and other organisms due to the high availability of nutrients and trophic resources in these areas[117]. Furthermore, equatorial areas favoured by mixing warm and cold-water masses typically show an increase in multicellular planktonic species richness, similar to what was observed in the protists in this study[118–120]. However, unlike unicellular eukaryotic cells, multicellular planktonic organisms often exhibit distributions more strongly structured by latitudinal gradients and show greater dependence on currents for dispersal[116,121]. Unicellular eukaryotic cells respond with greater sensitivity to changes in local environmental variables, such as mixed layer depth, nutrient, and pH concentrations (Fig. 7). In contrast, the distribution patterns of planktonic multicellular organisms are more predictably and continuously influenced by factors such as temperature and current dynamics[122,123].

Model	WAIC
N/Si ratio + Spatial effect	595.90
MDL + N/Si ratio + Spatial effect	596.64
pH + N/Si ratio + Spatial effect	597.64
Spatial effect	597.10
MDL + Spatial effect	598.32
pH + MDL + N/Si ratio + Spatial effect	597.82
pH + Spatial effect	598.61
pH + MDL + Spatial effect	599.18
pH + MDL + N/Si ratio	19,674.68
MDL + N/Si ratio	18,440.99
pH + N/Si ratio	4968.22
N/Si ratio	5896.79
pH + MDL	18,864.02
MDL	13,782.75
pH	3551.41
Intercept	3064.53

Table 2. Model comparison for the richness of eukaryotic cells based on Watanabe-Akaike Information Criterion (WAIC). The table compares various models incorporating combinations of environmental predictors (N/Si ratio, MLD, pH) and spatial effects. Lower values of WAIC indicate better model performance.

Ecological processes

According to the results based on WAIC values for model selection, different ecological processes in the SPO modulate the diversity of planktonic unicellular eukaryotes. The first process was the productivity, associated with N/Si ratio and upwelling index, followed by energy processes, measured with the Mixed Layer Depth (MLD) and finally, the ecological limits reflected by pH values (Table 2, Fig. 7). Several of the latitudes with the highest species richness are coincident with upwelling areas such as Mejillones Bay (23 °C), Punta Lengua de Vaca (30°S) and Concepcion Bay (36.5 °C) described by Thiel et al.[55], suggesting the upwelling events locally drive changes in the richness and species composition in the unicellular eukaryotic cells, concurrent with recent studies in high-productivity coastal areas[124,125]. This positive relationship between upwelling and diversity was corroborated by other microorganisms, such as nanoflagellates and bacterioplankton[126–128]. The increase in nutrients and food availability favours the proportion of primary producers and, in turn, the presence of heterotrophic eukaryotic cells. Thus, these upwelling conditions control the planktonic food webs.

On the other hand, the MLD reflects the thermal stratification pattern that indicates changes in temperature ranges in the water column in oceanic and coastal regions, as well as the vertical light available and as a heat source[129]. This variable drives the spatial diversity of plankton, such as unicellular eukaryotes, because many of these taxonomic groups (e.g. dinoflagellates, foraminifera) photosynthesize, and their photosynthetic rate increase as a function of available irradiance[130]. Shallower MDL, due to the warming of surface temperatures, restricts the passage of available light and nutrients for the planktonic organisms, decreasing their biomass[131].

The ecological limits reflected by pH values affect the diversity of unicellular eukaryotes, mainly in cells with calcareous exoskeletons (e.g. foraminifera, coccolithophores)[132], where acidic water conditions cause a shift in carbonate chemistry reflecting lower concentrations of carbonate ions concentrations. This could decrease calcium carbonate production by marine calcifying organisms and dissolve their exoskeletons, such as shells or plates[133]. Therefore, this causes a decrease in the survival of calcareous eukaryotic populations, being more fragile to environmental stress conditions (e.g. pollution, hypoxia and eutrophication)[132]. However, the calcification/dissolution rates fluctuate under different CO_2 concentrations, depending on taxa, nutritional status, photosynthetic capacity, life stages and pH regulation[134,135].

The presence of organelles in eukaryotic cells enhances their adaptability, allowing them to perform a diverse array of metabolic and biochemical reactions, maintain internal homeostasis, respond to environmental changes, and adapt to new conditions. This adaptability is particularly pertinent in coastal upwelling areas, where environmental conditions fluctuate significantly[124].

In examining coastal regions such as Peru and Chile, distinct environmental conditions emerged during upwelling events. In Peru, unfavourable conditions for species diversity were observed, characterized by decreased pH (7.87), a high N/Si ratio (> 4), and a shallow MLD (< 15 m). These conditions limit light availability in the water column and constrain cell growth, particularly for organisms with chloroplasts. Conversely, coastal areas of Chile exhibited conditions conducive to cell growth, including nutrient availability (N/Si = 1–4), higher pH (> 7.97), and a deeper MLD (40 m), resulting in increased species richness of unicellular eukaryotes (e.g. Hotspots, see Fig. 5). These findings underscore the importance of considering local environmental dynamics in understanding and predicting biological diversity patterns.

Therefore, this research suggests that the richness of planktonic unicellular eukaryotes is intricately linked to local and regional ecological processes related to productivity, energy dynamics and ecological limits, with specific environmental factors playing pivotal roles in shaping species diversity. Our study contributes to our

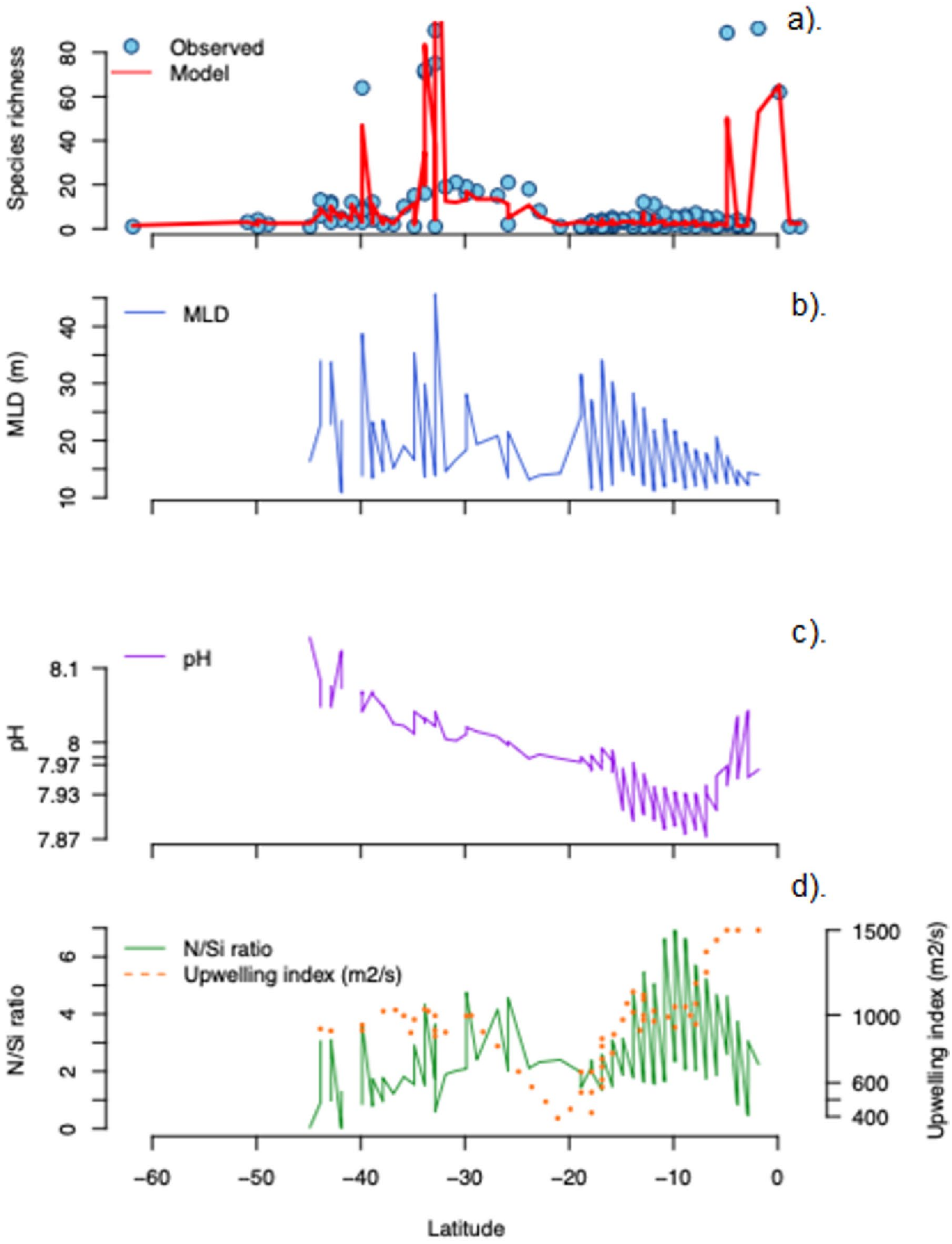

Fig. 7. Survey of unicellular eukaryote richness and environmental variables used as predictors in Species Distribution Models (SDMs) along latitudinal gradients in the coastal region of the Southeast Pacific Ocean. (**a**). Observed vs. Predicted Richness, (**b**). Mixed Layer Depth (MLD), (**c**) pH, and (**d**) N/Si Ratio and Upwelling Index (m^2/s).

broader understanding of marine ecosystem dynamics and biodiversity conservation efforts by elucidating the nuanced relationships between environmental variables and biological responses.

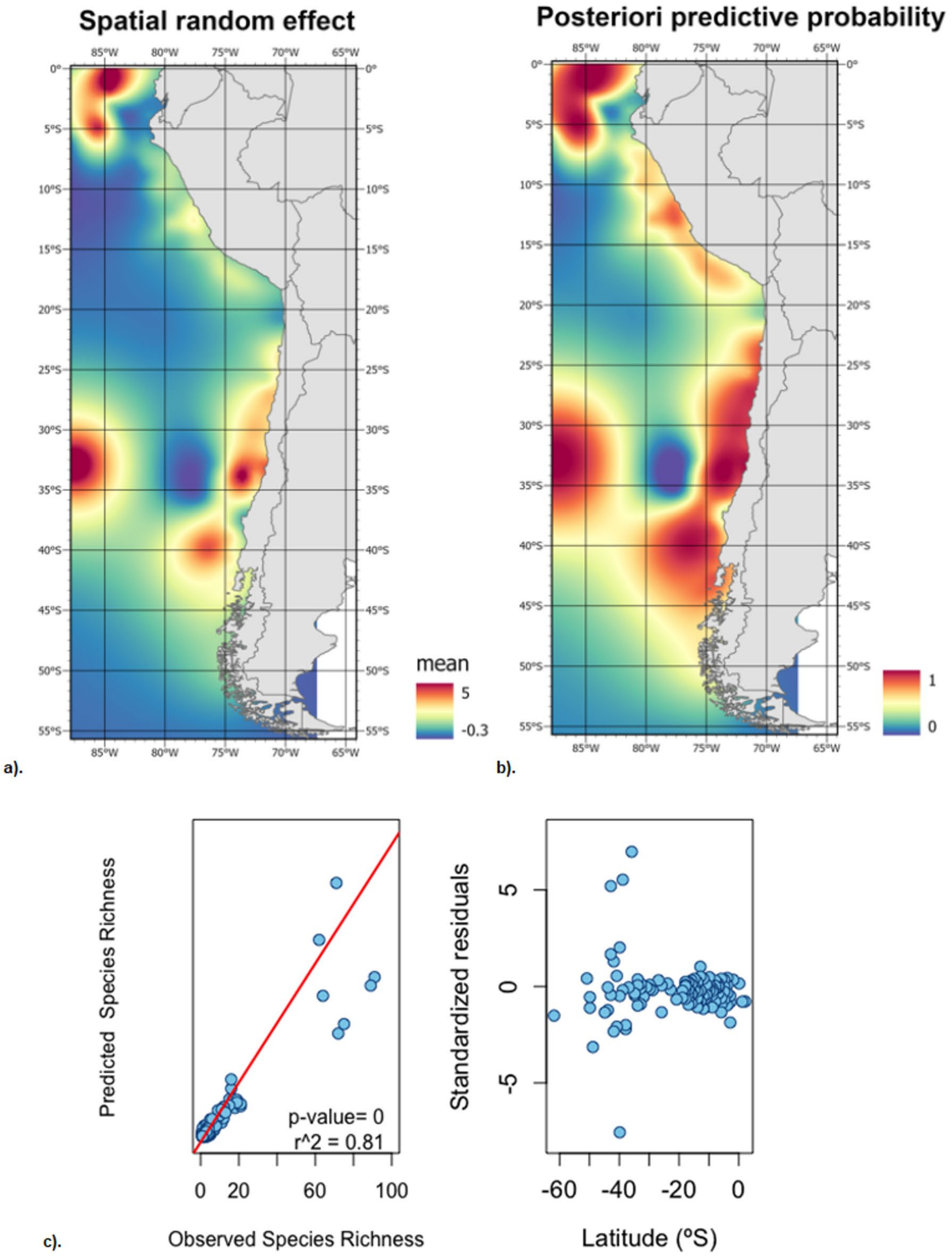

Fig. 8. Maps showing (**a**) the mean posterior distribution of the spatial random effect of the INLA model, (**b**) the mean posterior predictive probability of planktonic unicellular eukaryotes richness of coastal regions in the Southeast Pacific Ocean, and (**c**) the Pearson correlation between the observed species richness and the model predicted species richness.

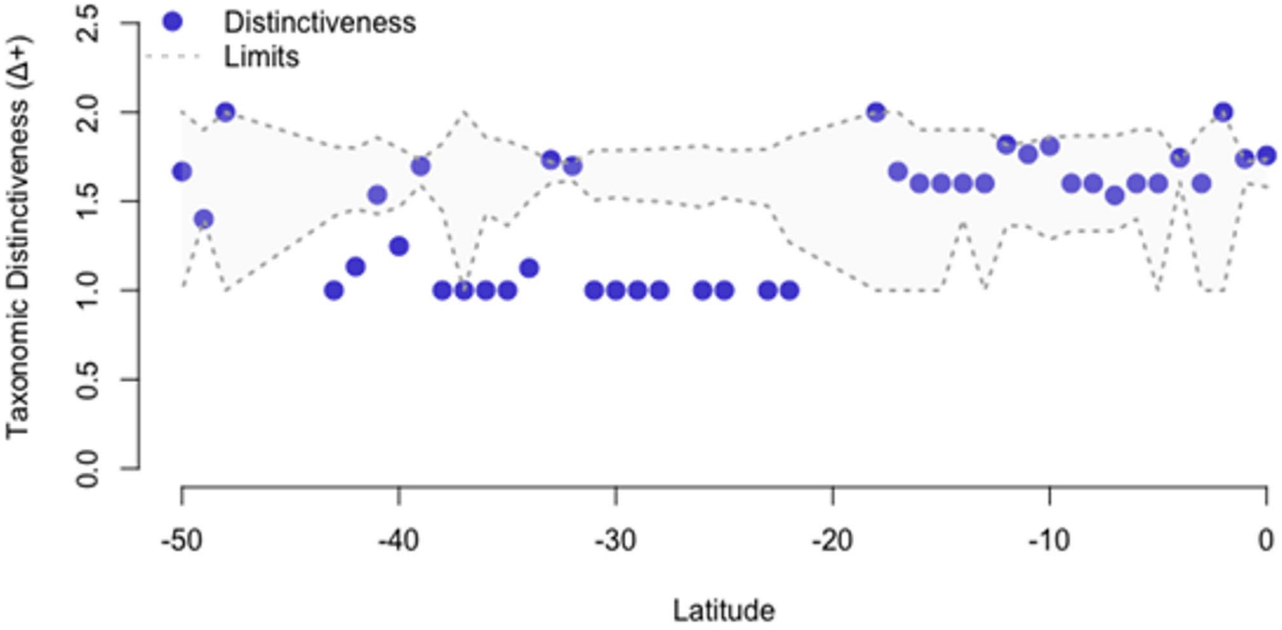

Fig. 9. Taxonomic distinctiveness obtained for planktonic unicellular eukaryotes of the Southeast Pacific Ocean. The dotted lines indicate the upper and lower limits of the null model.

Historical-evolutionary processes

According to the taxonomic distinctiveness index, the values between 0° and 19°S were inside expected according to the null model (Fig. 9), indicating the formation of communities with closely related species, where the signal of the phylogenetic novelties was attenuated or suggesting that historical-evolutionary processes are weak and contemporary ecological factors primarily shape diversity. In this area, the warm currents of equatorial origin enter, generating a transition zone represented by the interaction of subtropical and subantarctic origin species, reducing the latitudinal environmental gradient, and the community structure does not change considerably[110,111]. Furthermore, this area shows higher nestedness values (between 0.37 and 0.48), reflecting rare species as a subset of richness areas[37,136] and local extinction dynamic events modulating the species richness in this area. But, at the same time, these high values of nestedness in ecological systems can mean that a community is more stable and can coexist in a wider range of environmental conditions[137]. What could be caused by an environmental filter, where only a few species present wide ranges of environmental tolerance to the usual low oxygen conditions (<22 μmol kg^{-1}) in the northern Humboldt Current System between 4°S and 18°S, with anoxic episodes due to it being one of the most intense and shallow Oxygen Minimum Zone (OMZ)[138,139]. This is located in the poorly ventilated zone of the Tropical Southeastern Pacific, where OMZ waters have remained for a long time, and remineralising a high concentration of organic matter produces high oxygen consumption[140].

While south of 20°S to 44°S, we observed taxonomic distinctiveness lower than expected by the null model, pointing to a community with a loss of phylogenetic diversity or suggesting a narrower range of taxonomic groups likely due to niche conservatism and environmental filtering. The main taxonomic groups recorded belong to the same phylum (Foraminifera), whereas the other groups were less represented at these latitudes. Sampling bias probably affected these results or reflected a lack of records in the databases extracted for this study. However, the increase in turnover between 30°S and 44°S (from 0.48 to 0.78) indicates colonization dynamics of foraminifera species favored by the upwelling conditions. Where food is widely available, and resources are not limited, suggesting an attenuation of environmental barriers[12,141,142]. This could reflect low allopatric speciation with low extinction rates, corroborated by decreased nesting (between 0.17 and 0.11) and high local diversity[143]. Therefore, taxonomic distinctiveness did not detect the phylogenetic novelties, probably because the turnover was mainly between closely related species from similar phylogenetic clades. This allows us to recognize the importance of publishing and recording marine unicellular microorganisms' databases to improve and disentangle the biogeography of unicellular eukaryotes.

Our results show local colonization and extinction dynamics, in which local environmental fluctuations can shape these planktonic microorganisms' behaviour, ecology and distribution. Thus, the unicellular eukaryotes community in the SPO is consistent with other groups such as dinoflagellates[144], testate amoebae[145,146] and microbial taxa[147]. Whereby the ecological drivers determine the biogeography of these groups. Unlike other microorganisms such as bacteria[148,149], ciliates[143]; fungi[150], terrestrial protists[29,44,151] and marine eukaryotic cells[152,153], where historical-evolutionary processes drive their distribution.

Planktonic habitats could benefit colonization if environmental conditions are appropriate for their survival because marine currents more easily transport small organisms, and geographical barriers are less prominent, making long-distance dispersal of microorganisms more common. Historical geographic barriers in other habitats, such as soils and marine benthos, decrease dispersal and restrict distribution[47,152]. This will also depend on the morphological structures and behaviour of the microorganisms. For example, some eukaryotic cells can form cysts that help them to transport widely and support adverse conditions[143,154,155], while other species lack structures for aerial or aquatic dispersal[152,156]; others may not have latency stages or do not form resistance cysts[150,155,157].

Sampling bias

The spatial sampling effort across the study area was assessed using rarefaction/extrapolation curves, grouping cells according to the main surface current systems of the Southeast Pacific Ocean (Fig. S3). The species richness curve has passed the initial rapid discovery phase and shows a decelerating increase, indicating that the most of the species have been recorded, and therefore the sampling is adequate to describe the general macroecological patterns. However, additional sampling could reveal further diversity, and continued sampling is needed to ensure a comprehensive assessment of marine biodiversity in the study area.

Nonetheless, to address potential biases introduced by uneven sampling efforts, we applied the Integrated Nested Laplace Approximation (INLA) framework, which enables robust spatial modelling of species richness while explicitly accounting for spatial heterogeneity in sampling. Specifically, we incorporated a stochastic spatial effect using the Stochastic Partial Differential Equation (SPDE) approach[71]. This method models spatial autocorrelation in species distributions by representing the spatial process as a Gaussian Markov Random Field (GMRF) over a triangulated mesh. The mesh was designed to achieve an optimal balance between computational efficiency and spatial resolution, ensuring finer resolution in areas with higher sampling density while maintaining adequate coverage in under-sampled regions. By employing this spatially explicit approach, we accounted for the underlying spatial structure in the data, reducing bias in parameter estimation and enhancing predictions of species richness patterns. The spatial effect captures latent environmental gradients and spatial dependencies in species distributions that are not directly explained by the included environmental covariates. The predictions obtained with INLA help mitigate sampling biases and generate a more accurate and reliable spatial distribution of species richness. This method significantly improves our ability to assess marine biodiversity comprehensively and spatially explicitly, overcoming the limitations inherent in raw occurrence data and providing a more robust ecological interpretation.

Conclusion

Understanding spatial patterns of microbial biodiversity is crucial for conservation efforts, as these unicellular eukaryotes serve as a foundational food source for economically important planktonic metazoan groups, such as fish and invertebrates. The influence of historical-evolutionary processes in shaping the latitudinal diversity gradient in our model organisms appears attenuated in the SPO, with local ecological conditions predominantly driving their spatial distribution (productivity, availability of energy and ecological limits). Addressing gaps in the occurrence data across the latitudinal range in the SPO is recommended to build upon these findings and further validate the observed patterns. Complementary analyses focusing on specific taxonomic groups or clades—at the kingdom level (e.g., eukaryotic, chromists) or phylum level (e.g., Amoebozoa, Foraminifera, Radiozoa)—and habitat types (e.g. plankton, benthos, soils) could enhance our understanding of additional biogeographic processes, such as niche conservatism or historical drivers, that may also shape the distribution patterns in other microorganisms. Further research will clarify the role of speciation, extinction, and resource availability in driving microorganisms' macroecological patterns by evaluating the fundamental interrelationships among resource availability, community abundance, and species richness, as the recent equilibrium theory of biodiversity dynamics (ETBD) proposes[158].

Data availability

The supplementary material includes the validated occurrence database of taxa (Supplementary Table S1), taxonomic information (Supplementary Table S2), the values of the Taxonomic Distinctness Index (Supplementary Table S3), Beta diversity equation and its components (turnover and nestedness) estimation (Supplementary Table S4), correlation graphs (Supplementary Fig. S1), Pearson correlation matrix (Supplementary Fig. S2) and Rarefaction-extrapolation curve (Supplementary Fig. S3).

Received: 17 January 2025; Accepted: 19 May 2025

References

1. Sutherland, W. J. et al. Identification of 100 fundamental ecological questions. *J. Ecol.* **101**, 58–67 (2013).
2. Geisen, S. et al. Soil protistology rebooted: 30 fundamental questions to start with. *Soil Biol. Biochem.* **111**, 94–103 (2017).
3. Azeem Jadoon, W., Nakai, R. & Naganuma, T. Biogeographical note on Antarctic microflorae: Endemism and cosmopolitanism. *Geosci. Front.* **4**, 633–646 (2013).
4. Geisen, S. et al. Soil protists: A fertile frontier in soil biology research. *FEMS Microbiol. Rev.* **42**, 293–323 (2018).
5. Eme, L. & Tamarit, D. Microbial diversity and open questions about the deep tree of life. *Genome Biol. Evol.* **16**, evae053 (2024).
6. Xu, X. et al. Microbial macroecology: In search of mechanisms governing microbial biogeographic patterns. *Global Ecol. Biogeogr.* **29**, 1870–1886 (2020).
7. Dickey, J. R. et al. The utility of macroecological rules for microbial biogeography. *Front. Ecol. Evol.* **9**, 633155 (2021).
8. Bruni, E. P. et al. Global distribution modelling of a conspicuous Gondwanian soil protist reveals latitudinal dispersal limitation and range contraction in response to climate warming. *Divers. Distrib.* **30**, e13779 (2024).
9. Barberán, A. The microbial contribution to macroecology. *Front. Microbiol* https://doi.org/10.3389/fmicb.2014.00203 (2014).
10. Schiaffino, M. R. et al. Microbial eukaryote communities exhibit robust biogeographical patterns along a gradient of Patagonian and Antarctic lakes. *Environ. Microbiol.* **18**, 5249–5264 (2016).

11. Fernández, L. D., Hernández, C. E., Schiaffino, M. R., Izaguirre, I. & Lara, E. Geographical distance and local environmental conditions drive the genetic population structure of a freshwater microalga (Bathycoccaceae; Chlorophyta) in Patagonian lakes. *FEMS Microbiol. Ecol.* https://doi.org/10.1093/femsec/fix125 (2017).
12. Finlay, B. J. Global dispersal of free-living microbial eukaryote species. *Science* **296**, 1061–1063 (2002).
13. Fenchel, T. Biogeography for bacteria. *Science* **301**, 925–926 (2003).
14. Martiny, J. B. H. et al. Microbial biogeography: Putting microorganisms on the map. *Nat. Rev. Microbiol* **4**, 102–112 (2006).
15. Costello, M. J. & Chaudhary, C. Marine biodiversity, biogeography, deep-sea gradients, and conservation. *Curr. Biol.* **27**, R511–R527 (2017).
16. Fine, P. V. A. Ecological and evolutionary drivers of geographic variation in species diversity. *Annu. Rev. Ecol. Evol. Syst.* **46**, 369–392 (2015).
17. McClain, C. R. & Schlacher, T. A. On some hypotheses of diversity of animal life at great depths on the sea floor. *Mar. Ecol.* **36**, 849–872 (2015).
18. Cruz-Motta, J. J. et al. Latitudinal patterns of species diversity on South American rocky shores: Local processes lead to contrasting trends in regional and local species diversity. *J. Biogeogr.* **47**, 1966–1979 (2020).
19. Huggett, R. J. *Fundamentals of Biogeography* (Routledge, 2004).
20. Rabosky, D. L. & Hurlbert, A. H. Species richness at continental scales is dominated by ecological limits. *Am. Nat.* **185**, 572–583 (2015).
21. Bridle, J. & Hoffmann, A. Understanding the biology of species' ranges: When and how does evolution change the rules of ecological engagement?. *Phil. Trans. R. Soc. B* **377**, 20210027 (2022).
22. Rohde, K. Latitudinal gradients in species diversity: The search for the primary cause. *Oikos* **65**, 514 (1992).
23. Willig, M. R., Kaufman, D. M. & Stevens, R. D. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annu. Rev. Ecol. Evol. Syst.* **34**, 273–309 (2003).
24. Davies, T. J., Savolainen, V., Chase, M. W., Moat, J. & Barraclough, T. G. Environmental energy and evolutionary rates in flowering plants. *Proc. R. Soc. Lond. B* **271**, 2195–2200 (2004).
25. Evans, K. L., Warren, P. H. & Gaston, K. J. Species–energy relationships at the macroecological scale: A review of the mechanisms. *Biol. Rev.* **80**, 1–25 (2005).
26. Mittelbach, G. G. et al. Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. *Ecol. Lett.* **10**, 315–331 (2007).
27. Hurlbert, A. H. & Stegen, J. C. On the processes generating latitudinal richness gradients: identifying diagnostic patterns and predictions. *Front. Genet* https://doi.org/10.3389/fgene.2014.00420 (2014).
28. Rabosky, D. L., Title, P. O. & Huang, H. Minimal effects of latitude on present-day speciation rates in New World birds. *Proc. R. Soc. B.* **282**, 20142889 (2015).
29. Fernández, L. D. et al. Water–energy balance, past ecological perturbations and evolutionary constraints shape the latitudinal diversity gradient of soil testate amoebae in south-western South America. *Global Ecol. Biogeogr.* **25**, 1216–1227 (2016).
30. Jablonski, D., Roy, K. & Valentine, J. W. Out of the tropics: Evolutionary dynamics of the latitudinal diversity gradient. *Science* **314**, 102–106 (2006).
31. *Regionalización biogeográfica en Iberoamérica y tópicos afines: primeras jornadas biogeográficas de la Red Iberoamericana de biogeografia y entomología sistemática (RIBES XII.I-CYTED).* (Fac. des Ciencias, UNAM [u.a.], México, 2005).
32. Wiens, J. J. et al. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.* **13**, 1310–1324 (2010).
33. Li, F., Shao, L. & Li, S. Tropical niche conservatism explains the eocene migration from india to Southeast Asia in ochyroceratid spiders. *Syst. Biol.* **69**, 987–998 (2020).
34. Brown, J. H. Why are there so many species in the tropics?. *J. Biogeogr.* **41**, 8–22 (2014).
35. Atmar, W. & Patterson, B. D. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia* **96**, 373–382 (1993).
36. Wright, D. H., Patterson, B. D., Mikkelson, G. M., Cutler, A. & Atmar, W. A comparative analysis of nested subset patterns of species composition. *Oecologia* **113**, 1–20 (1997).
37. Baselga, A. Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* **19**, 134–143 (2010).
38. Beijerinck, M. W. De infusies en de ontdekking der backteriën. In *Jaarboek van de Koninklijke Akademie van Wetenschappen* (Müller, Amsterdam, 1913)
39. Baas Becking, L. B. *Geobiologie of Inleiding Tot de Milieukunde (in Dutch)* (WP Van Stockum & Zoon, 1934).
40. Holman, L. E. et al. Animals, protists and bacteria share marine biogeographic patterns. *Nat. Ecol. Evol.* **5**, 738–746 (2021).
41. Caracciolo, M. et al. Seasonal dynamics of marine protist communities in tidally mixed coastal waters. *Mol. Ecol.* **31**, 3761–3783 (2022).
42. Cabrerizo, M. J., Medina-Sánchez, J. M., González-Olalla, J. M., Sánchez-Gómez, D. & Carrillo, P. Microbial plankton responses to multiple environmental drivers in marine ecosystems with different phosphorus limitation degrees. *Sci. Total Environ.* **816**, 151491 (2022).
43. Haraguchi, L., Jakobsen, H. H., Lundholm, N. & Carstensen, J. Phytoplankton community dynamic: A driver for ciliate trophic strategies. *Front. Mar. Sci.* **5**, 272 (2018).
44. Fernández, L. D. et al. Niche conservatism drives the elevational diversity gradient in major groups of free-living soil unicellular eukaryotes. *Microb. Ecol.* **83**, 459–469 (2022).
45. Edgcomb, V. Marine protist associations and environmental impacts across trophic levels in the twilight zone and below. *Curr. Opin. Microbiol.* **31**, 169–175 (2016).
46. Chen, W., Pan, Y., Yu, L., Yang, J. & Zhang, W. Patterns and processes in marine microeukaryotic community biogeography from Xiamen coastal waters and intertidal sediments, Southeast China. *Front. Microbiol.* **8**, 1912 (2017).
47. Papke, R. T. & Ward, D. M. The importance of physical isolation to microbial diversification. *FEMS Microbiol. Ecol.* **48**, 293–303 (2004).
48. Vincent, W. F. Evolutionary origins of Antarctic microbiota: Invasion, selection and endemism. *Antart. Sci.* **12**, 374–385 (2000).
49. Danovaro, R. Understanding marine biodiversity patterns and drivers: The fall of Icarus. *Mar. Ecol.* **5**, e12814. https://doi.org/10.1111/maec.12814 (2024).
50. Costello, M. J. et al. A census of marine biodiversity knowledge, resources, and future challenges. *PLoS ONE* **5**, e12110 (2010).
51. Fiedler, P. C., Philbrick, V. & Chavez, F. P. Oceanic upwelling and productivity in the eastern tropical Pacific. *Limnol. Oceanogr.* **36**, 1834–1850 (1991).
52. Daneri, G. et al. Primary production and community respiration in the Humboldt Current System off Chile and associated oceanic areas. *Mar. Ecol. Prog. Ser.* **197**, 41–49 (2000).
53. Thiel, M. et al. The humboldt current system of northern and central Chile: Oceanographic processes, ecological interactions and socioeconomic feedback. In *Oceanography and Marine Biology* Vol. 20074975 (eds Gibson, R. et al.) 195–344 (CRC Press, 2007).
54. Hidalgo, P. & Escribano, R. Coupling of life cycles of the copepods *Calanus chilensis* and *Centropages brachiatus* to upwelling induced variability in the central-southern region of Chile. *Prog. Oceanogr.* **75**, 501–517 (2007).
55. Escribano, R., Hidalgo, P., Fuentes, M. & Donoso, K. Zooplankton time series in the coastal zone off Chile: Variation in upwelling and responses of the copepod community. *Prog. Oceanogr.* **97–100**, 174–186 (2012).

56. Escribano, R. & Schneider, W. The structure and functioning of the coastal upwelling system off central/southern Chile. *Prog. Oceanogr.* **75**, 343–347 (2007).
57. Strub, P. T., James, C., Montecino, V., Rutllant, J. A. & Blanco, J. L. Ocean circulation along the southern Chile transition region (38°–46°S): Mean, seasonal and interannual variability, with a focus on 2014–2016. *Prog. Oceanogr.* **172**, 159–198 (2019).
58. Kämpf, J. & Chapman, P. Seasonal wind-driven coastal upwelling systems. In *Upwelling Systems of the World* 315–361 (Springer International Publishing, 2016).
59. Iriarte, J. L., González, H. E. & Nahuelhual, L. Patagonian fjord ecosystems in southern Chile as a highly vulnerable region: Problems and needs. *Ambio* **39**, 463–466 (2010).
60. Corredor-Acosta, A. et al. Spatio-temporal variability of chlorophyll-A and environmental variables in the panama bight. *Remote Sens.* **12**, 2150 (2020).
61. Zuur, A. F., Ieno, E. N. & Elphick, C. S. A protocol for data exploration to avoid common statistical problems: *Data exploration. Methods Ecol. Evol.* **1**, 3–14 (2010).
62. Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K. & Toxopeus, A. G. Where is positional uncertainty a problem for species distribution modelling?. *Ecography* **37**, 191–203 (2014).
63. Williams, K. J., Belbin, L., Austin, M. P., Stein, J. L. & Ferrier, S. Which environmental variables should I use in my biodiversity model?. *Int. J. Geogr. Inf. Sci.* **26**, 2009–2047 (2012).
64. Jeng, C. C. Why a variance inflation factor of 10 is not an ideal cutoff for multicollinearity diagnostics. *J. Educ. Stud.* **57**, 067–092 (2023).
65. Van Proosdij, A. S. J., Sosef, M. S. M., Wieringa, J. J. & Raes, N. Minimum required number of specimen records to develop accurate species distribution models. *Ecography* **39**, 542–552 (2016).
66. MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G. & Franklin, A. B. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* **84**, 2200–2207 (2003).
67. Baselga, A. & Orme, C. D. L. betapart: An R package for the study of beta diversity. *Methods Ecol. Evol.* **3**, 808–812 (2012).
68. Melo, A. S., Cianciaruso, M. V. & Almeida-Neto, M. tree NODF: Nestedness to phylogenetic, functional and other tree-based diversity metrics. *Methods Ecol. Evol.* **5**, 563–572 (2014).
69. Swingland, I. R. Biodiversity, Definition of. In *Encyclopedia of Biodiversity* 399–410 (Elsevier, 2013). https://doi.org/10.1016/B978-0-12-384719-5.00009-5.
70. Getis, A. & Ord, J. K. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* **24**, 189–206 (1992).
71. Hsieh, T. C., Ma, K. H. & Chao, A. iNEXT: An R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods Ecol. Evol.* **7**, 1451–1456 (2016).
72. Dornelas, M. et al. BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecol. Biogeogr.* **27**, 760–786 (2018).
73. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian Inference for Latent Gaussian models by using Integrated Nested Laplace Approximations. *J. R. Stat. Soc. Ser. B Stat Methodol.* **71**, 319–392 (2009).
74. Lindgren, F., Rue, H. & Lindström, J. An explicit link between gaussian fields and gaussian markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B Stat Methodol.* **73**, 423–498 (2011).
75. Lindgren, F. & Rue, H. Bayesian spatial modelling with R—INLA. *J. Stat. Soft.* https://doi.org/10.18637/jss.v063.i19 (2015).
76. Poggio, L., Gimona, A., Spezia, L. & Brewer, M. J. Bayesian spatial modelling of soil properties and their uncertainty: The example of soil organic matter in Scotland using R-INLA. *Geoderma* **277**, 69–82 (2016).
77. Muñoz, F., Pennino, M. G., Conesa, D., López-Quílez, A. & Bellido, J. M. Estimation and prediction of the spatial occurrence of fish species using Bayesian latent Gaussian models. *Stoch. Environ. Res. Risk Assess* **27**, 1171–1180 (2013).
78. Dell'Apa, A., Maria Grazia, P. & Bonzek, C. Modeling the habitat distribution of spiny dogfish (*Squalus acanthias*), by sex, in coastal waters of the northeastern United States. *Fish. Bull.* **115**, 89–100 (2016).
79. Gelman, A., Hwang, J. & Vehtari, A. Understanding predictive information criteria for Bayesian models. *Stat Comput* **24**, 997–1016 (2014).
80. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *14th International Joint Conference on Artificial Intelligence (IJCAI)*, vol. 2, 1137–1143 (1995).
81. Elith, J. & Leathwick, J. R. Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* **40**, 677–697 (2009).
82. Pennino, M. G., Muñoz, F., Conesa, D., López-Quílez, A. & Bellido, J. M. Modeling sensitive elasmobranch habitats. *J. Sea Res.* **83**, 209–218 (2013).
83. Lezama-Ochoa, N. et al. Biodiversity and habitat characteristics of the bycatch assemblages in fish aggregating devices (FADs) and school sets in the Eastern Pacific Ocean. *Front. Mar. Sci.* **4**, 265 (2017).
84. Escalle, L. et al. Environmental factors and megafauna spatio-temporal co-occurrence with purse-seine fisheries. *Fish. Oceanogr.* **25**, 433–447 (2016).
85. Goldenberg, S. U. et al. Nutrient composition (Si:N) as driver of plankton communities during artificial upwelling. *Front. Mar. Sci.* **9**, 1015188 (2022).
86. Baumann, M. et al. Counteracting effects of nutrient composition (Si:N) on export flux under artificial upwelling. *Front. Mar. Sci.* **10**, 1181351 (2023).
87. Bode, A., Alvarez-Ossorio, M. T., Cabanas, J. M., Miranda, A. & Varela, M. Recent trends in plankton and upwelling intensity off Galicia (NW Spain). *Prog. Oceanogr.* **83**, 342–350 (2009).
88. Gonzalez-Nuevo, G., Gago, J. & Cabanas, J. M. Upwelling index: a powerful tool for marine research in the NW Iberian upwelling system. *J Op. Oceanogr.* **7**, 47–57 (2014).
89. Guo, X., Wu, L. & Huang, L. Spatiotemporal patterns in diversity and assembly process of marine protist communities of the Changjiang (Yangtze river) plume and its adjacent waters. *Front. Microbiol.* **11**, 579290 (2020).
90. Balzano, S., Abs, E. & Leterme, S. Protist diversity along a salinity gradient in a coastal lagoon. *Aquat. Microb. Ecol.* **74**, 263–277 (2015).
91. Li, S. et al. Ecological and evolutionary processes involved in shaping microbial habitat generalists and specialists in urban park ecosystems. *Msystems* **9**, e00469-e524 (2024).
92. Omand, M. M. et al. Eddy-driven subduction exports particulate organic carbon from the spring bloom. *Science* **348**, 222–225 (2015).
93. Freilich, M. A. et al. 3D intrusions transport active surface microbial assemblages to the dark ocean. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2319937121 (2024).
94. Moreles, E., Romero, E., Ramos-Musalem, K. & Tenorio-Fernandez, L. The global ocean mixed layer depth derived from an energy approach. EGUsphere https://doi.org/10.5194/egusphere-2024-4079 (2025).
95. Gill, A. E. *Atmosphere-Ocean Dynamics* (Academic Press, 1982).
96. Jetz, W. & Fine, P. V. A. Global gradients in vertebrate diversity predicted by historical area-productivity dynamics and contemporary environment. *PLoS Biol* **10**, e1001292 (2012).
97. Coogan, J., Dzwonkowski, B. & Lehrter, J. Effects of coastal upwelling and downwelling on hydrographic variability and dissolved oxygen in mobile bay. *JGR Oceans* **124**, 791–806 (2019).
98. Currie, D. J. Energy and large-scale patterns of animal- and plant-species richness. *Am. Nat.* **137**, 27–49 (1991).
99. Clarke, K. R. & Warwick, R. M. A taxonomic distinctness index and its statistical properties. *J. Appl. Ecol.* **35**, 523–531 (1998).

100. Rivadeneira, M. M., Thiel, M., González, E. R. & Haye, P. A. An inverse latitudinal gradient of diversity of peracarid crustaceans along the Pacific Coast of South America: Out of the deep south: An inverse gradient of latitudinal diversity. *Glob. Ecol. Biogeogr.* **20**, 437–448 (2011).
101. Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M. & Inouye, B. D. Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* **2**, art24 (2011).
102. Mattheeussen, R. et al. Habitat selection of aquatic testate amoebae communities on Qeqertarsuaq (Disko Island), West Greenland. *Acta Protozool* **44**(3), 253 (2005).
103. Sigala Regalado, I., Lozano García, S., Pérez Alvarado, L., Caballero, M. & Lugo Vázquez, A. Ecological drivers of testate amoeba diversity in tropical water bodies of central Mexico. *J. Limnol.* https://doi.org/10.4081/jlimnol.2018.1699 (2018).
104. Brandhorst, W. Condiciones oceanográficas estivales frente a la costa de Chile. *Rev. Biol. Mar.* **14**, 45–84 (1971).
105. Fonseca, T. & Farías, M. Estudio del proceso de surgencia en la costa chilena utilizando percepción remota. *Investigaciones Pesqueras* **34**, 33–46 (1987).
106. Strub, P. T., Mesias, J. M., Montecino, V., Ruttlant, J. & Salinas, S. Coastal ocean circulation off Western South America. In *The Global Coastal Ocean, Regional Studies and Syntheses.* 273–315 (Wiley, 1998).
107. Mohtadi, M., Hebbeln, D. & Marchant, M. Upwelling and productivity along the Peru-Chile Current derived from faunal and isotopic compositions of planktic foraminifera in surface sediments. *Mar. Geol.* **216**, 107–126 (2005).
108. Hansen, A., Ohde, T. & Wasmund, N. Succession of micro- and nanoplankton groups in ageing upwelled waters off Namibia. *J. Mar. Syst.* **140**, 130–137 (2014).
109. Bohata, K. *Microzooplankton of the northern Benguela upwelling system* (Doctoral dissertation, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky, 2015).
110. Brattström, H. & Johanssen, A. Ecological and regional zoogeography of the marine benthic fauna of Chile: Report no. 49 of the Lund University Chile Expedition 1948–49. *Sarsia* **68**, 289–339 (1983).
111. Camus, P. A. Biogeografía marina de Chile continental. *Rev. Chil. Hist. Nat.* https://doi.org/10.4067/S0716-078X2001000300008 (2001).
112. Saeedi, H., Dennis, T. E. & Costello, M. J. Bimodal latitudinal species richness and high endemicity of razor clams (Mollusca). *J. Biogeogr.* **44**, 592–604 (2017).
113. Rivadeneira, M. M. & Poore, G. C. B. Latitudinal gradient of diversity of marine crustaceans: TOWARDS a synthesis. In *Evolution and Biogeography* (eds Thiel, M. & Poore, G.) 389–412 (Oxford University Press, 2020). https://doi.org/10.1093/oso/9780190637842.003.0015.
114. Rivera, R., Escribano, R., González, C. E. & Pérez-Aragón, M. Latitudinal diversity of planktonic copepods in the Eastern Pacific: Overcoming sampling biases and predicting patterns. *Front. Ecol. Evol.* **12**, 1305916 (2024).
115. Hillebrand, H. On the generality of the latitudinal diversity gradient. *Am. Nat.* **163**, 192–211 (2004).
116. Tittensor, D. P. et al. Global patterns and predictors of marine biodiversity across taxa. *Nature* **466**, 1098–1101 (2010).
117. Escribano, R., Hidalgo, P. & Krautz, C. Zooplankton associated with the oxygen minimum zone system in the northern upwelling region of Chile during March 2000. *Deep Sea Res. Part II* **56**, 1083–1094 (2009).
118. Fernández-Álamo, M. A. & Färber-Lorda, J. Zooplankton and the oceanography of the eastern tropical Pacific: A review. *Prog. Oceanogr.* **69**, 318–359 (2006).
119. Hidalgo, P., Escribano, R. & Morales, C. E. Ontogenetic vertical distribution and diel migration of the copepod *Eucalanus inermis* in the oxygen minimum zone off northern Chile (20–21° S). *J. Plankton Res.* **27**, 519–529 (2005).
120. Fuhrman, J. A. et al. A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7774–7778 (2008).
121. Beaugrand, G., Reid, P. C., Ibañez, F., Lindley, J. A. & Edwards, M. Reorganization of North Atlantic marine copepod biodiversity and climate. *Science* **296**, 1692–1694 (2002).
122. Hernández-León, S. et al. Large deep-sea zooplankton biomass mirrors primary production in the global ocean. *Nat. Commun.* **11**, 6048 (2020).
123. Marañón, E., Cermeño, P., Latasa, M. & Tadonléké, R. D. Resource supply alone explains the variability of marine phytoplankton size structure. *Limnol. Oceanogr.* **60**, 1848–1854 (2015).
124. Gong, F. et al. Spatial shifts in size structure, phylogenetic diversity, community composition and abundance of small eukaryotic plankton in a coastal upwelling area of the northern South China Sea. *J. Plankton Res.* **42**(6), 650–667 (2020).
125. James, C. C. et al. Influence of nutrient supply on plankton microbiome biodiversity and distribution in a coastal upwelling region. *Nat. Commun.* **13**, 2448 (2022).
126. McManus, G. & Peterson, W. Bacterioplankton production in the nearshore zone during upwelling off central Chile. *Mar. Ecol. Prog. Ser.* **43**, 11–17 (1988).
127. Vargas, C., Contreras, P. & Iriarte, J. Relative importance of phototrophic, heterotrophic, and mixotrophic nanoflagellates in the microbial food web of a river-influenced coastal upwelling area. *Aquat. Microb. Ecol.* **65**, 233–248 (2012).
128. Figueiras, F. G., Arbones, B., Castro, C. G., Froján, M. & Teixeira, I. G. About pigmented nanoflagellates and the importance of mixotrophy in a coastal upwelling system. *Front. Mar. Sci.* **7**, 144 (2020).
129. Hou, L. et al. Effects of mixed layer depth on phytoplankton biomass in a tropical marginal ocean: A multiple timescale analysis. *Earth's Future* **10**, e2020EF001842 (2022).
130. Diaz, B. P. et al. Seasonal mixed layer depth shapes phytoplankton physiology, viral production, and accumulation in the North Atlantic. *Nat Commun* **12**, 6634 (2021).
131. Boyce, D. G., Lewis, M. R. & Worm, B. Global phytoplankton decline over the past century. *Nature* **466**, 591–596 (2010).
132. Charrieau, L. M. et al. Rapid environmental responses to climate-induced hydrographic changes in the Baltic Sea entrance. *Biogeosciences* **16**, 3835–3852 (2019).
133. Orr, J. C. et al. Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* **437**, 681–686 (2005).
134. Kroeker, K. J. et al. Impacts of ocean acidification on marine organisms: Quantifying sensitivities and interaction with warming. *Glob. Change Biol.* **19**, 1884–1896 (2013).
135. Ries, J. B., Ghazaleh, M. N., Connolly, B., Westfield, I. & Castillo, K. D. Impacts of seawater saturation state (ΩA= 0.4–4.6) and temperature (10, 25 °C) on the dissolution kinetics of whole-shell biogenic carbonates. *Geochim. Cosmochim. Acta* **192**, 318–337 (2016).
136. Moreno, R. A., Rivadeneira, M. M., Hernández, C. E., Sampértegui, S. & Rozbaczylo, N. Do Rapoport's rule, the mid-domain effect or the source–sink hypotheses predict bathymetric patterns of polychaete richness on the Pacific coast of South America?. *Glob. Ecol. Biogeogr.* **17**, 415–423 (2008).
137. Baumgartner, M. T. Connectance and nestedness as stabilizing factors in response to pulse disturbances in adaptive antagonistic networks. *J. Theor. Biol.* **486**, 110073 (2020).
138. Fuenzalida, R., Schneider, W., Garcés-Vargas, J., Bravo, L. & Lange, C. Vertical and horizontal extension of the oxygen minimum zone in the eastern South Pacific Ocean. *Deep Sea Res. Part II* **56**, 992–1003 (2009).
139. Espinoza-Morriberón, D. et al. Oxygen variability during ENSO in the tropical south eastern Pacific. *Front. Mar. Sci.* **5**, 526 (2019).
140. Paulmier, A. & Ruiz-Pino, D. Oxygen minimum zones (OMZs) in the modern ocean. *Prog. Oceanogr.* **80**, 113–128 (2009).

141. Finlay, B. J., Corliss, J. O., Esteban, G. & Fenchel, T. Biodiversity at the microbial level: The number of free-living ciliates in the biosphere. *Q. Rev. Biol.* **71**, 221–237 (1996).
142. Finlay, B. J. & Clarke, K. J. Ubiquitous dispersal of microbial species. *Nature* **400**, 828–828 (1999).
143. Agatha, S. Global diversity of *Aloricate Oligotrichea* (Protista, Ciliophora, Spirotricha) in Marine and Brackish Sea Water. *PLoS ONE* **6**, e22466 (2011).
144. Taylor, F. J. R., Hoppenrath, M. & Saldarriaga, J. F. Dinoflagellate diversity and distribution. *Biodivers. Conserv.* **17**, 407–418 (2008).
145. Mitchell, E. A. D. & Meisterfeld, R. Taxonomic confusion blurs the debate on cosmopolitanism versus local endemism of free-living protists. *Protist* **156**, 263–267 (2005).
146. Foissner, W. Dispersal and biogeography of protists: Recent advances. *Jpn. J. Protozool.* **40**(1), 1–16 (2007).
147. Fondi, M. et al. "Every gene is everywhere but the environment selects": Global geolocalization of gene sharing in environmental samples through network analysis. *Genome Biol Evol* **8**, 1388–1400 (2016).
148. Whitaker, R. J., Grogan, D. W. & Taylor, J. W. Geographic Barriers Isolate Endemic
149. González-Rocha, G. et al. Diversity structure of culturable bacteria isolated from the Fildes Peninsula (King George Island, Antarctica): A phylogenetic analysis perspective. *PLoS ONE* **12**, e0179390 (2017).
150. Foissner, W. Protist diversity and distribution: Some basic considerations. *Biodivers Conserv* **17**, 235–242 (2008).
151. Foissner, W. Ubiquity and cosmopolitanism of protists questioned. *Soc. Int. Limnol. News.* **43**, 6–7 (2004).
152. Foissner, W. Biogeography and dispersal of micro-organisms: A review emphasizing protists. *Acta Protozool.* **45**, 111–136 (2006).
153. De Vargas, C. et al. Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**, 1261605 (2015).
154. Fenchel, T. & Finlay, B. J. The ubiquity of small species: Patterns of local and global diversity. *Bioscience* **54**, 777 (2004).
155. Custer, G. F., Bresciani, L. & Dini-Andreote, F. Ecological and evolutionary implications of microbial dispersal. *Front. Microbiol.* **13**, 855859 (2022).
156. Foissner, W. Protist Diversity: Estimates of the Near-Imponderable. *Protist* **150**, 363–368 (1999). Populations of Hyperthermophilic Archaea. *Science* **301**, 976–978 (2003).
157. Pearson, L. A. & Neilan, B. A. Protozoan Diversity and Biogeography. in *Encyclopedia of Life Sciences* 1–7 (Wiley, 2021).
158. Okie, J. G. & Storch, D. The equilibrium theory of biodiversity dynamics: a general framework for scaling species richness and community abundance along environmental gradients. *Am. Nat.* **205**, 20–40 (2025).

Acknowledgements

This work was supported by the Scholarship of Vicerrectoría de Investigación y Desarrollo, Universidad de Concepción. Project number 219.112.112. ANID-FONDECYT Postdoctoral Projects Number 3210538 and 3230262. L.D.F. is funded by ANID-FONDECYT Regular Project Number 1220605. C.E.H. is funded by ANID-FONDECYT Regular Projects Numbers 1240219, 1201506 and 1220998.

Author contributions

SY developed conceptualization, methodology, software, data analysis, discussion, investigation, writing—original draft, and project administration. LT developed conceptualization, methodology (validation of species occurrences, ecology and historical-evolutionary processes), data analysis, discussion, investigation, and writing—original draft. GG methodology (validation of species occurrences), data analysis, discussion, investigation, writing—original draft. LF review, editing, discussion and supervision. NE developed conceptualization and methodology. MM review and supervision. CH developed conceptualization, review, editing, discussion and supervision.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-03220-3.

Correspondence and requests for materials should be addressed to L.T. or C.E.H.

Reprints and permissions information is available at www.nature.com/reprints.