Stream diatom biodiversity in islands and continents—A global perspective on effects of area, isolation and environment

Published Version:
Jamoneau A., Soininen J., Tison-Rosebery J., Boutry S., Budnick W.R., He S., et al. (2022). Stream diatom biodiversity in islands and continents—A global perspective on effects of area, isolation and environment. JOURNAL OF BIOGEOGRAPHY, 49(12), 2156-2168 [10.1111/jbi.14482].

Availability:
This version is available at: https://hdl.handle.net/11585/904161 since: 2022-11-19

Published:
DOI: http://doi.org/10.1111/jbi.14482

Jamoneau, A., Soininen, J., Tison-Rosebery, J., Boutry, S., Budnick, W.R., He, S., Marquié, J., Jyrkänkallio-Mikkola, J., Pajunen, V., Teittinen, A., Tupola, V., Wang, B., Wang, J., Blanco, S., Borrini, A., Cantonati, M., Valente, A.C., Delgado, C., Dörflinger, G., Gonçalves, V., Kennedy, B., Marques, H., Papatheodoulou, A., Raposeiro, P.M., Ritter, C., Serafim, A., Van de Vijver, B., Passy, S.I. Stream diatom biodiversity in islands and continents—A global perspective on effects of area, isolation and environment (2022) Journal of Biogeography, 49 (12), pp. 2156-2168

Stream diatom biodiversity in islands and continents – a global perspective on effects of area, isolation and environment

Short running title: Diatom in islands and continents

Aurélien Jamoneau[1†*], Janne Soininen[2†], Juliette Tison-Rosebery[1], Sébastien Boutry[1], William R. Budnick[3], Siwen He[2,4], Julien Marquié[1], Jenny Jyrkänkallio-Mikkola[2], Virpi Pajunen[2], Anette Teittinen[2], Vilja Tupola[2], Beixin Wang[4], Jianjun Wang[5,6], Saúl Blanco[7], Alex Borrini[8], Marco Cantonati[9], Adelaide Clode Valente[10], Cristina Delgado[11], Gerald Dörflinger[12], Vítor Gonçalves[13,14], Bryan Kennedy[15], Helena Marques[13], Athina Papatheodoulou[16], Pedro Miguel Raposeiro[13], Catarina Ritter[13], António Serafim[17], Bart Van de Vijver[18,19] and Sophia I. Passy[20].

[1] INRAE, EABX, 33612 Cestas, France.

[2] Department of Geosciences and Geography, University of Helsinki, Helsinki, FIN-00014, Finland.

[3] Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI 48824, USA.

[4] Laboratory of Aquatic Insects and Stream Ecology, Department of Entomology, Nanjing Agricultural University, Nanjing, China.

[5] State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China.

[6] University of Chinese Academy of Sciences, Beijing 100049, China.

[7] Department of Biodiversity and Environment Management, University of León, Spain.

[8] ERSE soc. coop. s.t.p. (Ecological Research and Services for the Environment), Viareggio, Italy.

[9] BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum, University of Bologna, Bologna, Italy.

[10] Secretaria Regional de Ambiente, Recursos Naturais e Alterações Climáticas - Governo Regional da Madeira, Funchal, Portugal.

[11] Department of Ecology and Animal Biology, Faculty of Science, University of Vigo, 36200 Vigo, Pontevedra, Spain.

[12] Water Development Department, 1646 Nicosia, Cyprus

[13] Centro de Investigação em Biodiversidade e Recursos Genéticos, Rede de Investigação em Biodiversidade e Biologia Evolutiva – Laboratório Associado, 9500-321 Ponta Delgada, Portugal.

[14] Faculty of Sciences and Technology, University of the Azores, 9500-321 Ponta Delgada, Portugal.

[15] Environmental Protection Agency, Castlebar, Co. Mayo, Ireland F23KT91.

[16] I.A.CO Environmental & Water Consultants Ltd, Nicosia 2035, Cyprus.

[17] AS.AMBI, Consultoria Ambiental, 7005-480 Évora, Portugal.

[18] Meise Botanic Garden, Meise, Belgium.

[19] Department of Biology, ECOSPHERE, University of Antwerp, Wilrijk, Belgium

[20] Department of Biology, University of Texas at Arlington, Arlington, TX 76019, USA

† equal contribution

* Corresponding author: Aurélien Jamoneau

Email: aurelien.jamoneau@inrae.fr

Acknowledgments

We thank OFB (Office Francais Biodiversité), all French Water Agencies and Asconit for data contribution and financial support and Anna Astorga and Riku Paavola for collecting the diatom data in New Zealand. We thank Hilaire Drouineau and Michel Coste for statistical and taxonomic discussion respectively. We thank colleagues at the Laboratory of Aquatic Insects and Stream Ecology of Nanjing Agricultural University for assistance with diatom sampling and processing in China. We thank Feiyan Pan, Han Xiao and Qing Liu for the field trip in Iceland. JJ. Wang was supported by National Natural Science Foundation of China (91851117, 41871048), CAS Key Research Program of Frontier Sciences (QYZDB-SSW-DQC043), and The National Key Research and Development Program of China (2019YFA0607100). We greatly thanks the Water Development Department of the Republic of Cyprus for providing data from Cyprus. We are grateful to RAS - Direzione generale Agenzia regionale del Distretto Idrografico della Sardegna

Servizio tutela e gestione delle risorse idriche, vigilanza sui servizi idrici e gestione delle siccità (STGRI) and to ARPAS - Agenzia Regionale per la Protezione dell’Ambiente della Sardegna for making available physical and chemical data collected during the institutional monitoring of surface waters for the streams in Sardinia. We thank the EPA (Environmental Protection Agency) for providing data from Ireland.

Abstract

Aim. The species-area relationship (SAR) is one of the most distinctive biogeographic patterns, but global comparisons of the SARs between island and mainland are lacking for microbial taxa. Here, we explore whether the form of the SAR and the drivers of species richness, including area, environmental heterogeneity, climate and physicochemistry, differ between islands and similarly sized areas on mainland, referred to as continental area equivalents (CAEs).

Location. Global.

Major taxa studied. Stream benthic diatoms.

Methods. We generated CAEs on six continental datasets and examined the SARs of CAEs and islands (ISAR). Then, we compared CAEs and islands in terms of total richness and richness of different ecological guilds. We tested the factors contributing to richness in islands and CAEs with regressions. We used structural equation models to determine the effects of area vs. environmental heterogeneity, climate and local conditions on species richness.

Results. We found a non-significant ISAR, but a significant positive SAR in CAEs. Richness in islands was related to productivity. Richness in CAEs was mainly dependent on area and climate, but not directly on environmental heterogeneity. Species richness within guilds exhibited inconsistent relationships with island isolation and area.

Main conclusions. Ecological and evolutionary processes shaping diatom island biogeography do not depend on area at the worldwide scale probably due to the presence of distinct species pool across islands. Conversely, area was an important driver of diatom richness in continents, and this effect could be attributed to dispersal. Continents had greater richness than islands, but this was a consequence of differences in environmental conditions such as specific island climatic conditions. We stress the need for more island data on benthic diatoms, particularly from archipelagos, to better understand the biogeography of this most speciose group of algae.

Keywords

ecological guilds, freshwater diatoms, island biogeography, macroecology, species-area relationship, streams

Main Text

Introduction

A fundamental ecological law that describes how the number of species increases with area is the species-area relationship (SAR, Arrhenius, 1921). The SAR belongs to a few, truly robust generalizations in ecology detected in a wide range of ecosystems and taxa (Connor & McCoy, 1979; Rosenzweig, 1995; Drakare et al., 2006). Islands represent perhaps the most straightforward study setting to explore the SAR because of their well-defined area. Unlike most mainland habitat patches, islands are surrounded by an inhospitable matrix for continental taxa, which cannot be colonized and, consequently, cannot serve as a source of immigrants. This peculiar feature of islands inspired MacArthur & Wilson to develop the theory of island biogeography (MacArthur & Wilson, 1967), which has contributed enormously to modern biodiversity theory (Chase & Leibold, 2003), metapopulation biology (Hanski & Gaggiotti, 2004), community ecology (Mittelbach & McGill, 2019), landscape ecology (Farina, 2008) and biodiversity conservation (Prugh et al., 2008).

Island biogeography investigates how species richness on islands varies spatially and through time (Whittaker & Fernandez-Palacios, 2007). It postulates that larger and less isolated islands host more species than small and remote islands because larger area decreases extinction and proximity to mainland increases immigration. Larger islands may also encompass more species because they provide a larger target for immigration, higher habitat diversity (Lack, 1976) and have higher speciation rates (Whittaker & Fernandez-Palacios, 2007). Lastly, since island age affects diversification and erosion, it may also determine species richness, which tends to be the highest in islands of intermediate age according to the general dynamic model (Whittaker et al., 2008, 2017).

In the light of this knowledge, Chase et al. (2019) recently presented a framework for the ecological mechanisms underlying the island SAR (ISAR). They suggested that passive sampling (i.e. larger islands passively sample more individuals and species from the regional pool than smaller islands), disproportionate effects (e.g. different colonization and extinction rates in larger vs. smaller islands) and habitat heterogeneity (greater in large islands) would be the main drivers of ISAR. Nevertheless, the major patterns and drivers of island vs. mainland SAR are still poorly understood, particularly for the species rich microorganisms.

Given the importance of environmental heterogeneity and dispersal on the SAR (Chase et al., 2019), functional groups varying in resource utilization and dispersal can have different SARs (Lomolino & Weiser, 2001; Báldi, 2008; Schrader et al., 2020). For example, the SAR slope was steeper for specialist than for generalist bird species (Matthews et al., 2014) and functional traits related to dispersal explained the SAR variation in plant communities (Schrader et al., 2020). Thus, evaluating the SAR of different ecological guilds may improve the knowledge of the niche- vs. dispersal-related processes behind the SAR patterns. As functional diversity may have a distinct (Jamoneau et al., 2018; Schrader et al., 2020) and even stronger response to environmental variation than species diversity (Krause et al., 2014; Abonyi et al., 2018), the SAR for different functional groups may elucidate how community assembly processes operate through space and time (Tilman et al., 1997).

The ISAR has been tested with larger-bodied organisms, including terrestrial arthropods (Simberloff & Wilson, 1969) and reptiles (Algar & Losos, 2011), birds (Kalmar & Currie, 2006, 2007), vascular plants (Kreft et al., 2008), and fish (Sandin et al., 2008). However, ISAR patterns are still poorly understood for microorganisms. Earlier microbial field studies that used microcosms (Smith et al., 2005), lakes (Reche et al., 2005), trees (Bell et al., 2005; Peay et al., 2007) or spring ecosystems (Teittinen & Soininen, 2015) as surrogates of islands, reported significantly positive ISARs in almost all systems (but see Teittinen & Soininen, 2015). However, investigations on microorganismal diversity in real islands at a global scale are, to our knowledge, still missing.

Rosenzweig (1995) hypothesized that islands should have lower local and regional species richness than similarly sized continental regions due to isolation (lower mass- and rescue effect), but steeper SAR slopes. This is because area tends to be a more critical factor for biota on islands than on continents due to its stronger effects on extinction and colonization (Kreft et al., 2008). However, in an extensive meta-analysis, Drakare et al. (2006) did not find evidence for steeper SARs on islands (ISARs) than on mainland across multiple species groups. The SAR patterns are typically explored within archipelagos due to the presence of a common species pool, allowing assessment of the pure area effect. However, there are also more general models for the SAR at the global scale, searching for broader influences on the SAR (Kalmar & Currie, 2006; Kreft et al., 2008; Triantis et al., 2015), including differences in evolutionary history (Rosenzweig, 1995).

Here, we adopted a similar perspective and investigated freshwater diatom SAR at a worldwide scale, given that diatoms have large distributions (Finlay, 2002) and are strongly controlled by environmental conditions (Soininen et al., 2016). We compared SARs, total species richness, and species richness drivers between islands and corresponding areas on five continents, referred to as continental area equivalents (CAEs). For this comparison, we devised a novel method based on island-mainland pairs (Fig. 1), assuming that terrestrial area is a good surrogate for area of freshwater habitat (see Appendix S1 in Supporting Information). The CAEs corresponded to the sampling area of 18 islands. We then examined (1) if SAR slopes differed between islands (ISAR) and continents, (2) whether islands showed overall lower diatom species richness than CAEs, (3) if species richness of island was related to environment, spatial isolation or island age, and (4) whether habitat diversity, passive sampling or disproportionate effects explained the SAR. We investigated these research questions separately for total diatom species richness and species richness of ecological guilds, differing in dispersal capacity and tolerance to nutrient limitation and disturbance (Passy, 2007, 2016), all expected to influence the SAR (Matthews et al., 2014; Schrader et al., 2020).

Materials and Methods

Biological and environmental datasets

In total, we included 18 island datasets (Corsica, Cyprus, Guadeloupe, Iceland, Ireland, Kauai, La Réunion, Martinique, Madeira, Majorca, Mayotte, New Caledonia, North New Zealand, Oahu, Possession, São Miguel, Sardinia and South New Zealand) and six continental datasets (China, Finland, France, French Guiana, Kenya and USA) in our study (see Appendix S2). Diatoms were sampled from hard substrates (typically stones) or macrophytes, generally during the low flow period (see Appendix S2 for details). Although diatoms in some datasets were collected over several years, we did not expect a substantial effect of interannual variation in our study, because we were interested in regional diversity patterns and included environmental variables to account for this potential variation.

Diatoms were cleaned with acid or hydrogen peroxide. A total of 400-700 diatom valves were counted for each sampling site, which is sufficient for reliable estimates of total diversity (Heino & Soininen, 2005). As the number of counted valves varied somewhat among the samples, we studied if this would affect our richness estimates. We estimated species richness with 300 valves and tested the correlation with the observed species richness. We observed a very strong relationship between the estimated and the observed species richness (R^2_{aj} = 0.98). Also, valve counts did not differ significantly between islands and continents (Cliff test difference for large dataset, delta=-0.15). We thus believe that the number of counted valves has only marginal impact on our richness results.

Diatoms were generally identified up to species level, except in some rare case where some of the valves were identified only to genus level (representing less than 5% of the entire dataset). Homogenization of the taxonomy among regions was performed using the OMNIDIA database (Lecointe et al., 1993, updated in November 2020). To ensure that we have a proper estimate of the diversity, we i) evaluated the proportion of observed species compared to the size of the species-pool in each region using basic Chao equation (Chao, 1987) and calculated a 'corrected' species richness according to this ratio (i.e. the observed species richness was increased relative to the proportion of missing species estimated from the species pool) and ii) calculated a genus-based richness assuming that genus level identification varies much less among diatomists than species identification. We then ran analyses with observed species richness, corrected species richness and genus richness (see Data analyses section). In total, our datasets comprised 1967 taxa, further classified into four ecological guilds: low profile (species of short stature), high profile (species of tall stature, typically filamentous, colonial or branched), motile (species moving freely in the biofilm) and planktonic species (species not innate to the benthos but originating from planktonic sedimentation) (Passy, 2007; Rimet & Bouchez, 2012; Soininen et al., 2016). Contrary to motile and high-profile species, low-profile species are tolerant to nutrient limitation and disturbance and exhibit wider distributions (Passy, 2016), suggesting potentially higher dispersal capabilities (Heino & Soininen, 2006). Planktonic species may indicate important features of the sites such as low current velocity and large rivers.

Physico-chemical data of each sampling site included pH, conductivity ($\mu S.cm^{-1}$), total phosphorus ($mg.l^{-1}$) and water temperature (°C), with the exception of Finland and Possession island (with no water temperature data) and Ireland, Kenya and New Zealand (with no total phosphorus data). Physico-chemical data were collected up to two months before the diatom sampling. Climate data were obtained from WorldClim database at 0.5 minutes resolution (Hijmans et al., 2005), including annual precipitation (mm), seasonality in precipitation (%), annual temperature (°C), and temperature seasonality (standard deviation of monthly mean temperatures). For each sampling site, we also extracted elevation from the Global Multi-resolution Terrain Elevation Data 2010 (Danielson & Gesch, 2011) and computed terrain slope as a proxy for current velocity.

For islands, we determined age of formation from the literature (see Appendix S2) and isolation using the isolation index of Dahl (Dahl, 1991, Gillespie et al., 2008). This index (equation 1) is based on the sum of square root distances to the nearest equivalent or larger island (d_i), the nearest island group or archipelago (d_a), and the nearest continent (d_c).

$$\text{Isolation index} = \sqrt{(d_i)} + \sqrt{(d_a)} + \sqrt{(d_c)} \quad (1)$$

Creation of continental area equivalents (CAEs)

For a reliable comparison of species-area relationships between islands and continents, which are vastly different in size, we generated CAEs, comparable in size to the islands by taking subsets of the continental data (see Algar & Losos, 2011) for a related approach). The method used to create these CAEs (Fig. 1) was as follows.

We first computed the geographical centroid of each island and calculated D_{c-i}, a vector representing the Euclidean distance between the centroid and each island sample site i. Second, for each continent, we calculated D_{jj} the Euclidean distance matrix between sample sites j. All Euclidean distances were calculated from geographical coordinates expressed in a projected geographical system adapted for each region (see Appendix S2). Third, we treated all continental sites as candidate CAE centroids and calculated $\mathrm{D_{v\text{-}j}}$ representing the Euclidean distance between the candidate CAE centroid v and all other j continental sites. We then computed a matrix $DD_{v-j,c-j}$ (equation 2), which represented the absolute difference between i) the distance between a candidate CAE centroid and all other sites in the focal continent (D_{v-j}, i.e. raw of the matrix D_{jj}) and ii) the distance between the island centroid and all other sites in the focal island (D_{c-i}).

$$DD_{v\text{-}j,c\text{-}i} = |D_{v\text{-}j} - D_{c\text{-}i}| \quad (2)$$

Note that the minimum value of $DD_{v\text{-}j,c\text{-}i}$ is theoretically 0, indicating that the distance between a centroid and an island site *i* is identical to the distance between a CAE centroid and a continent site *j*. Thus, smaller $DD_{v\text{-}j,c\text{-}i}$ equates to similar distances between an island centroid and island sites and the distances between a candidate CAE centroid and continental sites. We then assigned for each centroid-island site distance a unique corresponding CAE centroid-continent site distance ($\Delta_{c\text{-}i,v\text{-}j}$, i.e. the minimum value of the column of $DD_{v\text{-}j,c\text{-}i}$, equation 3).

$$\Delta_{c\text{-}i,v\text{-}j} = min(DD_{.,c\text{-}i}) \quad (3)$$

Then, we considered that the CAE centroid could be considered as the centroid of a CAE only if at least N=15 of the selected CAE centroid-continent site distances $\Delta_{c\text{-}i,v\text{-}j}$ were below a threshold value (θ) set to 5 km. Thus, theoretically, the number of sites in each CAE could vary between 15 and the total number of sites in each island. Note that because Kauai and Possession islands have less than 15 sites, N was set to 10 for the creation of their CAEs. Finally, to avoid pseudoreplication within the sites of CAEs, we selected for each continent-island pair only CAE separated by a distance of at least twice the mean distance between centroids and their corresponding sites. Due to this procedure, the size of CAEs could be, in some rare case, much smaller than the corresponding island size.

The CAEs, corresponding to the sampling area of an island, were successfully created in all continents (see appendix S3). Kenya was an exception because we were unable to create CAEs corresponding to Corsica, Iceland, Ireland, New Caledonia, Sardinia, North and South New Zealand, which were larger in size. Also, following our methodology, it was not possible to create CAEs corresponding to the island of Mayotte in USA and Finland, the island of São Miguel in Finland and USA and Possession Island in China and USA because their continental sites were more spread out than the island sites. Consequently, the total number of continent-island pairs for creating CAEs was 96.

Randomization procedure for calculation of species richness and other environmental variables

For each continent-island pair, we randomly selected 15 sites within the CAEs and 15 sites within each island (10 sites for Kauai and Possession and their respective CAEs; 20 iterations) to achieve comparable sampling effort for islands and CAEs. For each random subset, we calculated species richness as the total number of species observed among the 15 sites, and area from the convex hull around these 15 sites. We found that the areas in islands estimated using convex hulls were good surrogates for whole island areas (see appendix S4). We also calculated median values for each environmental variable for the 15 sites in the CAEs and islands and computed their environmental heterogeneity as the multivariate dispersion of all environmental variables using the average distance of all samples to the sample centroid in the multivariate space with the *betadisper* function in the *vegan* package. Environmental variables used in the analyses and computation of environmental heterogeneity were selected because they are known to be important for stream diatom distributions (Soininen, 2007; Soininen et al., 2016). For the computation of environmental heterogeneity in the Kenya and Ireland dataset, we respectively used total nitrogen and orthophosphate concentrations instead due to the lack of total phosphorus data (none of the nutrient concentration was used in the computation of heterogeneity for New Zealand and Possession islands due to missing data).

Data analyses

We conducted separate analyses for CAEs and islands to examine the relationship between species richness and area (SAR). We used linear mixed models (LMMs) for CAEs to account for continental influences that may underlie differences in species pools and the potential lack of independence among CAEs, given that multiple CAEs were created within a continent (i.e. continents were included as random factors). We performed traditional linear models for islands. We tested SAR with three commonly used models (DeMalach et al., 2019), including power (Arrhenius model), logarithmic (Gleason model) and Michaelis-Menten, and selected the best model based on the lowest Akaike Information Criterion. We also tested relationships between area and the 'corrected' species richness (according to the size of the species pool) and genus richness to ensure that the sampling effort or the taxonomic resolution did not influence our results.

To test for passive sampling, we estimated species richness from rarefaction curves based on species occurrence. For each CAE and island, we pulled at random 15 sites and randomly selected 130 species occurrences without replacement, thus ensuring that the maximum occurrence of each species did not exceed 15. Species richness was then estimated from the 130 occurrences and used to generate the SAR, which was fit with mixed models for CAEs and traditional linear models for islands. According to Chase et al. (2019), failure to detect SAR using this estimation of species richness would suggest that SAR is caused by passive sampling only. However, the reverse is not true, and significant SAR observed with this estimation of species richness does not necessarily prove the absence of passive sampling (Chase et al., 2019).

Then, to test for the effect of area on species richness after controlling for environmental variation, we first computed global LMMs for total and guild species richness and eight environmental variables (pH, conductivity, elevation, annual temperature, annual precipitation, temperature seasonality, precipitation seasonality and environmental heterogeneity). Models were constructed using the median values of species richness as the response variable and median environmental variables obtained from the subsampling procedure as explanatory variables (N = 851, i.e. one value for each 833 CAEs and each 18 island). Prior to analyses, explanatory variables were log-transformed to improve normality when necessary and standardized, but we did not treat for multicollinearity here, as this does not affect the fit of the model. Second, residuals from these regressions were regressed against log-transformed area with LMMs for CAEs and simple linear models for islands.

The number of islands in our study is comparable to the number of islands in many other SAR studies (see data used in Matthews et al., 2019) but admittedly not very high (N=18) for a study at the worldwide scale (Kalmar & Currie, 2006). Therefore, we performed a sensitivity test with our continental datasets to determine the number of CAEs required for observing a significant SAR, acknowledging that the number of islands and CAEs necessary to detect a SAR may be different. We used the median values of the species richness and area obtained from the randomization procedure for each continent-island pair (N=96), and randomly sampled (1000 times) K continent-island pairs. Each time we fit the SAR with the best SAR model (logarithmic) and extracted the probability (*P*) of observing a significant SAR, as well as the median values of model coefficients. We varied K from 11 to 96, i.e. the total number of continent-island pairs available in our dataset. We then identified the minimum number of 'islands' needed to observe a significant SAR with our data (*P*>95%). We performed these analyses with both traditional linear models and LMMs (e.g. assuming a common species pool).

We compared species richness of islands and species richness of their corresponding CAEs with Cliff's non-parametric effect size statistic (Romano et al., 2006; Tecchio et al., 2016), due to the large number of data points resulting from the randomization procedure (i.e., decreasing variance around the mean). We also used Cliff's tests to compare the species richness of each ecological guild between CAEs and islands.

To compare species richness of CAEs and islands after removing the effect of environment, we computed LMMs as above but also included all the values of random subsampling (x20, N=17020). We therefore used a nested design in the random factors of the models, so that subsampling values are nested within each continent/island. Residual richness values were then extracted from the models and compared between CAEs and islands with Cliff's tests.

We used linear mixed models for CAEs and traditional linear models for islands to examine the relationship between species richness, environmental heterogeneity, the median of all environmental variables and the median values of latitude and longitude. Environmental explanatory variables were log-transformed to improve normality when necessary and we run separate regression models with each environmental factor and species richness to avoid multicollinearity. We also tested for non-linear relationships separately with all environmental variables with the same procedure.

Finally, to disentangle the possible drivers of the SAR for continents, we implemented piecewise structural equation modeling (SEM, Lefcheck, 2016) using linear mixed models with continental dataset as a random factor. We could not implement such models for islands due to an insufficient number of data points. We assumed an *a priori* model (Fig. 2) predicting species richness as directly influenced by area, environmental heterogeneity (as defined above), local environmental conditions and climatic conditions. We used conductivity and elevation as predictors of local conditions, temperature seasonality and annual precipitations as predictors of climate, as they were significant predictors of species richness in global LMMs and exhibited low collinearity in pairwise correlations tests (see Appendix S5). We assumed that the effect of area on species richness could also be indirect through environmental heterogeneity, according to the habitat diversity hypothesis (Lack, 1976). Finally, we also assumed that temperature seasonality and precipitation are directly influenced by elevation. We included a correlation between temperature seasonality and precipitation as well as between conductivity and precipitation (see Appendix S5). We used the Fisher's *C* statistic to test the consistency of the theoretical model with the data. All analyses were run for total richness and separately for richness of each ecological guild.

All analyses were conducted with R (R Core Team, 2019) using packages 'vegan' (Oksanen et al., 2019), 'spatstat' (Baddeley et al., 2015), 'raster' (Hijmans, 2019), 'sf' (Pebesma, 2018), 'lmerTest' (Kuznetsova et al., 2017), 'lme4' (Bates et al., 2015), 'effsize' (Torchiano, 2020) and 'piecewiseSEM' (Lefcheck, 2016).

Results

SAR patterns

We found a significant positive SAR for total species richness in CAEs, but not in islands (Fig. 3a). The best model describing the SAR in CAEs was the logarithmic model (see Appendix S6). The observed R^2 values were relatively low compared to values usually observed for islands but comparable to those found in continental areas (Kreft et al., 2008). Similar results emerged with rarefied richness (see Appendix S7), 'corrected' species richness given the size of the species pool (see Appendix S8), genus richness (see Appendix S9) and also after removing the effect of environmental variation (see Appendix S10).

The sensitivity analysis revealed that a minimum of 52 continent-islands pairs is needed to observe a significant SAR with our data. This number dropped to 16 when using mixed models with continent (a surrogate for the species pool) as a random effect (see Appendix S11).

About half of the 1967 identified species belonged to the motile guild (see Appendix S12), followed by the high profile and low-profile guilds. Planktonic species and species with variable guilds represented a minor part of the communities. Species richness within all guilds was significantly and positively related to area in CAEs (Fig. 3b-e) and this relationship persisted for all but the high-profile guild after controlling for the environment (see Appendix S10).

Comparison of species richness of islands and CAEs

Overall, species richness was significantly lower in the islands than in the respective CAEs for more than 50% of all continent-island pairs (*N* = 96) (Fig. 5a). Similar results emerged for the guilds, especially for the planktonic guild, where over 70% of the comparisons had significantly higher species richness in CAEs. The only exception was the low-profile guild whose species richness tended to be higher in islands (ca. 60%). Importantly, however, when environmental variation was accounted for, the species richness differences between CAEs and islands disappeared in more than 80% of cases (Fig. 5b).

Ecological variables driving species richness

In islands, we found significant relationships between species richness and isolation for total, low profile species richness (U-shaped pattern) and planktonic guild (negative linear pattern and a weak non-linear pattern) (Fig. 4, see Appendix S13). There was no relationship between richness and age of island for the total community or any of the ecological guilds (see Appendix S13). Apart from isolation, total species richness in islands was significantly related only to phosphorus concentration. Species richness of guilds was also significantly related to some other environmental variables depending on the guild considered (see Appendix S13).

Total species richness in CAEs was significantly related to environmental heterogeneity, pH, conductivity, phosphorus concentration, all climatic variables and longitude (see Appendix S13). The piecewise SEM models (Lefcheck, 2016) disentangled the effects of the influencing factors and demonstrated that diatom species richness in CAEs was related to area, habitat heterogeneity, physicochemistry, elevation, and climate. The data fitted well the *a priori* model (Fig. 6) for total species richness and species richness of all ecological guilds. The marginal R^2 (variance explained by the fixed effects only) for total species richness was 0.71 and varied between 0.14 (for low-profile species) to 0.59 (for motile species). In the SEMs, area explained species richness independently, without any indirect effect through environmental heterogeneity, except for the motile species richness where the effect of environmental heterogeneity was negative. Indeed, total species richness and richness of the motile guild were mainly driven by area and temperature seasonality (Fig. 6a, d). Low-profile species richness was only explained by area and precipitation (Fig. 6b). High-profile species richness was mainly explained by area, elevation and conductivity, while climate had no direct effect (Fig. 6c). Finally, planktonic species richness was solely determined by elevation and was thus the only group without a significant relationship with area.

Discussion

Here, we conducted the first comparative analysis of island vs. mainland species-area relationship for microbes, providing insight into the roles of area, environmental heterogeneity, isolation and island age on species richness patterns. We showed for freshwater diatoms that: (i) there was a significant SAR in continents but not in islands (except for high profile), (ii) regional species richness was higher in continents than in islands, but this difference was explained entirely by environmental conditions (iii) the effect of isolation varied among diatom guilds and (iv) area and median environmental conditions but not environmental heterogeneity were significant predictors of diatom richness. Next, we will discuss the main findings in more detail and highlight our major conclusions about total community and guild richness.

Drivers of species richness in islands

The lack of a significant SAR in islands may be due to low sample size (N = 18) or may represent a real biogeographical pattern. Sensitivity analyses performed for CAEs revealed that 16 islands are needed to detect a SAR given a common species pool. This result is consistent with numerous studies on other organisms, reporting ISAR for a relatively small number of islands within archipelagos (Matthews et al., 2019). However, at a global scale, a much higher number of islands (N=52, Appendix S11) may be required for detection of diatom ISAR.

The absence of diatom ISAR may have evolutionary and ecological causes. First, diatoms may have distinct species pools across the globe (Soininen et al., 2016) and differences in island area may not be sufficient to predict richness on islands that differ greatly in species pool. As the size of the species pool influences the shape of the SAR (Catano et al., 2021), future analyses on archipelagos will be essential for determining whether ISAR exists for diatoms (but see Jüttner et al., 2018). Second, environmental heterogeneity, which increased with island size (Fig. 3, and see Appendix S5) and is recognized as an important driver of SAR (Lack, 1976; Chase et al., 2019), had no direct impact on island species richness. Third, island richness was related only to total phosphorus, suggesting that productivity is a key factor explaining island diatom species richness at this scale. Note however, that due to data availability, only phosphorus concentration was considered as a resource factor for explaining species richness. The consideration of other nutrient resources, known to influence diatom diversity (e.g. nitrogen, iron, Passy, 2007, Soininen, 2007), may improve the understanding of diatoms species richness in islands.

We found that isolation might have some effect on species richness in islands. Two of the most isolated islands (Oahu and New Caledonia) actually showed high species richness, resulting in a U-shaped relationship between species richness and isolation for total and low-profile species richness. Oahu and New Caledonia still had the highest species richness when the latter is corrected by species pool but the U-shaped relationship is only marginally significant ($p = 0.09$, see Appendix S8). Greater speciation in the most isolated islands, which have many endemic species, e.g. New Caledonia has been dubbed "Galapagos of diatoms" (Moser et al., 1998), may explain their higher richness considering that endemic and total species richness are typically correlated (Kallimanis et al., 2010). We could, however, not exclude the fact that some other unmeasured environmental factors, particularly related to islands conditions, may also be responsible for this pattern. Finally, our finding further suggests that the biogeographical drivers of diatom richness on real islands are trait dependent.

Following Rosenzweig (1995), we hypothesized that islands would harbour lower species richness than continents due to diminished dispersal and rescue effects. While species richness was indeed lower in islands compared to continents, this difference disappeared when we accounted for environmental differences. Thus, annual precipitation, higher in islands than continents, was associated with lower species richness (see Appendix S10), likely because of its positive effect on current velocity, and subsequently, shear stress (Heino & Soininen, 2007).

Drivers of species richness in continents

We tested whether SARs in continents could result from passive sampling and environmental heterogeneity, which are major drivers of the SAR (Lack, 1976; Rosenzweig, 1995; Stein et al., 2014; Chase et al., 2019). Surprisingly, species richness in continents was not directly explained by environmental heterogeneity in the SEM.

Although area was strongly related to environmental heterogeneity (but poorly related to other environmental variables, see Appendix S5), none of the SEM models showed a direct effect of habitat heterogeneity on either total or guild species richness, except for motile species. For the latter, the direct effect of habitat heterogeneity was negative, contrary to the results observed in univariate regressions (see Appendix 13) due to the strong collinearity between area and heterogeneity. Given that we still observed a significant SAR with species richness estimated from the rarefaction curves, passive sampling cannot be completely ruled out (Chase et al., 2019). However, the impact of area on richness in continents might also be due to disproportionate effects, including dispersal, extinction and speciation. While extinction and speciation have been less studied in diatoms, dispersal and mass effects (whereby species maintain their presence in unfavorable conditions via immigration, Shmida & Wilson, 1985) were shown to have a notable influence on regional to subcontinental diatom communities (Soininen, 2007; Jamoneau et al., 2018; Leboucher et al., 2020). For continental diatoms, larger areas may thus increase the probability of immigration from the surrounding landscape, particularly for species with high dispersal capabilities (mass-effect species), thereby increasing CAE's diversity.

Environmental factors, such as nutrients, climate and elevation, were also important predictors of total and guild species richness. Total species richness decreased with temperature seasonality, as did the species richness of motile species, which represented ca. 50% of the whole community (see Appendix S12). As motile species are generally warm-water species (Pound et al., 2021) and high seasonality occurs in colder areas, it is possible that motile guild richness was limited by unfavorable temperatures. Species richness of the high profile and planktonic guilds was the lowest at high elevation. For high-profile species, high elevation is stressful due to increased current velocity and probability for dislodgement. For planktonic species, high elevations do not provide sufficient habitat, given that these species require large water bodies. Species richness of the low-profile guild is positively influenced by annual precipitation probably because this guild is tolerant to physical disturbance (Passy, 2007), which should increase its richness in the community.

Conclusions

We examined diatom ISARs and compared them with the SARs of similarly sized continental area equivalent across five continents. Contrary to most previous studies, we did not find significant ISAR for total species richness but detected significant relationships of richness with total phosphorus. These results imply that diatom richness in islands is not related to area but is controlled by productivity. However, the lack of ISAR may be due to distinct species pool across islands in our study. Species richness was typically higher in continental areas than in similarly sized islands, most probably due to differences in climate and related environmental conditions, such as current velocity. The significant SAR for continents may originate from disproportionate effects, such as mass effect, but not from environmental heterogeneity. Isolation influenced the richness of the whole community and some diatom guilds in islands. These finding indicate that there are important differences in richness responses to island properties among ecological guilds and between the community level and the functional level. Finally, the proposed new method for species-area comparisons between islands and continental area equivalents will advance research on biogeography of islands vs.

mainland. We advocate obtaining global diatom data, particularly from archipelagos to better understand the drivers of island species diversity.

Figures legend

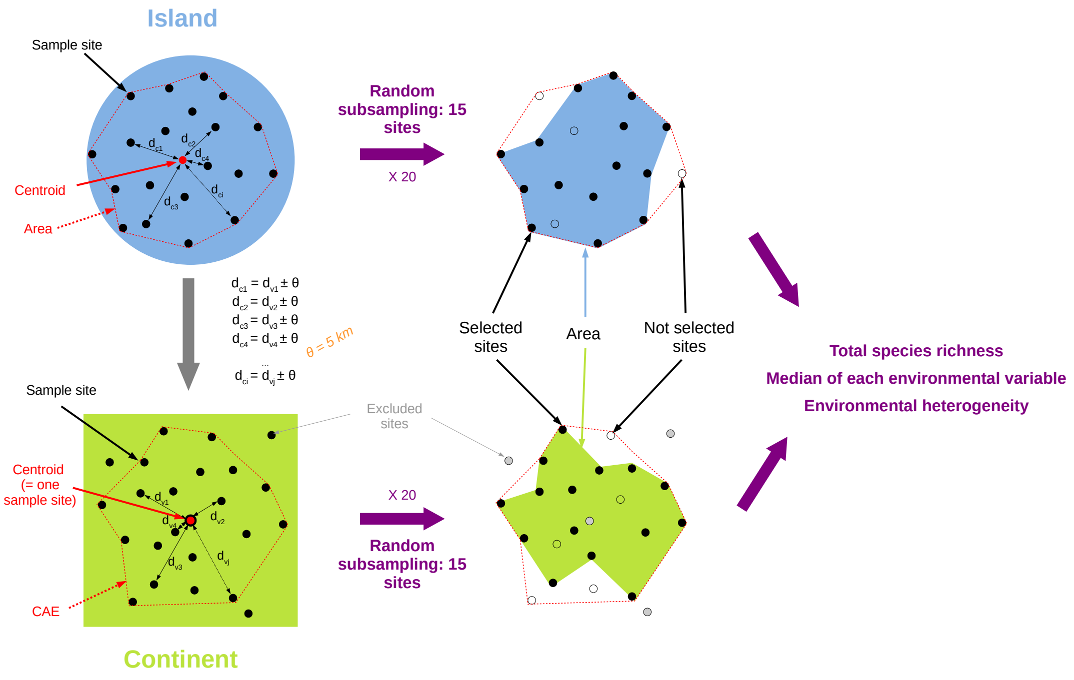

Figure 1: Descriptive diagram of the methods. Diagram describing the methodological process used for creation of continental area equivalents (CAE) and subsampling of both islands and CAEs.

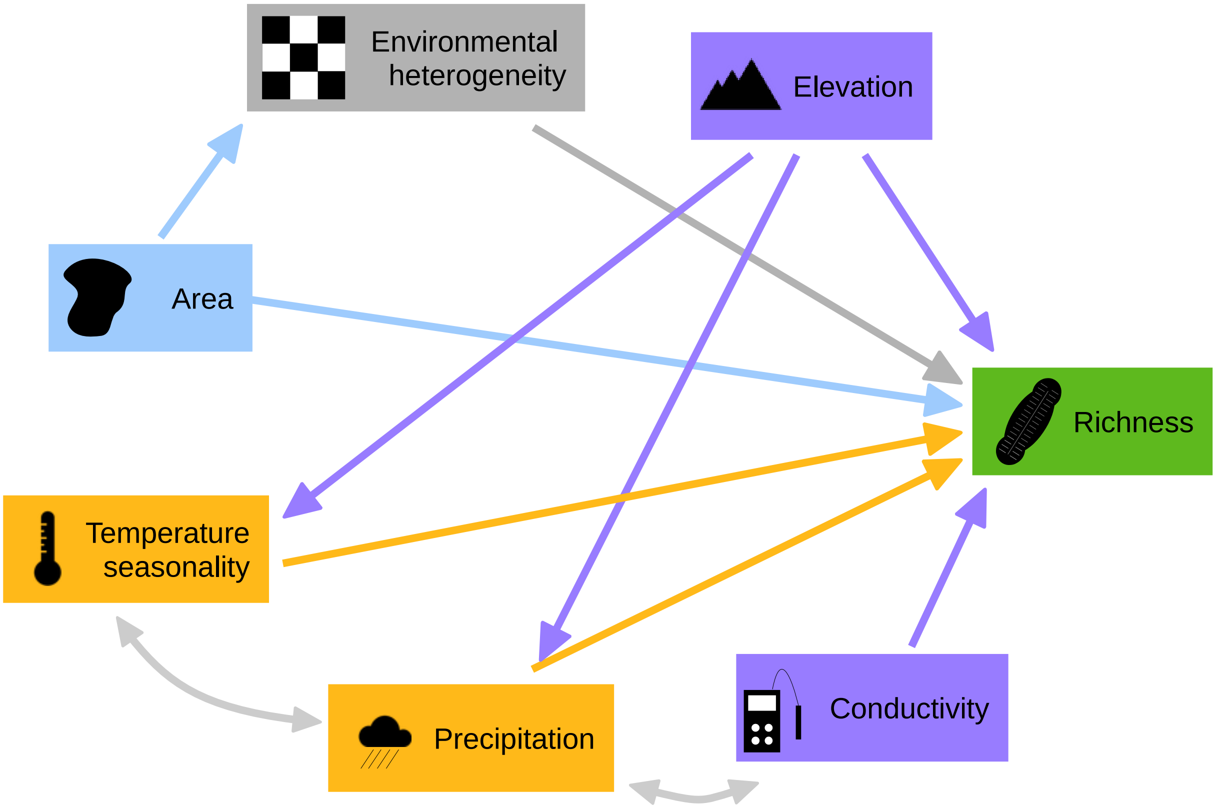

Figure 2. A priori model explaining diatom species richness. Species richness is modeled as a function of area, environmental heterogeneity, local environmental conditions (purple) and climate (orange).

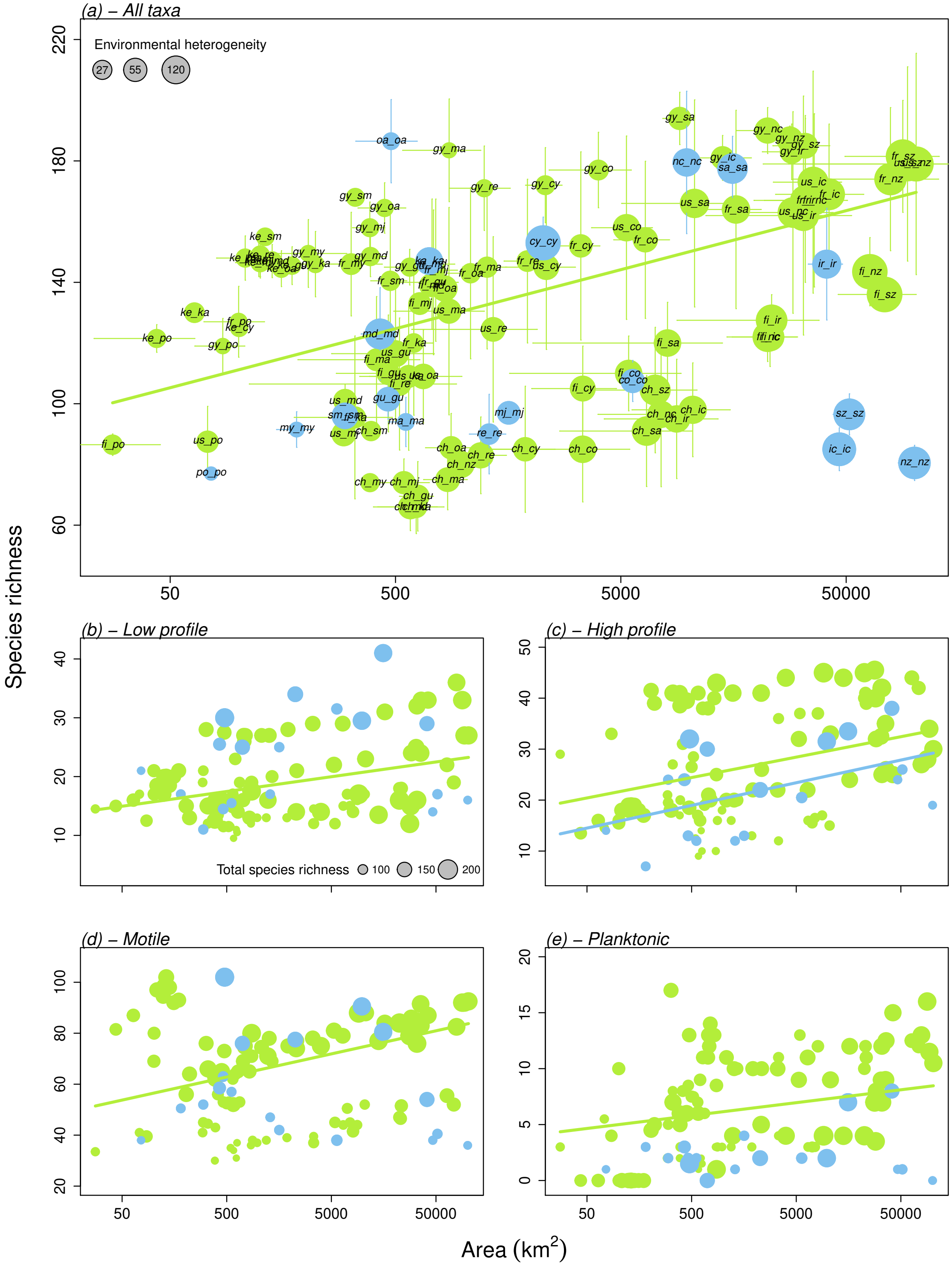
(a) – All taxa
Environmental heterogeneity
27
55
120
Species richness
220
180
140
100
60
50
500
5000
50000
oa_oa
gy_ma
gy_sa
gy_nc
gy_nz
gy_sz
gy_fr
gy_ic
nc_nc
sa_sa
gy_co
gy_re
gy_cy
gy_sm
gy_oa
gy_mj
gy_my
gy_md
gy_ka
ke_sm
fr_my
fr_sm
fr_cy
fr_co
fr_re
fr_oa
fr_ma
fr_mj
fr_gu
fr_sa
fr_nz
fr_sz
fr_ic
fr_ka
fr_po
us_co
us_sa
us_ic
us_ir
us_nc
us_sz
us_nz
us_cy
us_ma
us_re
us_gu
us_md
us_po
us_mj
cy_cy
ir_ir
fi_nz
fi_sz
fi_ir
fi_nc
fi_sa
fi_co
fi_cy
fi_mj
fi_ma
fi_gu
fi_re
fi_po
ke_ka
ke_po
ke_cy
gy_po
md_md
ka_ka
gu_gu
ma_ma
mj_mj
re_re
my_my
sm_sm
po_po
co_co
ch_sz
ch_nc
ch_ic
ch_ir
ch_sa
ch_oa
ch_re
ch_nz
ch_cy
ch_co
ch_ma
ch_mj
ch_my
ch_gu
ch_sm
ch_ka
sz_sz
ic_ic
nz_nz
(b) – Low profile
40
30
20
10
Total species richness
100
150
200
(c) – High profile
50
40
30
20
10
(d) – Motile
100
80
60
40
20
50
500
5000
50000
(e) – Planktonic
20
15
10
5
0
50
500
5000
50000
Area (km^2)

Figure 3. Species-area relationships for continents and islands. Species-area relationships for continental area equivalents (CAE) (in green, N=96) and islands (in blue, N=18) for total species richness (a), and richness of low profile (b), high profile (c), motile (d) and (e) planktonic species. Green regression lines represent significant linear fits in mixed models for CAEs: richness = 8.44x + 72.24, R^2_m = 0.22 for total species richness, 1.09x + 10.69, R^2_m = 0.13 for low-profile species, 1.74x + 13.61, R^2_m = 0.11 for high-profile species, 3.94x + 38.36, R^2_m = 0.11 for motile species and 0.50x + 2.69, R^2_m = 0.05 for planktonic species, where x = log(area). The blue regression line represents significant linear fit for high-profile species of islands: 1.93x + 6.95, R^2_{aj} = 0.23. Dot sizes are proportional to environmental heterogeneity (in log) for all taxa (a) and proportional to total species richness for functional groups (b-e). Error bars represent standard deviation estimated from the subsampling procedure. Text in dots indicate the dataset used for computing species richness and area. For example, 'fr_my' indicates the position of Mayotte CAE in France. Continental datasets are indicated by 'fr' for France, 'us' for US, 'fi' for Finland, 'ch' for China, 'ke' for Kenya and 'gy' for French Guiana and islands indicated by 'ic' for Iceland, 'co' for Corsica, 'gu' for Guadeloupe, 'ma' for Martinique, 're' for La Réunion, 'my' for Mayotte, 'nz' for North New Zealand, 'sz' for South New Zealand, 'nc' for New Caledonia, 'ka' for Kauai, 'oa' for Oahu, 'po' for Possession, 'cy' for Cyprus, 'ir' for Ireland, 'md' for Madeira, 'mj' for Majorca, 'sm' for São Miguel and 'sa' for Sardinia.

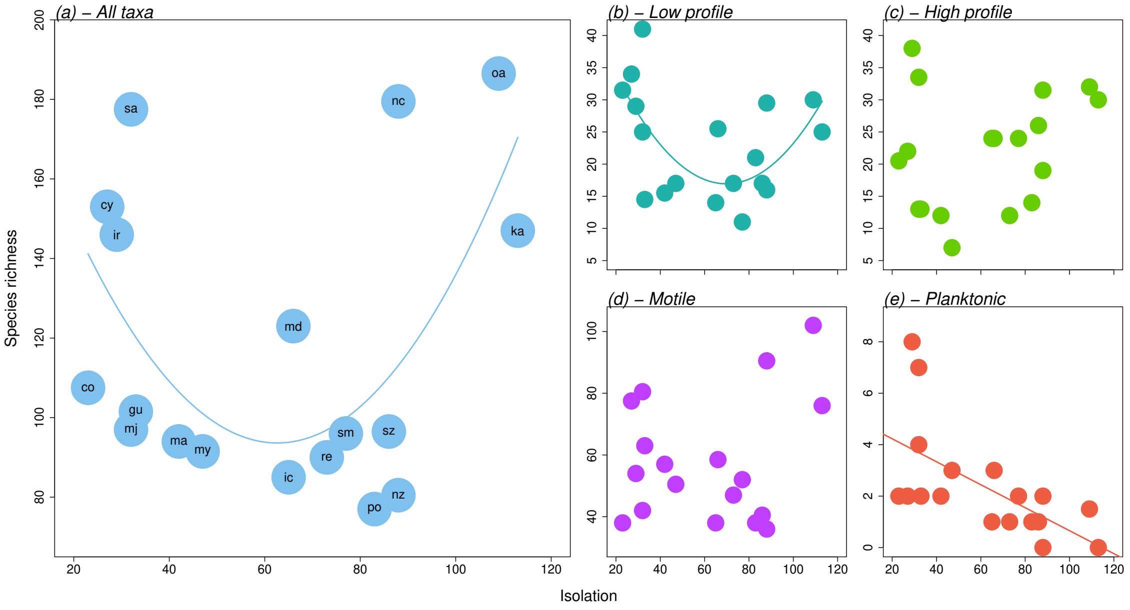

Figure 4. Relationships between island species richness and isolation. Relationship between total species richness (a) and species richness of each ecological guild (b-e) with island isolation for islands ($N = 18$). Significant linear and quadratic relationships ($p<0.05$) are shown by regression fits (only the fit with the lower AIC is shown if both are significant, see Appendix S13): $0.03x^2 - 3.79x + 212.32$, $R^2_{aj} = 0.26$ for total richness, $0.01x^2 - 0.96x + 50.44$, $R^2_{aj} = 0.32$ for low-profile and $-0.04x + 5.12$, $R^2_{aj} = 0.34$ for planktonic species. For island names, see Fig. 1. Isolation is based on index defined by Dahl (Dahl, 1991).

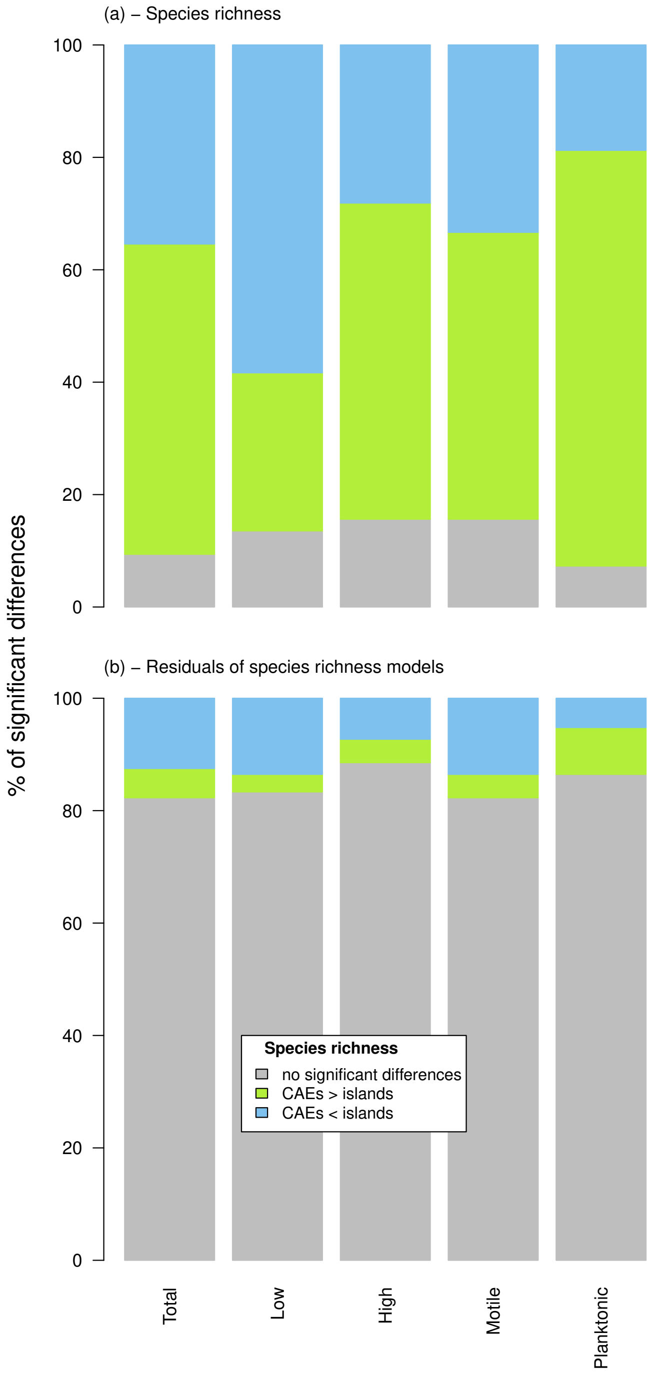

Figure 5. Comparison of species richness between continental area equivalents (CAEs) and islands. Percentage of significant and non-significant tests (*N* = 96 continent-island pairs) between CAEs and islands for species richness (a) and species

richness residuals (b). Tests were performed for total and guild species richness. Species richness residuals were estimated from linear mixed models with species richness as the dependent variable, and pH, conductivity, elevation, mean annual temperature and precipitation, temperature and precipitation seasonality and environmental heterogeneity as explanatory variables and continent as a random factor. Comparisons of values (i.e., species richness or residuals of species richness) were performed with Cliff's test, whereby tests with delta >0.33 indicated significant differences (Romano et al., 2006).

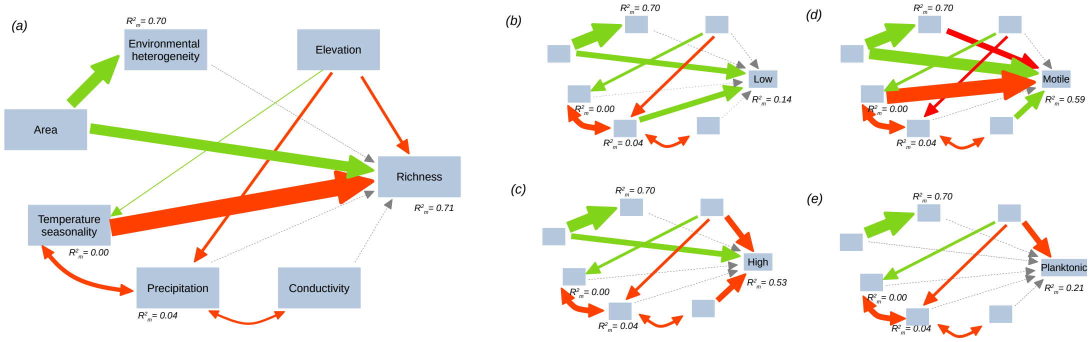

Figure 6. Structural equation models explaining species richness in continental area equivalents (CAEs). Structural equation models for total species richness (a), low profile (b), high profile (c), motile (d) and planktonic (e) species richness in continents (N = 96 CAEs). Green and red arrows represent significant positive and negative relationships, respectively, whereas gray-dashed arrows represent non-significant relationships. Arrow widths are proportional to the standardized regression coefficients and R^2_m values represent marginal R^2 from a linear mixed model. All models fitted well the *a priori* model, i.e. the model including all shown causal relationships (Fisher's C = 14.99, df = 14, p = 0.38 for all models).

Data Availability Statement

Data are available under the following link: https://doi.org/10.57745/ZPBSLT

References

Abonyi, A., Horváth, Z., & Ptacnik, R. (2018). Functional richness outperforms taxonomic richness in predicting ecosystem functioning in natural phytoplankton communities. *Freshwater Biology*, *63*(2), 178–186.

Algar, A. C., & Losos, J. B. (2011). Evolutionary assembly of island faunas reverses the classic island–mainland richness difference in Anolis lizards. *Journal of Biogeography*, *38*(6), 1125–1137.

Arrhenius, O. (1921). Species and area. *The Journal of Ecology*, 95–99.

Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press.

Báldi, A. (2008). Habitat heterogeneity overrides the species–area relationship. *Journal of Biogeography*, *35*(4), 675–681.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., & Grothendieck, G. (2015). *lme4: Linear Mixed-Effects Models using “Eigen” and S4* (1.1-8) [Computer software].

Bell, T., Ager, D., Song, J.-I., Newman, J. A., Thompson, I. P., Lilley, A. K., & Gast, C. J. van der. (2005). Larger Islands House More Bacterial Taxa. *Science*, *308*(5730), 1884–1884.

Catano, C. P., Grman, E., Behrens, E., & Brudvig, L. A. (2021). Species pool size alters species–area relationships during experimental community assembly. *Ecology*, *102*(1), e03231.

Chao, A. (1987). Estimating the Population Size for Capture-Recapture Data with Unequal Catchability. *Biometrics*, *43*(4), 783–791.

Chase, J. M., Gooriah, L., & May, F. (2019). A framework for disentangling ecological mechanisms underlying the island species–area. *Frontiers of Biogeography*.

Chase, J. M., & Leibold, M. A. (2003). *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press.

Connor, E. F., & McCoy, E. D. (1979). The Statistics and Biology of the Species-Area Relationship. *The American Naturalist*, *113*(6), 791–833.

Dahl, A. L. (1991). *Island Directory* (UNEP Regional Seas Directories and Bibliographies No. 35; p. 573). UNEP.

Danielson, J. J., & Gesch, D. B. (2011). *Global multi-resolution terrain elevation data 2010 (GMTED2010)* (U.S. Geo - Logical Survey Open-File Report No. 2011–1073; p. 26).

DeMalach, N., Saiz, H., Zaady, E., & Maestre, F. T. (2019). Plant species–area relationships are determined by evenness, cover and aggregation in drylands worldwide. *Global Ecology and Biogeography*, *28*(3), 290–299.

Drakare, S., Lennon, J. J., & Hillebrand, H. (2006). The imprint of the geographical, evolutionary and ecological context on species–area relationships. *Ecology Letters*, *9*(2), 215–227.

Farina, A. (2008). *Principles and Methods in Landscape Ecology: Towards a Science of the Landscape*. Springer.

Finlay, B. J. (2002). Global Dispersal of Free-Living Microbial Eukaryote Species. *Science*, *296*(5570), 1061–1063.

Gillespie, R. G., Claridge, E. M., & Roderick, G. K. (2008). Biodiversity dynamics in isolated island communities: Interaction between natural and human-mediated processes. *Molecular Ecology,* 17(1), 45–57.

Hanski, I. A., & Gaggiotti, O. E. (2004). *Ecology, Genetics and Evolution of Metapopulations*. Elsevier Academic Press.

Heino, J., & Soininen, J. (2005). Assembly rules and community models for unicellular organisms: Patterns in diatoms of boreal streams. *Freshwater Biology*, *50*(4), 567–577.

Heino, J., & Soininen, J. (2006). Regional occupancy in unicellular eukaryotes: A reflection of niche breadth, habitat availability or size-related dispersal capacity? *Freshwater Biology*, *51*(4), 672–685.

Heino, J., & Soininen, J. (2007). Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organisms? *Biological Conservation*, *137*(1), 78–89.

Hijmans, R. J. (2019). *raster: Geographic Data Analysis and Modeling*.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978.

Jamoneau, A., Passy, S. I., Soininen, J., Leboucher, T., & Tison-Rosebery, J. (2018). Beta diversity of diatom species and ecological guilds: Response to environmental and spatial mechanisms along the stream watercourse. *Freshwater Biology*, *63*(1), 62–73.

Jüttner, I., Vijver, B. V. de, Williams, D. M., Lange-Bertalot, H., & Ector, L. (2018). The genus Eunotia (Bacillariophyta) in the Falkland Islands and species-area relationships in sub-Antarctic islands. *Diatom Research*, *33*(4), 413–452.

Kallimanis, A. S., Bergmeier, E., Panitsa, M., Georghiou, K., Delipetrou, P., & Dimopoulos, P. (2010). Biogeographical determinants for total and endemic species richness in a continental archipelago. *Biodiversity and Conservation*, *19*(5), 1225–1235.

Kalmar, A., & Currie, D. J. (2006). A global model of island biogeography. *Global Ecology and Biogeography*, *15*(1), 72–81.

Kalmar, A., & Currie, D. J. (2007). A Unified Model of Avian Species Richness on Islands and Continents. *Ecology*, *88*(5), 1309–1321.

Krause, S., Le Roux, X., Niklaus, P. A., Van Bodegom, P. M., Lennon, J. T., Bertilsson, S., Grossart, H.-P., Philippot, L., & Bodelier, P. L. E. (2014). Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Frontiers in Microbiology*, *5*.

Kreft, H., Jetz, W., Mutke, J., Kier, G., & Barthlott, W. (2008). Global diversity of island floras from a macroecological perspective. *Ecology Letters*, *11*(2), 116–127.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(1), 1–26.

Lack, D. (1976). *Island biology: Illustrated by the land birds of Jamaica*. Blackwell Scientific Publications, Oxford, UK.

Leboucher, T., Tison-Rosebery, J., Budnick, William. R., Jamoneau, A., Vyverman, W., Soininen, J., Boutry, S., & Passy, S. I. (2020). A new metacommunity approach for detecting species influenced by mass effect. *Journal of Applied Ecology*.

Lecointe, C., Coste, M., & Prygiel, J. (1993). “Omnidia”: Software for taxonomy, calculation of diatom indices and inventories management. *Hydrobiologia*, *269–270*(1), 509–513.

Lefcheck, J. S. (2016). piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573–579.

Lomolino, M. V., & Weiser, M. D. (2001). Towards a More General Species-Area Relationship: Diversity on All Islands, Great and Small. *Journal of Biogeography*, *28*(4), 431–445. JSTOR.

MacArthur, R. H., & Wilson, E. O. 1967. (1967). *The theory of island biogeography*. Princeton University Press.

Matthews, T. J., Cottee‐Jones, H. E., & Whittaker, R. J. (2014). Habitat fragmentation and the species–area relationship: A focus on total species richness obscures the impact of habitat loss on habitat specialists. *Diversity and Distributions*, *20*(10), 1136–1146.

Matthews, T. J., Rigal, F., Triantis, K. A., & Whittaker, R. J. (2019). A global model of island species–area relationships. *Proceedings of the National Academy of Sciences*, *116*(25), 12337–12342.

Mittelbach, G. G., & McGill, B. J. (2019). *Community Ecology*. Oxford University Press.

Moser, G., Lange-Bertalot, H., & Metzeltin, D. (1998). *Insel der Endemiten. Geobotanisches Phänomen Neukaledonien—Island of Endemics. New Caledonia a geobotanical phenomenon* (J. Cramer). Schweizerbart'sche Verlagsbuchhandlung.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2019). *vegan: Community Ecology Package*.

Passy, S. I. (2007). Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. *Aquatic Botany*, *86*(2), 171–178.

Passy, S. I. (2016). Abundance Inequality in Freshwater Communities Has an Ecological Origin. *The American Naturalist*, *187*(4), 502–516.

Peay, K. G., Bruns, T. D., Kennedy, P. G., Bergemann, S. E., & Garbelotto, M. (2007). A strong species–area relationship for eukaryotic soil microbes: Island size matters for ectomycorrhizal fungi. *Ecology Letters*, *10*(6), 470–480.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446.

Pound, K. L., Larson, C. A., & Passy, S. I. (2021). Current distributions and future climate-driven changes in diatoms, insects and fish in U.S. streams. *Global Ecology and Biogeography*, *30*, 63–78.

Prugh, L. R., Hodges, K. E., Sinclair, A. R., & Brashares, J. S. (2008). Effect of habitat area and isolation on fragmented animal populations. *Proceedings of the National Academy of Sciences*, *105*(52), 20770–20775.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Reche, I., Pulido-Villena, E., Morales-Baquero, R., & Casamayor, E. O. (2005). Does Ecosystem Size Determine Aquatic Bacterial Richness? *Ecology*, *86*(7), 1715–1722.

Rimet, F., & Bouchez, A. (2012). Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. *Knowledge and Management of Aquatic Ecosystems*, *406*, 01.

Romano, J., Kromrey, J. D., Coraggio, J., & Skowronek, J. (2006). Appropriate statistics for ordinal level data: Should we really be using t-test and Cohen’s d for evaluating group differences on the NSSE and other surveys. *Annual Meeting of the Florida Association of Institutional Research*, 1–33.

Rosenzweig, M. L. (1995). *Species diversity in space and times*. Cambridge: Cambridge University Press.

Sandin, S. A., Vermeij, M. J. A., & Hurlbert, A. H. (2008). Island biogeography of Caribbean coral reef fish. *Global Ecology and Biogeography*, *17*(6), 770–777.

Schrader, J., König, C., Triantis, K. A., Trigas, P., Kreft, H., & Weigelt, P. (2020). Species–area relationships on small islands differ among plant growth forms. *Global Ecology and Biogeography*, *29*(5), 814–829.

Shmida, A., & Wilson, M. V. (1985). Biological determinants of species diversity. *Journal of Biogeography*, *12*(1), 1–20.

Simberloff, D. S., & Wilson, E. O. (1969). Experimental Zoogeography of Islands: The Colonization of Empty Islands. *Ecology*, *50*(2), 278–296.

Smith, V. H., Foster, B. L., Grover, J. P., Holt, R. D., Leibold, M. A., & deNoyelles, F. (2005). Phytoplankton species richness scales consistently from laboratory microcosms to the world’s oceans. *Proceedings of the National Academy of Sciences*, *102*(12), 4393–4396.

Soininen, J. (2007). Environmental and Spatial Control of Freshwater Diatoms—A Review. *Diatom Research*, *22*(2), 473–490.

Soininen, J., Jamoneau, A., Rosebery, J., & Passy, S. I. (2016). Global patterns and drivers of species and trait composition in diatoms. *Global Ecology and Biogeography*, *25*(8), 940–950.

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, *17*(7), 866–880.

Tecchio, S., Chaalali, A., Raoux, A., Tous Rius, A., Lequesne, J., Girardin, V., Lassalle, G., Cachera, M., Riou, P., Lobry, J., Dauvin, J.-C., & Niquil, N. (2016). Evaluating ecosystem-level anthropogenic impacts in a stressed transitional environment: The case of the Seine estuary. *Ecological Indicators*, *61, Part 2*, 833–845.

Teittinen, A., & Soininen, J. (2015). Testing the theory of island biogeography for microorganisms patterns for spring diatoms. *Aquatic Microbial Ecology*, *75*(3), 239–250.

Tilman, D., Knops, J., Wedin, D., Reich, P., Ritchie, M., & Siemann, E. (1997). The Influence of Functional Diversity and Composition on Ecosystem Processes. *Science*, *277*(5330), 1300–1302.

Torchiano, M. (2020). *effsize: Efficient Effect Size Computation*.

Triantis, K. A., Economo, E. P., Guilhaumon, F., & Ricklefs, R. E. (2015). Diversity regulation at macro-scales: Species richness on oceanic archipelagos. *Global Ecology and Biogeography*, *24*(5), 594–605.

Whittaker, R. J., & Fernandez-Palacios, J. M. (2007). *Island Biogeography: Ecology, Evolution, and Conservation*. OUP Oxford.

Whittaker, R. J., Fernández-Palacios, J. M., Matthews, T. J., Borregaard, M. K., & Triantis, K. A. (2017). Island biogeography: Taking the long view of nature's laboratories. *Science*, *357*(6354).

Whittaker, R. J., Triantis, K. A., & Ladle, R. J. (2008). A general dynamic theory of oceanic island biogeography. *Journal of Biogeography*, *35*(6), 977–994.

Biosketch

Aurélien Jamoneau is a researcher in community ecology at INRAE institute of Cestas-Gazinet in France. He is interested in aquatic community assemblages at multiple spatial and temporal scales. Janne Soininen in a professor in spatial environmental research at the University of Helsinki. He is interested in large-scale community ecology and especially in the distribution of small aquatic organisms.