

Environmental filtering drives assembly of diatom communities over evolutionary time-scales

Elena Jovanovska^{1,2}  | Torsten Hauffe^{1,3}  | Björn Stelbrink^{1,4}  |
Aleksandra Cvetkoska⁵  | Zlatko Levkov⁶  | Bernd Wagner⁷  | Jack H. Lacey⁸  |
Nadja Ognjanova-Rumenova⁹ | Paul B. Hamilton¹⁰  | Karen M. Brandenburg⁵  |
Christian Albrecht¹  | Thomas Wilke¹ 

¹Department of Animal Ecology and Systematics, Justus Liebig University Giessen, Giessen, Germany

²Department of Palaeoanthropology, Senckenberg Research Institute, Frankfurt am Main, Germany

³Department of Biology, University of Fribourg and Swiss Institute of Bioinformatics, Fribourg, Switzerland

⁴Zoological Institute, University of Basel, Basel, Switzerland

⁵Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, The Netherlands

⁶Institute of Biology, Faculty of Natural Sciences, Ss. Cyril and Methodius University, Skopje, Macedonia

⁷Institute of Geology and Mineralogy, University of Cologne, Cologne, Germany

⁸National Environmental Isotope Facility, British Geological Survey, Nottingham, UK

⁹Department of Paleontology, Stratigraphy and Sedimentology, Institute of Geology, Bulgarian Academy of Science, Sofia, Bulgaria

¹⁰Phycology Section, Research and Collections Division, Canadian Museum of Nature, Ottawa, Ontario, Canada

Correspondence

Elena Jovanovska, Department of Palaeoanthropology, Senckenberg Research Institute, Senckenberganlage 25, 60325 Frankfurt am Main, Germany. Email: jovanovska.eci@gmail.com

Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: AL 1076/9, JO 1731/1-1, STE 2460/2-1 and 2460/3-1 and WI 1902/13; Swiss National Science

Abstract

Aim: Ecological communities are structured through the interplay of deterministic assembly processes such as competition and environmental filtering. Whereas the drivers of spatial community structure are frequently studied in extant communities, little is known about the relative importance of assembly processes in response to environmental factors over evolutionary time-scales. Here, we use an integrative framework to unravel community assembly processes since the inception of a long-lived lake ecosystem.

Location: Lake Ohrid.

Time period: From lake formation 1.36 million years ago to the present.

Major taxa studied: Planktonic diatoms.

Methods: We constructed a dated phylogeny of extant and extinct diatoms and collected trait data for 380 fossil communities to quantify phylogenetic community structure and functional richness and to determine the relative importance of deterministic assembly processes over time. We then used regression analysis to correlate the phylogenetic community structure with palaeoenvironmental and intrinsic biological predictors and to identify primary drivers of assembly processes.

Results: Our results suggest a dense packing of niche space with higher species richness and co-occurrence of closely related species. There are only two short episodes in the very recent past dominated by distantly related taxa. We found distinct changes in phylogenetic community structure upon speciation or extinction events and an increase in mean community relatedness over time.

Main conclusions: Our finding of closely related co-occurring species implies environmental filtering as the primary assembly mechanism, with a minor but increasingly important role of competition towards the present, driven by evolutionary dynamics. Such an increase in the relative contribution of competition to the assembly of communities in relation to the aging of an insular ecosystem, together with a denser

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Global Ecology and Biogeography* published by John Wiley & Sons Ltd.

Foundation, Grant/Award Number:
PCEFP3_187012

Handling Editor: Petr Keil

packing of morphospace in the early phase of system ontogeny is compatible with ecological predictions according to the theory of island biogeography.

KEYWORDS

community assembly, diatoms, environmental change, functional richness, in situ evolutionary processes, insular ecosystems, molecular phylogenetics, trait evolution

1 | INTRODUCTION

Patterns of species diversity are strongly influenced by complex interactions between deterministic assembly processes such as competition and environmental filtering (Cavender-Bares et al., 2009; Webb et al., 2002). Our current understanding of the relative importance of these deterministic processes is largely derived from inferences in natural communities across spatial scales (Vamosi et al., 2009; Weber et al., 2017) or from experimental settings (Burns & Strauss, 2011; Narwani et al., 2013). However, very little is known about how these processes have changed over longer periods of time (e.g. Chang & Skipwith, 2021; Raia, 2010) and in response to geological, environmental, and/or climatic alterations (Mittelbach & Schemske, 2015; Stoof-Leichsenring et al., 2020). This is in part due to the lack of suitable model systems that enable the inference of temporal changes in deterministic assembly processes.

Insular ecosystems that have persisted for considerable time periods – such as oceanic islands, mountain tops, or ancient lakes – have long been identified as ideal settings to scrutinize in situ evolutionary and ecological patterns and to infer the processes underlying community assembly (Emerson & Gillespie, 2008; Losos & Ricklefs, 2009). Ancient lakes provide particularly powerful systems for studying changes in assembly processes over time as: (a) they can be hundreds of thousands to several millions of years old (Salzburger et al., 2014; Schön & Martens, 2004), enabling the reconstruction of assembly dynamics over long evolutionary time-scales (e.g. Wilke et al., 2016); (b) many of these lakes have high quality sediment records with high temporal resolution and good chronological control (e.g. Johnson et al., 2016; Wagner et al., 2019), facilitating the reconstruction of past environmental and climatic conditions; (c) their sediments often contain a well-preserved fossil record enabling the inclusion of extinct taxa in community assembly analyses through time (e.g. Fritz et al., 2013); (d) the relatively high degree of geographical and ecological isolation of ancient lakes limits immigration from regional species pools, allowing for a better understanding of the contribution of in situ evolutionary processes (e.g. Emerson & Gillespie, 2008; Fritz et al., 2013; Rominger et al., 2016); and (e) many of these lakes contain exceptionally taxonomically and phenotypically diverse radiations, so-called species flocks (e.g. Martens, 1997; Salzburger et al., 2014).

Ancient Lake Ohrid on the Balkan Peninsula is Europe's oldest lacustrine ecosystem with a high degree of endemism (Föller et al., 2015). The lake basin may have started to accumulate sediments approximately 1.9 million years ago (Ma) and has continuously

carried water for the past 1.4 Myr (Lindhorst et al., 2015; Wagner et al., 2017). In 2013, Lake Ohrid became the subject of a drilling project in the framework of the International Continental Scientific Drilling Program (ICDP), resulting in a continuous sediment record with a well-constrained chronology for the entire lacustrine phase (Wagner et al., 2019). The sediment succession contained an undisturbed and well-preserved fossil record of planktonic diatoms (i.e. single-celled siliceous algae; Cvetkoska et al., 2021). Previous studies focusing on the Late Pleistocene period have demonstrated that these photosynthetic organisms respond strongly to changes in key growth-limiting environmental and climatic factors related to nutrients and light availability (Jovanovska et al., 2016; Reed et al., 2010; Zhang et al., 2016). Although it has been suggested that the contemporary diatom community structure in Lake Ohrid is strongly driven by water depth and water chemistry (Cvetkoska et al., 2018), our recent work has shown that species competition intensified among endemic diatoms after the lake reached long-term environmental stability at around 1.15 Ma (Wilke et al., 2020). Both environmental filtering and biotic competitive interactions can thus play an important role in diatom community assembly; however, the contribution of the two processes over time remains poorly understood. Studies across spatial and shorter temporal scales have shown that either intraspecific resource-based competition (e.g. Carney et al., 1988) or environmental filtering (e.g. Stoof-Leichsenring et al., 2020) play a key role in structuring planktonic diatom communities. Island biogeography theory postulates that competition intensifies over time in insular ecosystems, because communities increase in taxonomic richness through immigration of ecologically pre-adapted species and in situ diversification (Emerson & Gillespie, 2008; Rominger et al., 2016; Warren et al., 2015). Accordingly, species should pack densely into the occupied regions of ecological niche space in the earlier stages and transition towards expansion into previously unoccupied niche space as the ecosystem becomes older (Borko et al., 2021; Rominger et al., 2016; Ronco et al., 2021). The extent to which these changes occur largely depends on the taxa and the environment (Emerson & Gillespie, 2008; Stoof-Leichsenring et al., 2020).

Here, we used the planktonic diatoms of Lake Ohrid as a model group to test the interplay of deterministic assembly processes in the context of island biogeography theory. We hypothesize that the importance of environmental filtering decreases over time and will subsequently become replaced by biotic competitive interactions. Under this assumption, we expect to observe a switch from packing of species in the niche space to niche space expansion over time. We

assume that changes in the relative importance of these assembly processes are largely affected by environmental and climatic fluctuations. To test these hypotheses, we integrated a time-calibrated phylogeny with extant and extinct species, fossil occurrence data of relative species abundances, ecological and morphological traits of the species, and palaeoenvironmental predictors from the lake's sediment succession (Figure 1). With this combined dataset, we assessed the phylogenetic structure of assemblages and the distribution of phenotypic traits to address three specific objectives. First, we inferred the relative contribution of environmental filtering versus species competition in governing diatom assemblage over the entire limnological history of Lake Ohrid by quantifying the mean phylogenetic distance (MPD) between species of each assemblage along a dated phylogeny. The MPD quantifies whether more closely or more distantly related taxa co-exist in an assemblage than expected by chance, related to environmental filtering (i.e. phylogenetic clustering) or competition (i.e. phylogenetic overdispersion), respectively (Webb et al., 2002). Second, we evaluated morphospace occupation patterns, and whether morphospace expansion and packing contributed to the accumulation of diatom diversity through the history of Lake Ohrid. Therefore, functional richness over time and the relationship between species and functional richness were quantified. Third, we tested whether the relative importance of assembly processes varied through time in response to local environmental changes (e.g. basin development and lake-level fluctuations), long-term climatic variations (e.g. glacial-interglacial cycles), and/or intrinsic biological changes (e.g. speciation, extinction, and population dynamics). For this purpose, we used a regression analysis to examine the influence of various palaeoenvironmental predictors and intrinsic changes on phylogenetic community structure.

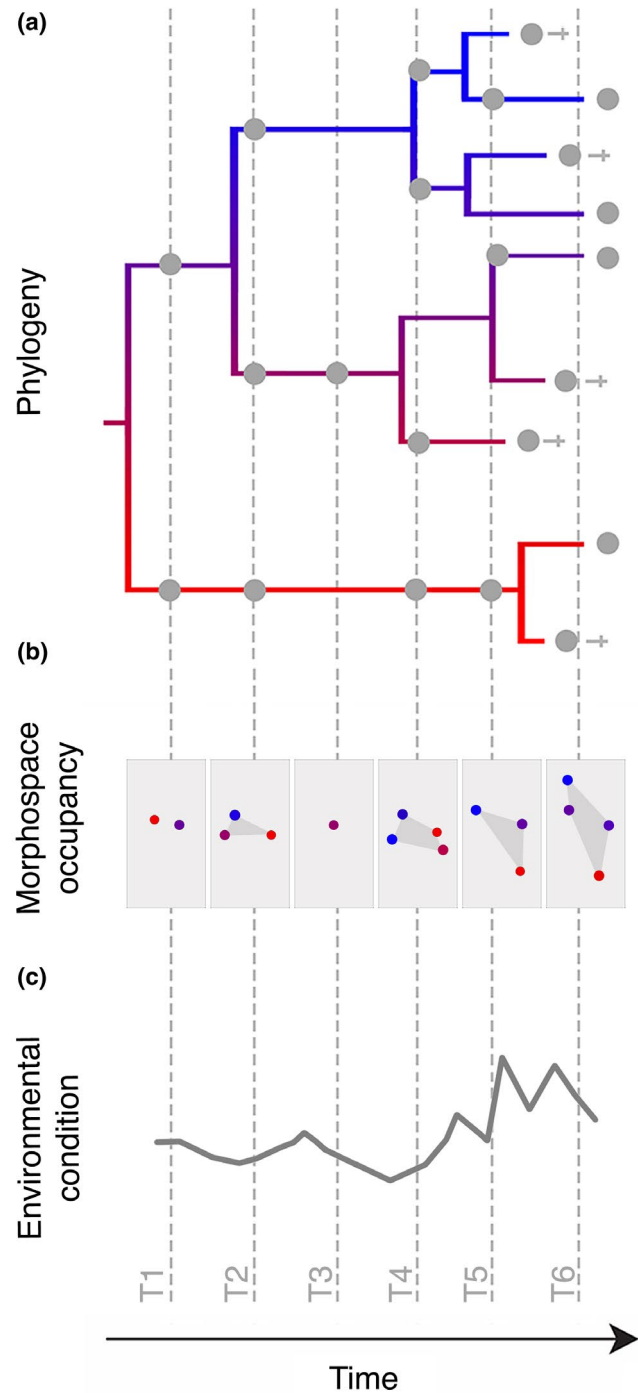
2 | MATERIALS AND METHODS

2.1 | Study site and sampling

2.1.1 | Diatom taxa in Lake Ohrid

Lake Ohrid (North Macedonia/Albania) is a tectonically formed oligotrophic lake with a tub-shaped morphology (Lindhorst et al., 2010, 2015; see Figure 2 and Supporting Information Appendix S1, Table S1). At least 201 extant and extinct endemic diatom species have been recorded for this lake, of which 184 are benthic and 17 planktonic (Wilke et al., 2020). With a total number of 873 diatom species, these algae arguably represent the most species-rich taxon in this lake (Levkov & Williams, 2011; Wilke et al., 2020).

For the present study, we focused on the dominant planktonic *Cyclotella sensu lato* (s.l.) group (Thalassiosirales), including the genera *Cyclotella sensu stricto* (s.s.), *Discostella*, *Pantocsekiella*, *Lindavia*, and *Cribrionella*, because planktonic diatoms respond sensitively to environmental and climatic changes (Saros & Anderson, 2015). This group is characterized by a rather low species richness, relatively high levels of endemism (Levkov et al., 2007) and a uniform distribution



in the nutrient-depleted open waters of Lake Ohrid between 0 and 80 m depth (Cvetkoska et al., 2018; Stanković, 1960). Hence, the silica frustule remains of *Cyclotella* s.l. are presumably evenly distributed over the lake's surface sediment (Vogel et al., 2010).

2.1.2 | Diatom sampling

We used 380 sediment samples that were obtained from the core DEEP-5045-1. The core is 584 m long and was taken from the central part of the lake basin at a water depth of 243 m (Francke

FIGURE 1 Community assembly through time as predicted by the theory of island biogeography, postulating that competition intensifies because communities increase taxonomic richness through immigration of pre-adapted species and in situ diversification as ecosystems age. (a) A hypothetical phylogeny including the occurrence of extinct (dagger symbol) and extant species (grey circles) and their traits (coloured branches in the phylogeny indicate trait values of a species). (b) Morphospace expansion of the assemblage through time (coloured dots represent different species). (c) Palaeoenvironmental conditions indicating potential environmental influence on community assembly. The time slices T1–T6 indicate phylogenetic community and trait structures, morphospace occupation and environmental conditions at given points of time. The time slices illustrate the expected change in community assembly from environmental filtering to species competition and of morphospace occupancy from packing to expansion due to ecosystem aging and environmental change.

et al., 2016; Wagner et al., 2014). Detailed descriptions of the core location and age-model construction are provided in Wagner et al. (2019). Our analyses are restricted to the upper 446.65 m composite depth (mcd), corresponding to 1.36 Ma, as this depth marks the onset of continuous lacustrine conditions at the drill site (Wagner et al., 2019). Diatom samples were taken at every 128 cm (corresponding to 2,000–4,000 years) between 0 and 406.96 mcd and every 64 cm (corresponding to c. 2,000 years) between 406.96 and 446.65 mcd. The higher sampling resolution during the early stage of Lake Ohrid was chosen because this lake phase was characterized by a particularly dynamic environment (Panagiotopoulos et al., 2020; Wilke et al., 2020).

Sediment samples were acid-cleaned following the method by Renberg (1990). The cleaned samples were mounted on glass microscope slides using Naphrax™ (Brunel Microscopes Ltd). For each slide, random transects were selected in which 200–400 diatom valves were counted and identified to the species level by two people that were trained in the same laboratory and used the same standards for species determination including cross-validation. Diatom counts were obtained using an Axioplan 2 microscope (Carl Zeiss) at 1,000× magnification or with a BX51 microscope (Olympus) at 1,500× magnification. We observed intraspecific morphological variability within species of *Cyclotella* s.l. (e.g. *Pantocsekiella ocellata*, *Cyclotella cavitata*, and *Cyclotella fottii*). Therefore, morphotypes, that is, entities that occurred for < 0.5 Myr and that showed highly variable morphological features, were combined to their nominal taxon. For a robust estimation of species richness and abundance, the number of valves was determined by performing a rarefaction analysis using the package 'vegan' 2.4.4 (Oksanen et al., 2017) for the R statistical environment 4.1.1 (R Core Team, 2021). All microscopic slides are stored at the Systematics and Biodiversity Collection of the University of Giessen, Germany (UGSB; Diehl et al., 2018).

In addition to fossil samples, fresh diatom material was obtained between 2015 and 2017 from the south-eastern part of Lake Ohrid as well as from its feeder springs near St. Naum, from neighbouring Lake Prespa, and from other water bodies in the lakes' surroundings (see Supporting Information Table S2). Samples were collected with a 5-µm mesh plankton net in the lakes' littoral and sublittoral zones

(0–60 m water depth) or from about 20 cm below the water surface in the springs (Figure 2). Samples were transferred alive to the laboratory for species determination and DNA isolation.

2.2 | MPD determination for assembly processes

2.2.1 | Molecular work and phylogenetic analyses

To assess phylogenetic relationships between different members of the planktonic diatoms in Lake Ohrid, we inferred a dated molecular phylogeny for Thalassiosirales on the basis of chloroplast (ribulose-1,5-bisphosphate carboxylase large subunit - *rbcl*) and nuclear (18S ribosomal RNA - 18S rRNA) DNA sequences obtained from GenBank and generated in the course of this study. In total, 23 specimens representing seven out of eight *Cyclotella* s.l. species from Lake Ohrid and Lake Prespa, as well as three species from the lakes' surroundings were newly sequenced (see Supporting Information Table S2). DNA was isolated following the single-cell DNA extraction method outlined by Hamilton et al. (2015). Prior to DNA extraction, each individual was photographed by light microscopy and voucher slides of oxidized specimens derived from the same samples as the ones used for DNA extraction were deposited in the UGSB collection. Standard and partly taxon-specific PCR primers and protocols were used to amplify the fragments of *rbcl* and 18S rRNA (for details see Supporting Information Appendix S2.1). The protein-coding *rbcl* sequences were aligned using the MULTIPLE SEQUENCE ALIGNMENT PROGRAM with default settings (Katoh & Standley, 2013). The 18S rRNA sequences, which have a pronounced secondary structure, were aligned using SSU-ALIGN 0.1.1 (Nawrocki, 2009).

For time-calibrating the phylogeny, seven fossil occurrences were used for node dating (see Supporting Information Tables S3, and S4; see also Alverson, 2014). We applied an offset that reflects the minimum age of each fossil and a gamma-distributed prior that covers its stratigraphic range. The phylogeny was inferred using an uncorrelated relaxed-clock model in BEAST 1.8.2 (Drummond et al., 2012). Four independent analyses were run on the cyberinfrastructure for phylogenetic research Science Gateway Portal (Miller et al., 2010) for 100 million generations sampling every 5,000th generation. Log and tree files of the independent runs were combined using LogCombiner 1.8.2 (BEAST package; 75% burn-in); the maximum clade credibility (MCC) tree was identified using TreeAnnotator 1.8.2 (BEAST package; no additional burn-in). The putatively extinct endemic species *Cyclotella cavitata*, *Cyclotella sollevata*, *Cyclotella* sp. 1, *Pantocsekiella* sp. 1, and *Cribrionella ohridana* were a posteriori incorporated in the time-calibrated molecular phylogeny based on their morphological similarities using the R package 'phytools' 0.7-80 (Revell, 2012). Branch lengths of the incorporated extinct species were defined relative to their stratigraphic range by using the time of the first occurrence in the lake sediment record as a maximum age (Figure 2). For the subsequent analyses related to trait evolution and phylogenetic community structure, we also placed a few *Pantocsekiella*

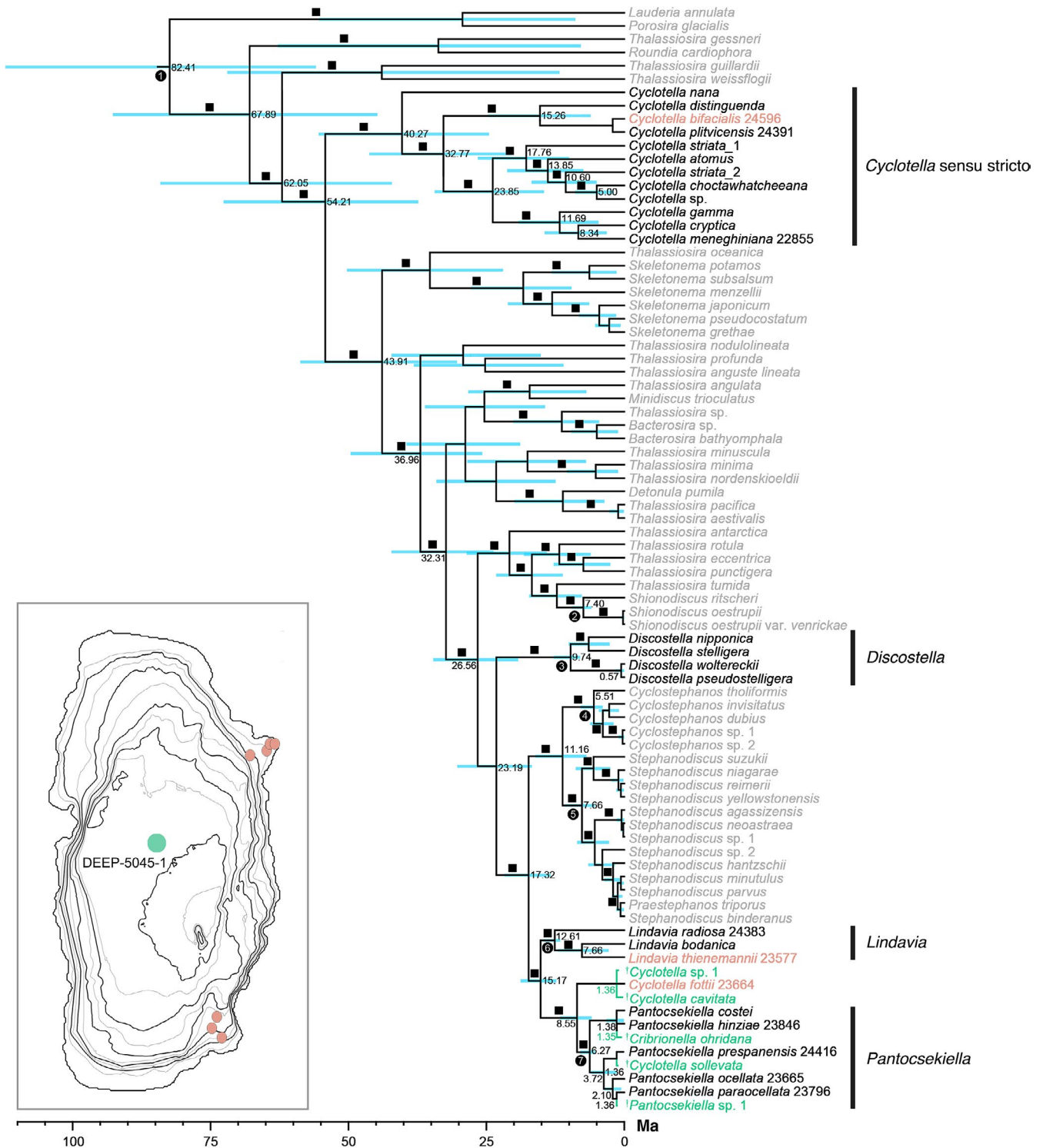


FIGURE 2 Time-calibrated phylogeny of Thalassiosirales based on seven fossil calibration points (numbered solid black circles, for details see Supporting Information Table S3). Numbers at nodes represent mean ages. Black squares on the branches show Bayesian posterior probabilities > 0.95. The blue bars represent 95% highest posterior densities of age estimates. *Cyclotella* sensu lato species are indicated in black. Extant and extinct *Cyclotella* s.l. endemic to Lake Ohrid are indicated in light pink and green, respectively. Other group members of the Thalassiosirales are indicated in grey. Inset map of Lake Ohrid showing the sampling locations of modern (light pink circles) and fossil (core DEEP-5045-1; green circle) specimens used in this study. Black lines indicate 50-m isobaths and grey lines indicate 25-m isobaths.

species that lack genetic information (*P. minuscula*, *P. delicatula*, *P. rossii*, and *P. polymorpha*) in a pruned molecular phylogeny using available taxonomic information as a guide (Figure 3a). The equal

branch lengths were constrained to be > 1.36 Myr because the four species originated before the formation of Lake Ohrid (Wilke et al., 2020).

2.2.2 | Assessing trait evolution and deterministic assembly processes

To infer the relative importance of species competition and environmental filtering, we assessed the phylogenetic structure of diatom communities (i.e. whether closely related species co-occur) and evaluated whether closely related species share similar traits (i.e. a high phylogenetic signal). Trait evolution in *Cyclotella* s.l. species was assessed via quantifying the phylogenetic signal using Pagel's λ (Pagel, 1999). This test indicates whether closely related species have similar trait values (i.e. $\lambda \approx 1$) or whether traits diverged so rapidly that the phylogenetic signal is concealed (i.e. $\lambda \approx 0$). Four continuous and nine categorical traits, representing morphological and ecological characteristics of taxonomic and/or functional importance, were scored (Figure 3b, Supporting Information Table S5). Utilizing the Akaike information criterion (AIC) to compare the fit (i.e. Δ AIC) between models of trait evolution with λ -transformed branch lengths versus models with untransformed branches provided the statistical support for the phylogenetic signal. Brownian motion models for continuous traits and Markov models with equal transition rates for the categorical traits were fitted using the R package 'GEIGER' 2.0.6 (Harmon et al., 2008).

The phylogenetic structure for each of the 380 diatom communities sampled at different time points in the past was quantified by calculating the MPD between species of a community along the dated phylogeny. A null model of random phylogenetic community structure assuming even species richness was compared to the observed communities to quantify the standardized effect size (SES) of the observed MPD. This measure evaluates whether a species community is a result of phylogenetic clustering (values < 0) or overdispersion (values > 0), reflecting the presence of either more closely or more distantly related species within the community than expected by the null model. For each time point, the null model involves shuffling of species identities 999 times across the tips of the phylogeny with probabilities proportional to their abundances in the previous time step. Standard null models of shuffling abundances in the community matrix

may be inappropriate here, because they do not yet consider the time structure of the data. Branch lengths were pruned to the age of the assemblage and the pairwise phylogenetic distance was weighted by the relative abundance of the respective species pair. The MPD was quantified with the R package 'picante' 1.6.2 (Kembel et al., 2010).

2.3 | Assessing morphospace occupation patterns

To test whether morphospace expands in a given assemblage and is more packed with higher species richness, we estimated functional richness over time and the correlation between functional richness and species richness. We quantified the functional richness index FRic (Villéger et al., 2008) based on all 13 traits for the same 380 communities, using the R package 'FD' 1.0–12 (Laliberté et al., 2014). The actual values of FRic were compared with the same null model as the observed MPD, resulting in the standardized effect size (SES.FRic). This allowed us to determine whether the phylogenetic community structure and species richness was associated with small (< 0) or large (> 0) occupancy of morphospace.

To evaluate the effect of taxon sampling on the species richness trajectory, we performed a rarefaction analysis. We randomly sampled 100 individuals from the entire diatom assemblage at a moment in time (extracted from Wilke et al., 2020), determined the number of *Cyclotella* s.l. species and repeated the random subsampling 1,000 times to calculate the mean richness and its variability over time.

2.4 | Local palaeoenvironmental, climatic, and intrinsic biological effects on phylogenetic community structure

To assess whether and to what extent local palaeoenvironmental, climatic, and intrinsic biological changes affect assembly processes, we used a regression analysis to examine the relationship of phylogenetic community

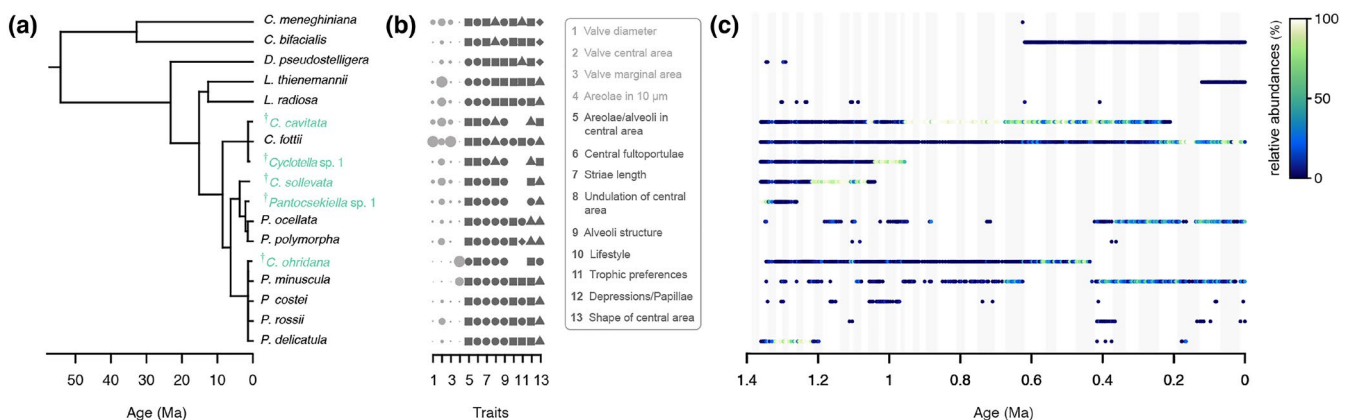


FIGURE 3 Trait evolution and abundancies of Lake Ohrid *Cyclotella* sensu lato species over time. (a) Dated phylogeny including extinct endemic species from Lake Ohrid (for details see Figure 2). (b) Light-grey symbols show trait expression with circle size proportional to the values of continuous traits and dark-grey symbols (i.e. triangle, square, and circle) the states of categorical traits. (c) Coloured dots represent relative abundances of *Cyclotella* s.l. species found in DEEP-5045–1 sediment record; vertical grey and white bars indicate glacial and interglacial periods, respectively.

structure (response variable) with three sets of predictor variables: local palaeoenvironmental changes, climatic changes, and intrinsic changes.

2.4.1 | Local palaeoenvironmental predictors

We derived the following local palaeoenvironmental predictors from the sediment record: total inorganic carbon (TIC), total organic carbon (TOC), potassium (K) counts from x-ray fluorescence scanning, relative sedimentary quartz content, percentages of arboreal pollen excluding *Pinus* pollen, percentage of pollen from deciduous oaks (Wagner et al., 2019), oxygen isotope composition of the lake water ($\delta^{18}\text{O}_{\text{lakewater}}$), and grain size (Wilke et al., 2020). In order to reduce the noise of these predictors in the regression analysis, we used Gaussian process smoothing, suggested by Simpson (2018) for sediment records. We used the implementation in the R package 'brms' 2.15.0 (Bürkner, 2018), which itself is an interface to the Bayesian inference machinery of stan 2.21 (Stan Development Team, 2021).

For the smoothing models, we specified the distribution according to the data type of the local palaeoenvironmental predictors (for details on the distribution see Supporting Information Appendix S4). To account for variations in sedimentation rate, we assigned weights to each sample of the palaeoenvironmental predictor representing the time captured by the sample (Simpson, 2018). All smoothing models were inferred through Hamiltonian Monte Carlo sampling with four parallel chains (details in Supporting Information Table S6). Residuals were tested for autocorrelation using the Moran's *I* index for irregularly spaced time series (see Legendre & Gauthier, 2014). Through this process we achieved the same temporal resolution for fossil and local palaeoenvironmental data. We then matched palaeoenvironmental predictors with MPD via their age.

2.4.2 | Climatic predictors

In addition to the predictors derived from the sediment record, we included the following climatic predictors in the regression analysis: the global benthic $\delta^{18}\text{O}$ isotope stack LR04 (Lisiecki & Raymo, 2005), $\delta^{18}\text{O}$ planktonic isotope ratios - Medstack (Wang et al., 2010), and Northern Hemisphere summer insolation at the latitude of Lake Ohrid (i.e. 41° N). The latter was calculated with the R package 'palin-sol' 0.93 (Crucifix, 2016) according to equations provided by Laskar et al. (2004). To match the MPD values with the age of climatic predictors, LR04 and Medstack were linearly interpolated, whereas the summer insolation was calculated with the ages of the assemblages.

2.4.3 | Intrinsic biological predictors

To assess the impact of intrinsic biological changes (e.g. speciation, extinction, and population dynamics) on assembly processes and to

avoid spurious effects of local palaeoenvironmental and climatic predictors, we included time as a predictor and allowed for discrete shifts in MPD at a given moment in time. The number of shifts was estimated with a change point analysis in the R package 'ecp' 3.1.0 (James & Matteson, 2015). The same number of shifts or fewer (depending on whether time shifts could be explained by palaeoenvironmental and/or climatic predictors) were allowed in the regression analysis. Ages of the shifts were not fixed as these might change when inferred together with palaeoenvironmental and/or climatic influence on MPD.

2.4.4 | Response of deterministic assembly processes to local palaeoenvironmental, climatic, and intrinsic biological predictors

For our regression model, we excluded colinear local palaeoenvironmental and climatic predictors from the analysis (Pearson correlation > |.7|) and retained the following predictors: grain size, isotope composition, potassium, TOC, deciduous oaks, Medstack, insolation and time and time shifts as biological predictors. The model structure is as follows:

$$MPD_{\tau} = \beta_1 X_{1,\tau} + \dots + \beta_p X_{p,\tau} + \beta_{S_1} \begin{cases} 1 & \text{if } \tau \leq S_1 \\ 0 & \text{if } \tau > S_1 \end{cases} + \dots + \beta_{S_k} \begin{cases} 1 & \text{if } \tau \geq S_k \wedge \tau < S_{k+1} \\ 0 & \text{if } \tau < S_k \vee \tau \geq S_{k+1} \end{cases} + \beta_{S_7} + \epsilon_{\tau} \quad (1)$$

where MPD_{τ} denotes the SES of the mean phylogenetic community distance at time τ , $\beta_p X_{p,\tau}$ the linear effect of the p th local palaeoenvironmental, climatic, or intrinsic predictor on MPD, β_k the parameters for the effects of the k th time shift S (with the maximum of $k = 6$), β_{S_7} the intercept, and ϵ_{τ} indicates the model error. The notation β_{S_7} was used for the intercept because the time most distant in the past was defined as the reference level (i.e. from 1.36 Ma until the earliest shift) and all effects β of the other shifts are relative to this level. Local palaeoenvironmental, climatic, and intrinsic predictors were centred to 0 and scaled to a variance of 1 to compare the strength of their influences on MPD (i.e. β).

We use Bayesian inference to obtain all parameters of the regression model through similar Hamiltonian Monte Carlo sampling using the 'brms' package (details in Supporting Information Appendix S5, Table S7). The influence of local palaeoenvironmental and climatic predictors on MPD was quantified by calculating a partial R^2 , which is defined as the Bayesian R^2 (Gelman et al., 2019) of the full regression model minus the R^2 of a model without the effect of these environmental predictors.

3 | RESULTS

3.1 | Temporal evolution of Lake Ohrid *Cyclotella* sensu lato

The time-calibrated phylogeny of Thalassiosirales revealed highly supported interspecific relationships for all major clades (posterior probabilities, PP > 0.95) and further indicated that the endemic *Cyclotella* s.l. species from Lake Ohrid form a paraphyletic group (Figure 2). Divergence time estimates suggested an age of the most recent common ancestor (MRCA)

for *Cyclotella* s.s. of c. 40 Myr, *Discostella* c. 10 Myr, *Lindavia* c. 13 Myr, and *Pantocsekiella* c. 6 Myr. The endemic *Cyclotella fottii* was revealed as sister to *Pantocsekiella* (PP = 1.0), which diverged about 8 Ma (Figure 2). The MRCA of the remaining endemic species *Cyclotella bifacialis* and *Lindavia thienemannii* and their respective sister groups probably originated in the Pliocene (c. 4 Ma) and Miocene (c. 7 Ma). The sister relationship between the endemic *L. thienemannii* and the non-endemic *Lindavia bodanica* is highly supported (PP = 1.0), whereas those between the endemic *C. bifacialis* and the non-endemic *Cyclotella distinguenda* and *Cyclotella plitvicensis* remain uncertain (PP = 0.60). The first occurrence of the endemic *C. fottii*, *C. bifacialis*, and *L. thienemannii* in the DEEP-5045-1 sediment record is estimated at c. 1.32, c. 0.62, and c. 0.12 Ma, respectively (Figure 3c).

3.2 | Relative importance of deterministic assembly processes

According to Pagel's λ and the phylogenetic model fit comparison, a strong phylogenetic signal was found for six ecological and morphological traits (e.g. areolae/alveoli in central area, central

fultoportulae, striae length, and undulation of central area), whereas a weak signal was found for four traits and no phylogenetic signal for three (valve central area, valve marginal area, and areolae in 10 μm) out of the total of 13 traits (Table 1).

The SESs of the MPD between species of the communities were mostly below zero, suggesting phylogenetic clustering. We found four distinct periods of phylogenetic clustering during the last 1.36 Myr (greenish dots in Figure 4a; period 1: 1.36–0.61 Ma, period 2: 0.42–0.17 Ma, period 3: 0.15–0.04 Ma, period 4: 0.01–0 Ma). In the time period 0.62–0.42 Ma, the phylogenetic community structure did not deviate substantially from zero, suggesting a random assembly scenario (blue dots in Figure 4a). The strongest overdispersions (MPD above zero) occurred between 0.17–0.14 and 0.03–0.01 Ma (reddish dots in Figure 4a).

3.3 | Temporal patterns of morphospace occupation

The observed *Cyclotella* s.l. species richness decreased until c. 0.9 Ma (Supporting Information Figure S1a). In contrast to the

TABLE 1 Phylogenetic signal in different ecological and morphological traits

Trait (mean)	Brownian motion				Pagel's λ transformation					
	σ^2	λ	Log-likelihood	AIC	σ^2	λ	log-likelihood	AIC	ΔAIC	Interpretation
1 Valve diameter	0.023	1.00	-23.44	50.88	0.004	.00	-22.05	50.10	0.77	WPS
2 Valve central area	0.064	1.00	-20.38	44.76	0.005	.00	-14.66	35.31	9.45	NPS
3 Valve marginal area	0.052	1.00	-22.64	49.27	0.004	.00	-14.30	34.60	14.67	NPS
4 Areolae in 10 μm	1.865	1.00	-43.87	91.78	0.064	.00	-32.73	71.46	20.29	NPS
	Markov model				Pagel's λ transformation					
	q	λ	Log-likelihood	AIC	q	λ	Log-likelihood	AIC	ΔAIC	
5 Areolae/alveoli in central area	0.050	1.00	-8.92	19.84	0.027	.95	-8.57	21.14	-1.30	PS
6 Central fultoportulae	0.079	1.00	-9.48	20.96	0.065	.99	-9.42	22.84	-1.88	PS
7 Striae length	0.008	1.00	-3.98	9.96	0.008	1.00	-3.98	11.96	-2.00	PS
8 Undulation of central area	0.041	1.00	11.88	25.76	0.041	1.00	11.88	27.76	-2.00	PS
9 Alveoli structure	0.018	1.00	-5.35	12.70	0.018	1.00	-5.35	14.70	-2.00	PS
10 Lifestyle	0.005	1.00	-4.12	10.33	0.002	.00	-3.44	10.88	-0.55	WPS
11 Trophic preferences	0.683	1.00	-15.83	37.66	0.107	.00	-14.89	37.78	-0.12	WPS
	0.327				0.015					
	10.37				0.022					
12 Depressions/papillae	0.118	1.00	-16.08	34.16	0.005	.00	-15.22	34.43	-0.27	WPS
13 Shape of central area	0.024	1.00	-18.52	39.04	0.004	.68	-16.08	36.15	2.89	MPS

Abbreviations: AIC, Akaike information criterion; NPS, no phylogenetic signal; WPS, weak phylogenetic signal; MPS, medium phylogenetic signal; PS, phylogenetic signal; σ^2 , rate of continuous trait evolution; q , transition rate between categorical traits. Phylogenetic signal (Pagel's λ) is the length transformation of phylogenetic branches with $\lambda = 1$ indicating no transformation and that all trait divergence occurs during anagenesis, whereas $\lambda = 0$ points to trait divergence only at speciation events. ΔAIC expresses the difference in fit between models of trait evolution without and with branch length transformation.

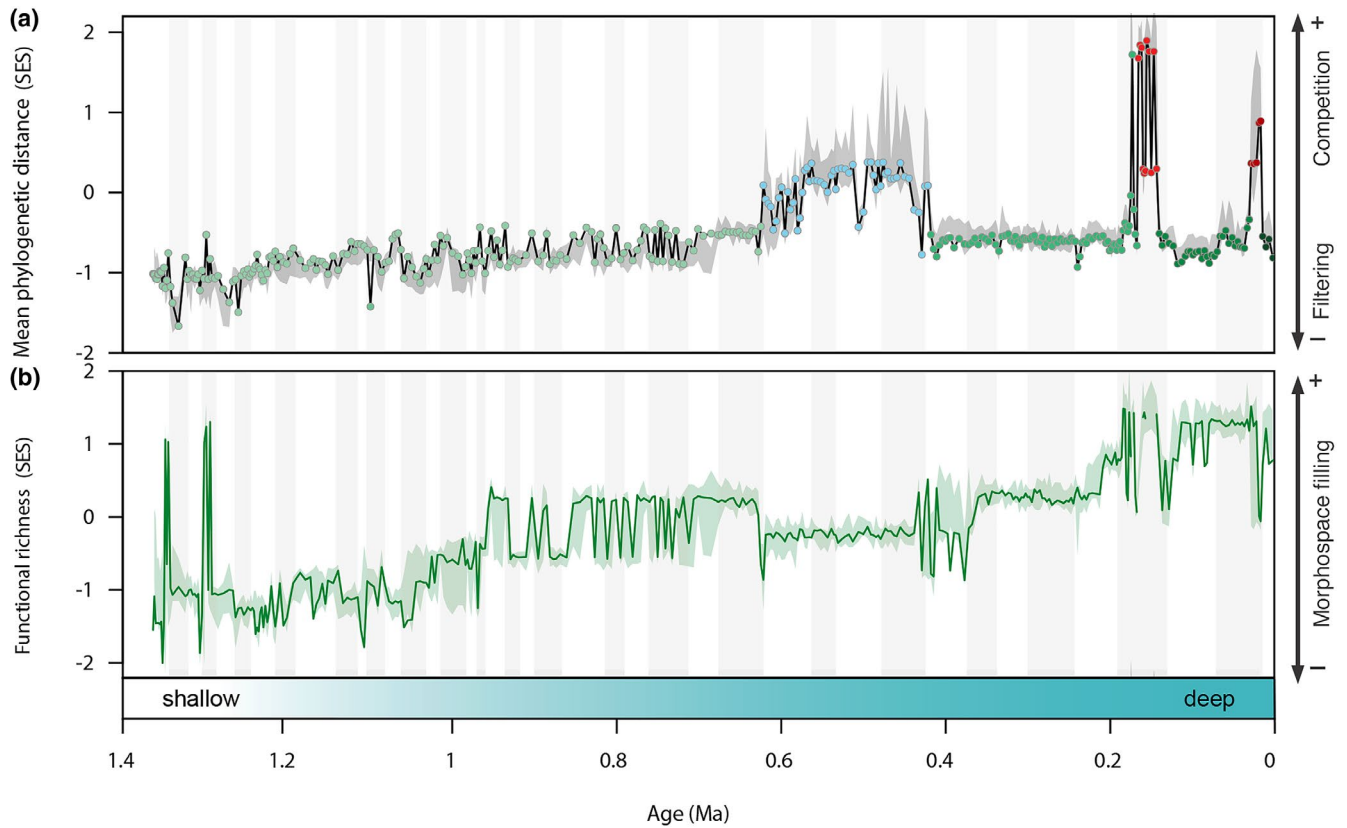


FIGURE 4 Phylogenetic community structure and functional richness in Lake Ohrid over time. (a) Phylogenetic community structure among *Cyclotella sensu lato* species; greenish dots showing distinct periods of environmental filtering, reddish dots phases of species interaction, and blue dots phase of random assembly. The shaded polygon indicates the effect of taxon sampling. (b) Green curve shows the accumulation of functional richness, indicating small (< 0) or large (> 0) occupancy of niche space. The shaded polygon displays the 95% range for a rarefied taxon sampling. Vertical grey and white bars indicate glacial and interglacial periods, respectively. Blue bar at the bottom of the figure indicates shallow- and deep-water phases of Lake Ohrid.

decreasing richness through time, the SES of the FRic (i.e. size of the occupied morphospace) progressively increased and reached a plateau at c. 0.6 Ma with a few exceptional positive spikes thereafter (Figure 4b), which is congruent with the phylogenetic community structure (Figure 4a). The negative correlation between species richness and SES.FRic (Supporting Information Figure S1b,c) indicates a denser packing of species inside the morphospace with higher species richness.

3.4 | Deterministic assembly processes in response to local palaeoenvironmental, climatic, and intrinsic biological changes

The regression model demonstrated that the mean phylogenetic distance of the community (SESs of the MPD) is mostly affected by intrinsic biological predictors ($R^2 = 0.765$; Supporting Information Figure S2; Figure 5a). The MPD linearly increased towards the present and shifted six times over the past 1.36 Myr (Figure 5a). Of the seven palaeoenvironmental predictors examined, the regression model indicated only a minor influence (partial $R^2 = 0.004$; limited influence of priors on inferred effects; Supporting Information

Figures S3–S5) of two local palaeoenvironmental predictors on the phylogenetic community structure: grain size and $\delta^{18}\text{O}_{\text{lakewater}}$ (Supporting Information Figure S6, Table S8). Response analyses on the individual local palaeoenvironmental predictors revealed a higher MPD with decreasing grain size and increasing $\delta^{18}\text{O}_{\text{lakewater}}$ (Figure 5b,c).

4 | DISCUSSION

4.1 | Relative importance of deterministic assembly processes

Following the prediction of island biogeography theory on community assembly, we sought to answer by which processes communities assemble over an extended time period after formation of a long-lived lake ecosystem. The strong phylogenetic signal revealed for most traits studied (Figure 3b, Table 1) suggests phylogenetic clustering of communities coincides with the co-existence of species with similar functional traits and that environmental filtering is the prevailing mechanism over the entire 1.36-Myr history of the lake (greenish dots in Figure 4a). The observed increase of mean

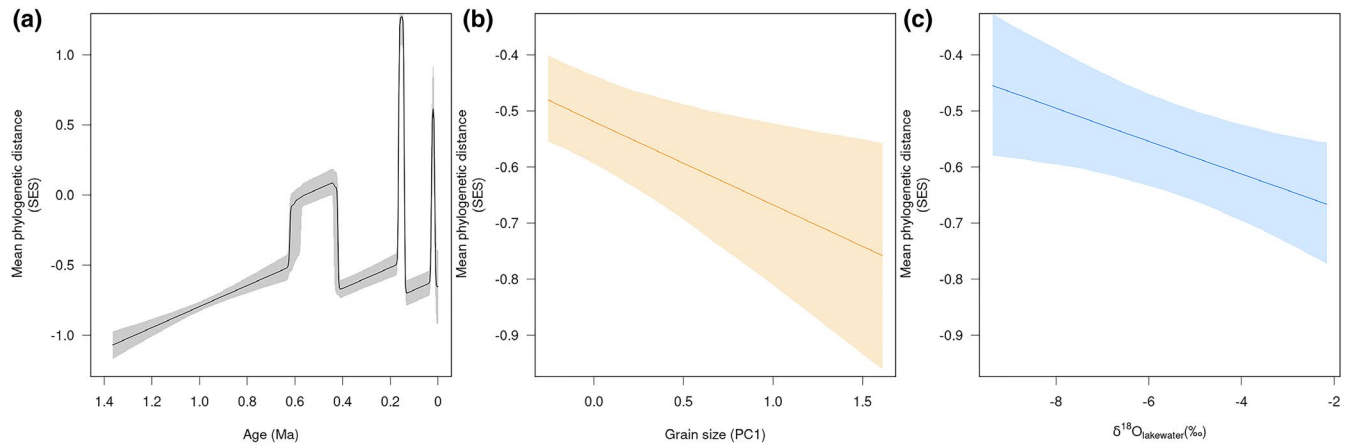


FIGURE 5 Response plots showing the relationship between phylogenetic community structure and (a) time with a linear increase and discrete shifts (i.e. intrinsic population dynamics of communities); predicted by the regression model, (b) grain size (high SES MPD = species competition; low SES MPD = environmental filtering), (c) $\delta^{18}\text{O}_{\text{lakewater}}$ (high SES MPD = species competition; low SES MPD = environmental filtering). The individual response plots were generated by keeping the other variables at their mean levels. The regression line is depicted by a solid line and the 95% credible intervals by shaded polygons. x-axis values were back-transformed to the observed scale after using values centered to a mean of 0 and scaled to a variance of 1 for the analysis. SES, standardized effect size; MPD, mean phylogenetic distance; PC1, principal component 1.

phylogenetic relatedness over time (Figure 5a) supports the prediction that competition should become increasingly important as the ecosystem ages (Emerson & Gillespie, 2008; Rominger et al., 2016; Warren et al., 2015). However, as the latter process dominated only in two relatively short time periods between 0.17–0.14 and 0.03–0.01 Ma (reddish dots in Figure 4a), it does not overturn the general prevalence of environmental filtering. A possible explanation for the slow change in community assembly processes towards increasing competition is the higher water depth during the interglacials as compared to the glacial periods and the ongoing deepening/widening of the lake at the early stage of its ontogeny, which in turn affected the available niche space. This interplay mitigated competition and therefore facilitated species' coexistence. Our findings are in agreement with those from previous empirical community assembly studies that suggest environmental filtering to be a key mechanism in shaping microbial, animal, and plant communities across spatial (Vamosi et al., 2009) and temporal scales (Pineda-Munoz et al., 2021; Raia, 2010; Stoof-Leichsenring et al., 2020).

The strong community–environment relationship in the Lake Ohrid diatoms might be due to a combination of factors, including environmental changes, in situ evolutionary processes, and the high dispersal abilities of planktonic diatoms. The latter are likely to have enabled non-endemic species to recolonize Lake Ohrid after local extinctions. Besides the environmentally very dynamic early lake phase when shallow water conditions prevailed (e.g. Panagiotopoulos et al., 2020), Lake Ohrid was also subject to environmental and climatic changes throughout its history, including glacial–interglacial cycles, volcanic ash depositions, earthquakes, lake-level fluctuations, and landslides, though no massive stressor events on taxa are documented (Wagner et al., 2017, 2019). All these factors influenced the lake and its biota via changes in physical-chemical conditions such as temperature, light, precipitation, and nutrient availability (Wagner

et al., 2017). The changing environmental conditions enabled and necessitated the co-existence of species that share common traits to proliferate under certain ecological conditions, probably through reduced competition, specific growth strategies, or efficient resource use. The long co-existence of closely related species such as *C. cavitata* and *C. fottii* might also be facilitated by the long-term decline of in situ speciation and extinction rates, and a transition to stable communities (Wilke et al., 2020). In contrast, the disappearance of some closely related species from the record during the last two glacial periods (see reddish dots in Figure 4a) can be explained by intensified competition (e.g. between *C. cavitata* and *C. fottii*) as a result of reduced nutrient supply due to the unusually dry and cold conditions during these time periods (Sadori et al., 2016; Wagner et al., 2017). Consequently, the few planktonic species present during these glacial periods were phylogenetically distinct and occupied different areas of the niche space (Figure 4b).

4.2 | Temporal patterns of morphospace occupation

Deterministic ecological processes affect variation in species richness and, hence, in functional richness in a given assemblage (Neves et al., 2019). Environmental filtering, for example, has been shown to decrease functional differences among species and lead to a more densely packed morphospace (e.g. Collar et al., 2014). While previous studies have shown that packing is the dominant pattern of morphospace evolution (McClain, Barry & Webb, 2018; McClain, Nunnally et al., 2018; Pigot et al., 2016), a change in the mode of morphospace occupancy has been identified to operate across evolutionary time-scales in birds (Cooney et al., 2017) and cichlid fishes (Ronco et al., 2021). Our findings of an ongoing expansion

contradict these studies and our prediction that species packing will be replaced by expansion over time. Niche space packing and environmental filtering are especially pronounced until c. 0.6 Ma when species richness is high, whereas the slowly increasing importance of competition over time (with two pronounced peaks during the last two glacial periods) resulted in a decrease of species richness via extinction and an unpacking of niche space towards the present (Figure 4b; Supporting Information Figure S1a). The decline in species richness over time, together with the negative relationship between species richness and occupied morphospace, suggests that niche space was indeed more densely packed at the initial stages of the ecosystem development as predicted by the island biogeography theory (Figure 4a; Supporting Information Figure S1a). This aligns well with the macroevolutionary trend for all Lake Ohrid endemic diatoms, where short-lived species were replaced by long-lived species as the lake achieved long-term environmental stability at around 1.1 Ma (Wilke et al., 2020). The replacement was particularly pronounced in planktonic diatoms, probably because their traits make them more vulnerable to extinction compared to other diatoms (Wilke et al., 2020).

4.3 | Deterministic assembly processes in response to environmental, climatic, and intrinsic biological changes

Environmental conditions influence the processes that determine the co-existence of species in an assemblage, for instance, through influencing resource availability (Oliveira et al., 2020; Stoof-Leichsenring et al., 2020). Although we incorporated a number of key local palaeoenvironmental and climatic predictors into our regression models, we found that intrinsic biological processes related to speciation or (local) extinction primarily influenced the assembly processes (Figure 5a). Population dynamics as well as stochastic disturbances related to dispersal limitation and evolutionary constraints on traits associated with morphospace occupancy may have also contributed to the planktonic diatom community assembly in Lake Ohrid.

Besides intrinsic biological processes, we only found a weak effect of environmental variables related to lake size variation and precipitation/evaporation fluctuations on assembly processes (Figure 5b,c). The interplay between ecosystem maturation, lake size and depth (especially during glacial periods) probably affected community assembly of diatoms, as is evident by the increased competition towards the present (Figure 4a). This interplay triggered variations in water level and affected mixing processes, nutrient content, turbidity, and light availability in the water column (Cvetkoska et al., 2021; Wilke et al., 2020), which are important factors controlling diatom planktonic communities. Hence, we assume that the changes in assembly processes might have been additionally determined by some unmeasured environmental factors, potentially related to nutrients (in particular silica, phosphorus, and nitrogen), light intensity, pH, and water mixis. These factors, in particular the availability of phosphorus, have been shown to play a key role in

shaping the contemporary diatom community structure (Cvetkoska et al., 2018).

5 | CONCLUSIONS

Assembly of planktonic diatoms in Lake Ohrid was mainly determined by environmental filtering, whereas species competition dominated only in very short and recent time periods. Island biogeography theory postulates that competition becomes increasingly important over time. However, our findings indicate that environmental filtering plays a greater role in structuring communities across the whole time period, with a tendency towards increasing competition at later stages. We also show that morphospace expanded and became less densely packed over time as a consequence of the increased competition. The shifts in the relative role of the assembly processes are primarily attributed to evolutionary dynamics related to species diversification or population growth. In contrast, these shifts were only weakly affected by local palaeoenvironmental predictors. The integrative analysis of community, morphological, genetic, and palaeoenvironmental data allowed us to obtain a more mechanistic understanding of how communities assemble over time and may thus serve as a basis for investigating the dynamic role of different assembly processes in other insular systems.

ACKNOWLEDGMENTS

This study was supported by funding from the Deutsche Forschungsgemeinschaft (DFG - German Research Foundation, grant numbers JO 1731/1-1 to EJ, WI 1902/13 to TW, AL 1076/9 to CA, and STE 2460/2-1 and 2460/3-1 to BS. The Swiss National Science Foundation (grant no. PCEFP3_187012) supported TH. Computational resources were provided by the de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI). We thank Luc Ector, Hans-Peter Grossart, Matteo Lattuada, Marija Gligora Udovič, and Petar Žutinić for collecting and maintaining live specimens. We also thank Matt Ashworth and Linda Medlin for the very useful advice during the molecular work. We further thank Walter Salzburger for insightful comments that helped us improve the manuscript. Finally, we thank Alexander Francke and the entire Science Team of the Scientific Collaboration on Past Speciation Condition in Lake Ohrid (SCOPSCO) project for the continuous support with data sharing and interpretations. Open access funding enabled and organized by ProjektDEAL.

CONFLICT OF INTEREST

Authors declare no competing interests.

AUTHOR CONTRIBUTIONS

EJ and TH designed the study. EJ collected and identified live specimens. EJ and BS performed DNA lab work and conducted the phylogenetic analyses in collaboration with PH and CA. EJ and AC generated the diatom fossil data in collaboration with ZL and NOR. JL and BW provided the palaeoenvironmental variables in

collaboration with the SCOPSCO Science Team. TH performed the community structure analyses. KB helped with specific parts in the methods. EJ wrote the manuscript with support from TW and TH and input from all authors.

DATA AVAILABILITY STATEMENT

R code, diatom, biogeochemical, and molecular phylogeny data are available from the GitHub repository: <https://github.com/thauffe/OhridDiatomAssembly>. Genetic data is available at GenBank.

ORCID

Elena Jovanovska  <https://orcid.org/0000-0002-3413-3683>

Torsten Hauffe  <https://orcid.org/0000-0001-5711-9457>

Björn Stelbrink  <https://orcid.org/0000-0002-7471-4992>

Aleksandra Cvetkoska  <https://orcid.org/0000-0002-9812-8816>

Zlatko Levkov  <https://orcid.org/0000-0002-1184-2356>

Bernd Wagner  <https://orcid.org/0000-0002-1369-7893>

Jack H. Lacey  <https://orcid.org/0000-0002-6329-2149>

Paul B. Hamilton  <https://orcid.org/0000-0001-6938-6341>

Karen M. Brandenburg  <https://orcid.org/0000-0002-9342-3349>

Christian Albrecht  <https://orcid.org/0000-0002-1490-1825>

Thomas Wilke  <https://orcid.org/0000-0001-8263-7758>

REFERENCES

- Alverson, A. J. (2014). Timing marine-freshwater transitions in the diatom order Thalassiosirales. *Paleobiology*, *40*, 91–101. <https://doi.org/10.1666/12055>
- Borko, Š., Trontelj, P., Seehausen, O., Moškrič, A., & Fišer, C. (2021). A subterranean adaptive radiation of amphipods in Europe. *Nature Communication*, *12*, 3688. <https://doi.org/10.1038/s41467-021-24023-w>
- Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*, 395–411. <https://doi.org/10.32614/RJ-2018-017>
- Burns, J. H., & Strauss, S. Y. (2011). More closely related species are more ecologically similar in an experimental test. *Proceedings of the National Academy of Sciences USA*, *108*, 5302–5307. <https://doi.org/10.1073/pnas.1013003108>
- Carney, H. J., Richerson, P. J., Goldman, C. R., & Richards, R. C. (1988). Seasonal phytoplankton demographic processes and experiments on interspecific competition. *Ecology*, *69*, 664–678. <https://doi.org/10.2307/1941015>
- Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, *12*, 693–715. <https://doi.org/10.1111/j.1461-0248.2009.01314.x>
- Chang, L. M., & Skipwith, P. L. (2021). Relatedness and the composition of communities over time: Evaluating phylogenetic community structure in the late Cenozoic record of bivalves. *Paleobiology*, *47*, 301–313. <https://doi.org/10.1017/pab.2020.38>
- Collar, D. C., Wainwright, P. C., Alfaro, M. E., Revell, L. J., & Mehta, R. S. (2014). Biting disrupts integration to spur skull evolution in eels. *Nature Communications*, *5*, 5505. <https://doi.org/10.1038/ncomm56505>
- Cooney, C. R., Bright, J. A., Capp, E. J. R., Chira, A. M., Hughes, E. C., Moody, C. J. A., Nouri, L. O., Varley, Z. K., & Thomas, G. H. (2017). Mega-evolutionary dynamics of the adaptive radiation of birds. *Nature*, *542*, 344–347. <https://doi.org/10.1038/nature21074>
- Crucifix, M. (2016). *palinsol: Insolation for Palaeoclimate studies*. R package version 0.93. Available at: <https://CRAN.R-project.org/package=palinsol>
- Cvetkoska, A., Jovanovska, E., Hauffe, T., Donders, T. H., Levkov, Z., Van de Waal, D. B., Reed, J. M., Francke, A., Vogel, H., Wilke, T., Wagner, B., & Wagner-Cremer, F. (2021). Drivers of phytoplankton community structure change with ecosystem ontogeny during the Quaternary. *Quaternary Science Reviews*, *265*, 107046. <https://doi.org/10.1016/j.quascirev.2021.107046>
- Cvetkoska, A., Pavlov, A., Jovanovska, E., Tofilovska, S., Blanco, S., Ector, L., Wagner-Cremer, F., & Levkov, Z. (2018). Spatial patterns of diatom diversity and community structure in ancient Lake Ohrid. *Hydrobiologia*, *819*, 197–215. <https://doi.org/10.1007/s10750-018-3637-5>
- Diehl, E., Jauker, B., Albrecht, C., Wilke, T., & Wolters, V. (2018). GIEßEN: University collections: Justus Liebig University Gießen. In L. A. Beck (Ed.), *Zoological collections of Germany, natural history collections* (pp. 73–381). Springer.
- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*, 1969–1973. <https://doi.org/10.1093/molbev/mss075>
- Emerson, B. C., & Gillespie, R. G. (2008). Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution*, *23*, 619–630. <https://doi.org/10.1016/j.tree.2008.07.005>
- Föller, K., Stelbrink, B., Hauffe, T., Albrecht, C., & Wilke, T. (2015). Constant diversification rates of endemic gastropods in ancient Lake Ohrid: Ecosystem resilience likely buffers environmental fluctuations. *Biogeosciences*, *12*, 7209–7222. <https://doi.org/10.5194/bg-12-7209-2015>
- Francke, A., Wagner, B., Just, J., Leicher, N., Gromig, R., Baumgarten, H., Vogel, H., Lacey, J. H., Sadori, L., Wonik, T., Leng, M. J., Zanchetta, G., Sulpizio, R., & Giaccio, B. (2016). Sedimentological processes and environmental variability at Lake Ohrid (Macedonia, Albania) between 637 ka and the present. *Biogeosciences*, *13*, 1179–1196. <https://doi.org/10.5194/bg-13-1179-2016>
- Fritz, S. A., Schnitzler, J., Eronen, J. T., Hof, C., Böhning-Gaese, K., & Graham, C. H. (2013). Diversity in time and space: Wanted dead and alive. *Trends in Ecology and Evolution*, *28*, 509–516. <https://doi.org/10.1016/j.tree.2013.05.004>
- Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian regression models. *The American Statistician*, *3*, 307–309. <https://doi.org/10.1080/00031305.2018.1549100>
- Hamilton, P. B., Lefebvre, K. E., & Bull, R. D. (2015). Single cell PCR amplification of diatoms using fresh and preserved samples. *Frontiers in Microbiology*, *6*, 1084. <https://doi.org/10.3389/fmicb.2015.01084>
- Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., & Challenger, W. (2008). GEIGER: Investigating evolutionary radiations. *Bioinformatics*, *24*, 129–131. <https://doi.org/10.1093/bioinformatics/btm538>
- James, N. A., & Matteson, D. S. (2015). ecp: An R package for nonparametric multiple change point analysis of multivariate data. *Journal of Statistical Software*, *62*, 7. <https://doi.org/10.18637/jss.v062.i07>
- Johnson, T. C., Werne, J. P., Brown, E. T., Abbott, A., Berke, M., Steinman, B. A., Halbur, J., Contreras, S., Grosshuesch, S., Deino, A., Scholz, C. A., Lyons, R. P., Schouten, S., & Damsté, J. S. S. (2016). A progressively wetter climate in southern East Africa over the past 1.3 million years. *Nature*, *537*, 220–224. <https://doi.org/10.1038/nature19065>
- Jovanovska, E., Cvetkoska, A., Hauffe, T., Levkov, Z., Wagner, B., Sulpizio, R., Francke, A., Albrecht, C., & Wilke, T. (2016). Differential resilience of ancient sister lakes Ohrid and Prespa to environmental disturbances during the Late Pleistocene. *Biogeosciences*, *13*, 1149–1161. <https://doi.org/10.5194/bg-13-1149-2016>

- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P., & Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26, 1463–1464. <https://doi.org/10.1093/bioinformatics/btq166>
- Laliberté, E., Legendre, P., & Shipley, B. (2014). *FD: Measuring functional diversity from multiple traits, and other tools for functional ecology*. R package version 1.0-12. Available at: <https://cran.r-project.org/web/packages/FD/>
- Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A. C. M., & Levrard, B. (2004). A long-term numerical solution for the insolation quantities of the Earth. *Astronomy & Astrophysics*, 428, 261–285. <https://hal.archives-ouvertes.fr/hal-00001603>
- Legendre, P., & Gauthier, O. (2014). Statistical methods for temporal and space–time analysis of community composition data. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20132728. <https://doi.org/10.1098/rspb.2013.2728>
- Levkov, Z., Krstić, S., Metzeltin, D., & Nakov, T. (2007). Diatoms of Lakes Prespa and Ohrid. About 500 taxa from ancient lake system. In: H. Lange-Bertalot (Ed.), *Iconographia diatomologica* (Vol. 16, pp. 1–603). A. R. G. Gartner Verlag.
- Levkov, Z., & Williams, D. M. (2011). Fifteen new diatom (Bacillariophyta) species from Lake Ohrid, Macedonia. *Phytotaxa*, 30, 1–41. <https://doi.org/10.11646/phytotaxa.30.1.1>
- Lindhorst, K., Krastel, S., Reicherter, K., Stipp, M., Wagner, B., & Schwenk, T. (2015). Sedimentary and tectonic evolution of Lake Ohrid (Macedonia/Albania). *Basin Research*, 27, 84–101. <https://doi.org/10.1111/bre.12063>
- Lindhorst, K., Vogel, H., Krastel, S., Wagner, B., Hilgers, A., Zander, A., Schwenk, T., Wessels, M., & Daut, G. (2010). Stratigraphic analysis of lake level fluctuations in Lake Ohrid: An integration of high resolution hydro-acoustic data and sediment cores. *Biogeosciences*, 7, 3531–3548. <https://doi.org/10.5194/bg-7-3531-2010>
- Lisiecki, L. E., & Raymo, M. E. (2005). A Pliocene–Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records. *Paleoceanography*, 20, PA1003. <https://doi.org/10.1029/2004PA001071>
- Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, 457, 830–836. <https://doi.org/10.1038/nature07893>
- Martens, K. (1997). Speciation in ancient lakes. *Trends in Ecology & Evolution*, 12, 177–182. [https://doi.org/10.1016/S0169-5347\(97\)01039-2](https://doi.org/10.1016/S0169-5347(97)01039-2)
- McClain, C. R., Barry, J. P., & Webb, T. J. (2018). Increased energy differentially increases richness and abundance of optimal body sizes in deep-sea wood falls. *Ecology*, 99, 184–195. <https://doi.org/10.1002/ecy.2055>
- McClain, C. R., Nunnally, C., Chapman, A. S. A., & Barry, J. P. (2018). Energetic increases lead to niche packing in deep-sea wood falls. *Biology Letters*, 14, 20180294. <https://doi.org/10.1098/rsbl.2018.0294>
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)* (pp. 1–8). <https://doi.org/10.1109/GCE.2010.5676129>
- Mittelbach, G. G., & Schemske, D. W. (2015). Ecological and evolutionary perspectives on community assembly. *Trends in Ecology & Evolution*, 30, 241–247. <https://doi.org/10.1016/j.tree.2015.02.008>
- Narwani, A., Alexandrou, M. A., Oakley, T. H., Carroll, I. T., & Cardinale, B. J. (2013). Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. *Ecology Letters*, 16, 1373–1381. <https://doi.org/10.1111/ele.12182>
- Nawrocki, E. P. (2009). *Structural RNA homology search and alignment using covariance models* (Ph.D. thesis). Washington University in Saint Louis, School of Medicine. <https://doi.org/10.7936/K78050MP>
- Neves, K., Moura, M. R., Maravalhas, J., Maravalhas, J., Pacheco, R., Pie, M. R., Schultz, T. R., & Vasconcelos, H. L. (2019). Functional richness shows spatial scale dependency in *Pheidole* ant assemblages from Neotropical savannas. *Ecology and Evolution*, 9, 11734–11741. <https://doi.org/10.1002/ece3.5672>
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2017). *vegan: Community ecology package*. R package version 2.4-4. <http://CRAN.R-project.org/package=vegan>, last accessed 27-11-2019.
- Oliveira, B. F., Flenniken, J. M., Guralnick, R. P., Williams, S. E., & Scheffers, B. R. (2020). Historical environmental stability drives discordant niche filling dynamics across phylogenetic scales. *Journal of Biogeography*, 47, 807–816. <https://doi.org/10.1111/jbi.13798>
- Pagel, M. D. (1999). The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic Biology*, 48, 612–622. <https://doi.org/10.1080/106351599260184>
- Panagiotopoulos, K., Holtvoeth, J., Kouli, K., Marinova, E., Francke, A., Cvetkoska, A., Jovanovska, E., Lacey, J. H., Lyons, E. T., Buckel, C., Bertini, A., Donders, T., Just, J., Leicher, N., Leng, M. J., Melles, M., Pancost, R. D., Sadori, L., Tauber, P., ... Wilke, T. (2020). Insights into the evolution of the young Lake Ohrid ecosystem and vegetation succession from a southern European refugium during the Early Pleistocene. *Quaternary Science Reviews*, 227, 106044. <https://doi.org/10.1016/j.quascirev.2019.106044>
- Pigot, A. L., Trisos, C. H., & Tobias, J. A. (2016). Functional traits reveal the expansion and packing of ecological niche space underlying and elevational diversity gradient in passerine birds. *Proceeding of the Royal Society B: Biological Sciences*, 283, 20152013. <https://doi.org/10.1098/rspb.2015.2013>
- Pineda-Munoz, S., Wang, Y., Lyons, S. K., Toth, A. B., & McGuire, J. L. (2021). Mammal species occupy different climates following the expansion of human impacts. *Proceedings of the National Academy of Sciences USA*, 118, e1922859118. <https://doi.org/10.1073/pnas.1922859118>
- R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Raia, P. (2010). Phylogenetic community assembly over time in Eurasian Plio–Pleistocene mammals. *Palaios*, 25, 327–338. <https://doi.org/10.2110/palo.2009.p09-154r>
- Reed, J. M., Cvetkoska, A., Levkov, Z., Vogel, H., & Wagner, B. (2010). The last glacial–interglacial cycle in Lake Ohrid (Macedonia/Albania): Testing diatom response to climate. *Biogeosciences*, 7, 3083–3094. <https://doi.org/10.5194/bg-7-3083-2010>
- Renberg, I. (1990). A procedure for preparing large sets of diatom slides from sediment cores. *Journal of Paleolimnology*, 4, 87–90. <https://doi.org/10.1007/BF00208301>
- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3, 217–223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>
- Rominger, A. J., Goodman, K. R., Lim, J. Y., Armstrong, E. E., Becking, L. E., Bennett, G. M., Brewer, M. S., Cotoras, D. D., Ewing, C. P., Harte, J., Martinez, N. D., O'Grady, P. M., Percy, D. M., Price, D. K., Roderick, G. K., Shaw, K. L., Valdovinos, F. S., Gruner, D. S., Gillespie, R. G., & Ricklefs, R. (2016). Community assembly on isolated islands: Macroecology meets evolution. *Global Ecology and Biogeography*, 25, 769–780. <https://doi.org/10.1111/geb.12341>
- Ronco, F., Matschiner, M., Böhne, A., Boila, A., Büscher, H. H., El Taher, A., Indermaur, A., Malinsky, M., Ricci, V., Kahmen, A., Jentoft, S., & Salzburger, W. (2021). Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature*, 589, 76–81. <https://doi.org/10.1038/s41586-020-2930-4>
- Sadori, L., Koutsodendris, A., Panagiotopoulos, K., Masi, A., Bertini, A., Combourieu-Nebout, N., Francke, A., Kouli, K., Joannin, S., Mercuri,

- A. M., Peyron, O., Torri, P., Wagner, B., Zanchetta, G., Sinopoli, G., & Donders, T. H. (2016). Pollen-based paleoenvironmental and paleoclimatic change at Lake Ohrid (south-eastern Europe) during the past 500 ka. *Biogeosciences*, *13*, 1423–1437. <https://doi.org/10.5194/bg-13-1423-2016>
- Salzburger, W., Bocxlaer, B. V., & Cohen, A. S. (2014). The ecology and evolution of the African Great Lakes and their faunas. *Annual Review of Ecology, Evolution, and Systematics*, *45*, 519–545. <https://doi.org/10.1146/annurev-ecolsys-120213-091804>
- Saros, J. E., & Anderson, N. J. (2015). The ecology of the planktonic diatom *Cyclotella* and its implications for global environmental change studies. *Biological Reviews*, *90*, 522–541. <https://doi.org/10.1111/brv.12120>
- Schön, I., & Martens, K. (2004). Adaptive, pre-adaptive and non-adaptive components of radiations in ancient lakes: A review. *Organisms Diversity & Evolution*, *4*, 137–156. <https://doi.org/10.1016/j.ode.2004.03.001>
- Simpson, G. L. (2018). Modelling palaeoecological time series using generalised additive models. *Frontiers in Ecology and Evolution*, *6*, 149. <https://doi.org/10.3389/fevo.2018.00149>
- Stan Development Team. (2021). *Stan modeling language users guide and reference manual*, 2.21. <https://mc-stan.org>
- Stanković, S. (1960). The Balkan Lake Ohrid and its living world. In F. S. Bodenheimer, & W. W. Weisbach (Eds.), *Monographiae biologicae* (Vol. 9, pp. 357). Uitgeverij, Dr. W. Junk, Den Haag.
- Stoof-Leichsenring, K. R., Pestryakova, L. A., Epp, L. S., & Herzschuh, U. (2020). Phylogenetic diversity and environment form assembly rules for Arctic diatom genera—A study on recent and ancient sedimentary DNA. *Journal of Biogeography*, *47*, 1166–1179. <https://doi.org/10.1111/jbi.13786>
- Vamosi, S. M., Heard, S. B., Vamosi, J. C., & Webb, C. O. (2009). Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, *18*, 572–592. <https://doi.org/10.1111/j.1365-294X.2008.04001.x>
- Villéger, S., Mason, N. W. H., & Moullot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, *89*, 2290–2301. <https://doi.org/10.1890/07-1206.1>
- Vogel, H., Wessels, M., Albrecht, C., Stich, H.-B., & Wagner, B. (2010). Spatial variability of recent sedimentation in Lake Ohrid (Albania/Macedonia). *Biogeosciences*, *7*, 3333–3342. <https://doi.org/10.5194/bg-7-3333-2010>
- Wagner, B., Vogel, H., Francke, A., Friedrich, T., Donders, T., Lacey, J. H., Leng, M. J., Regattieri, E., Sadori, L., Wilke, T., Zanchetta, G., Albrecht, C., Bertini, A., Combourieu-Nebout, N., Cvetkoska, A., Giaccio, B., Grazhdani, A., Hauffe, T., Holtvoeth, J., ... Zhang, X. (2019). Mediterranean winter rainfall in phase with African monsoons during the past 1.36 million years. *Nature*, *573*, 256–260. <https://doi.org/10.1038/s41586-019-1529-0>
- Wagner, B., Wilke, T., Francke, A., Albrecht, C., Baumgarten, H., Bertini, A., Combourieu-Nebout, N., Cvetkoska, A., D'Addabbo, M., Donders, T. H., Föller, K., Giaccio, B., Grazhdani, A., Hauffe, T., Holtvoeth, J., Joannin, S., Jovanovska, E., Just, J., Kouli, K., ... Zhang, X. S. (2017). The environmental and evolutionary history of Lake Ohrid (FYROM/Albania): Interim results from the SCOPSCO deep drilling project. *Biogeosciences*, *14*, 2033–2054. <https://doi.org/10.5194/bg-14-2033-2017>
- Wagner, B., Wilke, T., Krastel, S., Zanchetta, G., Sulpizio, R., Reicherter, K., Leng, M. J., Grazhdani, A., Trajanovski, S., Francke, A., Lindhorst, K., Levkov, Z., Cvetkoska, A., Reed, J. M., Zhang, X., Lacey, J. H., Wonik, T., Baumgarten, H., & Vogel, H. (2014). The SCOPSCO drilling project recovers more than 1.2 million years of history from Lake Ohrid. *Scientific Drilling*, *17*, 19–29. <https://doi.org/10.5194/sd-17-19-2014>
- Wang, P., Tian, J., & Lourens, L. J. (2010). Obscuring of long eccentricity cyclicity in Pleistocene oceanic carbon isotope records. *Earth and Planetary Science Letters*, *290*, 319–330. <https://doi.org/10.1016/j.epsl.2009.12.028>
- Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernández-Palacios, J. M., Hengl, T., Norder, S. J., Rijdsdijk, K. F., ... Thébaud, C. (2015). Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecology Letters*, *18*, 200–217. <https://doi.org/10.1111/ele.12398>
- Webb, C. O., Ackerly, D. D., McPeck, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, *33*, 475–505. <https://doi.org/10.1146/annurev.ecolsys.33.010802.150448>
- Weber, M. G., Wagner, C. E., Best, R. J., Harmon, L. J., & Matthews, B. (2017). Evolution in a community context: On integrating ecological interactions and macroevolution. *Trends in Ecology & Evolution*, *32*, 291–304. <https://doi.org/10.1016/j.tree.2017.01.003>
- Wilke, T., Hauffe, T., Jovanovska, E., Cvetkoska, A., Donders, T., Ekschmitt, K., Francke, A., Lacey, J. H., Levkov, Z., Marshall, C. R., Neubauer, T. A., Silvestro, D., Stelbrink, B., Vogel, H., Albrecht, C., Holtvoeth, J., Krastel, S., Leicher, N., Leng, M. J., ... Wagner, B. (2020). Deep drilling reveals unexpected decreasing extinction rate after formation of ancient ecosystem. *Science Advances*, *6*, eabb2943. <https://doi.org/10.1126/sciadv.abb2943>
- Wilke, T., Wagner, B., Van Bocxlaer, B., Albrecht, C., Ariztegui, D., Delicado, D., Francke, A., Harzhauser, M., Hauffe, T., Holtvoeth, J., Just, J., Leng, M. J., Levkov, Z., Penkman, K., Sadori, L., Skinner, A., Stelbrink, B., Vogel, H., Wesselingh, F., & Wonik, T. (2016). Scientific drilling projects in ancient lakes: Integrating geological and biological histories. *Global and Planetary Change*, *143*, 118–151. <https://doi.org/10.1016/j.gloplacha.2016.05.005>
- Zhang, X. S., Reed, J. M., Lacey, J. H., Francke, A., Leng, M. J., Levkov, Z., & Wagner, B. (2016). Complexity of diatom response to Lateglacial and Holocene climate and environmental change in ancient, deep and oligotrophic Lake Ohrid (Macedonia and Albania). *Biogeosciences*, *13*, 1351–1365. <https://doi.org/10.5194/bg-13-1351-2016>

BIOSKETCH

The author team is interested in geological and biological history of ancient lakes. The study was conducted as part of the deep-drilling project SCOPSCO, linking geology and biology to understand how a long-lived ecosystem and its biodiversity emerge and change over time.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Jovanovska, E., Hauffe T., Stelbrink B., Cvetkoska A., Levkov Z., Wagner B., Lacey J. H., Ognjanova-Rumenova N., Hamilton P. B., Brandenburg K. M., Albrecht C., & Wilke T. (2022). Environmental filtering drives assembly of diatom communities over evolutionary time-scales. *Global Ecology and Biogeography*, *31*, 954–967. <https://doi.org/10.1111/geb.13471>