Check for updates

Ecology and Evolution

WILEY

RESEARCH ARTICLE OPEN ACCESS

A Comparative Analysis of DNA Metabarcoding and Morphological Identification in Diatoms Reveals Similar Patterns of Environmental Response

Fernanda Gonzalez-Saldias[1,2] | Joan Gomà[1,2] | Sandra Garcés-Pastor[1,3] | Owen S. Wangensteen[1] | Albert Pèlachs[4] | Aaron Pérez-Haase[1,5]

[1]Department of Evolutionary Biology, Ecology, and Environmental Sciences, University of Barcelona, Barcelona, Spain | [2]Freshwater Ecology, Hydrology, and Management Research Group (FEHM-Lab), University of Barcelona, Barcelona, Spain | [3]Institute of Marine Sciences (ICM), CSIC, Barcelona, Catalonia, Spain | [4]Geography Department, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain | [5]Biodiversity Research Institute (IRBio), University of Barcelona, Barcelona, Spain

Correspondence: Fernanda Gonzalez-Saldias (fernanda_gonzalezs@ub.edu)

Received: 5 June 2025 | **Revised:** 8 October 2025 | **Accepted:** 27 November 2025

Keywords: 18S rRNA | Bacillariophyceae | COI | ecological gradient | mires | pH

ABSTRACT

Diatoms are widely used as environmental sentinels and are commonly studied through morphological identification. However, this method requires specialist knowledge, which is becoming increasingly scarce, resulting in species identification becoming costly, time-consuming, and difficult. DNA metabarcoding offers an alternative method to semiautomate some identification processes to circumvent these difficulties, thus minimizing identification costs while providing objective and reliable information. However, discrepancies between marker genes and morphological identification remain unsolved. This study investigated the 66 diatom communities on 26 high mountain mires in the Pyrenees by identifying morphological species and sequencing the Cytochrome c oxidase I (COI) and 18S rRNA markers. We assessed whether morphological and molecular approaches reveal the same patterns of alpha and beta diversity. Our results showed a strong correlation between beta diversity patterns of diatoms obtained by morphological identification and genetic markers. In contrast, alpha diversity calculated through molecular studies underestimated species richness and the Shannon index. The diatom community was chiefly driven by pH, Mg, and temperature, reflecting these patterns in morphological and molecular data. DNA metabarcoding also proved to represent diatom community patterns and the roles of the environmental drivers as clearly as in morphological identifications in high mountain mire diatom communities.

1 | Introduction

Diatoms are diverse and ubiquitous unicellular algae that play a pivotal role as primary producers in aquatic ecosystems (Mann and Vanormelingen 2013; Round et al. 1990; Zimmermann et al. 2015). Their species distribution is determined by various environmental variables, among which pH (Potapova and Charles 2002), electrical conductivity (Potapova and Charles 2003), temperature, and nutrient availability (Stevenson et al. 2010) stand out. This sensitivity to environmental changes makes diatoms valuable ecological indicators of ecosystem conditions and has often been used to assess eutrophication (Cantonati and Lowe 2014; Rimet 2012).

Traditional methods for identifying and counting diatoms depend on morphological features such as size, shape, and number of striae observed under light or electron microscopy. This dependence on detailed morphology renders the process

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Ecology and Evolution* published by British Ecological Society and John Wiley & Sons Ltd.

highly specialized, demanding, and time-consuming (Kahlert et al. 2009; Mann et al. 2010; Rimet et al. 2018), but important and relevant for ecological studies. Furthermore, cryptic species and phenotypic plasticity within populations complicate accurate species identification (Cox 2014; Mann and Evans 2008). Variations in taxonomists' expertise also contribute to discrepancies in morphological assessments, primarily due to the lack of intercalibration among experts (Vasselon et al. 2019). Consequently, these factors can lead to both underestimation or overestimation of species diversity (Bailet et al. 2020; Kahlert et al. 2009).

To address the challenge of morphological identification, molecular identification through metabarcoding has emerged as a promising alternative (Mann et al. 2010). Using DNA markers, this approach identifies multiple taxa regardless of their life stage (Rimet et al. 2018; Zimmermann et al. 2015). Using universal markers means that multiple taxa can be amplified simultaneously, eliminating the need to separate them in the samples. Such markers provide a good approximation of the maximum biodiversity in environmental DNA samples (Collins et al. 2019; Hadziavdic et al. 2014; Klepke et al. 2022). However, the results obtained from metabarcoding analyses can vary significantly depending on the gene marker used and the completeness and reliability of genetic databases. For instance, notable differences have been observed between the cytochrome oxidase I (COI) and 18S rRNA genes in community representation compared to morphological identification and the number of taxon groups (Kermarrec et al. 2013; Othman et al. 2021; Trobajo et al. 2010).

The COI-LerayXT marker (313-bp mitochondrial gene) has been widely used for animal identification due to its robustness and high taxonomic resolution (Mann et al. 2010; Moniz and Kaczmarska 2009; Wangensteen et al. 2018). This DNA region presents a high incidence of nucleotide substitution compared to other genes (Hebert et al. 2003). These characteristics have facilitated its application across various taxa, including protists (Nassonova et al. 2010), macroalgae (Saunders 2005), and diatoms (Evans et al. 2007; Guo et al. 2015; Kermarrec et al. 2013).

Similarly, the 18S rRNA gene has been proposed as a potential barcode marker for identifying eukaryotes, with different regions designated for various groups of organisms (Hadziavdic et al. 2014). For instance, the V1–V2 regions have been utilized for nematodes (Schenk et al. 2020), the V4, V5, and V9 regions for protists (Decelle et al. 2014; Taib et al. 2013), the V7 region for microbial eukaryotic communities (Capo et al. 2021), the V9 region for paleo-diatoms (Armbrecht et al. 2021), the V4 region for phytoplankton (Esenkulova et al. 2020), and both the V7 and V4 regions for diatoms (Garcés-Pastor et al. 2019; Zimmermann et al. 2015).

Diatoms, recognized for their sensitivity to environmental variables, serve as excellent models for ecological studies, particularly in diverse aquatic ecosystems like those present in the Pyrenees, which play a crucial role in biodiversity (Rivera Rondón and Catalan 2017, 2020). While numerous studies have investigated their responses to environmental changes, few have explored community patterns using metabarcoding techniques (Garcés-Pastor et al. 2019). Most research has concentrated so far on species-level representativeness and the use of diatoms in biomonitoring (Borrego-Ramos et al. 2021; Pérez-Burillo et al. 2022; Vidaković et al. 2024). Morphological species identification remains essential for validating metabarcoding studies. However, discrepancies between morphological identification and metabarcoding markers (Brown et al. 2022; Kim et al. 2024; Nistal-García et al. 2021) have resulted in inaccuracies in species representation (Bailet et al. 2020; Wang et al. 2024), underscoring the necessity for more integrated approaches in diatom research. Due to this low taxonomic representation, Molecular Operational Taxonomic Units (MOTUs) are often used in community and ecological studies.

To address these gaps, this study aims to utilize metabarcoding to provide a comprehensive understanding of diatom communities in the Pyrenees mires. Mire ecosystems are great environmental heterogeneities that can support diverse diatoms (Gonzalez-Saldias et al. 2025), being less studied than rivers or lakes. For this reason, the study of diatom communities in mires can provide relevant information on diversity and community patterns using universal markers. We hypothesized that genetic markers would provide accurate information on the diatom community patterns in high mountain mires. While alpha diversity describes local richness and diversity, community dissimilarity is a change between communities. To test this hypothesis, we employed two genetic markers, 18S rRNA and COI, and compared the results with morphological identification.

2 | Methodology

2.1 | Sample Drawing

A total of 66 plots on 26 mires located in the Central Pyrenees were sampled in the summers (July and August) of 2016, 2017, and 2018 under the LIFE+ Limnopirineus project (Figure 1). At each plot, five regularly spaced bites of topsoil (5 cm deep) samples (100 mL), including the bryophyte layer, were preserved in 96% ethanol. Half of each sample was allocated for morphological analysis, and the other half for molecular identification. The following variables were measured to characterize each sampling site: water table depth, vegetation cover, pH, and electrical conductivity (Table S1). The pH and electrical conductivity were measured using a WTW ProfiLine pH 3310 Portable pH Meter. We used a plot of 2.5 × 2.5 m to estimate vegetation cover. For each plot, we visually estimated the percentage cover of all plant species. We then classified the species into broad functional groups: vascular plants, bryophytes, sphagnum, pleurocarpic mosses, acrocarpic mosses, and liverworts (Marchantiophyta). To obtain the total cover of each functional group, we summed the percentage cover of the individual species within the group. Moreover, a sample of groundwater (100 mL) was taken at each plot, and the composition of metals and minerals (K, Mg, Mn, Na, S, and Si) was also analyzed by the Inductively Coupled Plasma (ICP-OES) method. As for climatic regional variables, we used the Pyrenean Digital Climatic Atlas with a 30-m resolution (Batalla et al. 2018). We retrieved mean annual precipitation, annual potential solar radiation, and growing degree days. The growing degree day value (GDD) for each site was estimated as follows:

$$\mathrm{GDD} = \sum \left(\left(\frac{\left(T_{\max} + T_{\min}\right)}{2} \right) - T_{\mathrm{base}} \right) \times N$$

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

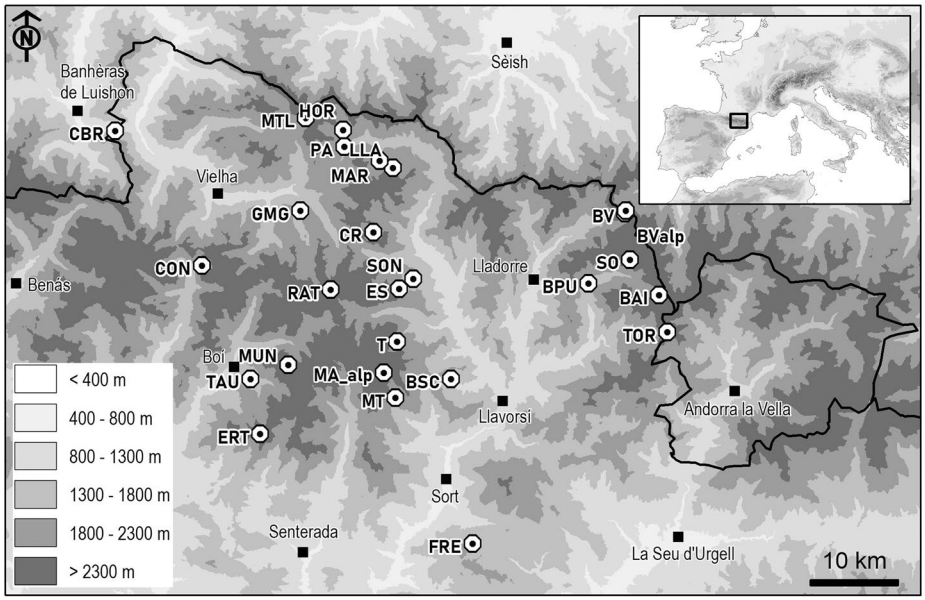


FIGURE 1 | Geographical location of the studied mires. Black dots correspond to the mires sampled in the Catalan Pyrenees in the summer of 2018, and the gray triangle represents municipalities. Code, location name, and coordinate in Table S6.

T_{max} is the day's maximum temperature, T_{min} is the day's minimum temperature, T_{base} equals zero, and N is the number of days/months.

2.2 | Morphological Identification of Diatoms

Half of each sample taken was used for morphological identification of diatoms, and organic matter was removed by treating the samples with 33% w/v hydrogen peroxide for 24 h at 100°C, followed by an additional 24 h at room temperature to remove all organic matter. After this treatment, the supernatant was carefully discarded, and the samples were rinsed with distilled water. The cleaned diatom valves were then mounted on permanent microscope slides using Naphrax. Identification and enumeration of diatoms were conducted under a POLYVAR microscope with Differential Interference Contrast with an objective of 100× magnification and 1.3 numerical aperture, counting 500 valves per sample. Taxonomic identification followed the monographs of Krammer and Lange-Bertalot (2004a, 2004b, 2004c, 2004d), Lange-Bertalot (2000, 2001), Lange-Bertalot et al. (2011, 2017), and Rivera-Rondón and Catalan (2020). The relative abundance of each diatom species was calculated for subsequent analysis.

2.3 | DNA Extraction

DNA extraction was performed at the genetic laboratories of the Department of Evolutionary Biology, Ecology, and Environmental Sciences at the University of Barcelona. All extraction procedures were conducted under a laminar-flow cabinet to prevent potential contamination. Between DNA sample extractions, all the equipment was thoroughly cleaned with 10% sodium hypochlorite and rinsed in Milli-Q water (Garcés-Pastor et al. 2019).

Peat samples were homogenized using a 600 W hand blender. Each homogenized sample (0.3 g) was extracted using the Soil DNA Isolation Plus Kit (Norgen Biotech Corp., Ontario, Canada, www.norgenbiotek.com). An extraction blank consisting of 300 μL of molecular biology-grade water was included in the DNA extraction batch, processed, and sequenced along with the rest of the samples (Garcés-Pastor et al. 2019).

2.4 | PCR and Sequencing

Two markers were used to identify a broad taxonomic range of eukaryotic taxa. The Leray-XT primer set, which amplifies a 313 bp fragment of the mitochondrial COI marker (miCOIintF-XT 5′-GGWACW RGWTGRACWITITAYCCYCC-3′; Wangensteen et al. 2018), and jgHCO2198 5′-TAIACYTCIGGRT GICCRAARAAYCA-3′ (Geller et al. 2013) were also used. Additionally, the V7 region of the nuclear-encoded ribosomal 18S rRNA gene was amplified using the 18S_allshorts primers (100–110 bp, 5′-TTTGTCTGSTTAATT SCG-3′ and 5′-TCACAGACCTGTTATTGC-3′; Guardiola et al. 2015).

PCR products were pooled by marker and purified using Minelute columns. Two Illumina libraries were prepared from the DNA pools using the NEXTflex PCR-free kit and tagged for identification. Both libraries were sequenced together on an Illumina MiSeq using v3 chemistry with 2×250 bp paired-end reads. Bioinformatics analyses were conducted using the OBITools software suite (Boyer et al. 2016) following the methods described by Garcés-Pastor et al. (2019). MOTUs' construction was conducted with SWARM 1.2.19 (Kopylova et al. 2016). For taxonomy annotation, representative MOTU sequences were BLAST (v2.9.0) against RSyst::diatom reference base v7 (Rimet et al. 2016) with an identity cut-off and minimum alignment length of 80%. Singletons and samples with < 50 reads were removed from the analysis. Rarefaction curves indicated

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

that sequencing depth approached saturation in the number of MOTUs for both markers at 53 reads and 103 reads per sample for 18S rRNA and COI, respectively.

2.5 | Statistical Analysis

MOTUs were used to calculate richness, the Shannon index, and Pielou's evenness for each sample in the molecular dataset. To evaluate significant differences in Shannon, Pielou, and species richness between the molecular and morphological identification methods, Kruskal–Wallis tests were performed because the data did not meet assumptions of normality and homoscedasticity. To quantify the differences in alpha diversity (Richness, Shannon, and Pielou evenness) between the morphological and molecular identification techniques, $\log_2$-fold changes were calculated. The $\log_2$ fold change is the $\log_2$ of the ratio between molecular and morphological alpha diversity. Then, positive $\log_2$-fold changes indicate higher alpha diversity in molecular approaches related to the morphological-derived estimations, while negative values indicate underestimation. To assess significant differences in alpha diversity (richness, Shannon, and Pielou) between morphological and molecular identification, we performed Student's *t*-tests for two samples, or Wilcoxon tests when assumptions were not met.

Rarefaction was applied to the reads to homogenize the sequencing depth. Reads were rarefied to the read count of the lowest sample in each marker. We used the "rarefy" function from the vegan package to MOTU rarefaction. Bray–Curtis dissimilarity was calculated with square root-transformed morphological abundance and molecular reads matrices to assess sample dissimilarities. The degree of correspondence between the ordinal results obtained for morphological and molecular datasets was evaluated with a Procrustean analysis followed by a PROTEST (Goodall 1991; Jackson 1995). Procrustean plots were generated from the ordinal results of morphological and molecular-based multidimensional scaling (MDS) based on the Bray–Curtis distance matrixes. Distance-based redundancy analyses (dbRDA) were performed on the Bray–Curtis dissimilarity matrix to elucidate the relationship between environmental variables and the diatom community composition. All statistical analyses were performed with MOTUs due to low taxonomic assignment.

All the statistical analyses were conducted in R (v4.2.3) using the packages vegan (Oksanen et al. 2022), dplyr (Wickham et al. 2023), and microbiome (Lahti and Shetty 2017).

3 | Results

3.1 | Environmental Variables

The high mountain mires studied displayed great variability within the mire ecological habitats (Table 1). The environmental variability related to altitudinal changes, such as temperature (1226–2967 GDD), and annual precipitation ranging from 943.6 to 1745.5 mm, presented a great range. The pH varied between 4.51 and 7.76, and K (1.480–21.769 mg/L) has a wide range. The vegetation cover of both vascular plants and bryophytes varied considerably. Within the bryophyte group, *Sphagnum* mosses and pleurocarpic mosses (0–119 and 0–94, respectively) presented a great cover range (Table 1).

TABLE 1 | Mean and range of environmental variables in the Pyrenees mires. The sum of species covers the different groups of *Sphagnum*, Brown mosses, Acrocarp mosses, Liverworts, Total bryophytes, and Vascular plants.

Environmental variables	Mean	Range
GDD (°C)	1871.561	1226 to 2967
Precipitation (mm)	1337.882	943.6 to 1745.5
pH	5.977	4.51 to 7.76
WTD (cm)	−6.117	−36.6 to 0.0
Radiation (kJ/m^2×day)	1871.561	1546 to 2174
Sphagnum	12.038	0 to 119
Pleurocarpic mosses	25.879	0 to 94
Acrocarp mosses	4.849	0 to 52
Liverworts	0.144	0 to 2
Total bryophytes	44.909	0 to 119
Vascular plants	106.940	0.055 to 285
K (mg/L)	4.069	1.480 to 21.769
Mg (mg/L)	1.018	0.174 to 7.379
Mn (mg/L)	0.122	0.005 to 3.233
Na (mg/L)	2.020	0.431 to 9.652
P (mg/L)	0.178	0.050 to 0.951
S (mg/L)	2.167	0.221 to 9.821
Si (mg/L)	0.519	0.050 to 2.001

3.2 | Morphological Identification

We identified a total of 383 species belonging to 65 genera using light microscopy (Table S2). The average number of species per sample was 42.33 ± 10.11. The most abundant species were represented by *Achanthidium minutissimum* (Kützing) Czarnecki, *Staurosira contruens* var. *venter* (Ehrenberg) P.B.Ham., *Eunotia incisa* W.Sm. ex W.Greg, *Eolimna minima* (Moser et al. 1998), and *Kobayasiella micropunctata* (H.Germ.) Lange-Bert. 1999 contributed to > 1% of total individuals across samples (Table S3).

3.3 | Diatom Molecular Identification

The DNA recovered from the peat samples yielded 4.9–31.2 ng/μL concentrations. For 18S rRNA sequences, 80,469 reads were retrieved (1219 ± 2772 reads per sample; mean ± SD). The final dataset for COI included a total of 153,315 reads (2322 ± 3419 reads per sample). The taxonomic assignment based on 18S rRNA resulted in a total of 166 MOTUs presented by 116 species belonging to 16 genera. In comparison, identification based on COI resulted in a total of 277 MOTUs corresponding to 18 species from six genera (Table S4). Taxonomic assignment at the genus

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and species level is 6.5% for COI and 69.9% for 18S rRNA (>95% percentage identity). For each marker, more than 1% of the total abundance contribution was made by 26 MOTUs. MOTUs were present at the genus and species level ($n=2$, >95% identity, and $n=5$, >97% identity, respectively). Conversely, 18S rRNA presented a higher number of MOTUs at the species level ($n=21$ at >97% identity) and four at the genus level (>95% identity). The 18S rRNA marker presents four species that are identical in morphology. The two most abundant species are *A. minutissimum* and *S. contruens*, but with lower relative abundance compared to morphological identification. Only one coincides with COI, but the level of identity does not exceed 90% (Table S3).

3.4 | Morphological and Molecular Alpha Diversity

Morphological diatom richness exhibited a positive and significant correlation with MOTUs richness for 18S rRNA (Spearman correlation 0.32, $p<0.01$) and COI (Spearman correlation 0.58, $p<0.001$), with COI-based estimations showing the highest correlation (Figure 2A). However, COI and 18S rRNA underestimated diatom richness compared to morphological data (Student's *t*-test, $p<0.001$, Figure 2B). The Shannon index indicated that COI and 18S rRNA had a similar correlation with morphological identification (Figure 2C). Furthermore, COI demonstrated greater similarity to the morphological data compared to 18S rRNA, as evidenced by a $\log_2$-fold change around 0 (Figure 2D). However, these relations were significantly when compared to morphological estimations (Wilcoxon test, $p<0.001$). For Pielou's evenness, we observed a positive correlation only between morphological data and 18S rRNA. In contrast, a negative correlation was noted with COI (Figure 2E). In both instances, the evenness calculated using the markers was found to be similar to the obtained from the morphological data (Wilcoxon test, $p>0.05$, Figure 2F).

3.5 | Morphological and Molecular Diatom Dissimilarity

Diatom communities have a high variability in species composition (Figure 3). Overall, Procrustes superimposition analysis showed a high similarity between morphological compositional variations and molecularly resolved communities for the

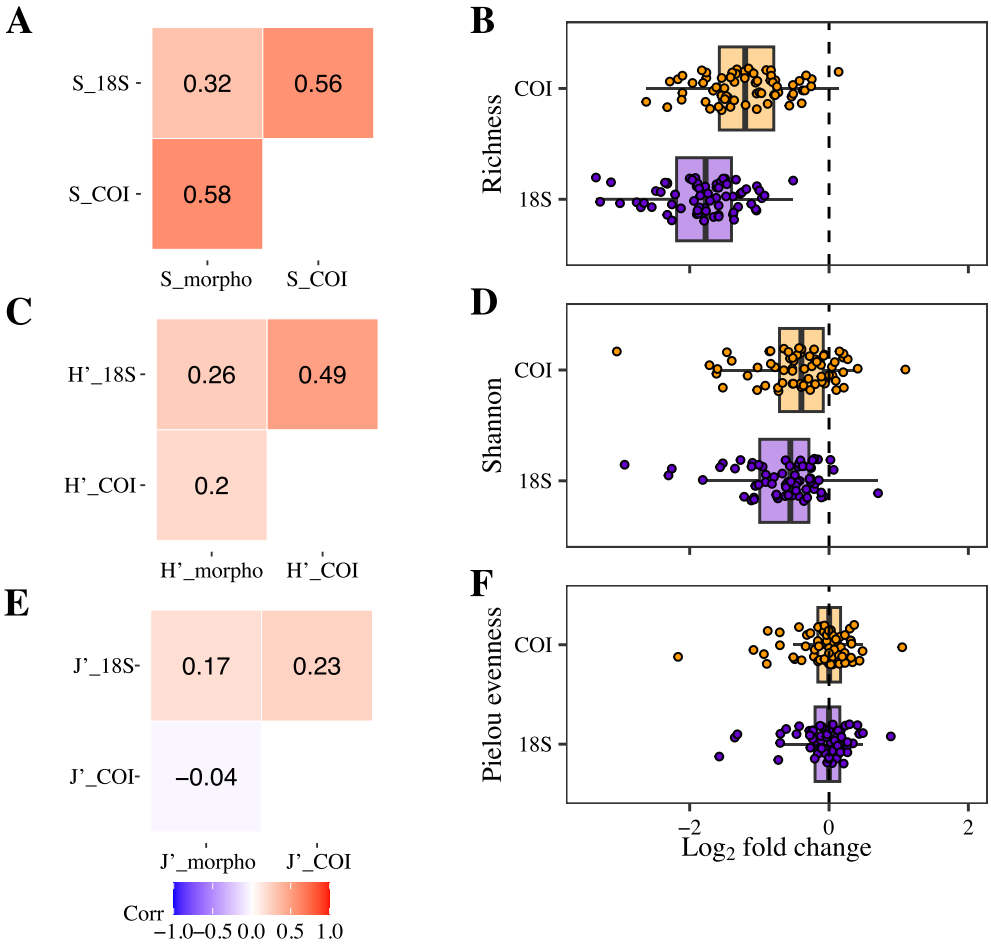


FIGURE 2 | Spearman correlation between molecular and morphological identification in richness (A), Shannon index (C), and Pielou evenness (E). Comparison of alpha diversity indices in the three approaches to diatom communities of high mountain mires in richness (B), Shannon index (D), and Pielou evenness (F), molecular with morphological identification. Log_2-fold change shows the magnitude of differences between the morphological and molecular identification techniques in terms of alpha diversity, using the morphological-derived indexes as a reference (dotted line). Positive $\log_2$-fold changes indicate overestimation of alpha diversity in the molecular approaches compared to morphological, while negative values indicate underestimation.

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

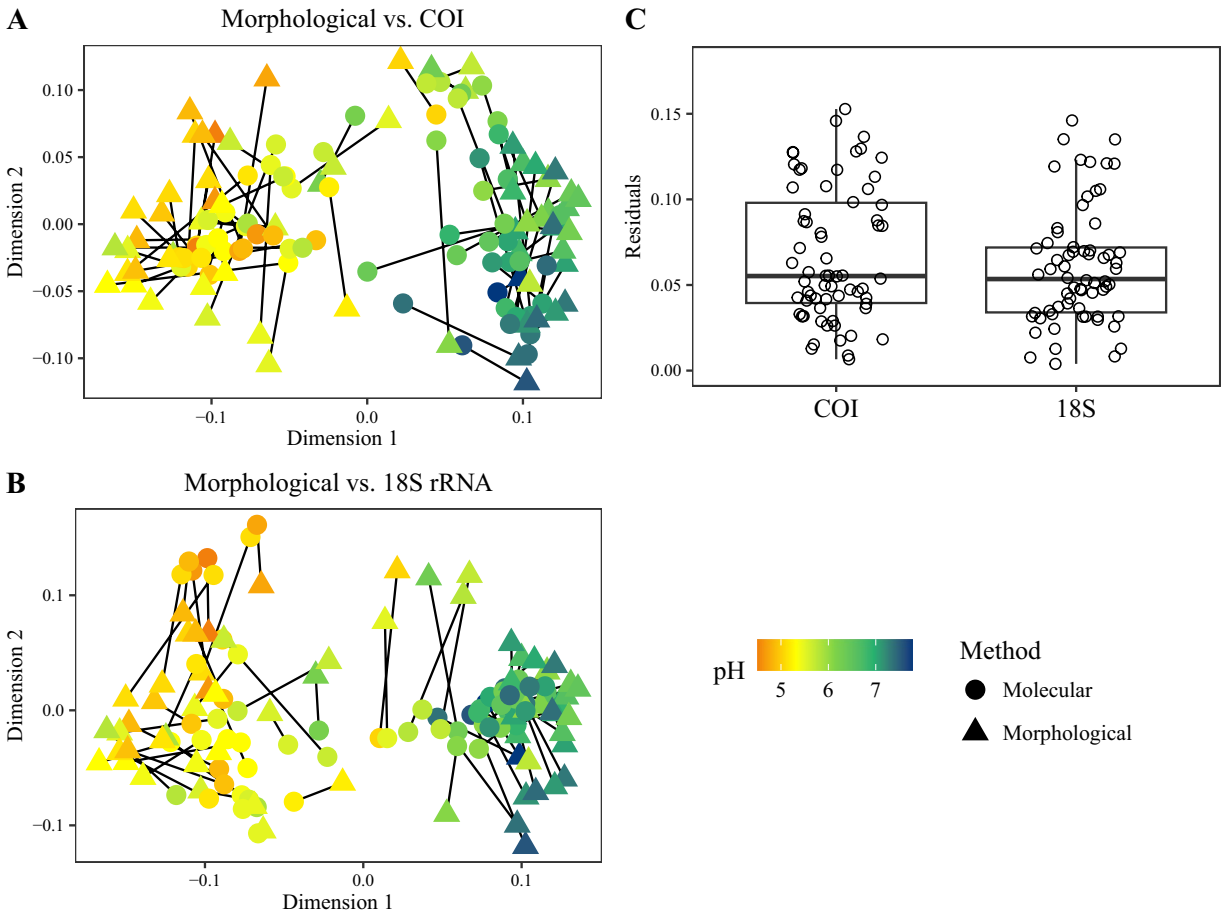


FIGURE 3 | Comparison of MDS through Procrustes analysis between morphological versus COI (A) and morphological versus 18S rRNA (B). Residual values from comparing morphological identification and the two molecular identifications (18S rRNA and COI) (C).

sampled sites ($r > 0.78$, $p < 0.001$; Figure 3). Both molecular markers significantly represented the diatom community pairwise dissimilarities, with 18S rRNA holding the higher similitude with morphological assignations than COI (morphological-COI PROTEST m1-2 = 0.40, $r = 0.78$, $p = 0.001$; morphological-18S PROTEST m1-2 = 0.32, $r = 0.83$, $p = 0.001$) (Figure 3A,B). Despite the observed variability, the residuals from Procrustes analysis for COI and 18S rRNA were not significantly different (Student's t-test, $p = 0.207$) (Figure 3C).

3.6 | Community Drivers

The dbRDA analysis indicated that the environmental drivers of community composition were similar for the morphological and molecular methods (COI and 18S rRNA) (Table 2). The environmental variables explain over 27% of the variability in both methods. The dbRDA1 explains 75.56%, 61.28%, and 63.82% for morphological, COI, and 18S rRNA identifications, respectively. The dbRDA2 were similar for all identifications, with values of 10.64% for morphological identification, 17.35% for COI, and 11.90% for 18S rRNA. Among these drivers, pH was the main structuring variable, followed by temperature (GDD) and Mg concentrations (Table 2).

Diatom richness exhibited a clear pattern along the pH gradient for both morphological and molecular marker identifications (Figure 4A–C). Morphological identification showed higher richness (20–61 species), followed by COI (9–37 MOTUs), and 18S rRNA (6–21 MOTUs). Similarly, the Shannon diversity changes along pH were similar in the morphological and molecular data, ranging from 1.26 to 3.53 in morphological, 0.27–3.42 for COI, and 0.27 to 2.83 for 18S rRNA (Figure 4D–F). Pielou's evenness estimations were similar between morphological and molecular data, morphological data varied between 0.42 and 0.86, while COI and 18S rRNA presented higher variability in the values of equitability from 0.15 to 0.92 for COI and 0.20 to 0.93 for 18S rRNA (Figure 4G–I).

4 | Discussion

COI and 18S rRNA markers showed a strong correspondence for diatom composition and response to environmental changes, despite not being specific markers for diatom identification. The most important variable structuring the diatom community was pH, which was observed in both morphological and molecular identification (Figures 3 and 4; Table 2). The importance of pH in diatom community structure has been reported previously, showing that it strongly influences diatom communities (Chen et al. 2014; Gonzalez-Saldias et al. 2025; Hargan et al. 2015). Both genetic markers are good for detecting these patterns and environmental variables affecting diatom communities. The pH 5.5 and 6.5 are critical for diatom

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Distance-based redundancy analysis (dbRDA) with forward selection based on the Bray–Curtis dissimilarity matrix, with environmental variables related to local factors, and vegetation structure for morphological data, COI, and 18S rRNA.

Data	Environmental variables	Forward selection		
		R^2	F	p
Morphological	pH	0.150	11.329	0.001
	GDD	0.027	2.058	0.002
	Mg	0.029	0.264	0.001
	Precipitation	0.019	1.499	0.036
COI	pH	0.131	9.636	0.001
	GDD	0.043	3.241	0.001
	Mg	0.031	2.418	0.001
	Sphagnum	0.021	1.617	0.016
	Acrocarp mosses	0.018	1.420	0.046
18S rRNA	pH	0.162	12.395	0.001
	GDD	0.031	2.393	0.002
	Mg	0.034	2.700	0.001
	S	0.022	1.779	0.014
	Pleurocarpic mosses	0.020	1.695	0.019
	Al	0.021	1.674	0.019

communities, causing compositional changes (Gonzalez-Saldias et al. 2025; Hargan et al. 2015), which can be seen in our results (Figure 4).

4.1 | Molecular Markers as Community Composition Indicators

Metabarcoding approaches for diatom identification at the community level provide information comparable to morphological-based analysis. Our results show that COI and 18S rRNA, both universal markers, are robust for detecting community (dis)similarities (b-diversity patterns). This is supported by the strong correlation in the Procrustean analysis along the environmental variables, in contrast to similar studies conducted in other aquatic habitats, such as ponds and rivers, where such correspondence is rarely observed (Brown et al. 2022; Nistal-García et al. 2021; Vasselon et al. 2019). These results highlight the potential of utilizing universal markers in facilitating the extensive labor associated with diatom identification, particularly in heterogeneous habitats with high biodiversity and rare species (Gonzalez-Saldias et al. 2025).

4.2 | Taxonomic Identification

Poor taxonomic assignment is a recurrent issue in diatom molecular identification (Rivera et al. 2018; Vasselon et al. 2017). We found that 62.16% of the total reads in 18S rRNA and only 4.97% in COI were assigned to the species level. Both markers underestimated species richness, a 2 and 4-fold for COI and 18S rRNA, respectively (Figure 3). Comparatively, 18S rRNA was more taxonomically accurate than COI. This finding contrasts with similar studies, which have generally reported that molecular identification tends to overestimate species richness relative to morphological assessments in rivers (Mora et al. 2019; Tapolczai et al. 2019; Wang et al. 2024). By contrast, a smaller number of taxa were identified using molecular techniques in less studied environments, such as glaciers (Chamorro et al. 2024). The poor taxonomic assignment may be related to the lack of studies on mire diatoms and, therefore, a poor representation of these species in the reference databases.

There were differences in the specificity of mitochondrial and ribosomal universal markers in Bacillariophyta. Our results indicate that the low species diversity captured by COI is due to the poor dataset and its poor performance in discriminating abundant and widespread diatom genera in mires, such as *Nitzschia*, *Eunotia*, and *Pinnularia* (Table S5), consistent with previous research (Hamsher et al. 2011; Moniz and Kaczmarska 2009; Trobajo et al. 2010). However, COI can effectively identify species affiliated with the *Sellaphora* genus, and the *Gomphonema parvulum* species, a recurrent dominant species in running waters (Evans et al. 2008, 2007; Mann et al. 2010). The V7 region of 18S rRNA has been successfully used to identify marine and mire eukaryotic organisms (Garcés-Pastor et al. 2019; Guardiola et al. 2015). However, the V7 region in our study showed limited resolution in richness and Shannon index (Figure 2). Some studies suggest the V4 region is a better alternative, as it encompasses numerous hypervariable regions, including inversions, insertions, and deletions, making it more suitable for diatom species identification (Alverson et al. 2006; Zimmermann et al. 2011).

Universal markers are aimed at identifying most of the total diversity of different groups of organisms within ecosystems. Markers such as COI and 18S rRNA allow the study of eukaryotic heterotrophs and autotrophs, which vary in size and lifestyle. However, both markers have several limitations when it comes to taxonomic identification. First, the 18S rRNA is characterized by conserved regions with limited polymorphisms and, consequently, is more prone to taxonomic inaccuracies (Kermarrec et al. 2013). While COI has been extensively studied in metazoans (Che et al. 2012; Elbrecht and Leese 2017; Sultana et al. 2018), its efficiency in resolving photosynthetic organisms such as diatoms in databases remains poor (Kermarrec et al. 2013; Trobajo et al. 2010). Second, the poor taxonomic coverage can be attributed to the limited diatom database size available for COI (293 sequences) and 18S rRNA (2646 sequences) in RSyst::diatom reference base v7 compared to for example bacterial databases, such as the SILVA 16S small subunit database, which contains approximately 500,000 representative sequences (Quast et al. 2013; Rimet et al. 2016). Furthermore, the species examined in this study originated from highly heterogeneous environments, making them difficult to cultivate and, consequently, underrepresented in public archives. An alternative to address this issue is using diatom-specific markers, such as the Rubisco *rbc*L (Kulaš et al. 2022). Therefore, COI and 18S rRNA appeared as not suitable identification markers for studies of biodiversity,

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

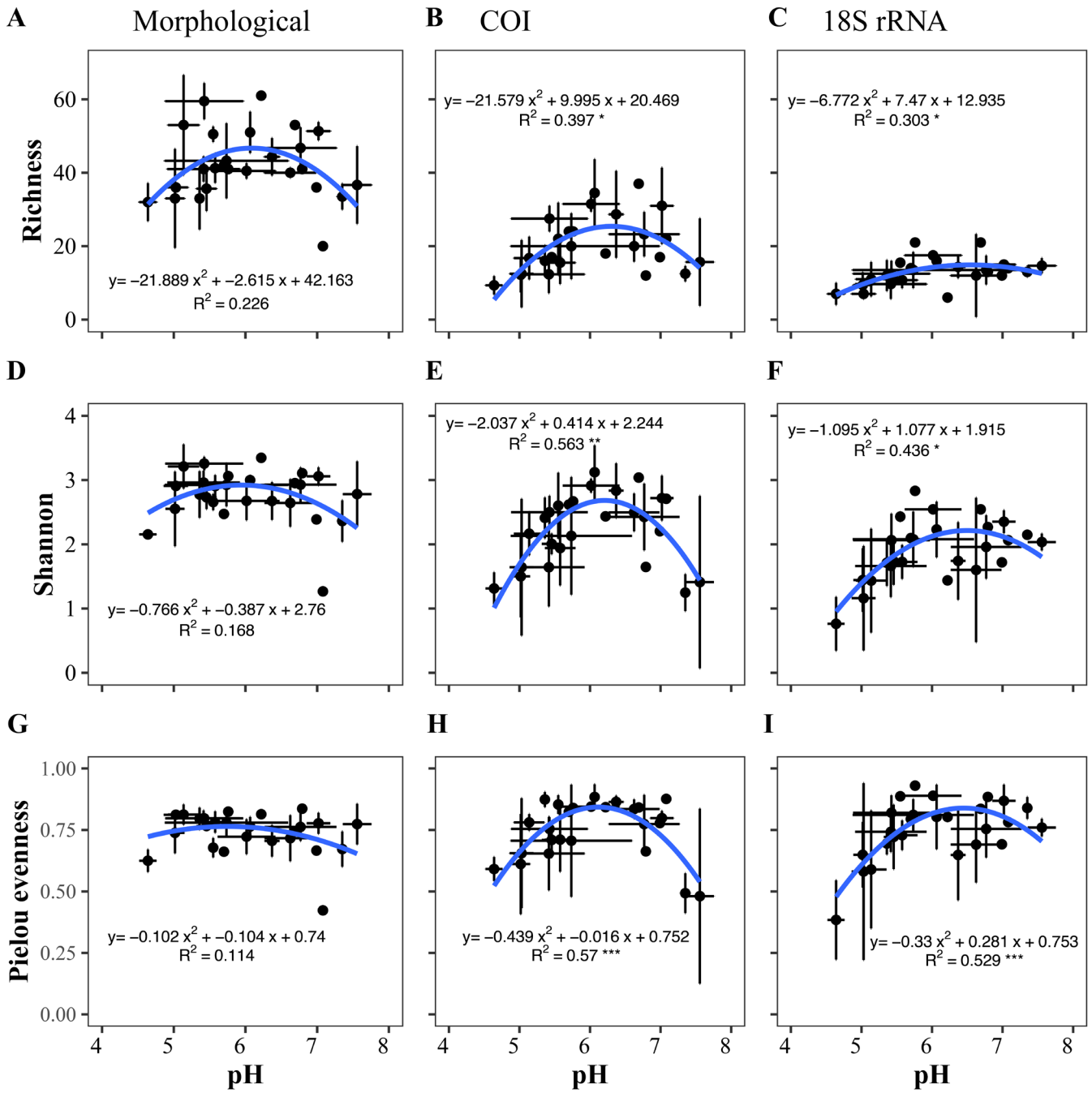


FIGURE 4 | Correlation between pH and alpha diversity variables (Richness, Shannon index, and Pielou) through morphological (A, D, G), COI (B, E, H), and 18S rRNA (C, F, I) datasets ordered by increasing pH. The dot is the average pH versus the average of the corresponding alpha diversity variable of each locality. Lines are the standard deviations of pH on the horizontal axis and the corresponding alpha on the vertical axis. The blue line is an adjusted polynomial model of order two.

because both underestimate alpha diversity (i.e., Richness and Shannon index), leading to data misinterpretation due to, for example, the poor coverage of the abundant representative genera in mires (Table S5).

4.3 | Challenges in Molecular Identification for Diatoms

This study observed differences in the taxonomic assignment and number of MOTUs identified for the two markers used. These differences may be due to species-specific gene copies, as observed in other studies with rbcL and 18S rRNA markers, where the copy number of different species is related to biovolume and therefore needs to be corrected (Kermarrec et al. 2013; Vasselon et al. 2019). Considering such a correction factor would help to improve the estimation of dominancy and abundance to match morphological-based estimations (Godhe et al. 2008; Vasselon et al. 2018). In addition, the presence of cryptic species, environmental DNA, and degraded DNA can also bias the taxonomic resolution (Vasselon et al. 2018).

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The variety of pipelines used for metabarcoding sequencing processing can generate differences in downstream analyses (Bailet et al. 2020). The bioinformatic analysis encompasses various processes, often including sequence primer clipping, quality check, and trimming, which may introduce variability in the taxonomic assignment and community representativeness (Bailet et al. 2020; Mora et al. 2019; Zimmermann et al. 2015). Another common issue in environmental DNA studies is the sequencing depth achieved with universal markers and untargeted organisms, which may affect the number of reads and MOTUs, directly impacting alpha diversity indices, as may be the case in our study.

5 | Conclusion

In conclusion, the universal markers COI and 18S rRNA showed a high correspondence for diatom beta diversity and its response to environmental changes. Taxonomic identification through metabarcoding represents a viable alternative for the study of diatom community patterns. This approach can reduce the time required for sample identification and data processing compared to traditional methods. However, these molecular markers are unsuitable for precise taxonomic assignment and diversity studies because they underestimate alpha diversity, leading to misinterpretation of data.

Author Contributions

Fernanda Gonzalez-Saldias: conceptualization (equal), data curation (lead), formal analysis (lead), funding acquisition (lead), investigation (equal), methodology (equal), software (lead), visualization (lead), writing – original draft (lead), writing – review and editing (lead). **Joan Gomà:** conceptualization (equal), funding acquisition (supporting), supervision (lead), writing – review and editing (equal). **Sandra Garcés-Pastor:** conceptualization (equal), resources (supporting), writing – review and editing (equal). **Owen S. Wangensteen:** methodology (equal), resources (supporting), software (equal), writing – review and editing (supporting). **Albert Pèlachs:** conceptualization (equal), resources (supporting), writing – review and editing (equal). **Aaron Pérez-Haase:** conceptualization (equal), funding acquisition (lead), investigation (equal), methodology (lead), project administration (lead), resources (lead), supervision (lead), writing – review and editing (lead).

Funding

F.G.-S. was funded by the National Agency for Research and Development (ANID)/Scholarship Program/DOCTORADO BECAS CHILE/2020–72210270. This work has been supported by coordinated project grants from: (i) the Ministerio de Economia y Competitividad (Spanish Ministry of Economics and Competitiveness): "Calibracion de indicadores de influencia humana y climatica para la (re)interpretacion de la expansion postglacial y de las dinamicas forestales en los ultimos 18.000 anos" (PID2019108282GB-I00/AEI/10.13039/501100011033); "Estudio biogeografico historico comparado (Montaña Cantabrica, Sistema Central y Pirineos): 18000 años de cambios climáticos y antrópicos sobre especies forestales indicadoras" (CSO2015-65216-C2-1-P); (ii) Organismo Autónomo de Parques Nacionales (OAPN) under the project BIOOCULT (OAPN 2413/2017); (iii) European Union LIFE programme under the project LIFE Limnopirineus (LIFE13 NAT/ES/001210) and the LIFE Resque Alpyr (LIFE20 NAT/ES/000369). S.G.-P. was also supported by the Beatriu de Pinós Programme (BP-2021-00131) and a fellowship from "la Caixa" Foundation (ID 100010434, fellowship code LCF/BQ/PI24/12040011). O.W. was also supported by the Blue DNA project (PID2023-1466307OB).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Sequence data used in this study are deposited in the European Nucleotide Archive (accession number PRJEB87341). The relative abundance data (Table S2), metadata (Table S1), percentage of identity of molecular identification (Table S4), and code will be available in GitHub: https://github.com/fergonzalezsaldias/diatom.morphology.dna.

References

Alverson, A. J., J. J. Cannone, R. R. Gutell, and E. C. Theriot. 2006. "The Evolution of Elongate Shape in Diatoms." *Journal of Phycology* 42: 655–668. https://doi.org/10.1111/j.1529-8817.2006.00228.x.

Armbrecht, L., R. Eisenhofer, J. Utge, et al. 2021. "Paleo-Diatom Composition From Santa Barbara Basin Deep-Sea Sediments: A Comparison of 18S-V9 and Diat-rbcL Metabarcoding vs Shotgun Metagenomics." *ISME Communications* 1: 66. https://doi.org/10.1038/s43705-021-00070-8.

Bailet, B., L. Apothéloz-Perret-Gentil, A. Baričević, et al. 2020. "Diatom DNA Metabarcoding for Ecological Assessment: Comparison Among Bioinformatics Pipelines Used in Six European Countries Reveals the Need for Standardization." *Science of the Total Environment* 745: 140948. https://doi.org/10.1016/j.scitotenv.2020.140948.

Batalla, M., M. Ninyerola, and J. Catalan. 2018. "Pyrenean Digital Climate Atlas (PDCA)." https://doi.org/10.5281/zenodo.1186639.

Borrego-Ramos, M., E. Bécares, P. García, A. Nistal, and S. Blanco. 2021. "Epiphytic Diatom-Based Biomonitoring in Mediterranean Ponds: Traditional Microscopy Versus Metabarcoding Approaches." *Water* 13: 1351. https://doi.org/10.3390/w13101351.

Boyer, F., C. Mercier, A. Bonin, Y. L. Bras, P. Taberlet, and E. Coissac. 2016. "Obitools: A Unix-Inspired Software Package for DNA Metabarcoding." *Molecular Ecology Resources* 16: 176–182. https://doi.org/10.1111/1755-0998.12428.

Brown, P. D., J. M. Craine, D. Richards, A. Chapman, and B. Marden. 2022. "DNA Metabarcoding of the Phytoplankton of Great Salt Lake's Gilbert Bay: Spatiotemporal Assemblage Changes and Comparisons to Microscopy." *Journal of Great Lakes Research* 48: 110–124. https://doi.org/10.1016/j.jglr.2021.10.016.

Cantonati, M., and R. L. Lowe. 2014. "Lake Benthic Algae: Toward an Understanding of Their Ecology." *Freshwater Science* 33: 475–486. https://doi.org/10.1086/676140.

Capo, E., S. Ninnes, I. Domaizon, et al. 2021. "Landscape Setting Drives the Microbial Eukaryotic Community Structure in Four Swedish Mountain Lakes Over the Holocene." *Microorganisms* 9: 355. https://doi.org/10.3390/microorganisms9020355.

Chamorro, S., M. Borrego-Ramos, C. Silva, et al. 2024. "A Protocol for the Characterization of Diatom Communities in Mountain Glaciers." *Water* 16: 3417. https://doi.org/10.3390/w16233417.

Che, J., H.-M. Chen, J.-X. Yang, et al. 2012. "Universal COI Primers for DNA Barcoding Amphibians." *Molecular Ecology Resources* 12: 247–258. https://doi.org/10.1111/j.1755-0998.2011.03090.x.

Chen, X., Y. Qin, M. A. Stevenson, and S. McGowan. 2014. "Diatom Communities Along pH and Hydrological Gradients in Three Montane Mires, Central China." *Ecological Indicators* 45: 123–129. https://doi.org/10.1016/j.ecolind.2014.04.016.

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Collins, R. A., J. Bakker, O. S. Wangensteen, et al. 2019. "Non-Specific Amplification Compromises Environmental DNA Metabarcoding With COI." *Methods in Ecology and Evolution* 10: 1985–2001. https://doi.org/10.1111/2041-210X.13276.

Cox, E. J. 2014. "Diatom Identification in the Face of Changing Species Concepts and Evidence of Phenotypic Plasticity." *Journal of Micropalaeontology* 33: 111–120. https://doi.org/10.1144/jmpaleo2014-014.

Decelle, J., S. Romac, E. Sasaki, F. Not, and F. Mahé. 2014. "Intracellular Diversity of the V4 and V9 Regions of the 18S rRNA in Marine Protists (Radiolarians) Assessed by High-Throughput Sequencing." *PLoS One* 9: e104297. https://doi.org/10.1371/journal.pone.0104297.

Elbrecht, V., and F. Leese. 2017. "Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment." *Frontiers in Environmental Science* 5: 11. https://doi.org/10.3389/fenvs.2017.00011.

Esenkulova, S., B. J. G. Sutherland, A. Tabata, N. Haigh, C. M. Pearce, and K. M. Miller. 2020. "Comparing Metabarcoding and Morphological Approaches to Identify Phytoplankton Taxa Associated With Harmful Algal Blooms." *Facets* 5: 784–811. https://doi.org/10.1139/facets-2020-0025.

Evans, K. M., A. H. Wortley, and D. G. Mann. 2007. "An Assessment of Potential Diatom 'Barcode' Genes (*cox*1, *rbc*L, 18S and ITS rDNA) and Their Effectiveness in Determining Relationships in *Sellaphora* (Bacillariophyta)." *Protist* 158: 349–364. https://doi.org/10.1016/j.protis.2007.04.001.

Evans, K. M., A. H. Wortley, G. E. Simpson, V. A. Chepurnov, and D. G. Mann. 2008. "A Molecular Systematic Approach to Explore Diversity Within the *Sellaphora pupula* Species Complex (Bacillariophyta)." *Journal of Phycology* 44: 215–231. https://doi.org/10.1111/j.1529-8817.2007.00454.x.

Garcés-Pastor, S., O. S. Wangensteen, A. Pérez-Haase, et al. 2019. "DNA Metabarcoding Reveals Modern and Past Eukaryotic Communities in a High-Mountain Peat Bog System." *Journal of Paleolimnology* 62: 425–441. https://doi.org/10.1007/s10933-019-00097-x.

Geller, J., C. Meyer, M. Parker, and H. Hawk. 2013. "Redesign of PCR Primers for Mitochondrial Cytochrome c Oxidase Subunit I for Marine Invertebrates and Application in All-Taxa Biotic Surveys." *Molecular Ecology Resources* 13: 851–861. https://doi.org/10.1111/1755-0998.12138.

Godhe, A., M. E. Asplund, K. Härnström, V. Saravanan, A. Tyagi, and I. Karunasagar. 2008. "Quantification of Diatom and Dinoflagellate Biomasses in Coastal Marine Seawater Samples by Real-Time PCR." *Applied and Environmental Microbiology* 74: 7174–7182. https://doi.org/10.1128/AEM.01298-08.

Gonzalez-Saldias, F., A. Pérez-Haase, E. Pladevall-Izard, and J. Gomà. 2025. "Environmental and Overgrazing Effects on Diatom Communities in High Mountain Mires." *Science of the Total Environment* 969: 178983. https://doi.org/10.1016/j.scitotenv.2025.178983.

Goodall, C. 1991. "Procrustes Methods in the Statistical Analysis of Shape." *Journal of the Royal Statistical Society: Series B: Methodological* 53: 285–321. https://doi.org/10.1111/j.2517-6161.1991.tb01825.x.

Guardiola, M., M. J. Uriz, P. Taberlet, E. Coissac, O. S. Wangensteen, and X. Turon. 2015. "Deep-Sea, Deep-Sequencing: Metabarcoding Extracellular DNA From Sediments of Marine Canyons." *PLoS One* 10: e0139633. https://doi.org/10.1371/journal.pone.0139633.

Guo, L., Z. Sui, S. Zhang, Y. Ren, and Y. Liu. 2015. "Comparison of Potential Diatom 'Barcode' Genes (The 18S rRNA Gene and ITS, COI, rbcL) and Their Effectiveness in Discriminating and Determining Species Taxonomy in the Bacillariophyta." *International Journal of Systematic and Evolutionary Microbiology* 65: 1369–1380. https://doi.org/10.1099/ijs.0.000076.

Hadziavdic, K., K. Lekang, A. Lanzen, I. Jonassen, E. M. Thompson, and C. Troedsson. 2014. "Characterization of the 18S rRNA Gene for Designing Universal Eukaryote Specific Primers." *PLoS One* 9: e87624. https://doi.org/10.1371/journal.pone.0087624.

Hamsher, S. E., K. M. Evans, D. G. Mann, A. Poulíčková, and G. W. Saunders. 2011. "Barcoding Diatoms: Exploring Alternatives to COI-5P." *Protist* 162: 405–422. https://doi.org/10.1016/j.protis.2010.09.005.

Hargan, K. E., K. M. Rühland, A. M. Paterson, et al. 2015. "The Influence of Water-Table Depth and pH on the Spatial Distribution of Diatom Species in Peatlands of the Boreal Shield and Hudson Plains, Canada." *Botany* 93: 57–74. https://doi.org/10.1139/cjb-2014-0138.

Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. "Biological Identifications Through DNA Barcodes." *Proceedings of the Royal Society of London, Series B: Biological Sciences* 270: 313–321. https://doi.org/10.1098/rspb.2002.2218.

Jackson, D. A. 1995. "PROTEST: A PROcrustean Randomization TEST of Community Environment Concordance." *Écoscience* 2: 297–303. https://doi.org/10.1080/11956860.1995.11682297.

Kahlert, M., R.-L. Albert, E.-L. Anttila, et al. 2009. "Harmonization Is More Important Than Experience—Results of the First Nordic–Baltic Diatom Intercalibration Exercise 2007 (Stream Monitoring)." *Journal of Applied Phycology* 21: 471–482. https://doi.org/10.1007/s10811-008-9394-5.

Kermarrec, L., A. Franc, F. Rimet, P. Chaumeil, J. F. Humbert, and A. Bouchez. 2013. "Next-Generation Sequencing to Inventory Taxonomic Diversity in Eukaryotic Communities: A Test for Freshwater Diatoms." *Molecular Ecology Resources* 13: 607–619. https://doi.org/10.1111/1755-0998.12105.

Kim, Y.-S., H.-S. Yun, J.-H. Lee, et al. 2024. "Comparison of Metabarcoding and Microscopy Methodologies to Analyze Diatom Communities in Five Estuaries Along the Southern Coast of the Korean Peninsula." *Microbial Ecology* 87: 95. https://doi.org/10.1007/s00248-024-02396-x.

Klepke, M. J., E. E. Sigsgaard, M. R. Jensen, K. Olsen, and P. F. Thomsen. 2022. "Accumulation and Diversity of Airborne, Eukaryotic Environmental DNA." *Environmental DNA* 4: 1323–1339. https://doi.org/10.1002/edn3.340.

Kopylova, E., J. A. Navas-Molina, C. Mercier, et al. 2016. "Open-Source Sequence Clustering Methods Improve the State of the Art." *mSystems* 1: 10–1128. https://doi.org/10.1128/msystems.00003-15.

Krammer, K., and H. Lange-Bertalot. 2004a. "Süßwasserflora von Mitteleuropa, Bd. 2/3: Bacillariophyceae. 3. Teil: Centrales, Fragilariaceae, Eunotiaceae, Achnanthaceae."

Krammer, K., and H. Lange-Bertalot. 2004b. "Süßwasserflora von Mitteleuropa, Bd. 02/4: Bacillariophyceae, 4. Teil: Achnanthaceae, Kritische Ergänzungen zu Navicula (Lineolatae) und Gomphonema."

Krammer, K., and H. Lange-Bertalot. 2004c. "Süßwasserflora von Mitteleuropa, Bd. 02/1: Bacillariophyceae, 1. Teil: Naviculaceae, A: Text; B: Tafeln."

Krammer, K., and H. Lange-Bertalot. 2004d. "Süßwasserflora von Mitteleuropa, Bd. 02/2: Bacillariophyceae, 2. Teil: Bacillariaceae, Epithemiaceae, Surirellaceae."

Kulaš, A., M. Gligora Udovič, K. Tapolczai, P. Žutinić, S. Orlić, and Z. Levkov. 2022. "Diatom eDNA Metabarcoding and Morphological Methods for Bioassessment of Karstic River." *Science of the Total Environment* 829: 154536. https://doi.org/10.1016/j.scitotenv.2022.154536.

Lahti, L., and S. Shetty. 2017. "microbiome R Package." Bioconductor. https://doi.org/10.18129/B9.bioc.microbiome.

Lange-Bertalot, H. 2000. *Diatoms of Europe: The Genus Pinnularia*. A.R.G. Gantner Verlag K.G.

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Lange-Bertalot, H. 2001. *Diatoms of Europe: Navicula Sensu Stricto, 10 Genera Separated From Navicula Sensu Lato, Frustulia*. A.R.G. Gantner Verlag K.G.

Lange-Bertalot, H., M. Bąk, W. Andrzej, and N. Tagliaventi. 2011. *Diatoms of Europe; Volume 6: Eunotia and Some Related Genera*, 740. Gantner Verlag.

Lange-Bertalot, H., G. Hofmann, M. Werum, and M. Cantonati. 2017. *Freshwater Benthic Diatoms of Central Europe. Over 800 common species used in ecological assessment. English edition with updated taxonomy and added species*, 908. Koeltz Botanical Books.

Mann, D., and K. Evans. 2008. "The Species Concept and Cryptic Diversity." In *Proceedings of the 12th International Conference on Harmful Algae*, 262–268. International Society for the Study of Harmful Algae and Intergovernmental Oceanographic Commission of UNESCO.

Mann, D., S. Sato, R. Trobajo, P. Vanormelingen, and C. Souffreau. 2010. "DNA Barcoding for Species Identification and Discovery in Diatoms." *Cryptogamie, Algologie* 31: 557–577.

Mann, D. G., and P. Vanormelingen. 2013. "An Inordinate Fondness? The Number, Distributions, and Origins of Diatom Species." *Journal of Eukaryotic Microbiology* 60: 414–420. https://doi.org/10.1111/jeu.12047.

Moniz, M. B. J., and I. Kaczmarska. 2009. "Barcoding Diatoms: Is There a Good Marker?" *Molecular Ecology Resources* 9: 65–74. https://doi.org/10.1111/j.1755-0998.2009.02633.x.

Mora, D., N. Abarca, S. Proft, et al. 2019. "Morphology and Metabarcoding: A Test With Stream Diatoms From Mexico Highlights the Complementarity of Identification Methods." *Freshwater Science* 38: 448–464. https://doi.org/10.1086/704827.

Nassonova, E., A. Smirnov, J. Fahrni, and J. Pawlowski. 2010. "Barcoding Amoebae: Comparison of SSU, ITS and COI Genes as Tools for Molecular Identification of Naked Lobose Amoebae." *Protist* 161: 102–115. https://doi.org/10.1016/j.protis.2009.07.003.

Nistal-García, A., P. García-García, J. García-Girón, M. Borrego-Ramos, S. Blanco, and E. Bécares. 2021. "DNA Metabarcoding and Morphological Methods Show Complementary Patterns in the Metacommunity Organization of Lentic Epiphytic Diatoms." *Science of the Total Environment* 786: 147410. https://doi.org/10.1016/j.scitotenv.2021.147410.

Oksanen, J., G. L. Simpson, F. G. Blanchet, et al. 2022. "vegan: Community Ecology Package."

Othman, N., H. Haris, Z. Fatin, et al. 2021. "A Review on Environmental DNA (eDNA) Metabarcoding Markers for Wildlife Monitoring Research." *IOP Conference Series: Earth and Environmental Science* 736: 012054. https://doi.org/10.1088/1755-1315/736/1/012054.

Pérez-Burillo, J., D. G. Mann, and R. Trobajo. 2022. "Evaluation of Two Short Overlapping *rbcL* Markers for Diatom Metabarcoding of Environmental Samples: Effects on Biomonitoring Assessment and Species Resolution." *Chemosphere* 307: 135933. https://doi.org/10.1016/j.chemosphere.2022.135933.

Potapova, M., and D. F. Charles. 2003. "Distribution of Benthic Diatoms in U.S. Rivers in Relation to Conductivity and Ionic Composition." *Freshwater Biology* 48: 1311–1328. https://doi.org/10.1046/j.1365-2427.2003.01080.x.

Potapova, M. G., and D. F. Charles. 2002. "Benthic Diatoms in USA Rivers: Distributions Along Spatial and Environmental Gradients." *Journal of Biogeography* 29: 167–187. https://doi.org/10.1046/j.1365-2699.2002.00668.x.

Quast, C., E. Pruesse, P. Yilmaz, et al. 2013. "The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools." *Nucleic Acids Research* 41: D590–D596. https://doi.org/10.1093/nar/gks1219.

Rimet, F. 2012. "Recent Views on River Pollution and Diatoms." *Hydrobiologia* 683: 1–24. https://doi.org/10.1007/s10750-011-0949-0.

Rimet, F., P. Chaumeil, F. Keck, et al. 2016. "R-Syst::Diatom: An Open-Access and Curated Barcode Database for Diatoms and Freshwater Monitoring." *Database* 2016: baw016. https://doi.org/10.1093/database/baw016.

Rimet, F., V. Vasselon, B. A-Keszte, and A. Bouchez. 2018. "Do We Similarly Assess Diversity With Microscopy and High-Throughput Sequencing? Case of Microalgae in Lakes." *Organisms Diversity & Evolution* 18: 51–62. https://doi.org/10.1007/s13127-018-0359-5.

Rivera Rondón, C. A., and J. Catalan. 2017. "Diatom Diversity in the Lakes of the Pyrenees: An Iconographic Reference." *Limnetica* 36: 127–395.

Rivera, S. F., V. Vasselon, S. Jacquet, A. Bouchez, D. Ariztegui, and F. Rimet. 2018. "Metabarcoding of Lake Benthic Diatoms: From Structure Assemblages to Ecological Assessment." *Hydrobiologia* 807: 37–51. https://doi.org/10.1007/s10750-017-3381-2.

Rivera-Rondón, C. A., and J. Catalan. 2020. "Diatoms as Indicators of the Multivariate Environment of Mountain Lakes." *Science of the Total Environment* 703: 135517. https://doi.org/10.1016/j.scitotenv.2019.135517.

Round, F. E., R. M. Crawford, and D. G. Mann. 1990. *Diatoms: Biology and Morphology of the Genera*. Cambridge University Press.

Saunders, G. W. 2005. "Applying DNA Barcoding to Red Macroalgae: A Preliminary Appraisal Holds Promise for Future Applications." *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 360: 1879–1888. https://doi.org/10.1098/rstb.2005.1719.

Schenk, J., N. Kleinbölting, and W. Traunspurger. 2020. "Comparison of Morphological, DNA Barcoding, and Metabarcoding Characterizations of Freshwater Nematode Communities." *Ecology and Evolution* 10: 2885–2899. https://doi.org/10.1002/ece3.6104.

Stevenson, R. J., Y. Pan, and H. van Dam. 2010. "Assessing Environmental Conditions in Rivers and Streams With Diatoms." In *The Diatoms: Applications for the Environmental and Earth Sciences*, edited by E. F. Stoermer and J. P. Smol, 57–85. Cambridge University Press. https://doi.org/10.1017/CBO9780511763175.005.

Sultana, S., M. E. Ali, M. A. M. Hossain, N. Asing Naquiah, and I. S. M. Zaidul. 2018. "Universal Mini COI Barcode for the Identification of Fish Species in Processed Products." *Food Research International* 105: 19–28. https://doi.org/10.1016/j.foodres.2017.10.065.

Taib, N., J.-F. Mangot, I. Domaizon, G. Bronner, and D. Debroas. 2013. "Phylogenetic Affiliation of SSU rRNA Genes Generated by Massively Parallel Sequencing: New Insights Into the Freshwater Protist Diversity." *PLoS One* 8: e58950. https://doi.org/10.1371/journal.pone.0058950.

Tapolczai, K., V. Vasselon, A. Bouchez, C. Stenger-Kovács, J. Padisák, and F. Rimet. 2019. "The Impact of OTU Sequence Similarity Threshold on Diatom-Based Bioassessment: A Case Study of the Rivers of Mayotte (France, Indian Ocean)." *Ecology and Evolution* 9: 166–179. https://doi.org/10.1002/ece3.4701.

Trobajo, R., D. G. Mann, E. Clavero, K. M. Evans, P. Vanormelingen, and R. C. McGregor. 2010. "The Use of Partial cox1, rbcL and LSU rDNA Sequences for Phylogenetics and Species Identification Within the *Nitzschia palea* Species Complex (Bacillariophyceae)." *European Journal of Phycology* 45: 413–425. https://doi.org/10.1080/09670262.2010.498586.

Vasselon, V., A. Bouchez, F. Rimet, et al. 2018. "Avoiding Quantification Bias in Metabarcoding: Application of a Cell Biovolume Correction Factor in Diatom Molecular Biomonitoring." *Methods in Ecology and Evolution* 9: 1060–1069. https://doi.org/10.1111/2041-210X.12960.

Vasselon, V., I. Domaizon, F. Rimet, M. Kahlert, and A. Bouchez. 2017. "Application of High-Throughput Sequencing (HTS) Metabarcoding to Diatom Biomonitoring: Do DNA Extraction Methods Matter?" *Freshwater Science* 36: 162–177. https://doi.org/10.1086/690649.

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Vasselon, V., F. Rimet, I. Domaizon, O. Monnier, Y. Reyjol, and A. Bouchez. 2019. "Assessing Pollution of Aquatic Environments With Diatoms' DNA Metabarcoding: Experience and Developments From France Water Framework Directive Networks." *Metabarcoding and Metagenomics* 3: 101–115. https://doi.org/10.3897/mbmg.3.39646.

Vidaković, D., N. Albert Serge Mayombo, A. Burfeid Castellanos, M. Kloster, and B. Beszteri. 2024. "Diatom Metabarcoding as a Tool to Assess the Water Quality of Two Large Tributaries of the Danube River." *Ecological Indicators* 168: 112793. https://doi.org/10.1016/j.ecolind.2024.112793.

Wang, B., R. Li, X. Lan, D. Kong, X. Liu, and S. Xie. 2024. "Benthic Diatom eDNA Metabarcoding for Ecological Assessment of an Urban River: A Comparison With Morphological Method." *Ecological Indicators* 166: 112302. https://doi.org/10.1016/j.ecolind.2024.112302.

Wangensteen, O. S., C. Palacín, M. Guardiola, and X. Turon. 2018. "DNA Metabarcoding of Littoral Hard-Bottom Communities: High Diversity and Database Gaps Revealed by Two Molecular Markers." *PeerJ* 6: e4705. https://doi.org/10.7717/peerj.4705.

Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan. 2023. "dplyr: A Grammar of Data Manipulation." R Package Version 1.1.4.

Zimmermann, J., G. Glöckner, R. Jahn, N. Enke, and B. Gemeinholzer. 2015. "Metabarcoding vs. Morphological Identification to Assess Diatom Diversity in Environmental Studies." *Molecular Ecology Resources* 15: 526–542. https://doi.org/10.1111/1755-0998.12336.

Zimmermann, J., R. Jahn, and B. Gemeinholzer. 2011. "Barcoding Diatoms: Evaluation of the V4 Subregion on the 18S rRNA Gene, Including New Primers and Protocols." *Organisms, Diversity and Evolution* 11: 173–192. https://doi.org/10.1007/s13127-011-0050-6.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Tables S1–S6:** ece372644-sup-0001-TablesS1-S6.zip.

20457758, 2026, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.72644 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License