





Cyphoderia ampulla (Cyphoderiidae: Rhizaria), a tale of freshwater sailors: The causes and consequences of ecological transitions through the salinity barrier in a family of benthic protists

Rubén González-Miguéns¹  | Carmen Soler-Zamora¹  | Fernando Useros¹ | Sandra Nogal-Prata¹ | Cédric Berney^{2,3} | Andrés Blanco-Rotea⁴ | María Isabel Carrasco-Braganza⁴ | David de Salvador-Velasco⁴ | Antonio Guillén-Oterino⁴ | Daniel Tenorio-Rodríguez⁴ | David Velázquez⁵ | Thierry J. Heger⁶  | Isabel Sanmartín¹ | Enrique Lara¹ 

¹Real Jardín Botánico de Madrid (RJB-CSIC), Madrid, Spain

²Université de la Sorbonne CNRS, Station Biologique de Roscoff, ECOMAP, Roscoff, France

³Research Federation for the Study of Global Ocean Systems Ecology and Evolution, FR2022/Tara GOSSE, Paris, France

⁴Estación Biológica Internacional Duero-Douro (EUROPARQUES-EBI), Buque hidrográfico Helios-Cousteau en el Lago de Sanabria, Castilla y León, Spain

⁵Department of Biology, Universidad Autónoma de Madrid, Madrid, Spain

⁶Soil Science and Environment Group, CHANGINS, University of Applied Sciences and Arts Western Switzerland, Nyon, Switzerland

Correspondence

Rubén González-Miguéns, Real Jardín Botánico de Madrid (RJB-CSIC), Madrid, Spain.
Email: ruben.gonzalez.miguens@gmail.com

Enrique Lara, Real Jardín Botánico de Madrid (RJB-CSIC), Madrid, Spain.
Email: enrique.lara@rjb.csic.es

Abstract

The salinity barrier that separates marine and freshwater biomes is probably the most important division in biodiversity on Earth. Those organisms that successfully performed this transition had access to new ecosystems while undergoing changes in selective pressure, which often led to major shifts in diversification rates. While these transitions have been extensively investigated in animals, the tempo, mode, and outcome of crossing the salinity barrier have been scarcely studied in other eukaryotes. Here, we reconstructed the evolutionary history of the species complex *Cyphoderia ampulla* (Euglyphida: Cercozoa: Rhizaria) based on DNA sequences from the nuclear SSU rRNA gene and the mitochondrial *cytochrome oxidase subunit I* gene, obtained from publicly available environmental DNA data (GeneBank, EukBank) and isolated organisms. A tree calibrated with euglyphid fossils showed that four independent transitions towards freshwater systems occurred from the mid-Miocene onwards, coincident with important fluctuations in sea level. Ancestral trait reconstructions indicated that the whole family Cyphoderiidae had a marine origin and suggest that ancestors of the freshwater forms were euryhaline and lived in environments with fluctuating salinity. Diversification rates did not show any obvious increase concomitant with ecological transitions, but morphometric analyses indicated that species increased in size and homogenized their morphology after colonizing the new

Rubén González-Miguéns and Carmen Soler-Zamora should be considered as joint first authors.

Isabel Sanmartín and Enrique Lara should be considered as joint senior authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

Funding information

Ministerio de Economía, Industria y Competitividad, Gobierno de España, Grant/Award Number: PGC2018-094660-B-I00/10.13039/501100011033/; Consejería de Educación, Juventud y Deporte, Comunidad de Madrid, Grant/Award Number: 2017-T1/AMB-5210; Ministerio de Ciencia e Innovación, Grant/Award Number: PID2019-108109GB-I00/10.13039/501100011033/; National Geographic Grant, Grant/Award Number: EC-86022R-21; Spanish Ministry of Science, Grant/Award Number: FPU19/06077

Handling Editor: Suhelen Egan

environments. This suggests adaptation to changes in selective pressure exerted by life in freshwater sediments.

KEYWORDS

adaptive morphology, *Cyphoderia*, ecological transition, parallel evolution, preadaptation, salinity barrier

1 | INTRODUCTION

The major types of ecosystems on Earth shape the distributions and evolutionary histories of the organisms inhabiting them. Indeed, crossing the boundary between two ecosystems characterized by sharp differences in climatic (temperature, humidity, seasonality) or physiological parameters (pH, salinity) demands profound adaptations (Simpson, 1984). This explains why ecological transition events have only occurred on a few occasions in the evolutionary history of taxa. Despite its apparent permeability, the interface between marine and freshwater is one of these barriers that only relatively few groups have been able to overcome (Hutchinson, 1975). Crossing this salinity barrier implies adapting to new ionic concentrations and fundamental changes in the mechanism of osmotic pressure regulation. Accordingly, certain ancient lineages such as Echinodermata, Ctenophora (Gray, 1988) or polycystine radiolarians (Boltovskoy, 2003) seem to have never managed to achieve this transition. Other groups have been more successful; the first adaptations in plants (Mansouri et al., 2019) and animals (Hong et al., 2019; Zhang et al., 2017) can be witnessed at transcriptomic level before populations became separated and speciation occurred.

The selective pressures exerted on the organisms that overcame the salinity barrier led to population diversification and, eventually, to ecologically driven speciation in the new free niches (Rautsaw et al., 2021). These speciation events occurring during ecological transitions, together with massive extinctions, are considered as the main drivers for the disruption of diversification rates across the Tree of Life (Schluter, 2000). However, the outcome of these transitions does not follow a single rule and differs deeply among taxa. Selective pressures can vary sharply in the newly colonized environment, and act differently on the organisms depending on their own ecology and life history (Pallarés et al., 2017). For example, the effect of marine-fresh water transitions on the rate of cladogenesis and morphological evolution (i.e., changes in phenotypic traits) varied depending on taxa, as exemplified in several bony fish families: needlefishes (Belontiidae)

evolved towards dwarfism (Kolmann et al., 2020) when colonizing freshwater; pufferfishes (Tetraodontidae), in contrast, increased body size heterogeneity (Santini et al., 2013) and underwent lower rates of cladogenesis; therapontid grunners (Therapontidae) increased both cladogenesis rate and morphological diversification in parallel to dietary diversification (Davis et al., 2012, 2016). Thus, there is no single scenario describing the effect of crossing the salinity barrier on the rate of diversification in fishes. However, while this question is still open to debate in relatively well-known groups such as vertebrates, it has never been investigated in other groups of eukaryotes. There is therefore a clear a necessity to build hypotheses that may be generalized across all groups of organisms, or at least across the eukaryotes.

Eukaryotic diversity is composed to a large extent of unicellular, microbial-sized organisms collectively called protists. This paraphyletic assemblage includes all eukaryotes with the exception of fungi, animals and plants (Taylor, 2003). Many major protist clades have been able to cross the salinity barrier, a phenomenon that is considered infrequent in plants, animals and microbes (Logares et al., 2009). The phenotypic changes brought by this transition have remained understudied because many of these protist groups have a relatively conserved morphology, with few variable traits, which make it difficult to analyse patterns of morphological diversification. One exception is testate amoebae: they possess a self-secreted shell or test, whose morphology varies between species (Kosakyan et al., 2016). Their shape and composition have been thought to have an adaptive value to environmental factors (González-Miguéns et al., 2021). Amongst them, family Cyphoderiidae (Rhizaria) is one of the few clades that managed to cross the salinity barrier (Heger et al., 2010). These organisms are typically found in the benthic fraction of freshwater ecosystems among sediments and aquatic plants (Todorov et al., 2009), as well as in the supralittoral zone of sandy marine ecosystems (Golemansky, 1970). The family comprises six genera with marked interspecific morphological differences (Heger et al., 2011), which makes it an excellent model for studying transitions through the salinity barrier in protists and in eukaryotes in general.

Cyphoderia is the largest genus of the family with 14 species, eleven of which have been reported from freshwater ecosystems. *Cyphoderia ampulla* (Ehrenberg, 1840) is the most widespread species. This freshwater species possesses tests that are typically round in cross-section with a characteristic aperture forming an angle like a "bota bag" (Figure S1). *Cyphoderia ampulla* has a worldwide distribution in freshwater sediments or amongst wet peat mosses (Todorov et al., 2009). However, molecular studies showed that, despite a homogeneous shell outline and ecology, this taxon should be considered as a paraphyletic species complex (Heger et al., 2010, 2011). This homogeneity in shape and ecology contrasts with the variable morphology and ecology reported among other genera within Cyphoderiidae, and among species within the genus *Cyphoderia* (Todorov et al., 2009).

Here, we reconstruct the evolutionary history of Cyphoderiidae, focusing on the paraphyletic group *C. ampulla*, but also including its closest relatives within *Cyphoderia*. We based our inferences on molecular sequence data (nuclear SSU rRNA gene and mitochondrial cytochrome oxidase subunit I gene [COI]), ecological data (inferred from the habitat sampled) and morphological data obtained from the bibliography and from our own isolates from the Iberian Peninsula. In addition, we integrated environmental sequence data obtained from the EukBank public repository (Berney et al., 2017). Based on this information, we aimed at: (1) finding the original habitat of Cyphoderiidae and the direction of the ecological transitions, (2) inferring whether these ecological transitions were synchronized in geological time or occurred independently, (3) evaluating whether these transitions generated a significant change in the tempo of molecular diversification, and (4) testing if salinity barrier transitions were correlated with changes in the shell morphology, as it could be expected if shells have an adaptive value.

2 | MATERIALS AND METHODS

2.1 | Data acquisition

2.1.1 | Sampling and specimen preparation

We sampled different freshwater sites from the Iberian Peninsula, a well-known diversity hotspot, where we expected to have a high probability of increasing the number of lineages and the taxonomic representation of *Cyphoderia ampulla*. The localities sampled are detailed in Figure 1 and Table S1. The most active collected cells, those showing a high mobility under the microscope, were isolated for molecular analyses and deposited individually in Eppendorf tubes containing 100 µl of buffer for DNA extraction (Chomczynski & Sacchi, 1987). DNA was extracted from single cells as described in Duckert et al. (2018). The remaining cells were deposited on stubs for scanning electron microscopy (SEM) analyses.

DNA extraction and amplification and microscopical observation are described in "DNA extraction and amplification" and "Microscopical observation" of Appendix 1.

2.2 | Data analyses

2.2.1 | Phylogenetic analyses

We built the COI database by combining our single cell sequences with publicly available sequences obtained from Heger et al. (2011) (Table S1), ending up with a total of 52 sequences from *Cyphoderia* and three from *Pseudocorythion acutum*, which were used to root the tree. The SSU rRNA gene database included sequences from Heger

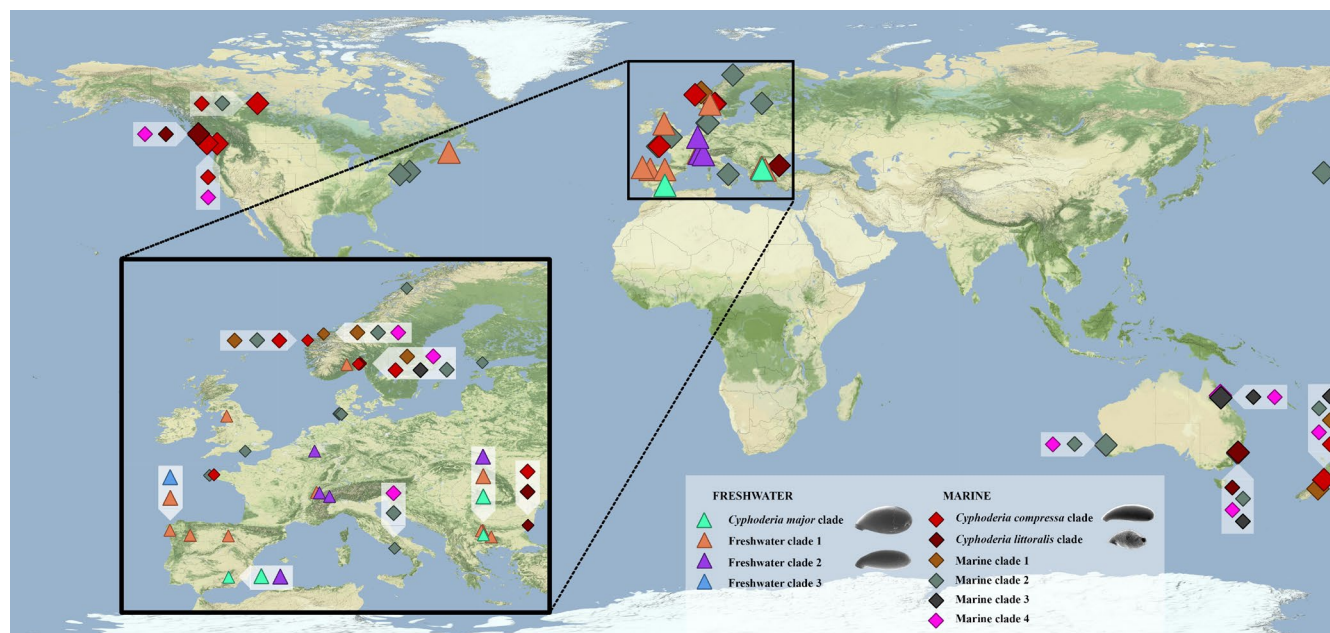


FIGURE 1 Maps showing the localities from where Cyphoderiidae COI and SSU rRNA sequences were retrieved. Triangles and diamonds represent freshwater and marine environments, respectively. The colours of the symbols represent the different molecular clades

et al. (2010) and Wylezich et al. (2002), as well as environmental clones from Bass and Cavalier-Smith (2004) (Table S1). To obtain the most complete picture of the global diversity of Cyphoderiidae, we included environmental sequences retrieved from the EukBank public repository, which contains eukaryotic metabarcodes based on the V4 region of the SSU rRNA gene, generated from about 13,000 samples distributed worldwide, involving 122 projects from over 90 research teams (Berney et al., 2017). Based on this database, we performed a blastn analyses (Altschul et al., 1990) to extract the amplicon sequence variants (ASVs) belonging to Cyphoderiidae, obtaining 64 sequences that were added to our initial data set. The tree was rooted with 26 sequences of Euglyphina (Kosakyan et al., 2016). In all, we analysed 146 SSU rRNA gene sequences; sensitivity analyses were performed with and without environmental sequences to check if the topology of the phylogenetic tree was affected (phylogenetic trees and nexus files are available at TreeBASE permalink <http://purl.org/phylo/treebase/phyloids/study/TB2:S29006>).

Tree topologies and node supports were evaluated with both Bayesian inferences (BI) and maximum likelihood (ML) for each marker/gene individually. ML analyses were conducted using IQ-TREE (Nguyen et al., 2015) and Bayesian inference (BI) analyses were conducted using MRBAYES 3.2.7a (Ronquist et al., 2012). Phylogenetic analyses are described in Tree analyses of Appendix 1.

2.3 | Divergence time estimation

Lineage divergence times were estimated in BEAST v.1.8.4 (Drummond et al., 2012), as implemented in the CIPRES Science Gateway version 3 (Miller et al., 2010) and using the SSU rRNA gene data set. We implemented the general time reversible (GTR) substitution model with gamma across-site rate variation and an invariant portion (I + G). We calibrated lineage divergence times using two external points from sister groups belonging to Euglyphina, based on fossil evidence: (1) *Trinema* spp. Penard, 1890: Fossils from this genus consist in *Trinema* shells found in Borneo and dated from the late Miocene (Schiller, 1997), and a second fossil resembling a modern *Trinema lineare* from the French Pliocene (Boeuf & Gilbert, 1997). We decided to constrain the age of the most recent common ancestor (MRCA) of *Trinema* (the node grouping *T. lineare* and *Trinema enchelys*) with this second fossil, using a lognormal prior to reflect the uncertainty in the fossil calibration, an offset = 5 million years ago (Ma) and a standard deviation = 1. (2) *Scutiglypha* node: several fossils have been identified as belonging to genus *Scutiglypha* (represented as *Euglyphina*). These fossils date back to the Eocene (Barber et al., 2013; Schiller, 1999) and Miocene (Foissner & Schiller, 2001; Schiller, 1997) periods, respectively. They are characterized by strongly scutiform scales without spines, which correspond with the morphology of both *Scutiglypha cristata* and *Scutiglypha tuberculata* but not with that of *Scutiglypha acanthophora*. Thus, we considered *S. acanthophora* and *S. cristata* as the crown group, and constrained the stem node of this group (*S. acanthophora*, *S. cristata* and *S. tuberculata*) with a lognormal prior with offset = 40 Ma and standard deviation = 1.

Although some *Cyphoderia* fossils exist (Waggoner, 1996), the preservation state of this material does not allow comparisons with modern taxa morphology. Therefore, we decided not to use these fossils as calibration points in our analyses. Sensitivity analyses were carried out to evaluate the effect of tree priors on the molecular clock. We performed several runs, with the parameters described above, and using alternative tree (Yule process, birth-death and birth-death incomplete sampling) and molecular clock (strict, uncorrelated, relaxed, random and fixed) priors (phylogenetic trees and nexus files are available at TreeBASE permalink <http://purl.org/phylo/treebase/phyloids/study/TB2:S29003>). Next, marginal likelihood values were calculated for each model, using path sampling (PS) and stepping-stone (SS) sampling (Baele et al., 2013). We then compared the model with the lowest marginal likelihood estimated, as by ps and ss (Table S2) against the remaining models in pairwise Bayes Factors comparison (Kass & Raftery, 1995) (Table S3). Finally, we also extracted the time of the most recent common ancestor (tMRCA) for the nodes of interest (freshwater clades 1 and 2, see Section 3) from the BEAST posterior distribution of each run using TREESTAT (Vargas et al., 2014, 2020), in order to capture the uncertainty in divergence time estimates for those nodes.

Our sensitivity analysis indicated that the most adequate priors for our database was birth-death (Gernhard, 2008) as tree prior and the uncorrelated lognormal distribution, for the relaxed molecular clock. We repeated three times the analysis under this model, each with two MCMC chains for 20×10^7 generations, sampling every 1000th generation, and then used TRACER version 1.7.1 to monitor convergence and adequate mixing (effective sample size [ESS] >200) (Rambaut et al., 2018). We combined the three BEAST runs with Log Combiner, resampling states every 20,000th, with a burnin of 25%. A maximum clade credibility (MCC) tree was constructed in TREEANNOTATOR version 1.8.4 (Drummond et al., 2012).

Additionally, due to the difficulties in establishing correspondences between fossils and nodes in the phylogenetic tree, we performed a second sensitivity analysis with the parameters described above but with different node positions for the calibration points, to see how these affect the estimation of nodal ages (more information in "BEAST sensitivity analysis" in Appendix 1).

2.4 | Ancestral state reconstructions (ASR)

To reconstruct the ancestral habitat state, we used the combined MCC tree based on the SSU data set and the three BEAST runs, as described above. This tree was pruned to remove the outgroups, using the function drop.tip in the R package APE version 5.5 (Paradis et al., 2004). We used the PHYTOOLS R package (Revell, 2012) to infer the habitat history of Cyphoderiidae. We estimated the posterior probability of the discrete characters marine and freshwater for each ancestral node, using an empirical Bayesian inference method, that is, stochastic character mapping (Huelsenbeck et al., 2003). We used the function make.simmap (Bollback, 2006) to generate 500 stochastic character maps from our data set, under two different

models: all rates different (ARD) model, which allows different rates of transition between all states, and equal rates (ER), which enforces the same rate of transition between states. We then summarized these sets of stochastic maps with the function `densityMap`, plotting the posterior probability of being in each state across all the edges and nodes of the tree.

For the COI data set, we used the Bayesian tree, as ML and BI recovered identical topologies. We first transformed this nonultrametric consensus tree into an ultrametric one with the function `chronos` in *ape* R package (Paradis et al., 2004) and then resolved the basal polytomy in the *MRBAYES* tree by outcome assigning a branch length of 0.00001. To reconstruct ancestral states on this tree, we followed the same procedure described above.

2.5 | Diversification analyses

Our aim here was to perform an exploratory analysis of the diversification dynamics of Cyphoderiidae considering all the data we collected. To define these OTUs, we used the criterion proposed by Heger et al. (2011) to delimit species, which is based on support from molecular markers (both SSU rRNA gene and COI gene grouped conspecific sequences unambiguously together), as well as on morphological and ecological consistency. Missing species in a phylogeny can mislead estimates of diversification rates (Höhna, 2014; Sanmartín & Meseguer, 2016); although we cannot correct our analyses for possible missing species, as the number of lineages considered as independent taxonomic units or OTUs in our analysis is considerably higher than any of the estimates given in the literature or based on Cyphoderiidae systematics. However, given the size of the database explored, we can reasonably assume that we have a good representation of the groups diversity. We thus performed our diversification analyses assuming that we have retrieved all the diversity in Cyphoderiidae, that is., assuming a sampling fraction of 1.

We used the MCC tree described above, without outgroups, and further removed the intraspecific variability by pruning the tree to leave only one single terminal per species (74 tips). We then used the function `litt.plot` in the package *APE* version 5.5 (Paradis et al., 2004) to plot the number of lineage diversification events through time. These lineage-through-time (LTT) plots were drawn for each of the individual *BEAST* runs, the combined MCC tree, and two subtrees representing the trajectories of the freshwater lineages, and the marine lineages, separately. For each tree, we also estimated the gamma statistics (Pybus & Harvey, 2000), and assessed its significance by comparing the observed value against a distribution of values from 1000 random trees with 74 tips, generated under a Yule model with the empirical speciation rate estimated in *ape*.

We used a reversible jump Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) approach implemented in the Bayesian analysis of macroevolutionary mixtures (*BAMM* version 2.5; Rabosky, 2014; Rabosky, Donnellan, et al., 2014), following the *BAMM* tutorials (Mitchell & Rabosky, 2017), to model the dynamics

of speciation and extinction over time and across branches in the combined MCC tree above (with every tip representing a single species). The prior block parameter values were chosen using *BAMMTOOLS* package (Rabosky, Grudler, et al., 2014), with a Poisson rate prior of 1.0. We ran four chains with 10×10^6 generations, sampling every 1000th, the final posterior distribution was estimated from 1000 trees. MCMCMC operators were used under default settings. *BAMM* has been criticized recently because of its biased modelling of the probability of rate shifts occurring in extinct lineages (i.e., a lineage undergoes a rate shift and then goes extinct), leading in some cases to incorrect likelihood estimation (Laudanno et al., 2021; Moore et al., 2016). We used here an alternative lineage-specific birth-death shift model (*LSBDS*), which uses the incremental-time, state-dependent differential equations developed by Maddison et al. (2015) to model shifts in diversification rates dynamically (Höhna et al., 2019), and which is implemented in the Bayesian software *REVBAYES* version 1.1.1 (Höhna et al., 2014; Höhna, Landis, et al., 2016). We followed the tutorial (https://revbayes.github.io/tutorials/divrate/branch_specific.html), using four rate categories (i.e., allowing for three rate shifts in the speciation rate), $\rho = 1$ (global sampling fraction), and a chain length of 10×10^6 generations. We then obtained the branch-specific diversification rate estimates for each branch of the tree with a rejection-free stochastic rate mapping algorithm and dividing the phylogeny into 500 time slices (Freyman & Höhna, 2019). We also performed two additional analyses to explore the effect of assuming varying levels of missing species in the phylogeny: using a global sampling fraction of $\rho = .5$ and $\rho = .25$.

Both *BAMM* and *LSBDS* model changes in diversification rates over time that are lineage-dependent. To explore the possibility of a rate shift that affected all lineages in the phylogeny simultaneously (e.g., a mass extinction event), we ran an episodic birth-death model to estimate the timing and magnitude of discrete shifts in diversification and relative extinction rates. We used the compound Poisson process of mass extinction (*COMET*) model (May et al., 2016) implemented in the R package *TESS* (Höhna, May, et al., 2016), with a log-normal empiricalHyperPriorForm and sampling probability = 1, a chain length of 20×10^6 iterations, sampling every 100th generation, and a minimum number of effective sample size of 1000 to stop the run. *COMET* uses independent compound Poisson process priors to estimate the posterior probability of the timing and magnitude of rate shifts in speciation and extinction; mass extinction events that deplete the standing diversity by a certain fraction (ρ) can also be estimated by integrating out speciation and extinction rate shifts via Bayes factor comparisons (May et al., 2016). Additionally, we ran an episodic birth death model that is implemented in *REVBAYES* and uses auto correlated Brownian rates across time slices, and a horseshoe estimator for sparse signals (Magee et al., 2020). We followed the tutorial (<https://revbayes.github.io/tutorials/divrate/ebd.html#Hoehn2015a>) with $\rho = 1$ and a chain length of 10×10^6 generations. We used the R package *REVGADGETS* version 1.0 (Tribble et al., 2021) to visualize the rates. For both *COMET* and *REVBAYES*, we performed two additional analyses changing the global sampling fraction to $\rho = .5$

and $\rho = .25$, considering that we only sampled the 50% and 25% of the total diversity of Cyphoderiidae, respectively.

To explore if diversification rate shifts across branches detected by BAMM and LSBDs (results) could be explained by transitions between the marine and fresh-water habitat, we performed state-dependent extinction and speciation diversification rate analyses. We first ran binary state speciation and extinction (BiSSE) (Maddison et al., 2007) analysis using the R package diversitree (FitzJohn, 2012) and two states: (1) marine and (2) freshwater. We performed empirical Bayesian estimation of the parameters using the MCMC option. As our tree only included 74 tips, we used an exponential prior with a rate of $1/(2r)$ (where r is the character independent diversification rate) as recommended in the diversitree tutorial; the step size (argument w) was set as in Revell and Betancur (2017). We performed two runs of 30,000 steps, with a final burnin of 20% for all runs. We also repeated the BiSSE analysis in REVBAYES (Höhna, Landis, et al., 2016), which implements a fully Bayesian approach, following the procedures in (López-Estrada et al., 2021) and (https://revbayes.github.io/tutorials/morph/morph_more.html) (Figure 3c). We ran the analysis with a log-standard deviation of $= 0.587405$ and a chain length of 40,000 generations; we used 500 time slices to estimate the timing and number of state changes along branches using stochastic character mapping. The BiSSE model has been criticized because it assumes a null model for statistical testing, the constant-rate birth death process (i.e., heterogeneity in diversification rates occurs virtually in any phylogeny; Rabosky & Goldberg, 2015, 2017). To test for the existence of diversification rate shifts that are unrelated to state transitions in the focal trait (marine-freshwater habitats), we compared four different SSE models (explained in Appendix 1) in the R package HiSSE (Beaulieu & O'Meara, 2016) using the Akaike information criterion (AIC).

2.6 | Morphometric analyses

We used five continuous morphometric traits for the morphometric analyses, as described in Todorov et al. (2009): (1) Length of the shell, (2) width of the shell, (3) large pseudostome axis, (4) length-large pseudostome axis ratio, and (5) length-width ratio. We combined measurements from Todorov et al. (2009) with our new data from the Iberian Peninsula, in total 719 cells (the data are available in Morphobank permalink <http://morphobank.org/permalink/?P4082>). We followed the same groups created by Todorov et al. (2009), as the monophyly of each of these lineages was supported by both SSU rRNA gene and COI gene sequences (Heger et al., 2011). Our cells were ascribed to these groups based on their COI and SSU rRNA gene phylogenetic position.

We carried out a principal component analysis (PCA) to observe the relationships between the different phylogroups (lineages recovered monophyletic with COI and SSU), using the *stats* package, with the *prcomp* function with the five continuous morphometric traits described above. Finally, we plotted the results using the

package *GGPLOT2* (Wickham, 2016) and used the function *factoextra* (Kassambara & Mundt, 2020) to see the contribution in the analysis of each variable used.

2.7 | Hypervolume

We calculated the overlap between each phylogroup using the R package *hypervolume* (Blonder et al., 2014), allowing the quantification of the phylogroups morphospace as determined by the measured of five continuous morphometric traits: shell length, shell width, large pseudostome axis, length-large pseudostome axis ratio and length-width ratio. We used the first three principal components from the previously described PCA, to obtain a limited number of linearly uncorrelated variables. We used a Silverman bandwidth estimator (Silverman, 2018) with a quantile threshold of 0.05 to infer the Kernel density estimations (KDEs). Then, we created a multidimensional space for each phylogroup, calculating its size and the extent of overlap with other phylogroups, using the function *hypervolume_overlap_statistics* and *hypervolume_distance*, respectively (Blonder et al., 2014).

2.8 | Combining phylomorphospace and hypervolume

We tested how far ecological transitions are correlated with changes in the morphology by mapping the morphological diversification history of the freshwater and marine clades, respectively. As not all measured specimens yielded a DNA barcode, but all morphology groups were recovered at least with two monophyletic sequences (both COI and SSU), we consider here each lineage-morphology group as an independent evolutionary unit. We therefore pruned the SSU rRNA MCC, leaving only one tip per morphology group to link the molecular and the morphometric data. Then, we took the centroids of the hypervolume described above (Table S4) and mapped the morphological diversification history of each lineage by plotting the tree tips in a two-dimensional space. We reconstructed the ancestral states along morphospace axes (hypervolume centroid), using the function *phylomorphospace* in the *phytools* package (Revell, 2012; Sidlauskas, 2008). The hyperphylomorphospace resulting from this combination of approaches allows the evaluation of both lineage morphological similarity (distance between each lineage) and the trajectory of the phenotype (directionality of tree branches). The procedure followed is described in Appendix 1.

2.9 | Morphological similarity and equivalence: Testing the possible morphological adaptive value

Finally, we wanted to test how far the shell morphology of the different species was a response to selective pressures exerted by the new environments found after the transition across the salinity barrier.

Assuming that there is a selective pressure of the environment (freshwater) on shell morphology, we would expect to see a homogenization of the shells. If this was the case, morphologies would be more similar between freshwater clades than with their phylogenetic sister marine clades. Thus, to statistically test the morphological patterns observed in *Cyphoderia*, we calculated two indexes: (1) "morpho-niche" equivalence: whether the morphologies of resembling species are effectively indistinguishable; and (2) "morpho-niche" similarity, whether the morphology model in one species predict the morphology of a second species better than expected by chance (Warren et al., 2008).

We first defined a background morphology, a "potential morphology" that can represent each morphology group, by drawing a hypervolume (as described above) with five continuous morphometric traits (see morphometric analyses). As the hypervolume generates random points, inferred by a Silverman bandwidth estimator (described above), we considered these estimated points as the potential morphological background, which were used to calibrate our variables (aka the native area in a niche equivalence test). To do this calibration, we performed

the PCA-env (Broennimann et al., 2012) with the function `dui.pca`, using the random points of the hypervolume and the measures of the organism. Then, we performed niche similarity and equivalence tests (Warren et al., 2008), in the R package `ECOSPAT` (Broennimann et al., 2018), using the functions `ecospat.niche.similarity.test` and `ecospat.niche.equivalency.test`, respectively, with 1000 replications in both tests. We also plotted the Niche overlap between each freshwater phylogroup with the function `ecospat.plot.niche.dyn`.

3 | RESULTS

3.1 | Molecular phylogenetic and divergence time estimation analyses

The nomenclature of the clades (freshwater clades 1 and 2) and the taxonomic identifications of the species were given after Heger et al. (2011). The COI tree included 52 sequences (Figure 2;

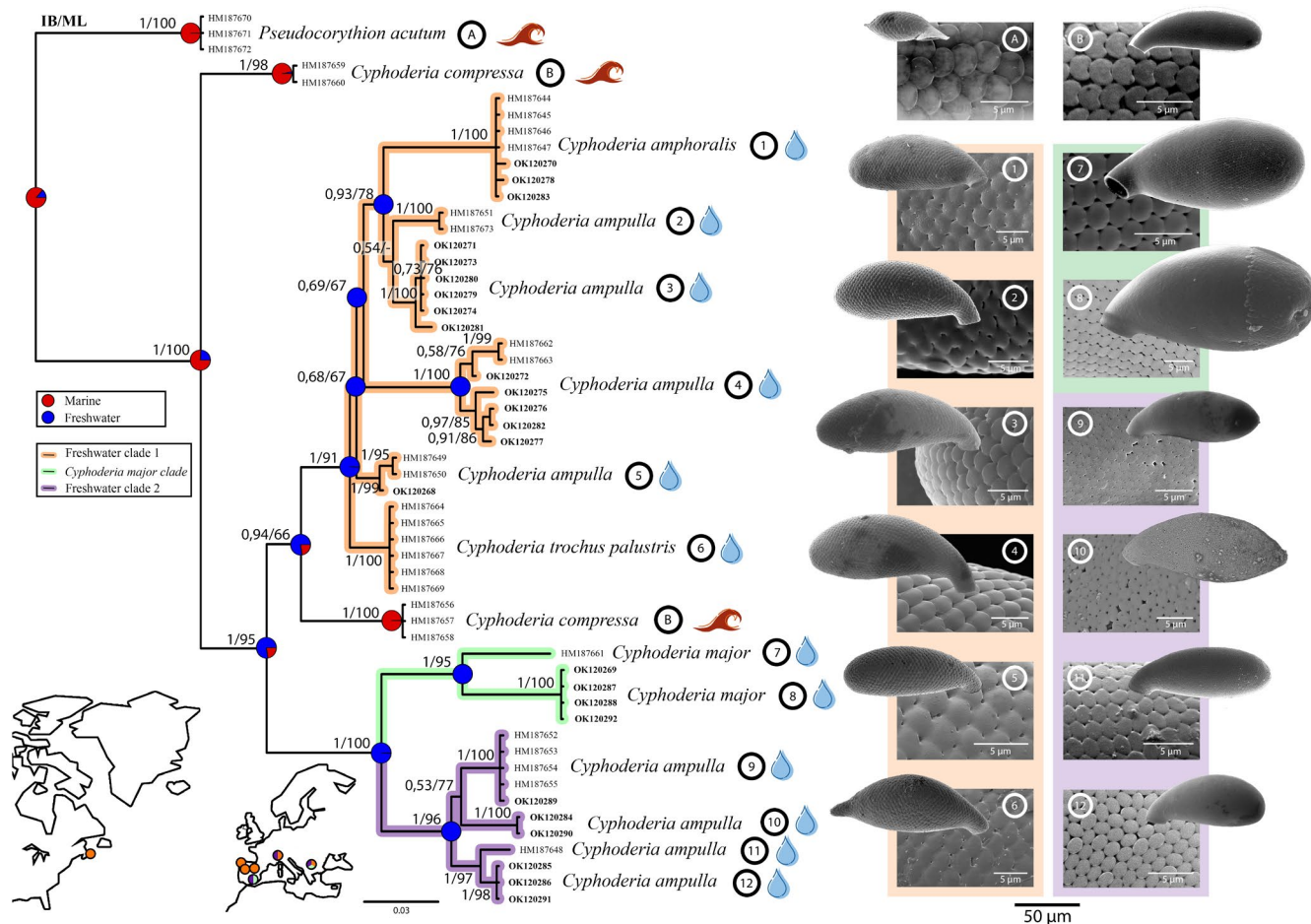


FIGURE 2 Bayesian phylogenetic tree based on COI sequences of family Cyphoderiidae. Posterior probability values (Bayesian analysis) and bootstrap values (maximum likelihood analysis) are represented at each node. Pie charts at nodes represent marginal probabilities for ancestral state reconstructions obtained in *phytools* (red colour for marine environment, blue for freshwater). Branch colours represent the molecular clades. GenBank accession codes are given on the tips. Species names correspond to the taxonomic assignment of the organisms based on morphology for each mitochondrial lineage; indices correspond to illustrations. Symbols represent the original habitat of the organism. The images on the right are scanning electron microscopy photographs showing the shell morphology at the same size scale, together with details of the scaling patterns of each phylogroup. The map represents the localities sampled for each clade

Table S1). Our phylogenetic analyses recovered the monophyly of *Cyphoderia* with a Bayesian posterior probability (PP) of 1 and a maximum likelihood bootstrap (ML) of 100. The freshwater clades 1, 2 and the *Cyphoderia major* (Penard, 1890) clade were recovered with a PP = 1/ML = 91, PP = 1/ML = 96 and PP = 1/ML = 95, respectively. The freshwater clade 1 grouped six mitochondrial lineages (1–6 orange colour), the *Cyphoderia major* clade, two lineages (7 and 8 green colour), and the freshwater clade 2, four lineages (9–12 violet colour), all supported by PP = 1 and ML = 95. *Cyphoderia compressa* (Golemansky, 1979) was recovered as sister to freshwater clade 1 with a PP = 0.94 and ML = 66; the addition of sequences from our new isolates is not congruent with the topology in Heger et al. (2011), where *C. compressa* is sister to freshwater clade 2 (PP = 0.54 and ML = 60 values). *Cyphoderia compressa* appears as paraphyletic in the COI tree, which strongly suggests that both mitochondrial lineages are distinct species (and potentially a novel clade not recovered by SSU).

For the SSU rRNA, we generated 146 sequences, in which 120 represent Cyphoderiidae (Figure S2; Table S1). In our phylogenetic analyses, we recovered the monophyly of Cyphoderiidae and *Cyphoderia* (freshwater 1, 2, marine 1, 2, *C. compressa*, *C. littoralis* and *C. major* clades) with a PP = 1/ML = 100 and PP = 1/ML = 100, respectively. The freshwater clade 1, 2 and *Cyphoderia major* clade were recovered with a PP = 1/ML = 100, PP = 1/ML = 100 and PP = 1/ML = 100, respectively. Sequences from the EukBank database grouped into a new freshwater clade (3) with a PP = 1/ML = 100, coming from a single geographic location (Lake Sanabria, Spain). The freshwater clade 1 is composed of four nuclear lineages, *Cyphoderia major* clade by one, and freshwater clade 2 by two all supported by PP = 1 and ML <85. *Cyphoderia compressa* is the sister group to freshwater clade 1 with a PP = 0.99 and ML = 98; recovering the same topology as in our COI tree. EukBank sequences also highlighted the existence of new clades, only present in marine environments (marine clades 1, 2, 3; Figure 3a and Figure S2).

The maximum clade credibility (MCC) tree obtained from BEAST (Figure 3a, Figures S3–S5) recovered the same clades as the nonclock trees obtained with IQ-TREE and MRBAYES (Figure S2). The most recent common ancestor (MRCA) of Cyphoderiidae was estimated to have been living during the Mesozoic, around 121 million years ago (Ma) in the Early Cretaceous (95% highest posterior density [HPD] = 57.8–223.2 Ma). This node is where the largest discrepancy between sensitivity analyses was found (Figures S3–S5). More fossil data, if possible, from Cyphoderiidae themselves, will be needed to resolve more accurately the age of this clade. However, the nodes including both freshwater clades 1 and 2 (and *Cyphoderia ampulla*) were more homogeneous among sensitivity analyses (Figure S4). They support the existence of a MRCA in both clades during the Miocene. The MCC tree indicates that the MRCA of the freshwater clade 1 probably occurred around 14.2 Ma (95% HPD = 5.7–28.4 Ma), and for the freshwater clade 2, at 6.3 Ma (95% HPD = 1.6–14.6 Ma) obtained similar results with other models (Figure S6).

3.2 | Ancestral state reconstruction and diversification analyses

Ancestral state reconstruction results obtained with phytools for the COI and SSU trees were congruent in recovering that the most probable habitat for the common ancestor of family Cyphoderiidae and genus *Cyphoderia* was marine (Figures 2 and 3a and Figure S7). The larger SSU tree recovered four branches in which the ecological transitions occurred from marine to freshwater, and we did not detect any transition in the opposite direction (Figure 3a). The lineages of these four freshwater clades (*Cyphoderia major* clade, freshwater clades 1, 2 and 3), are named in the following analyses as “freshwater lineages”, for short.

The LTT plots of the combined MCC tree (Figure 3a) show a nonsignificant positive gamma (γ) value ($\gamma = 0.5469$, $p = .5845$) indicating that the speciation rate was constant over time; while the random Yule trees showed a significant negative gamma ($\gamma = -0.1490$, $p = .033$). The LTTs for the marine ($\gamma = 0.2625$, $p = .793$) and freshwater ($\gamma = 0.7702$, $p = .4412$) lineages gave also a nonsignificant positive γ (Pybus & Harvey, 2000).

BAMM and REVBayes branch-specific diversification rate analyses gave congruent results (Figure 4a,b), showing an increase in the diversification rates in the clade formed by genus *Cyphoderia* and the marine clade 3. When experimentally decreasing the sampling fraction ($\rho = .5$ and $\rho = .25$), the overall diversification rate increased but branch diversification patterns were similar to those obtained with $\rho = 1$ in each of the two analyses (Figure 4 and Figure S8).

The episodic birth-death model showed differences between the REVBayes and COMET analyses (Figure 4 and Figure S8). REVBayes recovered an upward trend in the net diversification rate, which is mainly explained by an initially high extinction rate (between 120 and 110 Ma), and a gradual, slow increase in the speciation rate (Figure 4c). COMET inferred more constant net-diversification rates over time ending with a rapid decrease around 5 Ma (Figure 4b), caused by a significant downward shift in the speciation rate (Figure 4d, BF > 2). There were no significant shifts in the extinction rate; however, the COMET analysis supported a mass extinction event between 120 and 110 Ma in the Bayes factor comparisons (BF ~2), in agreement with the REVBayes signal (Figure 4c). Experimentally decreasing the sampling fraction ($\rho = .5$ and $\rho = .25$) recovered patterns that were very similar to those obtained with the full sampling by the two analyses, showing an overall increase in the net diversification rate (Figure S8).

BiSSE with Diversitree showed that both speciation and extinction rates are higher in freshwater lineages ($\lambda = 0.07702517$; $\mu = 0.03678705$) than in marine lineages ($\lambda = 0.0458489$; $\mu = 0.01513946$) (Figure 3c). BiSSE with REVBayes found similar results: ($\lambda = 0.06372$; $\mu = 0.02799$) in freshwater and ($\lambda = 0.0466$; $\mu = 0.01666$) in marine lineages (Figure 3c). The speciation rate was higher in the freshwater lineages than in the marine lineages in 92.31% of the post-burn posterior distribution, suggesting that speciation rate increases when transitions occur but that these differences are not significant (Figure 3c). “Null” was the best model

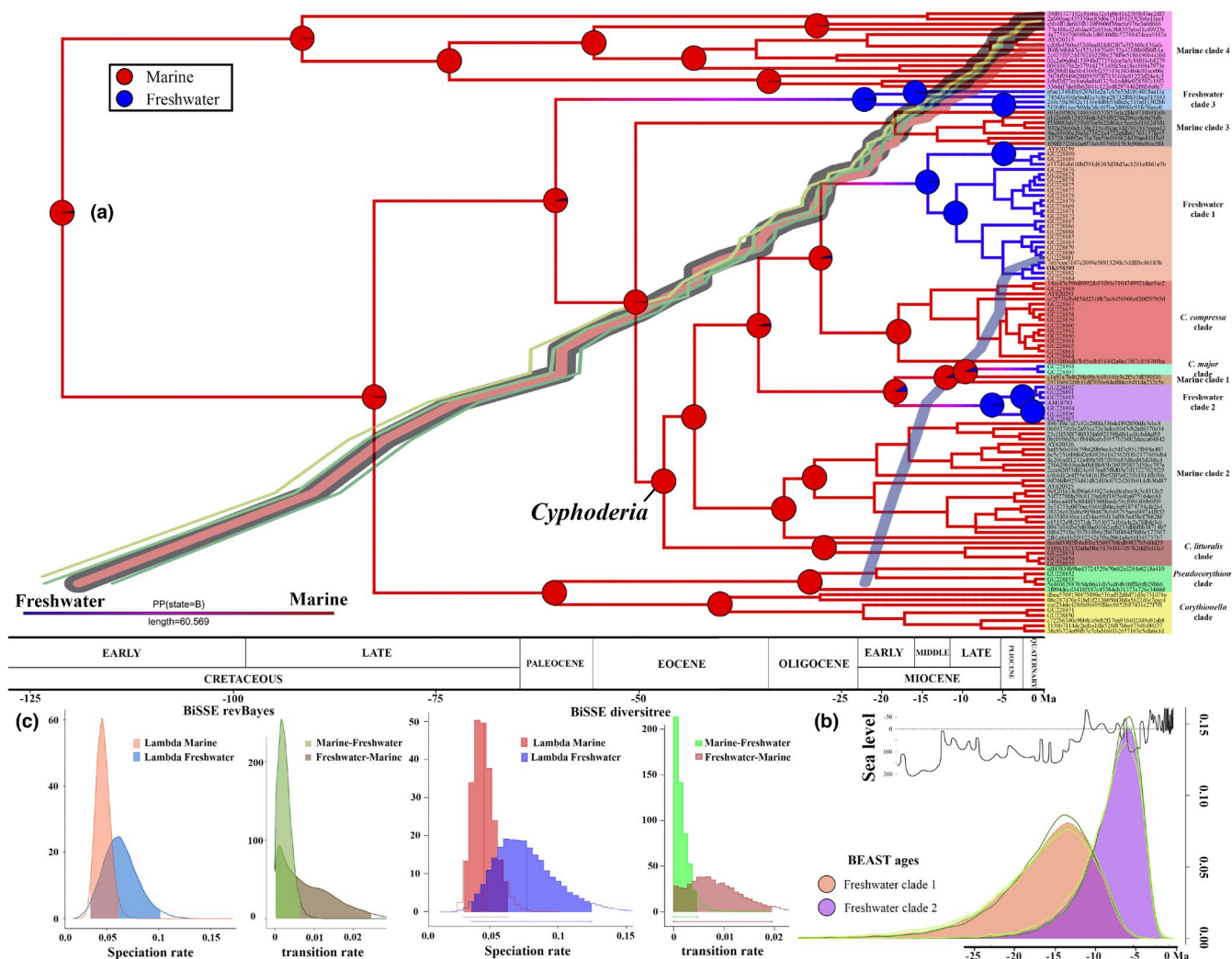


FIGURE 3 (a) Time-calibrated maximum clade credibility (MCC) tree based on the SSU rRNA gene data set of family Cyphoderiidae. Branches are proportional to absolute time, with the x-axis representing geological epoch boundaries. Branch colours and pie charts at nodes indicate marginal probabilities for ancestral state reconstructions and stochastic mappings for marine (red) and freshwater (blue) habitats. Lineage through time plots are shown for different MCC trees: the thin green lines represent each of the independent BEAST runs for the entire phylogeny; gray/grey line represents the combined MCC tree; red line represents the marine lineages only, and the blue line, the freshwater lineages only. (b) Marginal densities for the age of the most recent common ancestor (tMRCA) for the freshwater clade 1 (orange) and the freshwater clade 2 (purple). A schematic illustration of sea level fluctuations over geological epochs (Eberli, 2000) is shown on the top. The three green lines represent the marginal densities of the tMRCA of each independent BEAST run. (c) Posterior distribution of speciation and transition rates for marine and freshwater habitats, and transition rates between obtained them, with the BiSSE model in diversitree and REVBayes

to explain the observed diversification patterns, according to AIC comparisons (Table S5). This indicates that the “habitat” trait marine-freshwater does not explain the diversification rate heterogeneity among clades in the phylogeny.

3.3 | Morphometric and hyperphylomorphospace analyses

We delimited 10 morphogroups, or independent taxonomic units (Heger et al., 2010; Todorov et al., 2009), for the morphometrics analysis. These morphogroups were defined, based on the molecular results, as: *Cyphoderia littoralis* (Golemansky, 1973) (dark red in

Figure 5), *Cyphoderia compressa* (light red in Figure 5), *Cyphoderia major* (green in Figure 5), freshwater clade 1 (orange in Figure 5) and 2 (violet in Figure 5).

The first two axes in the PCA (PC1 = 65.96%; PC2 = 25.87%) explained 91.83% of the total variation of the morphological data set (Figure 5b). The variables with the highest contribution were shell width and length (Figure S9A). Both PCA and hypervolume plots showed that *C. littoralis*, *C. compressa* and *C. major* groups have specific morphospaces that do not overlap (Table S6). In turn, freshwater clades 1 and 2 do overlap, with their centroids close to each other (Table S6). This morphological resemblance contrasts with their phylogenetic position, indicating that these two groups are distantly related (Figure 5c).

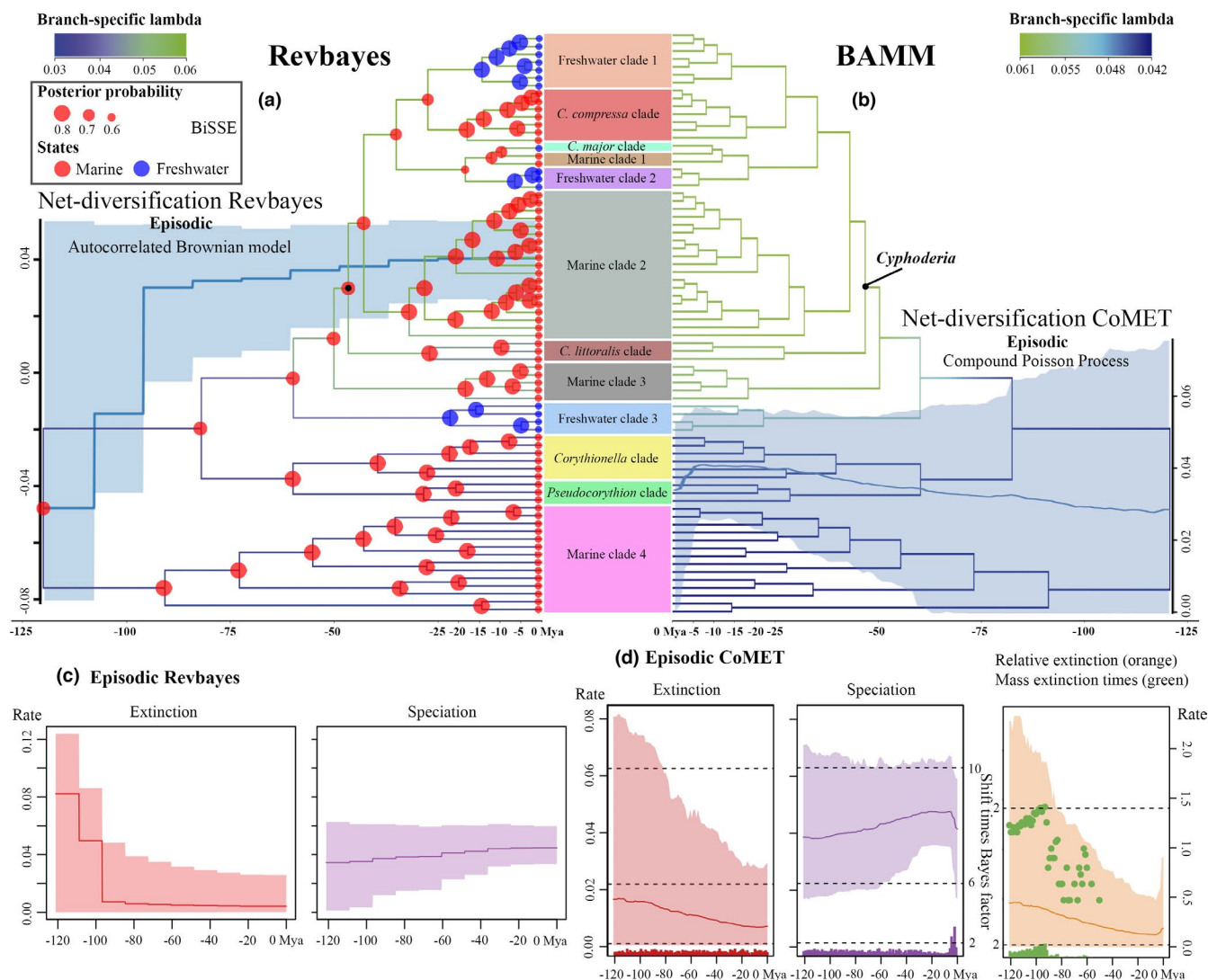


FIGURE 4 Diversification analyses: (a) **REVBAYES**: Estimated branch-specific speciation rates in the pruned SSU rRNA tree, each tip representing an independent taxonomic unit; circles at nodes are marginal probabilities for ancestral character states with BiSSE (red: marine; blue: freshwater; branch colors/present marginal stochastic character mapping); the net-diversification graph shows results from the episodic birth-death model implemented in **REVBAYES** (autocorrelated Brownian model). (b) **BAMM**: Estimated branch-specific speciation rates in the pruned SSU rRNA tree; the net-diversification graph shows results from the episodic birth-death model in **CoMET** (compound Poisson prior model). (c) Estimates of speciation and extinction rates from the episodic birth-death model implemented in **REVBAYES**. (d) Estimates of speciation and extinction rates, and relative extinction rates, from the episodic birth-death model implemented in **CoMET**; the extinction plot depicts also the probability over time of mass extinction events; significance was assessed using Bayes Factor comparisons (BF > 2, positive support Kass & Raftery, 1995).

The first two axes (axis1 = 63.24%; axis2 = 24.88%) of the PCA-env analysis, used to calculate morphological similarity and equivalence, explains 88.12% of the morphological data set variation (Figure S9B). The Schoener's D metric presents similar values with the hypervolume overlap (Tables S6 and S7), showing overlap only between freshwater clades 1 and 2 (Table S7). *Cyphoderia trochus palustris* (Penard, 1902), *Cyphoderia ampulla* 2 and *C. ampulla* 5 (freshwater clade 1) present a morphology more similar than expected by chance with *C. ampulla* 11 (freshwater clade 2) p -value > .05 (Warren et al., 2008), and equivalent between *C. ampulla* 2 and *C. ampulla* 11 p -value > .005 (Figure 6). This also contrasts with their respective phylogenetic positions (Figure 5c).

4 | DISCUSSION

4.1 | Marine origins of family Cyphoderiidae and *Cyphoderia*, and their parallel preadaptation to freshwater

Ancestral state reconstruction analyses inferred the marine habitat as the most likely original habitat for both genus *Cyphoderia* and family Cyphoderiidae. This result is supported by ASR analyses on both mitochondrial (COI) (Figure 2 and Figure S7) and nuclear (SSU rRNA gene) (Figure 3a) markers, and by state-dependent models (Figure 4). In line with this finding, our survey of the EukBank

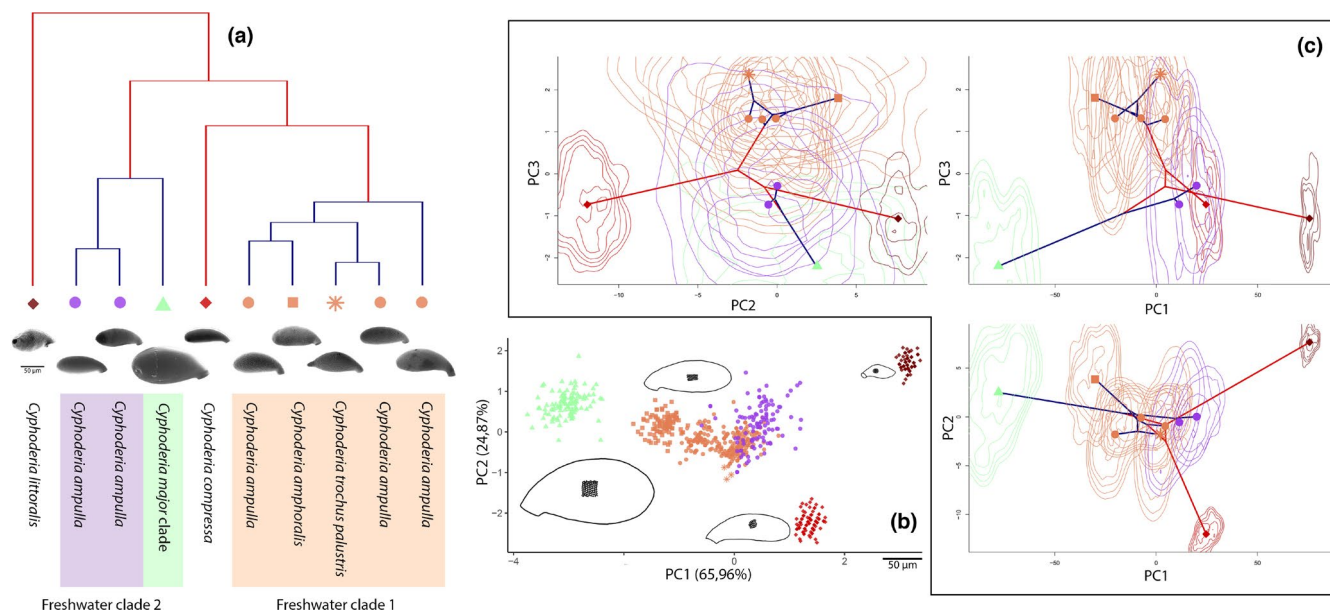


FIGURE 5 (a) Bayesian phylogenetic tree based on the SSU rRNA data set, pruned from Figure 2, leaving one tip per species; branch colours represent the habitat (red, marine; blue, freshwater), (b) Principal component analysis of five continuous morphological traits. (c) Hyperphylomorphospace of the phylogeny (a) with the combinations of the three principal components (b); symbols represent the hypervolume centroids of each phylogroup, and the lines, the Kernel density of the random points generated in the hypervolume

environmental SSU rRNA V4 sequence data set revealed that the majority of *Cyphoderiidae* amplicon sequence variants (ASVs) stored there come from marine coastal sediments (Figure 1 and Figure S2). Furthermore, all sequences from deep environmental clade come from these environments (Figures 2 and 3 and Figure S7). This stands in contrast with the fact that most species described in *Cyphoderia* (14) are from freshwater environments, while only three species (*Cyphoderia littoralis*, *Cyphoderia compressa*, and *Cyphoderia truncata* [Schulze, 1875]) have been described from the marine environment. This knowledge gap demands more research on testate amoebae from marine ecosystems, a surprising fact given that current hypotheses based on fossils and phylogenomic reconstructions support a marine origin for the group (Porter et al., 2003; Porter & Knoll, 2000).

Most marine *Cyphoderiidae* specimens have been discovered in shallow areas near the coast (Figure 1; Bass & Cavalier-Smith, 2004). Some observations describe their presence in the supralittoral zone of sandy beaches, and especially abundant several metres inland from the coastline (Golemansky, 2007). Interestingly, in our study, the *Cyphoderia compressa* clade, a euryhaline group, was reconstructed as sister to freshwater clade 1, while marine clade 1 (sequences retrieved from sandy beaches, probably also euryhaline) appeared as sister to the freshwater species *Cyphoderia major* (Figures 2 and 3a). This suggests that euryhaline environments possibly hosted the ancestors of each of these freshwater cyphoderiid clades.

The euryhaline environment is exposed to quick changes in salinity (Golemansky, 2007), and fluctuating osmotic pressures. These fluctuations imply physiological stresses to which marine organisms need to be adapted. Euryhaline environments possibly

acted as a transition environment in the preadaptation of marine organisms to freshwater ecosystems, as these transitions, and vice versa, from freshwater to the marine habitat, are known to require more than one transitional step (Coates, 1993; Lee et al., 2012). Therefore, many species that crossed the salinity barrier seem to have descended from euryhaline ancestors that lived in naturally fluctuating environments; examples of this evolutionary pathway can be found in bony fishes, such as freshwater terapontid grunters (Davis et al., 2012), killifishes (Whitehead, 2010), or in the three-spined stickleback (Raeymaekers et al., 2005). Other freshwater animals exhibit similar patterns (Lee & Bell, 1999), as well as some clades of eukaryotes, such as cryptomonads (Shalchian-Tabrizi et al., 2008), dinoflagellates (Logares et al., 2007), diatoms (Alverson, 2014; Alverson et al., 2007), and even plants (Dittami et al., 2017). These dynamic euryhaline ecosystems have sometimes facilitated the reverse colonization of marine systems by freshwater organisms (for example, in some groups of insects such as the marine caddisflies, Riek, 1977), but this event is rare (Lee & Bell, 1999). For example, our analysis, based on a large data set of sequences retrieved from the EukBank repository, did not reconstruct any instance of recolonization of the marine environment by freshwater *Cyphoderiidae* (Figure 4). This asymmetry in the salinity barrier transition is common in eukaryotes (Betancur et al., 2015), and has been explained by a decreased probability of metabolic adaptation to the higher osmotic pressures that characterize the marine environment, or by competitive exclusion of freshwater species by marine resident communities (Bloom & Lovejoy, 2012; Whitehead, 2010). In *Cyphoderia*, one explanation is that the lack of salinity fluctuations in freshwater ecosystems prevented the colonization of euryhaline zones, due to either

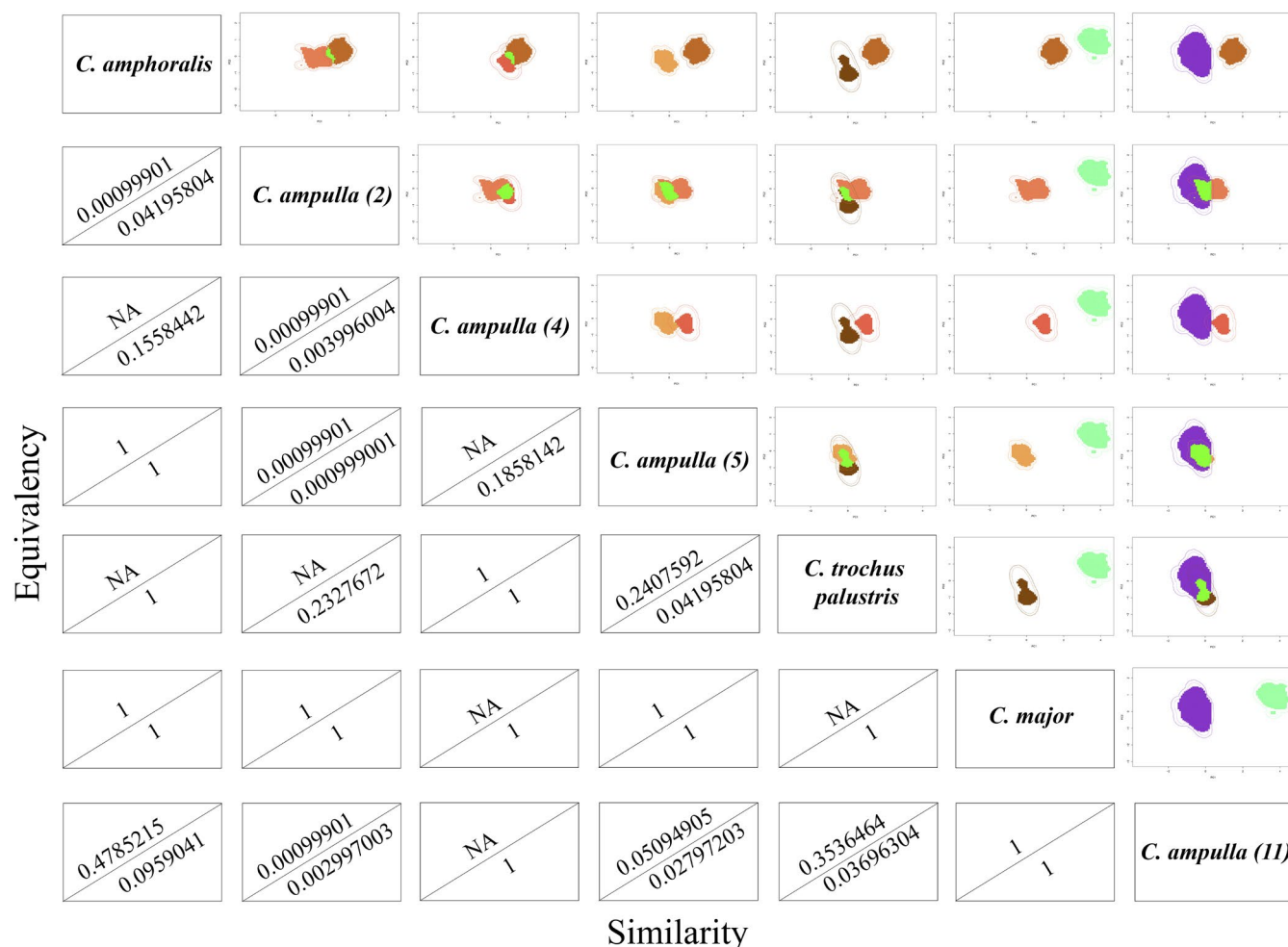


FIGURE 6 Top: composite represents the morphology overlaps between the phylogroups inferred from COI and SSU results (names in the diagonal); the x- and y-axes correspond to PC1 and PC2, respectively, and the green colour represents the overlap between each group. Bottom: morphology overlap comparisons between each morphogroups in base to the p-values obtained in the morphology equivalency/similarity (representing only A to B groups) tests

exclusive competition with marine clades or the loss of key adaptations to high salinity environments, or both.

From a common marine ancestor, our phylogenetic reconstruction indicates that there have been at least four independent events of transition from marine to freshwater environments in Cyphoderiidae: in the *Cyphoderia major* clade and in the freshwater clades 1, 2 and 3 (Figure 2a). This suggests that those key innovations that facilitated adaptation to freshwater environments in the family evolved independently from distantly related and, presumably, euryhaline ancestors. This potential case of parallel preadaptive evolution (Anker et al., 2006) provides strong evidence for a key adaptation in the evolutionary history of Cyphoderiidae (Futuyma, 1998), which could facilitate further adaptive radiation (Mayr, 1960). Similar parallel ecoevolutionary pathways have been found in other eukaryotic groups, such as fishes (Whitehead, 2010) and crustaceans (Mamos et al., 2016), as well as in some groups of protists like diatoms (Alverson et al., 2007) and dinoflagellates (Logares et al., 2007).

4.2 | Sea level fluctuations favour ecological transitions through the salinity barrier

Ecological transitions through the salinity barrier can occur at any time, but they are more likely to happen when marine incursions (transgressions) expand the contact zones between freshwater systems and the sea, often isolating landlocked populations in zones with fluctuating salinity (euryhaline) (Schultz & McCormick, 2012). Fluctuations in the sea level may drive and structure species distributions, connectivity between populations, and can also drive speciation-extinction patterns. As described above, the creation of euryhaline zones during marine transgressions provided an opportunity for marine organisms to colonize freshwater environments. However, these saltwater incursions could also have triggered extinction events in other, already-adapted freshwater lineages, opening ecological opportunities for new colonizers from the marine and euryhaline environments to diversify via competition release; the latter has been often posited as a driver of adaptive radiation (Losos,

2010; Yoder et al., 2010) in marine fishes (Betancur et al., 2012). For this reason, the glacial-interglacial cycles of the Pleistocene have been considered as key periods in the diversification of continental organisms. Eustatic changes in the sea level (Lambeck et al., 2002) during this epoch probably facilitated the ecological transition from marine to freshwater ecosystems in crustaceans (Dooh et al., 2006), and, in general, led to accelerated diversification rates in freshwater species (Hewitt, 2000).

Miocene ecosystems underwent even higher sea level fluctuations than the Pleistocene (Kominz et al., 2008; Westerhold et al., 2005), and this geological epoch was probably one of accelerated diversification rates (Miller et al., 2005). Indeed, our divergence time estimation analysis, and the sensitivity analysis exploring the uncertainty in lineage divergence times, suggest that ecological transitions through the salinity barrier in Cyphodermiidae date back to the mid-late Miocene (Figure 3a, Figure S6), a period of massive sea level decreases. We postulate here that the cyclic fluctuations of sea level during the mid-late Miocene generated new niches for marine Cyphodermiidae in periods of marine transgression (rise in sea level), allowing them to colonize freshwater ecosystems (via the euryhaline transition habitat) and favoured by the extinction of freshwater competitors. Subsequent speciation would have taken place in periods of marine regression (decrease in sea level) via competitive exclusion, thus generating parallel ecological transitions cyclically. The correspondence between Miocene sea-level fluctuations, transitions through the salinity barrier, and diversification have been extensively studied in fishes (Bloom & Lovejoy, 2017; Lovejoy et al., 2006; Tsigonopoulos et al., 2003). This pattern has also been recurrently observed in other eukaryotes, such as amphipods (Mamos et al., 2016; Yang et al., 2013), molluscs (Vermeij & Wesselingh, 2002), plants (Triest & Sierens, 2014) or diatoms (Hayashi et al., 2018). Again, cyphodermiids seem to follow evolutionary patterns that had been already described in macroscopic organisms, suggesting similarities in ecological transition pathways.

It has been demonstrated that diversification rates increase substantially with every transition through the salinity barrier. This event has been mostly described in clades of fishes (Bloom et al., 2013) and diatoms (Alverson, 2014). The explanation invoked is the lower connectivity of freshwater systems in comparison to the sea, along with the higher variability in abiotic factors. In our analysis, both state-dependent models (BiSSE in diversitree and REVSTATES) and the gamma test results support higher rates of cladogenesis in freshwater than marine lineages (Figure 3). However, a better fit to the null model (Table S5) suggests that the heterogeneity in diversification rates observed in Cyphodermiidae cannot be explained by transitions towards the freshwater habitat. Instead, the branch-specific diversification rate models (BAMM, REVSTATES BDS) suggest that speciation rates may be clade-dependent rather than state-dependent: for example, the group formed by *Cyphoderia* and Marine clade 3 exhibit higher diversification rates than other marine or freshwater clades in Cyphodermiidae, such as Freshwater clade 3 or Marine clade 4 (Figure 4).

The most recent common ancestor of genus *Cyphoderia* was dated at c. 50 Ma, during the Eocene, and the episodic birth-death

models in COMET and REVSTATES (Figure 4) show a gradual increase in the diversification rates from the Early Cenozoic onwards (Figure 4). Silica is a limiting factor in aquatic systems, and its increased bioavailability during the Eocene has been linked to diversification events in organisms that use this resource to build mineral structures (Lazarus et al., 2009; Rabosky & Sorhannus, 2009). It is possible that the higher diversification rate exhibited by genus *Cyphoderia* and the Marine clade 3 was favoured by the ecological opportunity offered by the increase in silica bioavailability during this epoch. Competition between cyphodermiids and other benthic groups (e.g., diatoms) for the water dissolved silica might have caused an acceleration of speciation events in these clades; similar hypotheses have been posited for planktonic groups that construct siliceous structures, such as radiolarians or diatoms (Falkowski et al., 2004; Katz et al., 2004; Sims et al., 2006). Furthermore, we observed a decrease in speciation rates in Cyphodermiidae starting during the Pliocene that is temporally congruent with similar patterns in diatoms and radiolarians (Katz et al., 2004; Lazarus et al., 2009). Though other explanations cannot be discarded (lack of resolution in the selected molecular markers, incomplete taxon sampling, etc), the observed similarity in diversification patterns between such phylogenetically distant groups as testate amoeba, diatoms, and radiolarians, suggest a clade-wide effect of abiotic factors (long-term environmental changes, geological events), acting at the biome level (Ezard et al., 2011). Further studies will help determine the drivers behind these patterns.

4.3 | Concomitant changes in morphology and adaptive value of the shell

Mineral elements have played a major role in the shaping of evolutionary patterns in many groups of microeukaryotes, such as coccolithophorids, diatoms and foraminiferans (Katz et al., 2004). When comparing freshwater and marine Cyphodermiidae, Golemansky (2007) noticed the larger size of shells in the freshwater forms, and hypothesized that the small and compressed shells of marine species were an adaptation to life between sand grains. Our study confirms that all surveyed freshwater species are larger than the marine isolates (Figures 5 and 6). All transitions inferred in Cyphodermiidae occurred from the marine environment towards the freshwater habitat (Figure 3). We observed that, alongside a general increase in the shell size, there was a homogenization of the shape after the transition, that is, freshwater forms became ballooned or expanded in comparison with marine forms (Figure 5, PC1 axis). Interestingly, other testate amoebae living in marine supralittoral systems (e.g., Arcellinida) are smaller than their freshwater counterparts (Golemansky & Todorov, 2004).

We thus hypothesize, reversing Golemansky's hypothesis, that the increase in size in freshwater species compared with marine lineages of Cyphodermiidae was an adaptive process acting on the freshwater lineages, and working directly or indirectly through the shell morphology. Adaptive selection in response to similar selective

pressures may also explain the polyphyletic nature of the complex *Cyphoderia ampulla sensu lato* (specimens named *C. ampulla* in Figure 2; and those in Freshwater clades 1 and 2 in Figure 3). Our phylomorphospace and principal component analysis show that Freshwater clades 1 and 2 are morphologically more similar to each other than to other phylogenetically closer marine species (i.e., *C. compressa* clade, Figures 5 and 6). However, while size and shape are conserved between specimens with *C. ampulla* morphology, the shape, size and distribution of the siliceous scales that form their shells are highly variable (Husnot, 1943). This morphological diversification of the shell scales was probably correlated with molecular diversification. Heger et al. (2011) noticed that *Cyphoderia ampulla* clades could be easily discriminated based on the overlapping or nonoverlapping of the scales, a finding that remains valid after adding new isolate sequences (Figure 2). In Euglyphida, in general, scale size has been shown to be highly variable and probably species-specific (Chatelain et al., 2013; Wylezich et al., 2002). Altogether, this suggests that scale shape and size are less prone to selective pressures than shell shape and size, and should therefore constitute better traits for clade delimitation in Cyphoderiidae systematics.

5 | CONCLUSIONS

Cyphoderiidae is one of the few groups of benthic marine protists that successfully managed the transition from marine to freshwater ecosystems. These transitions are inferred to have occurred during the mid-late Miocene, a period of frequent fluctuations in the sea level. The ancestors of present-day freshwater species were most probably marine coastal organisms that colonized freshwater systems with the euryhaline environment as a transitional step. This probably took place during marine transgressions, associated to ecological release, which was in turn favoured by the extinction of already-adapted freshwater clades, and the ecological opportunity offered by the increasing levels of water-dissolved silica. Finally, our study confirms that microscopic organisms such as benthic protists, may exhibit similar evolutionary patterns to those observed in macroscopic organisms, supporting the idea that global patterns apply to all life on Earth, no matter the organisms' size.

ACKNOWLEDGEMENTS

We express our gratitude to Dr I. García-Cunchillos, M. Blázquez, M. Rincón, M. Villar-DePablo for helpful discussions on the methodology used in the manuscript. We also acknowledge the help of Y. Ruiz-León (electron microscopy), E. Cano and M. García-Gallo (molecular biology laboratory), M. Mazuecos and S. Martín-Herranz (data analysis), as well as C. Zamora-Muñoz and J. J. Soler (fieldwork). This study was funded by the Spanish Government PGC2018-094660-B-I00 (MCIU/AEI/FEDER,UE) to E.L. and an "Atracción de Talento Investigador" grant awarded by the Consejería de Educación, Juventud y Deporte, Comunidad de Madrid (Spain) (2017-T1/AMB-5210) to R.G.M.; a National Geographic Grant EC-86022R-21 to R. G.

M. and E. L.; and a FPU fellowship FPU19/06077 from the Spanish Ministry of Science awarded to F.U.; I.S. was funded by a MICIN-AEI grant PID2019-108109GB-I00.

AUTHOR CONTRIBUTIONS

Rubén González-Miguéns and Enrique Lara designed the study; Rubén González-Miguéns, Carmen Soler-Zamora, Andrés Blanco-Rotea, María Isabel Carrasco-Braganza, David de Salvador-Velasco, Antonio Guillén-Oterino, and Daniel Tenorio-Rodríguez carried out the fieldwork; Rubén González-Miguéns and Carmen Soler-Zamora performed the laboratory work; Cédric Berney analysed the EukBank data set; Carmen Soler-Zamora prepared the images; Rubén González-Miguéns and Isabel Sanmartín performed the diversification analyses and subsequent interpretation of the results; Rubén González-Miguéns performed the rest of the analyses, interpreted and discussed the data together with Enrique Lara; Rubén González-Miguéns, Carmen Soler-Zamora, Fernando Useros, Sandra Nogal-Prata, Cédric Berney, David Velázquez, Thierry J. Heger, Isabel Sanmartín and Enrique Lara wrote the manuscript.

CONFLICT OF INTEREST

The authors have no conflicts of interest.

BENEFIT-SHARING STATEMENT

Benefits generated: There are no benefits outlined in the Nagoya protocol associated with this study to report. All data produced in this research are deposited in public databases as described above.

DATA AVAILABILITY STATEMENT

The genetic data generated in this article are available in GenBank Nucleotide Database at <https://www.ncbi.nlm.nih.gov/genbank/> and can be accessed with the accession number OK120268–OK120292 and OK058389. The morphological data is deposited in Morphobank and can be accessed at <http://morphobank.org/permalink/?P4082>. The nexus alignment and sensitive phylogenetic analyses for SSU without eDNA sequences, SSU and COI are deposited in TreeBASE by accessing through the links <http://purl.org/phylo/treebase/phyloids/study/TB2:S29006>, <http://purl.org/phylo/treebase/phyloids/study/TB2:S29003> and <http://purl.org/phylo/treebase/phyloids/study/TB2:S29007>, respectively.

ORCID

Rubén González-Miguéns  <https://orcid.org/0000-0002-8274-266X>

Carmen Soler-Zamora  <https://orcid.org/0000-0003-3777-6470>

Thierry J. Heger  <https://orcid.org/0000-0003-3614-0964>

Enrique Lara  <https://orcid.org/0000-0001-8500-522X>

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)

- Alverson, A. J. (2014). Timing marine–freshwater transitions in the diatom order Thalassiosirales. *Paleobiology*, 40(1), 91–101. <https://doi.org/10.1666/12055>
- Alverson, A. J., Jansen, R. K., & Theriot, E. C. (2007). Bridging the Rubicon: Phylogenetic analysis reveals repeated colonizations of marine and fresh waters by thalassiosirid diatoms. *Molecular Phylogenetics and Evolution*, 45(1), 193–210. <https://doi.org/10.1016/j.ympev.2007.03.024>
- Anker, A., Ah Yong, S. T., Noel, P. Y., & Palmer, A. R. (2006). Morphological phylogeny of alpheid shrimps: Parallel preadaptation and the origin of a key morphological innovation, the snapping claw. *Evolution*, 60(12), 2507–2528. <https://doi.org/10.1111/j.0014-3820.2006.tb01886.x>
- Baele, G., Li, W. L. S., Drummond, A. J., Suchard, M. A., & Lemey, P. (2013). Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Molecular Biology and Evolution*, 30(2), 239–243. <https://doi.org/10.1093/molbev/mss243>
- Barber, A., Siver, P. A., & Karis, W. (2013). Euglyphid testate amoebae (Rhizaria: Euglyphida) from an Arctic Eocene waterbody: Evidence of evolutionary stasis in plate morphology for over 40 million years. *Protist*, 164(4), 541–555. <https://doi.org/10.1016/j.protis.2013.05.001>
- Bass, D., & Cavalier-Smith, T. (2004). Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of Cercozoa (Protozoa). *International Journal of Systematic and Evolutionary Microbiology*, 54(6), 2393–2404. <https://doi.org/10.1099/ijs.0.63229-0>
- Beaulieu, J. M., & O'Meara, B. C. (2016). Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology*, 65(4), 583–601. <https://doi.org/10.1093/sysbio/syw022>
- Berney, C., Ciuprina, A., Bender, S., Brodie, J., Edgcomb, V., Kim, E., ... de Vargas, C. (2017). UniEuk: Time to speak a common language in protistology! *Journal of Eukaryotic Microbiology*, 64(3), 407–411. <https://doi.org/10.1111/jeu.12414>
- Betancur, R., Ortí, G., & Pyron, R. A. (2015). Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. *Ecology Letters*, 18(5), 441–450. <https://doi.org/10.1111/ele.12423>
- Betancur-R., R., Ortí, G., Stein, A. M., Marceniuk, A. P., & Alexander Pyron, R. (2012). Apparent signal of competition limiting diversification after ecological transitions from marine to freshwater habitats. *Ecology Letters*, 15(8), 822–830. <https://doi.org/10.1111/j.1461-0248.2012.01802.x>
- Blonder, B., Lamanna, C., Violle, C., & Enquist, B. J. (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography*, 23(5), 595–609. <https://doi.org/10.1111/geb.12146>
- Bloom, D. D., & Lovejoy, N. R. (2012). Molecular phylogenetics reveals a pattern of biome conservatism in New World anchovies (family Engraulidae). *Journal of Evolutionary Biology*, 25(4), 701–715. <https://doi.org/10.1111/j.1420-9101.2012.02464.x>
- Bloom, D. D., & Lovejoy, N. R. (2017). On the origins of marine-derived freshwater fishes in South America. *Journal of Biogeography*, 44(9), 1927–1938. <https://doi.org/10.1111/jbi.12954>
- Bloom, D. D., Weir, J. T., Piller, K. R., & Lovejoy, N. R. (2013). Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of New World silversides (Atherinopsidae). *Evolution*, 67(7), 2040–2057. <https://doi.org/10.1111/evo.12072>
- Boeuf, O., & Gilbert, D. (1997). Présence de Thécamoebiens du genre *Trinema*, au Pliocène supérieur, découverte à Chilhac (Haute-Loire, France). *Comptes Rendus de l'Académie des Sciences – Series IIA – Earth and Planetary Science*, 325(8), 623–627. [https://doi.org/10.1016/S1251-8050\(97\)89464-5](https://doi.org/10.1016/S1251-8050(97)89464-5)
- Bollback, J. P. (2006). SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics*, 7(1), 88. <https://doi.org/10.1186/1471-2105-7-88>
- Boltovskoy, D. (2003). First record of a brackish radiolarian (Polycystina): *Lophophaena rioplatensis* n. sp. in the Rio de la Plata estuary. *Journal of Plankton Research*, 25(12), 1551–1559. <https://doi.org/10.1093/plankt/fbg107>
- Broennimann, O., Cola, V. D., Petitpierre, B., Breiner, F., Scherrer, D., D. Amen, M., ... Guisan, A. (2018). *Ecospat: spatial ecology miscellaneous methods. Package "ecospat"* (R Package Version 3.0). <http://cran.r-project.org/web/packages/ecospat/ecospat.pdf>
- Broennimann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N. G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N. E., Graham, C. H., & Guisan, A. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, 21(4), 481–497. <https://doi.org/10.1111/j.1466-8238.2011.00698.x>
- Chatelain, A. P., Meisterfeld, R., Roussel-Delif, L., & Lara, E. (2013). Sphenoderiidae (fam. nov.), a new clade of euglyphid testate amoebae characterized by small, round scales surrounding the aperture. *Protist*, 164(6), 782–792. <https://doi.org/10.1016/j.protis.2013.08.001>
- Chomczynski, P., & Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analytical Biochemistry*, 162(1), 156–159. <https://doi.org/10.1006/abio.1987.9999>
- Coates, D. (1993). Fish ecology and management of the Sepik-Ramu, New Guinea, a large contemporary tropical river basin. *Environmental Biology of Fishes*, 38(4), 345–368. <https://doi.org/10.1007/BF00007528>
- Davis, A. M., Unmack, P. J., Pusey, B. J., Johnson, J. B., & Pearson, R. G. (2012). Marine–freshwater transitions are associated with the evolution of dietary diversification in terapontid grunners (Teleostei: Terapontidae). *Journal of Evolutionary Biology*, 25(6), 1163–1179. <https://doi.org/10.1111/j.1420-9101.2012.02504.x>
- Davis, A. M., Unmack, P. J., Vari, R. P., & Betancur-R., R. (2016). Herbivory promotes dental disparification and macroevolutionary dynamics in grunners (Teleostei: Terapontidae), a freshwater adaptive radiation. *The American Naturalist*, 187(3), 320–333. <https://doi.org/10.1086/684747>
- Dittami, S. M., Heesch, S., Olsen, J. L., & Collén, J. (2017). Transitions between marine and freshwater environments provide new clues about the origins of multicellular plants and algae. *Journal of Phycology*, 53(4), 731–745. <https://doi.org/10.1111/jpy.12547>
- Dooh, R. T., Adamowicz, S. J., & Hebert, P. D. N. (2006). Comparative phylogeography of two North American 'glacial relict' crustaceans. *Molecular Ecology*, 15(14), 4459–4475. <https://doi.org/10.1111/j.1365-294X.2006.03095.x>
- 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(8), 1969–1973. <https://doi.org/10.1093/molbev/mss075>
- Duckert, C., Blandinier, Q., Kupferschmid, F. A. L., Kosakyan, A., Mitchell, E. A. D., Lara, E., & Singer, D. (2018). En garde! Redefinition of *Nebela militaris* (Arcellinida, Hyalospheniidae) and erection of *Alabasta* gen. nov. *European Journal of Protistology*, 66, 156–165. <https://doi.org/10.1016/j.ejop.2018.08.005>
- Eberli, G. P. (2000). The record of Neogene sea-level changes in the prograding carbonates along the Bahamas Transect—Leg 166 Synthesis. In *Proceedings of the Ocean Drilling Program, 166 scientific results* (Vol. 166, pp. 167–177). Ocean Drilling Program. doi: <https://doi.org/10.2973/odp.proc.sr.166.128.2000>
- Ezard, T. H. G., Aze, T., Pearson, P. N., & Purvis, A. (2011). Interplay between changing climate and species' ecology drives macroevolutionary dynamics. *Science*, 332, 349–351. <https://doi.org/10.1126/science.1203060>
- Falkowski, P. G., Katz, M. E., Knoll, A. H., Quigg, A., Raven, J. A., Schofield, O., & Taylor, F. J. R. (2004). The evolution of modern eukaryotic phytoplankton. *Science*, 305(5682), 354–360. <https://doi.org/10.1126/science.1095964>

- FitzJohn, R. G. (2012). Diversitree: Comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, 3(6), 1084–1092. <https://doi.org/10.1111/j.2041-210X.2012.00234.x>
- Foissner, W., & Schiller, W. (2001). Stable for 15 million years: Scanning electron microscope investigation of Miocene euglyphid thecamoebians from Germany, with description of the new genus *Scutiglypha*. *European Journal of Protistology*, 37(2), 167–180. <https://doi.org/10.1078/0932-4739-00012>
- Freyman, W. A., & Höhna, S. (2019). Stochastic character mapping of state-dependent diversification reveals the tempo of evolutionary decline in self-compatible Onagraceae lineages. *Systematic Biology*, 68(3), 505–519. <https://doi.org/10.1093/sysbio/syy078>
- Futuyama, D. J. (1998). *Evolutionary biology* (3rd ed.). Sinauer Associates.
- Gernhard, T. (2008). The conditioned reconstructed process. *Journal of Theoretical Biology*, 253(4), 769–778. <https://doi.org/10.1016/j.jtbi.2008.04.005>
- Golemansky, V. (1970). Rhizopodes nouveaux du psammon littoral de la Mer Noire (Note préliminaire). *Protistologica*, 6(4), 365–371.
- Golemansky, V. (1979). *Cyphoderia compressa* sp. n. (Rhizopoda: Arcellinida) – Un nouveau thécamoebien psammobionte de supralittoral des mers. *Acta Protozoologica*, 18(3), 429–434.
- Golemansky, V. (2007). Testate amoebae and monothalamous foraminifera (protozoa) from the Bulgarian Black Sea Coast. In V. Fet & A. Popov, (Eds.), *Biogeography and ecology of Bulgaria* (pp. 555–570). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5781-6_20
- Golemansky, V., & Todorov, M. (2004). Shell morphology, biometry and distribution of some marine interstitial testate amoebae (Sarcodina: Rhizopoda). *Acta Protozoologica*, 43(2), 147–162.
- González-Miguéns, R., Soler-Zamora, C., Villar-Depablo, M., Todorov, M., & Lara, E. (2021). Multiple convergences in the evolutionary history of the testate amoeba family Arcellinidae (Amoebozoa: Arcellinida: Sphaerothecina): When the ecology rules the morphology. *Zoological Journal of the Linnean Society*, 1–28. in press. <https://doi.org/10.1093/zoolinnean/zlab074>
- Gray, J. (1988). Evolution of the freshwater ecosystem: The fossil record. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 62(1–4), 1–214. [https://doi.org/10.1016/0031-0182\(88\)90054-5](https://doi.org/10.1016/0031-0182(88)90054-5)
- Hayashi, T., Krebs, W. N., Saito-Kato, M., & Tanimura, Y. (2018). The turnover of continental planktonic diatoms near the middle/late Miocene boundary and their Cenozoic evolution. *PLoS One*, 13(6), e0198003. <https://doi.org/10.1371/journal.pone.0198003>
- Heger, T. J., Mitchell, E. A. D., Todorov, M., Golemansky, V., Lara, E., Leander, B. S., & Pawlowski, J. (2010). Molecular phylogeny of euglyphid testate amoebae (Cerczoa: Euglyphida) suggests transitions between marine supralittoral and freshwater/terrestrial environments are infrequent. *Molecular Phylogenetics and Evolution*, 55(1), 113–122. <https://doi.org/10.1016/j.ympev.2009.11.023>
- Heger, T. J., Pawlowski, J., Lara, E., Leander, B. S., Todorov, M., Golemansky, V., & Mitchell, E. A. D. (2011). Comparing potential COI and SSU rDNA barcodes for assessing the diversity and phylogenetic relationships of cyphoderiid testate amoebae (Rhizaria: Euglyphida). *Protist*, 162(1), 131–141. <https://doi.org/10.1016/j.protis.2010.05.002>
- Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789), 907–913. <https://doi.org/10.1038/35016000>
- Höhna, S. (2014). Likelihood inference of non-constant diversification rates with incomplete taxon sampling. *PLoS One*, 9(1), e84184. <https://doi.org/10.1371/journal.pone.0084184>
- Höhna, S., Freyman, W. A., Nolen, Z., Huelsenbeck, J. P., May, M. R., & Moore, B. R. (2019). A Bayesian approach for estimating branch-specific speciation and extinction rates. *BioRxiv*. <https://doi.org/10.1101/555805>
- Höhna, S., Heath, T. A., Boussau, B., Landis, M. J., Ronquist, F., & Huelsenbeck, J. P. (2014). Probabilistic graphical model representation in phylogenetics. *Systematic Biology*, 63(5), 753–771. <https://doi.org/10.1093/sysbio/syu039>
- Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., Huelsenbeck, J. P., & Ronquist, F. (2016). RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic Biology*, 65(4), 726–736. <https://doi.org/10.1093/sysbio/syw021>
- Höhna, S., May, M. R., & Moore, B. R. (2016). TESS: An R package for efficiently simulating phylogenetic trees and performing Bayesian inference of lineage diversification rates. *Bioinformatics*, 32(5), 789–791. <https://doi.org/10.1093/bioinformatics/btv651>
- Hong, M., Jiang, A., Li, N., Li, W., Shi, H., Storey, K. B., & Ding, L. (2019). Comparative analysis of the liver transcriptome in the red-eared slider *Trachemys scripta elegans* under chronic salinity stress. *PeerJ*, 7(3), e6538. <https://doi.org/10.7717/peerj.6538>
- Huelsenbeck, J. P., Nielsen, R., & Bollback, J. P. (2003). Stochastic mapping of morphological characters. *Systematic Biology*, 52(2), 131–158. <https://doi.org/10.1080/10635150390192780>
- Husnot, P. T. (1943). In P. Lechevalier (Ed.), *Les Cyphoderia de la vallée du Gouedic, en Saint-Brieuc*, (p. 144). Paul Lechevalier.
- Hutchinson, G. E. (1975). *A treatise on limnology* (Vol. III). Wiley.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773–795. <https://doi.org/10.1080/01621459.1995.10476572>
- Kassambara, A., & Mundt, F. (2020). *factoextra: Extract and visualize the results of multivariate data analyses*. Package version 1.0.7. R package version website <https://cran.r-project.org/package=factoextra>
- Katz, M. E., Finkel, Z. V., Grzebyk, D., Knoll, A. H., & Falkowski, P. G. (2004). Evolutionary trajectories and biogeochemical impacts of marine eukaryotic phytoplankton. *Annual Review of Ecology, Evolution, and Systematics*, 35(1), 523–556. <https://doi.org/10.1146/annurev.ecolsys.35.112202.130137>
- Kolmann, M. A., Burns, M. D., Ng, J. Y. K., Lovejoy, N. R., & Bloom, D. D. (2020). Habitat transitions alter the adaptive landscape and shape phenotypic evolution in needlefishes (Belontiidae). *Ecology and Evolution*, 10(8), 3769–3783. <https://doi.org/10.1002/ece3.6172>
- Kominz, M. A., Browning, J. V., Miller, K. G., Sugarman, P. J., Mizintseva, S., & Scotese, C. R. (2008). Late Cretaceous to Miocene sea-level estimates from the New Jersey and Delaware coastal plain coreholes: An error analysis. *Basin Research*, 20(2), 211–226. <https://doi.org/10.1111/j.1365-2117.2008.00354.x>
- Kosakyan, A., Goma, F., Lara, E., & Lahr, D. J. G. (2016). Current and future perspectives on the systematics, taxonomy and nomenclature of testate amoebae. *European Journal of Protistology*, 55(Part B), 105–117. <https://doi.org/10.1016/j.ejop.2016.02.001>
- Lambeck, K., Esat, T. M., & Potter, E.-K. (2002). Links between climate and sea levels for the past three million years. *Nature*, 419(6903), 199–206. <https://doi.org/10.1038/nature01089>
- Laudanno, G., Haegeman, B., Rabosky, D. L., & Etienne, R. S. (2021). Detecting lineage-specific shifts in diversification: A proper likelihood approach. *Systematic Biology*, 70(2), 389–407. <https://doi.org/10.1093/sysbio/syaa048>
- Lazarus, D. B., Kotrc, B., Wulf, G., & Schmidt, D. N. (2009). Radiolarians decreased silicification as an evolutionary response to reduced Cenozoic ocean silica availability. *Proceedings of the National Academy of Sciences of the United States of America*, 106(23), 9333–9338. <https://doi.org/10.1073/pnas.0812979106>
- Lee, C. E., & Bell, M. A. (1999). Causes and consequences of recent freshwater invasions by saltwater animals. *Trends in Ecology & Evolution*, 14(7), 284–288. [https://doi.org/10.1016/S0169-5347\(99\)01596-7](https://doi.org/10.1016/S0169-5347(99)01596-7)
- Lee, C. E., Posavi, M., & Charmantier, G. (2012). Rapid evolution of body fluid regulation following independent invasions into freshwater habitats. *Journal of Evolutionary Biology*, 25(4), 625–633. <https://doi.org/10.1111/j.1420-9101.2012.02459.x>
- Logares, R., Bråte, J., Bertilsson, S., Clasen, J. L., Shalchian-Tabrizi, K., & Rengefors, K. (2009). Infrequent marine–freshwater transitions in the microbial world. *Trends in Microbiology*, 17(9), 414–422. <https://doi.org/10.1016/j.tim.2009.05.010>

- Logares, R., Shalchian-Tabrizi, K., Boltovskoy, A., & Renfjors, K. (2007). Extensive dinoflagellate phylogenies indicate infrequent marine-freshwater transitions. *Molecular Phylogenetics and Evolution*, 45(3), 887–903. <https://doi.org/10.1016/j.ympev.2007.08.005>
- López-Estrada, E. K., Sanmartín, I., Uribe, J. E., Abalde, S., & García-París, M. (2021). Diversification dynamics of hypermetamorphic blister beetles (Meloidae): Are homoplastic host shifts and phoresy key factors of a rushing forward strategy to escape extinction? *BioRxiv*. <https://doi.org/10.1101/2021.01.04.425192>
- Losos, J. B. (2010). Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American Naturalist*, 175(6), 623–639. <https://doi.org/10.1086/652433>
- Lovejoy, N. R., Albert, J. S., & Crampton, W. G. R. (2006). Miocene marine incursions and marine/freshwater transitions: Evidence from Neotropical fishes. *Journal of South American Earth Sciences*, 21(1–2), 5–13. <https://doi.org/10.1016/j.jsames.2005.07.009>
- Maddison, W. P., & FitzJohn, R. G. (2015). The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology*, 64(1), 127–136. <https://doi.org/10.1093/sysbio/syu070>
- Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, 56(5), 701–710. <https://doi.org/10.1080/10635150701607033>
- Magee, A. F., Höhna, S., Vasylyeva, T. I., Leaché, A. D., & Minin, V. N. (2020). Locally adaptive Bayesian birth-death model successfully detects slow and rapid rate shifts. *PLoS Computational Biology*, 16(10), e1007999. <https://doi.org/10.1371/journal.pcbi.1007999>
- Mamos, T., Wattier, R., Burzyński, A., & Grabowski, M. (2016). The legacy of a vanished sea: A high level of diversification within a European freshwater amphipod species complex driven by 15 My of Paratethys regression. *Molecular Ecology*, 25(3), 795–810. <https://doi.org/10.1111/mec.13499>
- Mansouri, M., Naghavi, M. R., Alizadeh, H., Mohammadi-Nejad, G., Mousavi, S. A., Salekdeh, G. H., & Tada, Y. (2019). Transcriptomic analysis of *Aegilops tauschii* during long-term salinity stress. *Functional & Integrative Genomics*, 19(1), 13–28. <https://doi.org/10.1007/s10142-018-0623-y>
- May, M. R., Höhna, S., & Moore, B. R. (2016). A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods in Ecology and Evolution*, 7(8), 947–959. <https://doi.org/10.1111/2041-210X.12563>
- Mayr, E. (1960). The emergence of evolutionary novelties. In S. Tax, (Ed.), *Evolution after Darwin, Vol. 1, The evolution of life: Its origin, history, and future*, (pp. 349–380). Chicago: The University of Chicago Press.
- Miller, K. G., Kominz, M. A., Browning, J. V., Wright, J. D., Mountain, G. S., Katz, M. E., Sugarman, P. J., Cramer, B. S., Christie-Blick, N., & Pekar, S. F. (2005). The Phanerozoic record of global sea-level change. *Science*, 310(5752), 1293–1298. <https://doi.org/10.1126/science.1116412>
- Mitchell, J. S., & Rabosky, D. L. (2017). Bayesian model selection with BAMM: Effects of the model prior on the inferred number of diversification shifts. *Methods in Ecology and Evolution*, 8(1), 37–46. <https://doi.org/10.1111/2041-210X.12626>
- Moore, B. R., Höhna, S., May, M. R., Rannala, B., & Huelsenbeck, J. P. (2016). Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences of the United States of America*, 113(34), 9569–9574. <https://doi.org/10.1073/pnas.1518659113>
- Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1), 268–274. <https://doi.org/10.1093/molbev/msu300>
- Pallarés, S., Arribas, P., Bilton, D. T., Millán, A., Velasco, J., & Ribera, I. (2017). The chicken or the egg? Adaptation to desiccation and salinity tolerance in a lineage of water beetles. *Molecular Ecology*, 26(20), 5614–5628. <https://doi.org/10.1111/mec.14334>
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20(2), 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Porter, S. M., & Knoll, A. H. (2000). Testate amoebae in the Neoproterozoic Era: Evidence from vase-shaped microfossils in the Chuar Group, Grand Canyon. *Paleobiology*, 26(3), 360–385. [https://doi.org/10.1666/0094-8373\(2000\)026<0360:TAITNE>2.0.CO;2](https://doi.org/10.1666/0094-8373(2000)026<0360:TAITNE>2.0.CO;2)
- Porter, S. M., Meisterfeld, R., & Knoll, A. H. (2003). Vase-shaped microfossils from the Neoproterozoic Chuar Group, Grand Canyon: A classification guided by modern testate amoebae. *Journal of Paleontology*, 77(3), 409–429. <https://doi.org/10.1017/S0022336000044140>
- Pybus, O. G., & Harvey, P. H. (2000). Testing macro-evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1459), 2267–2272. <https://doi.org/10.1098/rspb.2000.1278>
- Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One*, 9(2), e89543. <https://doi.org/10.1371/journal.pone.0089543>
- Rabosky, D. L., Donnellan, S. C., Grundler, M., & Lovette, I. J. (2014). Analysis and visualization of complex macroevolutionary dynamics: An example from Australian scincid lizards. *Systematic Biology*, 63(4), 610–627. <https://doi.org/10.1093/sysbio/syu025>
- Rabosky, D. L., & Goldberg, E. E. (2015). Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology*, 64(2), 340–355. <https://doi.org/10.1093/sysbio/syu131>
- Rabosky, D. L., & Goldberg, E. E. (2017). FiSSE: A simple nonparametric test for the effects of a binary character on lineage diversification rates. *Evolution*, 71(6), 1432–1442. <https://doi.org/10.1111/evo.13227>
- Rabosky, D. L., Grundler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W., Huang, H., & Larson, J. G. (2014). BAMMtools: An R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, 5(7), 701–707. <https://doi.org/10.1111/2041-210X.12199>
- Rabosky, D. L., & Sorhannus, U. (2009). Diversity dynamics of marine planktonic diatoms across the Cenozoic. *Nature*, 457(7226), 183–186. <https://doi.org/10.1038/nature07435>
- Raeymaekers, J. A. M., Maes, G. E., Audenaert, E., & Volckaert, F. A. M. (2005). Detecting Holocene divergence in the anadromous-freshwater three-spined stickleback (*Gasterosteus aculeatus*) system. *Molecular Ecology*, 14(4), 1001–1014. <https://doi.org/10.1111/j.1365-294X.2005.02456.x>
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901–904. <https://doi.org/10.1093/sysbio/syy032>
- Rautsaw, R. M., Schrammer, T. D., Acuña, R., Arick, L. N., DiMeo, M., Mercier, K. P., Schrum, M., Mason, A. J., Margres, M. J., Strickland, J. L., & Parkinson, C. L. (2021). Genomic adaptations to salinity resist gene flow in the evolution of Floridian watersnakes. *Molecular Biology and Evolution*, 38(3), 745–760. <https://doi.org/10.1093/molbev/msaa266>
- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3(2), 217–223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>
- Revell, L. J., & Betancur, R. (2017). Exercise 14: State dependent diversification analysis using diversitree. <http://www.phytools.org/Cordoba2017/ex/14/SSE-models.html>
- Riek, E. F. (1977). The marine caddisfly family Chathamidae (Trichoptera). *Australian Journal of Entomology*, 15(4), 405–419. <https://doi.org/10.1111/j.1440-6055.1976.tb01724.x>
- Sanmartín, I., & Meseguer, A. S. (2016). Extinction in phylogenetics and biogeography: From timetrees to patterns of biotic assemblage. *Frontiers in Genetics*, 7, 35. <https://doi.org/10.3389/fgene.2016.00035>

- Santini, F., Nguyen, M. T. T., Sorenson, L., Waltzek, T. B., Lynch Alfaro, J. W., Eastman, J. M., & Alfaro, M. E. (2013). Do habitat shifts drive diversification in teleost fishes? An example from the pufferfishes (Tetraodontidae). *Journal of Evolutionary Biology*, 26(5), 1003–1018. <https://doi.org/10.1111/jeb.12112>
- Schiller, W. (1997). Kieselige thekamöben aus der miozänen Kieselgur von Beuern/Vogelsberg im vergleich mit rezentem material von Borneo (Malaysia). *CFS Courier Forschungsinstitut Senckenberg*, 201, 385–392.
- Schiller, W. (1999). Kieselige Thekamöben aus dem Mittel-Eozän des Eckfelder Maeres in der Eifel. *Mainzer Naturwissenschaftliches Archiv*, 37, 55–62.
- Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford University Press.
- Schultz, E. T., & McCormick, S. D. (2012). Euryhalinity in an evolutionary context. In S. D. McCormick, A. P. Farrell & C. J. Brauner, (Eds.), *Fish physiology*, 32, 477–533. Academic Press. <https://doi.org/10.1016/B978-0-12-396951-4.00010-4>
- Shalchian-Tabrizi, K., Minge, M. A., Espelund, M., Orr, R., Ruden, T., Jakobsen, K. S., & Cavalier-Smith, T. (2008). Multigene phylogeny of Choanozoa and the origin of animals. *PLoS One*, 3(5), e2098. <https://doi.org/10.1371/journal.pone.0002098>
- Sidlauskas, B. (2008). Continuous and arrested morphological diversification in sister clades of characiform fishes: A phylomorphospace approach. *Evolution*, 62(12), 3135–3156. <https://doi.org/10.1111/j.1558-5646.2008.00519.x>
- Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge. <https://doi.org/10.1201/9781315140919>
- Simpson, G. G. (1984). *Tempo and mode in evolution*. Columbia University Press. <https://doi.org/10.7312/simp93040>
- Sims, P. A., Mann, D. G., & Medlin, L. K. (2006). Evolution of the diatoms: Insights from fossil, biological and molecular data. *Phycologia*, 45(4), 361–402. <https://doi.org/10.2216/05-22.1>
- Taylor, F. J. R. M. (2003). The collapse of the two-kingdom system, the rise of protistology and the founding of the International Society for Evolutionary Protistology (ISEP). *International Journal of Systematic and Evolutionary Microbiology*, 53(6), 1707–1714. <https://doi.org/10.1099/ijls.0.02587-0>
- Todorov, M., Golemansky, V., Mitchell, E. A. D., & Heger, T. J. (2009). Morphology, biometry, and taxonomy of freshwater and marine interstitial Cyphoderia (Cercaria: Euglyphida). *Journal of Eukaryotic Microbiology*, 56(3), 279–289. <https://doi.org/10.1111/j.1550-7408.2009.00394.x>
- Tribble, C. M., Freyman, W. A., Landis, M. J., Lim, Y., Barido-Sottani, J., Kopperud, B. T., ... May, M. R. (2021). RevGadgets: An R Package for visualizing Bayesian phylogenetic analyses from RevBayes. *BioRxiv*. <https://doi.org/10.1101/2021.05.10.443470>
- Triest, L., & Sierens, T. (2014). Seagrass radiation after Messinian salinity crisis reflected by strong genetic structuring and out-of-Africa scenario (Ruppiales). *PLoS One*, 9(8), e104264. <https://doi.org/10.1371/journal.pone.0104264>
- Tsigenopoulos, C. S., Durand, J. D., Önlü, E., & Berrebi, P. (2003). Rapid radiation of the Mediterranean *Luciobarbus* species (Cyprinidae) after the Messinian salinity crisis of the Mediterranean Sea, inferred from mitochondrial phylogenetic analysis. *Biological Journal of the Linnean Society*, 80(2), 207–222. <https://doi.org/10.1046/j.1095-8312.2003.00237.x>
- Vargas, P., Jiménez-Mejías, P., & Fernández-Mazuecos, M. (2020). 'Endangered living fossils' (ELFs): Long-term survivors through periods of dramatic climate change. *Environmental and Experimental Botany*, 170, 103892. <https://doi.org/10.1016/j.envexpbot.2019.103892>
- Vargas, P., Valente, L. M., Blanco-Pastor, J. L., Liberal, I., Guzmán, B., Cano, E., Forrest, A., & Fernández-Mazuecos, M. (2014). Testing the biogeographical congruence of palaeofloras using molecular phylogenetics: Snapdragons and the Madiran-Tethyan flora. *Journal of Biogeography*, 41(5), 932–943. <https://doi.org/10.1111/jbi.12253>
- Vermeij, G. J., & Wesselingh, F. P. (2002). Neogastropod molluscs from the Miocene of western Amazonia, with comments on marine to freshwater transitions in molluscs. *Journal of Paleontology*, 76(2), 265–270. <https://doi.org/10.1017/s002233600004169x>
- Waggoner, B. M. (1996). The first fossil cyphoderiid testate amoeba, in Dominican Republic amber (Eocene-Oligocene). *Paleobios*, 17(1), 17–19.
- Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution*, 62(11), 2868–2883. <https://doi.org/10.1111/j.1558-5646.2008.00482.x>
- Westerhold, T., Bickert, T., & Röhl, U. (2005). Middle to late Miocene oxygen isotope stratigraphy of ODP site 1085 (SE Atlantic): New constraints on Miocene climate variability and sea-level fluctuations. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 217(3–4), 205–222. <https://doi.org/10.1016/j.palaeo.2004.12.001>
- Whitehead, A. (2010). The evolutionary radiation of diverse osmotolerant physiologies in killifish (*Fundulus* sp.). *Evolution*, 64(7), 2070–2085. <https://doi.org/10.1111/j.1558-5646.2010.00957.x>
- Wickham, H. (2016). *ggplot2 elegant graphics for data analysis (use R!)*. Springer.
- Wylezich, C., Meisterfeld, R., Meisterfeld, S., & Schlegel, M. (2002). Phylogenetic analyses of small subunit ribosomal RNA coding regions reveal a monophyletic lineage of euglyphid testate amoebae (order Euglyphida). *The Journal of Eukaryotic Microbiology*, 49(2), 108–118. <https://doi.org/10.1111/j.1550-7408.2002.tb00352.x>
- Yang, L., Hou, Z., & Li, S. (2013). Marine incursion into East Asia: A forgotten driving force of biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, 280(1757), 20122892. <https://doi.org/10.1098/rspb.2012.2892>
- Yoder, J. B., Clancey, E., Des roches, S., Eastman, J. M., Gentry, L., Godsoe, W., Hagey, T. J., Jochimsen, D., Oswald, B. P., Robertson, J., Sarver, B. A. J., Schenk, J. J., Spear, S. F., & Harmon, L. J. (2010). Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, 23(8), 1581–1596. <https://doi.org/10.1111/j.1420-9101.2010.02029.x>
- Zhang, X., Wen, H., Wang, H., Ren, Y., Zhao, J., & Li, Y. (2017). RNA-Seq analysis of salinity stress-responsive transcriptome in the liver of spotted sea bass (*Lateolabrax maculatus*). *PLoS One*, 12(3), e0173238. <https://doi.org/10.1371/journal.pone.0173238>

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: González-Miguéns, R., Soler-Zamora, C., Useros, F., Nogal-Prata, S., Berney, C., Blanco-Rotea, A., Carrasco-Braganza, M. I., de Salvador-Velasco, D., Guillén-Oterino, A., Tenorio-Rodríguez, D., Velázquez, D., Heger, T. J., Sanmartín, I., & Lara, E. (2022). *Cyphoderia ampulla* (Cyphoderiidae: Rhizaria), a tale of freshwater sailors: The causes and consequences of ecological transitions through the salinity barrier in a family of benthic protists. *Molecular Ecology*, 31, 2644–2663. <https://doi.org/10.1111/mec.16424>

APPENDIX 1

MATERIAL AND METHODS EXTENDED

DNA extraction and amplification

Total DNA was extracted from 25 organisms from genus *Cyphoderia* (family Cyphoderiidae); four of them were assigned to *Cyphoderia major*, three to *Cyphoderia amphoralis* and 18 to *Cyphoderia ampulla* (Table S1) based on morphological identification.

For each cell, we amplified a portion of the *cytochrome oxidase subunit I* (COI) into a final reaction volume of 23 µl containing 6 µl of distilled water, 12 µl MyTaq Red DNA polymerase Mix (BioLine), 1 µl of each primer (10 µmol), and 5 µl of DNA template. We applied a two-step protocol using the primers designed by Heger et al. (2011); Eucox1F (5'-GAYATGGCKTTNCCAAGATTTAA-3') and Euglycox1R (5'-AGCACCCATTGAHAAACRTAATG-3') or Euglycox6R (5'-GTTGGWACWGCTATWATCATWGT-3') in the first PCR and a semi-nested PCR with Euglycox1R or Euglycox6R in combination with Eucox5F (5'-ACAGGTGGRACYRTTATC-3') and Eucox6F (5'-GAAGCWGGWGTWGGDACAGG-3'). We used the same PCR protocol for both PCRs, which included an initial denaturation at 96°C for 5 min followed by 40 cycles of 96°C for 30 s; 40°C for 30 s; and 72°C for 90 s and a final extension at 72°C for 10 min (Heger et al., 2011).

The SSU rRNA gene was amplified in a final reaction volume of 20 µl containing 6 µl of distilled water, 12 µl MyTaq Red DNA polymerase Mix (BioLine), 1 µl of each primer (10 µmol), and 2 of DNA template. We used the broadscale eukaryotic primers EK555F (5'-AGTCTGGTGCCAGCAGCCGC-3') and EK1498R (5'-CACCTACGGAAACCTGTGA-3'). The PCR profile was the following: initial denaturation at 96°C for 5 min, followed by 40 cycles at 96°C for 30 s, 58°C for 30 s and 72°C for 90 s, and a final extension step at 72°C for 10 min. PCR products were directly sequenced without any additional purification step.

PCR products were sequenced using Sanger dideoxy-technology in both directions by the company Macrogen Inc. (Macrogen Europe). Raw sequence quality control and the assembling were performed by using the GENEIOUS PRIME software (version 2019.0.4). Finally, best match the identity of the sequences was performed by blastn analysis (Altschul et al., 1990). NCBI accession for COI OK120268-OK120292 and for SSU OK058389 (Table S1).

Microscopical observation

Sampled cells were isolated and documented using an inverted light microscope (Leica DMI8) mounted with DIC lenses, and a Leica MC170 HD camera controlled by the Leica application suite (version 4.12.0). For electron microscopy, shells were placed on stubs and coated in 8 nm gold with a Balzers SCD 004 sputter coater with a tension of 15 kV, and then observed with a Hitachi S-3000N scanning electron microscope (SEM). We used the software IMAGEJ (version 1.52) (Schneider et al., 2012) to obtain the shell measurements from the images.

Tree analyses

COI and SSU rRNA sequences databases were aligned using MAFFT (Katoh et al., 2002) and, resulted in a total alignment of 837 bp for COI and 1889 bp for SSU, respectively. Tree topologies and node supports were evaluated with both Bayesian inferences (BI) and maximum likelihood (ML) for each marker/gene individually.

ML analyses were conducted using IQ-TREE (Nguyen et al., 2015). Best substitution models were selected with MODELFINDER (Kalyaanamoorthy et al., 2017) under the Bayesian information criterion (BIC). Node support was assessed with 1000 nonparametric bootstrap replicates.

Bayesian inference (BI) analyses were conducted using MRBAYES 3.2.7a (Ronquist et al., 2012) implemented in the CIPRES Science Gateway version 3 (Miller et al., 2010). Posterior probabilities were calculated with the Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) method by sampling trees (Huelsenbeck & Ronquist, 2001; Larget & Simon, 1999). MCMCMC settings consisted of two independent runs, with four chains for each run and 20×10^6 generations. Trees were sampled every 1000th generation; the first 25% was discarded as burnin. Substitution models were selected with the reversible-jump MCMC method (Huelsenbeck et al., 2004), including parameters for among-site rate variation, according to the ML results, with an estimated proportion of invariable sites and a gamma shaped distribution of variable sites (I+G). Convergence of the different runs was evaluated with TRACER version 1.7.1 (Rambaut et al., 2018), ensuring that all parameters reached values of effective sample size (ESSs) larger than 200. The resulting trees were summarized in a 50% majority rule consensus tree. The trees obtained were edited in FIGTREE version 1.4.3 (Rambaut, 2012).

Sensitivity analysis beast

We ran two MCMC chains for 10×10^7 generations, sampling every 1000th generation per BEAST run. For path sampling (PS) and stepping-stone (SS) sampling we used 100 path steps and a 10^6 chain length, sampling every 1000 trees. Fossils: (1) Constraining the basal node for all Euglyphidae at 40 Ma, the age of the most ancient *Scutiglypha* fossils encountered (Figure S4), with a standard deviation = 1. (2) Constraining the basal node for all Euglyphida at 40 Ma (Figure S5). This is the least realistic approximation, but it will allow to see how the calibration of these nodes affects the ages of our nodes of interest (*Cyphoderia ampulla* clades, freshwater clades 1 and 2).

HISSE models used

(1) A Null model where all rates are equal across the tree; (2) a BiSSE model; (3) a hidden state speciation-extinction (HiSSE) model in which variation in diversification rates is correlated with the trait "habitat" (marine or freshwater) but also with additional unobserved, "hidden-traits" (Beaulieu & O'Meara, 2016); for the latter, we implemented a HiSSE model with two hidden states; (4) a character-independent (CID) model, in which diversification rate shifts are decoupled from

the focal trait states on any coevolving trait (Caetano et al., 2018); we implemented a CID model with two hidden states (CID2), that is, these hidden states represent diversification rate heterogeneity that is not structured as a coevolving trait.

Hyperphylomorphospace

We extracted the "RandomPoints" of each morphological group as generated in the hypervolume to graphically represent the "hyperphylomorphospace" and estimate the two-dimensional kernel density estimation per clade with the function "kde2d" (to give a clearer picture) in the R package "MASS" (Venables & Ripley, 2002). Then, we added the centroids of the hypervolume, where the colours represent the molecular clade and the symbols the species as identified through morphology. For the MCC tree, we first pruned the tree leaving only one tip per group (10 tips), matching with the centroid of the hypervolume. We then coloured the branches after the ancestral habitat (red = marine and blue = freshwater). Finally, we performed the phylomorphospace analysis with the pruned tree and the hypervolume centroids and added one by one the kernel density of the groups and centroids points.

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Beaulieu, J. M., & O'Meara, B. C. (2016). Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology*, 65(4), 583–601. <https://doi.org/10.1093/sysbio/syw02>
- Caetano, D. S., O'Meara, B. C., & Beaulieu, J. M. (2018). Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution*, 72(11), 2308–2324. <https://doi.org/10.1111/evo.13602>
- Heger, T. J., Pawlowski, J., Lara, E., Leander, B. S., Todorov, M., Golemansky, V., & Mitchell, E. A. D. (2011). Comparing potential COI and SSU rDNA barcodes for assessing the diversity and phylogenetic relationships of cyphodermiid testate amoebae (Rhizaria: Euglyphida). *Protist*, 162(1), 131–141. <https://doi.org/10.1016/j.protis.2010.05.002>
- Huelsenbeck, J. P., Larget, B., & Alfaro, M. E. (2004). Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Molecular Biology and Evolution*, 21(6), 1123–1133. <https://doi.org/10.1093/molbev/msh123>
- Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17(8), 754–755. <https://doi.org/10.1093/bioinformatics/17.8.754>
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589. <https://doi.org/10.1038/nmeth.4285>
- Katoh, K., Misawa, K., Kuma, K. I., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), 3059–3066. <https://doi.org/10.1093/nar/gkf436>
- Larget, B., & Simon, D. L. (1999). Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molecular Biology and Evolution*, 16(6), 750–759. <https://doi.org/10.1093/oxfordjournals.molbev.a026160>
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *2010 Gateway Computing Environments Workshop (GCE)* (pp. 1–8). IEEE. <https://doi.org/10.1109/GCE.2010.5676129>
- Rambaut, A. (2012). *FigTree, molecular evolution, phylogenetics and epidemiology*, v.1.4. Accessed 14 September, 2021.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901–904. <https://doi.org/10.1093/sysbio/syy032>
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3), 539–542. <https://doi.org/10.1093/sysbio/sys029>
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 Years of image analysis. *Nature Methods*, 9(7), 671–675. <https://doi.org/10.1038/nmeth.2089>
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer. www.stats.ox.ac.uk/pub/MASS4/VR4stat.pdf