MBMG
Metabarcoding & Metagenomics

Metabarcoding and Metagenomics 3: 21–35
DOI 10.3897/mbmg.3.34002

Research Article

Molecular versus morphological data for benthic diatoms biomonitoring in Northern Europe freshwater and consequences for ecological status

Bonnie Bailet[1], **Agnes Bouchez**[2], **Alain Franc**[3], **Jean-Marc Frigerio**[3], **François Keck**[1,2], **Satu-Maaria Karjalainen**[4], **Frederic Rimet**[2], **Susanne Schneider**[5], **Maria Kahlert**[1]

1 *Swedish University of Agricultural Sciences, Department of Aquatic Sciences and Assessment, PO Box 7050, SE – 750 07 Uppsala, Sweden.*

2 *CARRTEL, French National Institute for Agricultural Research (INRA), University of Savoie Mont Blanc, 75 bis avenue de Corzent, 74200 Thonon-les-Bains, France.*

3 *BioGeCo, French National Institute for Agricultural Research (INRA), 69 route d'Arcachon 33610 Cesta, France.*

4 *Finnish Environment Institute, University of Oulu, P. O. Box 413, 90014 Oulu, Finland.*

5 *Norwegian Institute of Water Research, Gaustadalleen 21, 0349 Oslo, Norway.*

Corresponding author: *Bonnie Bailet* (bonnie.bailet@slu.se)

Academic editor: *Hugo De Boer* | **Received** 20 February 2019 | **Accepted** 8 May 2019 | **Published** 12 June 2019

Abstract

Diatoms are known to be efficient bioindicators for water quality assessment because of their rapid response to environmental pressures and their omnipresence in water bodies. The identification of benthic diatoms communities in the biofilm, coupled with quality indices such as the Indice de polluosensibilité spécifique (IPS) can be used for biomonitoring purposes in freshwater. However, the morphological identification and counting of diatoms species under the microscope is time-consuming and requires extensive expertise to deal with a constantly evolving taxonomy. In response, a molecular-based and potentially more cost-effective method has been developed, coupling high-throughput sequencing and DNA metabarcoding. The method has already been tested for water quality assessment with diatoms in Central Europe. In this study, we applied both the traditional and molecular methods on 180 biofilms samples from Northern Europe (rivers and lakes of Fennoscandia and Iceland). The DNA metabarcoding data were obtained on two different DNA markers, the 18S-V4 and rbcL barcodes, with the NucleoSpin Soil kit for DNA extraction and sequenced on an Ion Torrent PGM platform. We assessed the ability of the molecular method to produce species inventories, IPS scores and ecological status class comparable to the ones generated by the traditional morphology-based approach. The two methods generated correlated but significantly different IPS scores and ecological status assessment. The observed deviations are explained by presence/absence and abundance discrepancies in the species inventories, mainly due to the incompleteness of the barcodes reference databases, primer bias and strictness of the bioinformatic pipeline. Abundance discrepancies are less common than presence/absence discrepancies but have a greater effect on the ecological assessment. Missing species in the reference databases are mostly acidophilic benthic diatoms species, typical of the low pH waters of Northern Europe. The two different DNA markers also generated significantly different ecological status assessments. The use of the 18S-V4 marker generates more species inventories discrepancies, but achieves an ecological assessment more similar to the traditional morphology-based method. Further development of the metabarcoding method is needed for its use in environmental assessment. For its application in Northern Europe, completion and curation of reference databases are necessary, as well as evaluation of the currently available bioinformatics pipelines. New indices, fitted for environmental biomonitoring, should also be developed directly from molecular data.

Key Words

Metabarcoding, environmental assessment, 18S-V4, rbcL, Bacillariophyta, water quality

Introduction

Diatom communities are excellent bioindicators of water quality because of their rapid response to environmental changes, such as eutrophication and pollution (Rimet and Bouchez 2012) and their ubiquitous distribution in all types of water bodies. Due to their species sensitivity to pollution and eutrophication, the identification of benthic diatoms is accepted as part of the monitoring programme for the Water Framework Directive in Europe (WFD; European Parliament 2000) in many European countries. Based on their relative abundance and tolerance characteristics, diatom assemblages are used to calculate biological indices, as part of the assessment of the ecological status of a water body (Kelly et al. 2009, 2014). However, the current methodology, based on morphological taxonomic identification, is time-consuming (counting of 400 valves per sample under microscope), and requires extensive expertise, due to a constantly evolving taxonomy (Kahlert et al. 2012). Moreover, there are taxonomic discrepancies between laboratories, hampering the sharing of data (Kahlert et al. 2009).

High Throughput Sequencing based metabarcoding, provides an alternative for diatom monitoring (Kermarrec et al. 2014; Zimmermann et al. 2015). This method is more and more cost-effective, thanks to quickly evolving technology (Stein et al. 2014) and has the ability to facilitate the monitoring programmes at large spatiotemporal scales, even if there are still many challenges (Hajibabaei et al. 2011; Keck et al. 2017; Pawlowski et al. 2018). It also potentially has a greater sensitivity for rare species detection (Zhan et al. 2013) and cryptic species (Rimet et al. 2018) thus enabling the detection of more species than the morphological identification of a standardised number of valves per samples. Applications and improvement of metabarcoding methods for diatoms in water quality assessment are currently being extensively undertaken, especially for European rivers and lakes (Kelly et al. 2018; Rivera et al. 2018; Vasselon et al. 2017a) but also in tropical regions (Vasselon et al. 2017b).

In the scope of developing diatom metabarcoding for Fennoscandia, we need to compare the results of metabarcoding quality assessment to the classical morphological approach in use. Currently, both Finland and Sweden are using diatoms for routine environmental assessment of lakes and streams in the framework of the WFD (Kelly et al. 2009; Kelly et al. 2014) and one of the important diatom indices in use is the Indice de polluosensibilité spécifique (IPS) which has been established to detect eutrophication and pollution (Cemagref 1982) and has been intercalibrated within Europe under the WFD (Kelly et al. 2009; Kelly et al. 2014). The IPS is based on a large number of diatom taxa with indicator values which provide a solid taxonomical and ecological base (Lecointe et al. 1993). Studies comparing molecular and morphological assessment of diatoms communities have been done before in Europe and have shown that, in general, the two methods generate comparable results (Kermarrec et al. 2014; Visco et al. 2015; Zimmermann et al. 2014). However, these studies had worked usually with one biomarker and at a small scale (typically low number of samples and small geographical area). Vasselon et al. (2017b) were the first to scale up the number of samples (80 samples from Mayotte islands, French overseas department), followed by Rivera et al. (2018) (66 samples from Lake Bourget, France).

The present study is now comparing the diatom metabarcoding approach for environmental assessment in Fennoscandia with the established morphological approach on a large scale, using 180 samples of benthic biofilm from both lakes and rivers, covering a broad environmental gradient across Finland, Sweden, Norway and Iceland. We aimed to compare the calculation of a diatom index and the subsequent assessment of ecological status and qualitative and quantitative species identification of benthic diatoms.

Previous studies (Rivera et al. 2018; Vasselon et al. 2017a; Vasselon et al. 2017b) have shown that one of the main problems when using metabarcoding, along with the marker polymorphism, is the taxonomic coverage of the DNA reference library. The quality of the reference database is the most crucial factor for diatom identification at species level (Kermarrec et al. 2013). To our knowledge, the most complete accessible database for diatoms barcodes is R-syst::diatom. This database has been set up as the micro-algal component of a database shared between several taxonomic groups of interest (R-Syst, available at https://www6.inra.fr/r-syst_eng/) (Chaumeil et al. 2018)). The database is regularly curated and updated (Rimet et al. 2016; Rimet et al. 2018). However, the R-syst::diatom database was constructed with a focus on French monitoring needs and contains, until now, mainly barcodes from samples taken in temperate regions (Kermarrec et al. 2014; Rimet et al. 2016).

The water bodies of Fennoscandia (Finland, Sweden and Norway) are different from the ones of France and Central Europe. Fennoscandia freshwater have, on average, lower pH values, higher amount of humic substances and lower nutrient concentrations than freshwaters in Central and Southern Europe (Erlandsson et al. 2008; Johansson and Persson 2001; Ramsay 1898). Considering that diatom assemblages are affected by the local water chemistry (Smol and Stoermer 2010), we would in turn expect different diatom communities in Fennoscandia (Gottschalk and Kahlert 2012). Both Fennoscandia water chemistry itself and the diatom flora are different from Central Europe, where the existing metabarcoding methods have been developed. Thus a lower efficiency when extracting and amplifying the DNA barcodes using these methods may be expected, due to untested water chemistry, possibly including inhibitors. A lower number of detected diatom taxa, mainly due to missing taxa in the established reference database and lower primer suitability are expected as well.

Another challenge that could affect the results given by molecular analysis is the choice of the DNA marker. The two main markers, currently used for diatoms in Europe, are the 18S-V4 region (SSU rRNA) and rbcL (from the chloroplast genome). The rbcL-region has proved to be more efficient for diatom communities from temperate and tropical

rivers of French territories (Kermarrec et al. 2014; Vasselon et al. 2017a) while the 18S-V4 was used in Germany and Switzerland (Visco et al. 2015; Zimmermann et al. 2015) because of the great amount of reference barcodes available. Contrary to the 18S marker which has potential to assess non-photosynthetic organisms, the rbcL marker allows focusing and putting all sequencing effort directly on photosynthetic organisms. Both of these markers are represented in the R-syst::diatom database. Since different markers possess different discriminatory power (Kermarrec et al. 2013) and species-specific affinities (Vasselon et al. 2017a) we decided to work with both the 18S-V4 and the rbcL markers.

Finally, the suitability of the molecular approach strongly depends on the quality of the generated taxonomic assignments (Keck et al. 2017) and a great diversity of 'bioinformatics pipelines' have been developed in an attempt to produce the best taxonomic assignation while keeping up with the high amount of HTS data currently produced. When selecting a 'pipeline', one must consider the balance between the accuracy of the supervised clustering method and the calculation power required. Preferably, the calculation steps should include as few heuristics as possible because, while they speed up calculation time, those shortcuts can significantly lower the quality of the taxonomic assignment (Frigerio et al. 2016).

In the scope of this first application of metabarcoding for environmental assessment of Fennoscandia's water bodies using diatom communities, we aim to test four hypothesis:

1. Ecological status assessment of Fennoscandia lakes and rivers based on diatom indices will be similar using either metabarcoding or morphological methods for diatom identification.
2. Due to its detection power towards rare and cryptic species, the metabarcoding approach will detect species missed by the traditional approach.
3. However, because of the dependence on reference databases that are potentially incomplete regarding the Fennoscandia diatom flora, the metabarcoding approach will not cover all taxonomic diversity. Some species will be missing and specific genera will be less covered than others.
4. The use of two different DNA markers, rbcL and 18S-V4, will lead to dissimilarities in the ecological assessment caused by differences in the taxonomic coverage of the associated reference databases and by primers specificity.

Material and methods

Sampling sites

A total of 180 environmental samples were collected from benthic biofilms in 65 streams and 43 lakes in Fennoscandia (Sweden, Finland, Norway) and two sites in Iceland (Fig. 1). The sites cover a broad environmental gradient

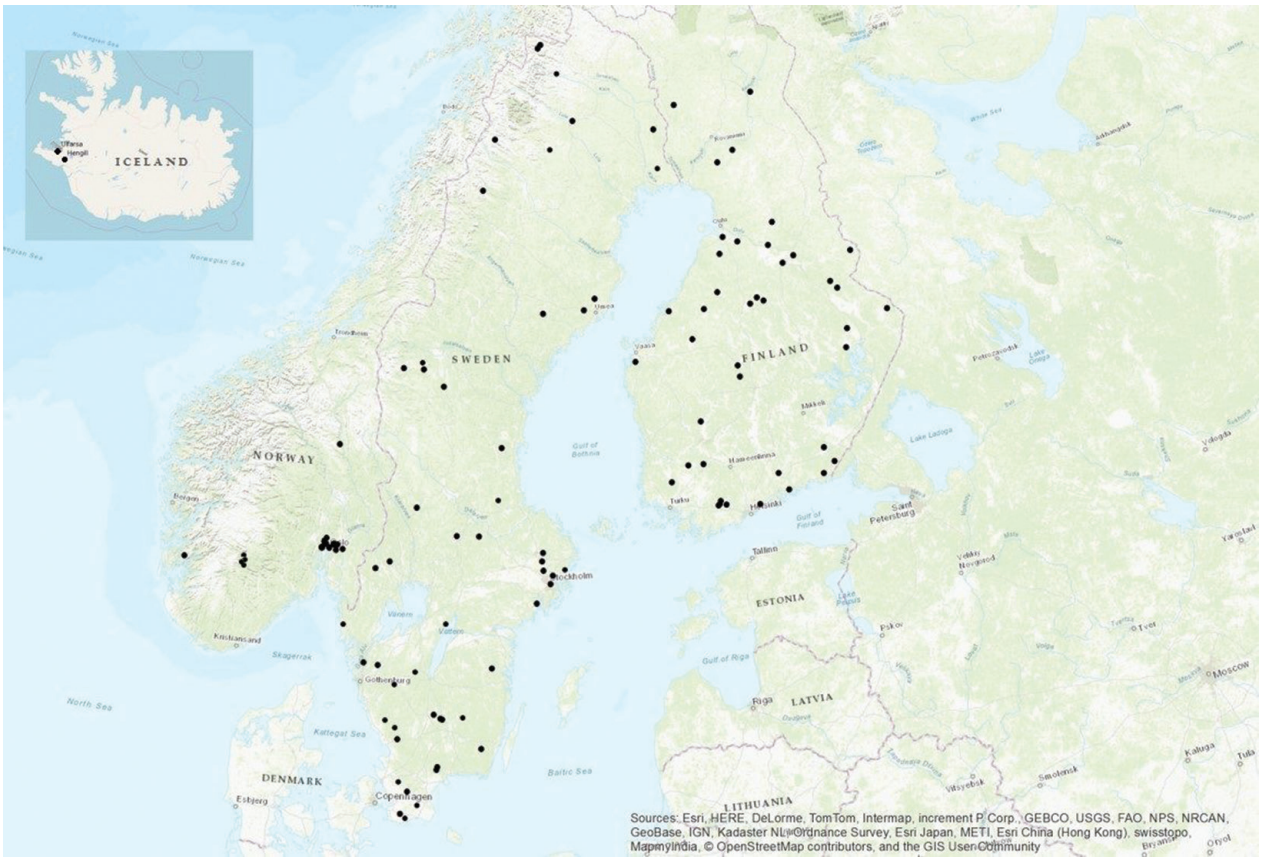

Figure 1. Sampling sites location.

Table 1. Water chemistry characteristics of the 110 sites included in this study.

		Mean	Range
Alkalinity (Alk)	mEq/l	0.543	0.01–6.03
Conductivity (Kond)	mS/m	12	0.5–131
pH		6.7	4.6–8.6
Total organic Carbon (TOC)	mg/l	12	0.9–35
Total Nitrogen (TotN)	µg/l	777	7–3801
Total Phosphorus (TotP)	µg/l	39	0.2–433

of agriculture, alpine and boreal catchment areas and water chemistry parameters (Table 1) were extracted from the Swedish National database (https://miljodata.slu.se/mvm/) for the Swedish sites, from the Hertta system version 5.7 of Finnish Environmental Administration (http://www.syke.fi/en-US/Open_information) for the Finnish sites and from (Friberg et al. 2009) and (Ólafsson et al. 2010) for the two Iceland sites. Water chemical analyses for the sites in Norway were performed by NIVAs accredited laboratory according to Norwegian standard methods.

Sample analysis

All samples were collected in autumn from submerged hard substrate following the European standard for diatom sampling (EN 13946:2014, CEN (2014)). The samples were preserved with 97% ethanol (final concentration approximately 70%) to protect the DNA from degradation in long term storage (Stein et al. 2013). The samples were kept in the dark and at room temperature, where they later were divided for morphological analysis and DNA extraction.

Morphological analysis

Preparation, identification and counting for the morphological diatom analysis were performed using European and Swedish standards (SS-EN 13946:2014; SS-EN 14407:2014; (Jarlman et al. 2016). Briefly, the samples were oxidised with hydrogen peroxide ($H_2O_{2)}$, and the cleaned diatom valves were mounted with Naphrax (Brunel Microscope Ltd) to permanently fix the sample material on glass slides. Identification to the lowest taxonomic level possible was done under light microscope with interference contrast (1000× magnification). At least 400 valves per sample were counted and identified using standard literature (Jarlman et al. 2016). Taxonomic names were harmonised in between countries by using OMNIDIA codes (Lecointe et al. 1993) instead of binomial nomenclature.

DNA extraction and PCR amplification

Water samples from Fennoscandia can have high humic acid and TOC concentrations, so the DNA was extracted from 4–8 ml of sample using the NucleoSpin Soil Kit (Macherey-Nagel), following the recommendations of the manufacturer with one modification: the centrifugation time for washing the silica membrane was changed from 30 seconds to 1 minute for each of the 4 washing steps. DNA quality of the samples was assessed by spectrophotometry using the 260/280 nm ratio on the NanoDrop ND-1000 (Thermo Fisher Scientific) in order to estimate the water dilution factor needed to achieve 25 ngDA/µl for PCR. For the PCR amplification, we targeted two different markers: a 312 bp barcode on the *rbc*L plastid gene using the modified Diat_rbcL_708F and R3 primer pair, with increased degeneracy to match a broader diversity of diatoms (Vasselon et al. 2017b) and the hypervariable region V4 of the 18S rRNA gene using the modified DIV4for and DIV4rev3 (Visco et al. 2015).

To achieve sufficient DNA concentration, each sample for PCR amplification was done in triplicate. Each PCR mix was composed by 7.5 µl of DNA extract (for the *rbc*L marker) or 2 µl of DNA extract (for the 18S-V4 marker), 2.5 µl of 10× buffer, 2 µl of dNTP (2.5 mM), 1.25 µl of the respective forward primer mix (10 pmol/µl) and 1.25 µl of the respective reverse primer mix (10 pmol/µl), 1.25 µl of BSA (10 mg/ml), 0.15 µl of TakaraLa Taq polymerase and completed with molecular biology grade water. For the rbcL marker, the PCR reactions conditions were 5 min at 95 °C for initial denaturation, followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 54 °C for 1 min and extension at 72 °C for 1 min. For the 18S-V4 marker, the PCR reactions conditions were 2 min at 94 °C for initial denaturation, followed by 35 cycles of denaturation at 94 °C for 45 sec, annealing at 50 °C for 45 sec and extension at 72 °C for 1 min. The product of the PCR amplification was tested by electrophoresis on 1.5% agarose gel before further purification of the DNA.

Preparation and High-Throughput sequencing of DNA

The triplicates of PCR products of each sample were pooled together in 1 DNA LoBind 1.5 ml tubes (Eppendorf) and purified with Agencourt AMPure beads (Beckman Coulter) following the manufacturer's instruction but with an adjusted volume ratio of 1.5× beads/DNA. Repair of amplicons fragments, ligation of tags to amplicons and library preparation were done using the NEBNext FastDNA Library Prep set for Ion Torrent (Biolabs) as described in (Vasselon et al. 2017a). The quantification and quality of the resulting libraries was checked using the 2200 TapeStation (Agilent Technologies) with D1000 High Sensitivity screen tape and reagents. The libraries were pooled together in 2 mix for each markers (98 samples per mix, 4 mix in total) at a final concentration of 100 pM per mix. Each mix was sequenced independently by the Platform Genome Transcriptome (PGTB, Bordeaux, France) using an Ion 318 Chip Kit V2 (Life technologies) on an Ion Torrent Personal Genome Machine (PGM).

Sequence data processing

The sequencing platform provides fastq file for each sample after demultiplexing and removing the tags se-

quences. The first filtering step of the sequences excluded too short and too long reads, so that only reads between 300–315 bp were left for the rbcL marker and between 320–340 bp for the 18S-V4 marker.

Diatoms molecular inventories were obtained with the R-syst::diatom database (Rimet et al. 2016) using the 18-02-2016: R-Syst::diatom (version 5) for the rbcL barcode (1625 reference sequences covering 180 genera and 605 species of diatoms) and 20-03-2017: R-Syst::diatom (version 6) for the 18S barcode (2652 reference sequences covering 222 genera and 844 species of diatoms). The taxonomic identification of DNA reads at species level (or at genus level when species level was not possible) was done following a two-step process. First, a matrix of pair-wise distances between the unknown read and the reads from the reference database was computed, for each sample and each database, using the programme MPI-disseq. The MPI-disseq programme runs under Python as a parallel implementation of a Smith-Waterman algorithm with Message Passing Interface (MPI) and computations were done at the French National Computing Center (IDRIS) on a Blue Gene Q hyper-parallel machine (1024 cores). Secondly, each matrix was processed with the Python programme Diagno-syst (Frigerio et al. 2016) which selects, for each read, the references barcodes that are at a distance lower than a given gap and assigns the taxonomy to the unknown read only if all references at the same distance or lower have the same taxonomic name. We chose a 10 bp gap to make sure to encompass the similarity levels expected for species (strict 99% similarity with gaps 0–4 and relaxed 97% similarity with gaps 4–10). A distance of more than 10bp difference translates as an identification above species level and has a higher possibility of being a mismatch and thus was not kept for further analysis. However, if a reference barcode for a genus (e.g. the reference barcodes KU179117 to KU179119 identify *Caloneis* sp. with the rbcL marker) was assigned to a read at a distance smaller than 10 bp, the taxonomical assignment was kept in the dataset.

In this study, to compare valve counts and reads number, we used two comparable abundance thresholds following Frigerio et al. (2016): low abundance in the morphological assessments was defined as ≤ 1% relative abundance and high abundance was > 1%. Identification at ≤ 0.0025% relative abundance was not taken into account. In the molecular assessments, low abundance was defined as number of amplicon reads ≤ 1000 and high abundance then was defined as > 1000 reads. For reads number, detection rate < 10 reads was removed from the final dataset.

Ecological status class assessment

We used the diatoms' inventories produced by morphological and molecular methods to assess the ecological status class of our studied lakes and streams, as required by the WFD, by calculating the intercalibrated diatoms index IPS (Kelly et al. 2009; Kelly et al. 2014) and the intercalibrated class boundaries. Intercalibrated IPS values exist for each species of the Fennoscandia taxa list (Kahlert et al. 2017), valid at the moment of the study and any species not included in this taxa list has no assigned IPS value in this study. We used the IPS index as it is one of the most widely used diatoms' index in Europe, is currently used in Sweden and previously also in Finland. No diatom index is used yet in Norway or Iceland. The IPS takes into account both presence and relative abundance and has indicator values for almost all freshwater diatoms species. The species encountered in our samples, without an IPS indicator value, were included in the calculation as "unidentified taxa" with an IPS score of 0.

Statistical analysis

To test if the metabarcoding method returns comparable results regarding ecological status class when using benthic diatoms communities, we tested first if the ecological status classes were significantly different between the two identification methods with a Friedman's test (Hollander and Wolfe 1973) and a pair-wise Wilcoxon test (Bauer 1972). To understand which of the rbcL or 18S marker gave more similar results compared to the established method, based on morphological diatom identification, we calculated how many sites were classified differently when using either barcodes. We also compared the IPS scores calculated on the taxa lists generated by morphology and metabarcoding using Student paired sample t-test (Student 1908; Zabell 2008). We also used the correlation of the IPS scores between morphological and molecular methods to assess similarity of results.

To understand the causes for eventual deviations between the morphological and metabarcoding methods and to focus future development studies, we used multiple methods. We began by analysing which environmental variables (amongst Alk, TOC, TotP, TotN, Kond and pH, cf. Table 1), were mainly correlated to the observed deviations in IPS scores, using a Partial least squares regression (PLS regression, Chambers and Pope (1992)). Then we used two analyses to determine if morphological and metabarcoding methods generated different species lists and what species mainly caused the observed IPS scores deviations. For these, we tested if diatom diversity, represented by the Shannon index, was different between the two methods and the two DNA markers. After that, for more depth, we investigated the causes of the species list differences between the morphological and molecular assessments and labelled them using the codes described in Table 2. We then assessed if specific causes were linked to specific sites and ecology using Canonical Correspondence Analysis. For this analysis, the occurrences across the 180 samples for each mismatch code was extracted from the dataset as binary data (1= occurring problem, 0=no problem occurring).

Additionally, we carried out a SIMPER analysis (Clarke 1993) to understand which species are likely to account for most of the observed IPS values deviations.

Table 2. Codes for species list deviations.

MOL		Species found by molecular method only
	H	Species not in the Fennoscandia taxa list
MOR		Species found with morphological method only
	ND	Species represented in the DNA database but no identification
	NR	Species not in the respective DNA database
AB1		Species found with both techniques but with higher abundance in morphological inventory
AB2		Species found with both techniques but with higher abundance in molecular inventory
GM		Species found by both techniques with the same abundance
G		Taxonomic identification stopped at the genus level

A SIMPER analysis assesses which taxa are predominant in different groups of samples. Since a deviation of ± 1 in the IPS scores is considered method-bound uncertainty for a class (Swedish Environmental Protection Agency 2007) we used the following three clustering groups to define deviation from expected value: "positive deviation > 1", "negative deviation > 1" and "no deviation". Then, we analysed, on species relative abundance data, which of the taxa were most abundant in each group, i.e. most likely responsible for the observed positive or negative deviations in IPS scores.

Finally, each analysis mentioned previously was run on two datasets (one with the rbcL marker results and one for the 18S-v4 marker results) in order to compare the effect of the DNA marker choice. To evaluate the comparability with the established environmental assessment, we compared the ecological status class derived from the different markers directly and also analysed the extension and the impact of the lacks in the respective databases on IPS calculation and species assignment.

Software

We performed the statistical analyses using the R 3.3.1 software (R-CoreTeam 2013) in RStudio (version 1.0.136), using the following packages: Packages utils (version 3.3.2), base (version 3.3.2), readxl (version 1.0.0) and dplyr (version 0.7.4) for basic data handling, packages stats (version 3.3.2), relaimpo (version 2.2-3), pls (version 2.6.0) and vegan (version 2.4-2) for statistical analyses and package ggplot2 (version 2.2.1) for all graphic representations. The Simper analysis was performed in the software PAST, version 2.15b (Hammer et al. 2001).

Results

Comparison of methods for environmental assessment – Ecological status class

The IPS index calculations done on morphological inventories, ranged from 12 to 20. The calculations done on the molecular inventories ranged from 7.9 to 20 and from 7.8 to 20 for the 18S and rbcL markers, respectively.

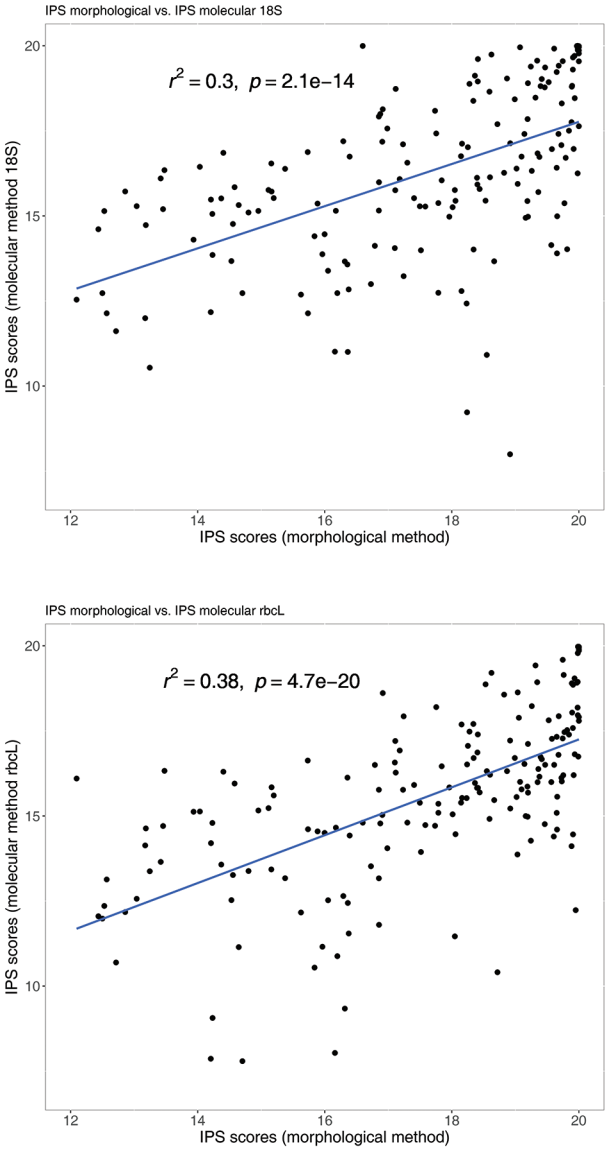

Figure 2. IPS scores correlation. The two axes show the IPS scores values of the samples assessed by the molecular (y-axis) or morphological method (x-axis). Increasing IPS scores values show increasingly good ecological status (IPS ≥ 17.5 = high, IPS ≤ 8 = bad).

We found that, even if both the established and the molecular method indicated the same trend regarding the assessment of water quality (Fig. 2), the ecological status classes were significantly different between the methods. Indeed, only 48% of our samples with the 18S marker and 37.5% with the rbcL marker gave the same ecological status (Table 3) as the morphological analysis. The Friedman's test confirmed a significant difference (p-value < 0.001) between the ecological status classes obtained by the morphological and molecular methods and each used marker. A pair-wise Wilcoxon test indicated that the two markers generated significantly different ecological status class from one another as well (p-value < 0.05).

Table 3. Percentages of overestimation and underestimation of ecological status class in the samples.

	Overestimate	Exact	Underestimate			
Magnitude of quality class alteration	1		1	2	3	4
rbcL	6%	38%	45%	8%	1%	2%
18S	11%	49%	34%	5%	1%	0%

Overall, 56% of the rbcL samples and 40% of the 18S samples were associated with an ecological status lower than the one found using morphological assessment (Table 3), while 6% and 11% of the samples, for the rbcL and 18S markers respectively, were associated with a higher ecological status than the one found using the morphological assessment (Table 3).

Comparison of methods for environmental assessment – Polluosensibilité spécifique index (IPS)

In general, IPS scores calculated on taxa lists generated from molecular markers were correlated with those calculated with the morphological method (18S marker: $r^2 = 0.29$, $p < 0.001$ and rbcL marker: $r^2 = 0.38$, $p < 0.001$, R 3.3.2 package pls) (Fig. 2). However, correlations were weak and there was a significant difference between the IPS scores calculated on taxa lists generated by morphological and molecular assessment, for both the 18S and rbcL markers (Student's paired T-test, $p < 0.001$ and $p < 0.001$, respectively).

Deviations link to Environmental variables

The PLS regression showed that TotP ($p < 0.01$) and conductivity ($p < 0.01$) were significant predictors of the deviation of IPS scores from morphological to 18S communities (Table 4). Indeed, there was a larger difference between the morphological and molecular assessments on the 18S marker for water bodies with a low TotP and low Conductivity, indicated by the negative estimate values of the PLS (Table 4). Conductivity ($p < 0.05$) was also found significant to predict the deviation of IPS scores from morphological communities to the rbcL communities and, thus, there was a larger difference between the morphological and molecular assessments on the rbcL marker for water bodies with low conductivity.

General deviations in species lists

In total, the morphological analysis identified 585 species within 87 genera across all 4 Fennoscandia countries included in the analyses. Species richness per sample varied from 4 to 103 species, being lowest in the Norwegian samples and highest in Swedish samples. The average number of species per sample was 62 in Swedish sites, 63 in Finnish sites and 49 and 21 in Icelandic and Norwegian sites, respectively. The dominant genera across all samples were *Achnanthidium*, *Eunotia* and *Fragilaria* and the dominant species were *Achnanthidium minutissimum*, *Tabellaria flocculosa*, *Fragilaria gracilis* and *Eunotia incisa*.

Table 4. Estimates and p-values of the PLS regression on environmental variables and the deviations (Δ) of IPS scores between morphological and molecular assessments.

	Δ IPS scores 18S marker/morphology		Δ IPS scores rbcL marker/morphology	
	Estimate	p-value	Estimate	p-value
(Intercept)	-0.9338	0.66	2.14	0.25
Conductivity	-0.0339	**p < 0.01**	-0.0225	**p < 0.05**
pH	0.4339	0.14	0.0028	0.99
TOC	-0.0064	0.85	0.0077	0.79
TotP	-0.0137	**p < 0.01**	-0.0063	0.17
TotN	0.0004	0.23	0.0003	0.28

Shannon diversity and taxa missing from reference databases

The Shannon scores of the morphological and molecular taxa inventories were significantly different (Student's paired-sample t-test, $p < 0.001$ both for the use of 18S and rbcL). As expected, the two references databases are lacking a significant amount of species used in Fennoscandia water quality assessment, with only 15.4% of all Fennoscandian taxa represented in the 18S database and 17.8% in the rbcL database. Many genera had only few species represented in the reference databases, especially *Achnanthidium*, *Eunotia*, *Gomphonema*, *Navicula* and *Nitzschia*. However, regarding our own study, actually 70% of the species found by the morphological method were represented in the reference databases.

Causes for differences in the species list between morphological and molecular assessments

When calculating the probability for a correct identification, including also the comparison of abundance, on average, 5% of all taxa showed a good match between the molecular and morphological techniques (Fig. 3). However, about 95% showed a mismatch, either in species presence or abundance or both (Fig. 3). We present our proportions for every mismatch between high abundance or low abundance, except the abundance mismatches which are a combination of high and low abundance between the morphological and molecular assessments (cf. Table 2, Fig. 3). Across all our samples, the most common difference between the species lists was when a species is found by the morphological assessment only (MOR mismatch). This mismatch was either associated with a lack of reference barcode in the DNA database (NR, in red in Fig. 3) or either by a lack of DNA detected by the molecular method, despite having a reference barcode for the species in the reference database (ND, in orange in Fig. 3). The probability of getting these MOR-NR or MOR-ND mismatches in the species list was high and

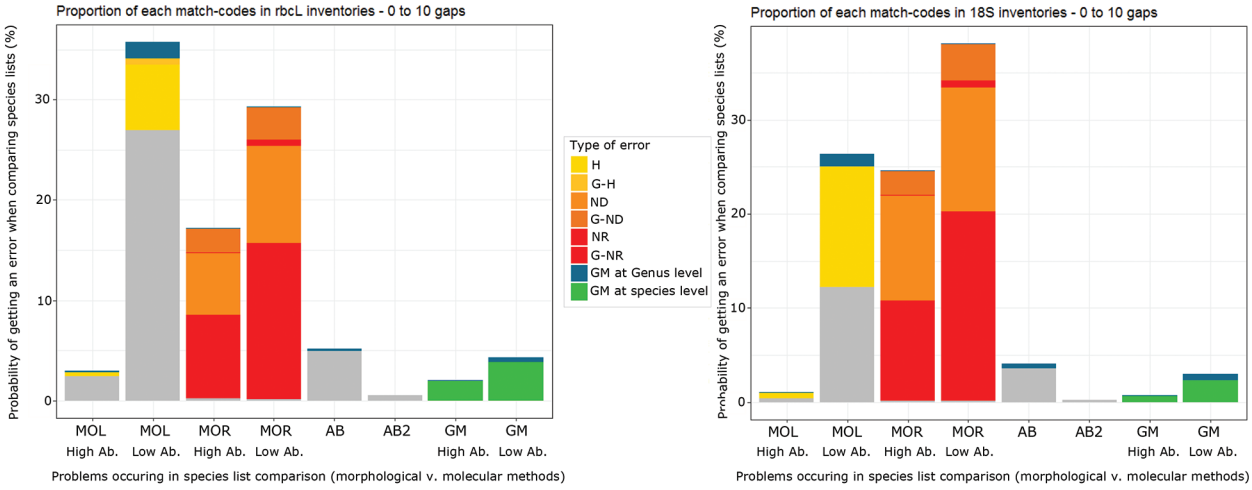

Figure 3. Average probability, per analysis, of a correct match or mismatch for a species identification and abundance assessment when comparing molecular and morphological methods. The code G represents a reference barcode stopping at the genus level; the yellow portion represents a species (H) or a genus (G-H) not expected in Fennoscandia; the orange portion represents the case when no DNA was detected for a species (ND) or a genus (G-ND) despite having a reference barcode in the database; the red portion represents the case when the database lacked a reference barcode for a species (NR) or a genus (G-NR). The green portion represents a good match both in presence and abundance of species, between the morphological and molecular assessments.

very similar for both the 18S and rbcL markers (MOR-NR: 50% and 53%, respectively and MOR-ND: 49% and 46%, respectively) (Fig. 3).

Another, less common, mismatch was an abundance discrepancy (AB1 and AB2). This mismatch happens when a species is found by both techniques but in significantly different abundances. In the case of a higher abundance in morphological than in molecular species lists (AB1), 44% of the concerned species were found to be causing the error with both markers. Similarly, 25% of the species, found in higher abundance in the molecular species list (AB2), were common to the two markers. The abundance mismatch thus seems to be linked to specific taxa (Ex: *Achnanthidium minutissimum*).

We also found that the hypothesis of a better taxonomic coverage achieved with molecular technique is supported by a significant number of species, most of them rare species, detected only when using the metabarcoding technique (MOL mismatch): 56 species and 8 genera when using the 18S marker, representing 27% of our diatoms' communities and 122 species and 7 genera with the rbcL marker which represented 38% of our diatoms' communities. Furthermore, some species, detected only by the molecular technique, were not included in the taxonomy used for morphological identification in Fennoscandia (H, in yellow). More precisely, 8% of the identifications for the rbcL marker (52 taxa amongst 34 genera) and 6% for the 18S marker (58 taxa amongst 45 genera) were not included.

To track the origin (link to specific sites and ecology) of the observed discrepancies in the species lists, we used Canonical Correspondence Analysis (CCA) on the occurrences across the 180 samples, for each mismatch code against our five environmental parameters. The CCA analysis showed that low pH and high TOC explained most of the variability observed when a species is not identified with the molecular technique due to a lack in the DNA reference database (NR), in agreement with the fact that many of the species lacking from the databases are acidophilic diatoms. None of the tested environmental variables explained the occurrence of abundances mismatch (AB1 or AB2) with either of the markers. The occurrence of a good match between molecular and morphological methods (GM code) was correlated to a high pH. We found no differences between the two studied DNA markers regarding how environmental variables could explain the different types of matches and mismatches between molecular and morphological species lists.

Main species causing differences in IPS

The SIMPER analysis highlighted the species that are most likely to contribute to the IPS deviation. When using the 18S or rbcL marker, *Achnanthidium minutissimum*, *Eolimna minima*, *Amphora pediculus*, *Rhoicosphenia abbreviata*, *Nitzschia dissipata* and *Eunotia incisa* were the main species contributing to an overestimation of the IPS values with molecular assessment compared to the morphological assessment. In the case of underestimation of the IPS values with molecular assessment, the main contributing species were *Achnanthidium minutissimum*, *Tabellaria flocculosa*, *Fragilaria gracilis*, *Aulacoseira ambigua*, *Cocconeis placentula*, *Staurosira pinnata* and *Fragilaria capucina* for both markers, as well as *Eunotia incisa* for the rbcL and *Eunotia minor* for the 18S marker.

When looking back at the mismatch codes, the majority of these species were represented in both reference data-

bases, but showed significant discrepancy in their relative abundance when assessed by morphological or by molecular techniques (AB1 and AB2). *Achnanthidium minutissimum* was found occurring with a higher abundance with the morphological assessments when using either markers but also, in some cases, found by molecular assessment only. *Fragilaria gracilis* was found with higher abundance with the morphological method than when using metabarcoding with the rbcL marker. With the 18S marker, it was only identified below the threshold we used for our analysis (< 10 reads). *Amphora pediculus* was found with a higher abundance in the morphological method than by metabarcoding with the rbcL marker but with higher abundance in metabarcoding when using the 18S marker. *Tabellaria flocculosa* was always found with higher abundance with the morphological method when using the rbcL marker and was not found at all by the molecular technique when using the 18S marker (despite being represented in the reference database). Only *Eunotia incisa*, amongst the species most important for the IPS deviations, was actually missing from the reference databases, the reason why it could not be found by the molecular method.

Comparative use of the rbcL and 18S markers

The gaps in the two reference databases were overlapping: out of 416 species and 409 species missing across 74 genera for the 18S and the rbcL databases, respectively, 388 species were missing in both databases (across 69 genera). The IPS scores obtained with the two markers were correlated ($r^2 = 0.30$, $p < 0.001$, Fig. 4). However, our analyses still confirmed significant discrepancies between the results produced using the two markers. The IPS scores were significantly different from one another (Student paired t-test, $p < 0.01$). The rbcL marker gave more results of "Poor" and "Bad" water quality class, none of them being the same sites as with the 18S marker (Fig. 4, 9 dots versus 4 triangles). Student's paired t-test on the Shannon scores also revealed a significant difference in the species communities obtained with the rbcL marker and the ones obtained with the 18S marker ($p < 0.001$). The proportion of good match between the species lists generated by the molecular and morphological methods was higher with the rbcL marker (6.2%) than with the 18S marker (3.6%).

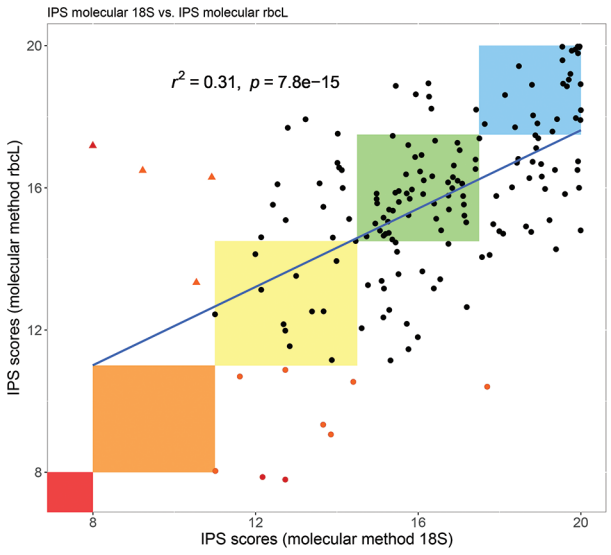

Figure 4. Correlation between the IPS scores obtained with the 18S and the rbcL markers. The boundaries for the ecological status classes defined by the IPS are indicated by the coloured squares (red=very bad, orange=bad, yellow=moderate, green=good, blue=very good). Red dots highlight "very bad" ecological status samples when using the rbcL marker and red triangles when using the 18S marker. Orange dots and triangles highlight "bad" ecological status samples using the rbcL and 18S markers, respectively.

Discussion

Contrary to our first hypothesis, we found significant differences between the ecological status class generated from the morphological and the molecular assessments, rejecting our first hypothesis. We could confirm our second hypothesis as we found some taxa with the metabarcoding approach which were not detected by the morphological assessment. On the other hand, there were even more taxa which were not found by the metabarcoding approach, even if the morphological method detected them, which confirmed our third hypothesis. Finally, we found differences between the two markers used leading to discrepancies in the ecological assessment and confirming our fourth hypothesis.

Different ecological status resulted from the morphological and the molecular assessments

The linear relationship of the ecological assessment results of both methods found in our study confirms that metabarcoding has the potential to be used for biomonitoring , as previously shown by other studies (Apothéloz-Perret-Gentil et al. 2017; Kermarrec et al. 2014; Vasselon et al. 2017a; Visco et al. 2015). However, we found many discrepancies between the IPS index scores calculated from the diatoms taxa lists generated by the molecular and by morphological methods, despite the fact that they were significantly correlated. Similar discrepancies have been observed in previous studies (Rivera et al. 2018; Vasselon et al. 2017b).

The ecological status class boundary between "good" and "moderate" ecological status is especially important for decision-makers, since the WFD defines that water bodies below this "good" ecological status are in need of remediation. Consequently, the discrepancies in ecological status class are the major concern when applying metabarcoding for monitoring purposes. The IPS index and, in turn, ecological status classes, are based on the presence and abundance of species of the diatom community.

In this study, we found evidence of both differences in species presence and in species abundances in the species lists derived from the two methods, as well their impact on the IPS index calculations.

The reference database incompleteness leads both to non-identification and misidentification of species

Even with the current effort to complete and curate the diatom reference databases, many species are still lacking barcode information. Visco et al. (2015) estimated that no more than 30% of European species are currently represented in reference databases. The main species coverage in the Rsyst::diatom database is for temperate species and lack an important number of species from the Fennoscandian communities. For example, our results showed that species list discrepancies from a lack of reference barcodes were strongly correlated to the low pH levels and high TOC values, which are common in Northern European waters. Similarly, the highest IPS deviations, found in oligotrophic sites, highlight a lack of representation of acidophilic species in the reference database. This is also supported by the CCA results, showing that samples, not affected by the databases gaps, were mostly correlated to a high pH, in agreement with the fact that many of the species lacking from the databases are acidophilic diatoms, such as *Eunotia incisa* and *Encyonema neogracile*. With more than half of the Fennoscandian taxa missing a reference barcode, the database is the main reason for the non-identification, with metabarcoding, of species otherwise identified with the conventional morphological method. There have been efforts to replace missing species information using phylogenetic information derived from higher taxonomical level. However, this approach is not refined yet and might lead to wrong indicator values of ecologically different but closely related species (Keck et al. 2018).

Incorrect identification of the reference barcodes can, of course, lead to important discrepancies, hence the importance of a continuous database verification and curation. However, even with the benefit of being constantly updated and curated, some taxonomic discrepancies were also detected when we used using Rsyst::diatom. Some taxonomic names differed from the ones used for the same taxon in morphological assessment leading to an artificial mismatch between morphological and metabarcoding species inventories. Sometimes, taxonomic information came on different taxonomic levels which required the merging of taxa for comparison (e.g. *Ulnaria ulna var. acus* and *Ulnaria acus)*. The Rsyst::diatom version v5 for the rbcL barcode also included some sequences without a reliable taxonomic name assigned (e.g. "*Nitzschia aff. dissipata*"), although those mainly concerned reference barcodes included in Rsyst from the NCBI nucleotide database. Sequences assigned to such reference barcodes were removed from our dataset to avoid incorrect identification. Completion and curation efforts of the rbcL marker database is now being undertaken at European level and the latest Rsyst::diatom version v7 (in February 2018, renamed "Diat.barcode") was released with most of those taxonomic discrepancies removed. On the other hand, the curation of the Rsyst::diatom database for the 18S-V4 barcode is still only partial.

The detection of species is also affected by primer bias and the choice of the pipeline

Primer bias is often found to be a major source of variation (Pawlowski et al. 2018) and PCR primers efficiency differ between species (Kermarrec et al. 2013), i.e. some primers lead to a preferential amplification of one taxon over another. We used strongly degenerated primers for both the 18S and the rbcL markers to match a wider range of species needed for assessment of Fennoscandian diatom communities (Vasselon et al. 2017a; Visco et al. 2015). However, in our study, a significant number of taxa were not detected at all with metabarcoding assessment: the DNA extraction and amplification may have been hampered because of primer specificity (Elbrecht and Leese 2015). For example, the species *Tabellaria flocculosa* was not well detected by the 18S marker and its absence greatly affected the calculation of IPS scores. Additionally, we found evidence of amplification of green algae in some samples when using the 18S marker, leading to a low share of diatoms reads per sample. This, in turn, might have prevented the amplification and sequencing of some non-dominant diatom taxa in those samples. A way to circumvent this problem is to multiplex less samples together for the sequencing, generating more reads per sample to be assigned to diatom taxa (Zimmermann et al. 2015).

Another well-known limitation in using metabarcoding for ecological assessment is the clustering method used before the taxonomic assignment, often leading to massive loss of genetic information (Keck et al. 2017). Diagnosyst is a pipeline without clustering of sequences into Molecular Operational Taxonomic Units (MOTUS), which allows the avoidance of problems linked to arbitrary threshold for clustering (Frigerio et al. 2016). In the Diagnosyst process, a simple sequences filtering on size is followed by a strict taxonomic assignment without heuristics: every query is compared against the each barcode in the reference database. While this strict assignment might produce a more precise and clean taxonomic assignment than most currently used pipelines, it in turn requires the computer power for dealing with a massive amount of molecular data. The strict taxonomic assignment should remove chimera sequences (artificial read created by two distinct portions of the genome) and prevent most misidentifications of taxa. However, a strict assignment also leads to a massive proportion of discarded reads and thus, to the loss of sequencing data. It is also possible that the strict size selection of sequences might have removed valuable genetic data (specific species may produce shorter sequences than 300 bp (for the rbcL) or 320 bp (for the 18S)). The only reliable way to assess the effects of different clustering and different taxonomic assignment processes is to compare the species list generated by different pipelines when using the same reference database. This has yet to be done for the variety of pipelines used in molecular assessment of diatom communities in Europe.

Finally, the limitation of taxonomic resolution (when identification at species level was not possible) can also lower the quality of the ecological assessment using the IPS index because the genus level includes the pooling of closely related species, some of which can exhibit very different ecological preferences (Keck et al. 2018). However, even if the Rsyst::diatom database (rbcL and 18S V4) contains a number of reference barcodes identified only at genus level, we judge that this factor had only a relatively small impact on the ecological assessment because only about 10% of our taxa were assigned to genus level (rbcL and 18S markers).

The metabarcoding detect additional species

Naturally, taxonomic discrepancies can also arise from the morphological assessment, with the possible misidentification of small forms, as well as omission of rare species. For example, *Entomoneis* sp.'s silica skeleton is easily dissolved in the routine process of morphological diatom identification using oxidised samples. This omission of species in morphological identification creates species list discrepancies when comparing with the metabarcoding method. Moreover, a higher number of identified taxa with metabarcoding is expected because of the high number of sequences taken into account compared to only a 400 valves count with light microscopy (Zimmermann et al. 2015). However, if the species detected by metabarcoding is missing from the taxonomy list of Fennoscandia species (Kahlert et al. 2017) and, thus, has no associated IPS values (H codes in Fig. 3), it will be accounted for as unidentified species and not used for index calculation until the list is updated and completed with indices. This specific scenario represented only 8% of the identifications with the rbcL marker and 6% with the 18S marker. However, even if unidentified species have a stronger impact on Shannon scores than on IPS scores, when the majority of the reads in a sample are unidentified, the ecological assessment becomes unreliable. In this way, the molecular method can be used to detect taxa that are currently missing from the ecological assessment based on IPS calculations.

Not only the presence of species had an impact on the IPS scores, species abundances were also affecting the index calculations

We found that, even if abundance discrepancies occurred less often than presence/absence disparities, they affect the ecological assessment much more than any other type of mismatch. Especially when using the rbcL marker, the species mainly responsible for overestimation and underestimation of the IPS scores were found with significantly higher abundance in one or the other type of assessment (e.g. the species *Achnanthidium minutissimum* and *Tabellaria flocculosa*).

One explanation of a mismatch in relative abundance might be the known problem that the number of generated sequences by HTS does not directly correspond to the number of specimens or biomass (Pawlowski et al. 2018; Zimmermann et al. 2015) and that different species can produce different amount of reads, e.g. due to differences in the chloroplast size with the rbcL marker. However, even if relative and absolute abundance in valve count or in read numbers are difficult to compare directly, this approach has been successfully used previously (Kermarrec et al. 2014) and we think that our choice of comparing "low" and "high" abundances using a threshold, based on experience, effectively allowed us to highlight which species are most often found with abundance discrepancies.

This specific limitation of the metabarcoding method has few known solutions. However, the SIMPER analysis which we performed to assess which species accounted for most of the observed IPS scores deviations, highlighted several problematic species such as *Achnanthidium minutissimum*, *Cyclotella meneghiniana*, *Nitzschia palea* and *Ulnaria ulna*. These species are known for significant abundance discrepancies when assessed by morphological or metabarcoding methods: they were also found to be either under-represented or over-represented in the study by Vasselon et al. (2018) when using the rbcL barcode and were included in a biovolume correction factor. For example, *A. minutissimum* has a small biovolume and, thus, will generate less copies of the rbcL fragment (located in the chloroplast) than larger species. As a result, this taxon is less abundant in molecular assessments than in morphological ones, which will lead to underestimation of the IPS score. Applying Vasselon et al. (2018)'s biovolume correction factor might thus improve the ecological assessment with metabarcoding on the rbcL marker in Fennoscandia. Similarly, a variable number of gene copies have also been observed when using the 18S-V4 marker (Godhe et al. 2008).

Another source of abundance discrepancy is the possible assignation to a similar reference barcode when the correct taxon is not represented in the database. As mentioned previously, closely related species can have different ecological preferences, such as shown for *Halamphora veneta* and *H. oligotraphenta* by Keck et al. (2018). However, the strictness of Diagnosyst taxonomic assignment should limit this type of error: a sequence assigned to several reference taxa will be discarded, thus, an erroneous assignation could only happen in the event of a very close match to a single similar reference barcode.

Additionally, the cryptic diversity can create abundance disparities coupled with presence/absence discrepancies: where limited morphological identification under a light-microscope may result in assignment to a single taxon, metabarcoding works at a finer taxonomic level and can split the specimen into several taxa. In that case, we obtain a higher richness of species, at lower abundance, with the molecular technique. A similar problem arises if one of the method's identification stops at genus level, when the other method splits the identification into several species. With the morphological assessment under a light-microscope, a genus level taxonomic identifi-

cation may result from lack of taxonomic expertise or too small specimens. When assessed with the metabarcoding method, a limited taxonomic level may be due to a lack of higher level reference barcodes or to low primer specificity. The former case was more represented in our dataset but still rarely occurring.

Different markers give different ecological status assessment

The DNA barcodes rbcL and 18S-V4 were chosen in this study because of their power to discriminate diatom communities, covering the three major diatom divisions and for their balance between variability and conservation of the primer binding sites (Kermarrec et al. 2014; Zimmermann et al. 2011). The V4 region is the largest and most complex of the highly variables region in the 18S locus (Zimmermann et al. 2011) but the rbcL barcode has a higher polymorphism than the 18S one.

When looking at the species lists, the Shannon scores show greater deviations from the one calculated on morphological communities when using the 18S marker than when using the rbcL marker. The presence of green algae barcodes in the 18S dataset, which are completely absent from the rbcL dataset, is most likely responsible for that trend and a greater difference was found between the Shannon scores of the two markers in the molecular analysis, rather than between the two methods. The rbcL marker also had a better proportion of "good match" between the species lists generated by the morphological and by the molecular methods.

The inverse trend is observed for the ecological assessment: even though the results were very similar between the two markers and, in both cases, significantly correlated to the morphological method, the 18S marker achieves more exact ecological status classes than the rbcL marker, as well as less underestimation. Furthermore, a greater deviation was found between the IPS scores calculated with the rbcL marker and the morphological communities than between the IPS scores calculated on the 18S marker and morphological communities. Additionally, the IPS scores generated by the two markers were significantly different but less than when compared to the morphological assessment, highlighting that both markers produce more similar results than expected.

As mentioned before, the abundance discrepancies were present when using the 18S marker but less important than with the rbcL marker. Indeed, the majority of the species strongly contributing to the IPS deviations with the 18S were actually reflecting presence/absence discrepancies rather than abundances. Contrary to the rbcL marker, which exhibits a clear correlation between the species biovolume and the gene copy number (Vasselon et al. 2018), fluctuations of gene copy number for the 18S marker does not seem to be species specific. However, no study has tried to untangle the reasons behind it yet. The number of molecular data, generated by a known number of valves, should be assessed for the 18S-V4 marker, similarly to that which Vasselon et al. (2018) did on the rbcL marker. The presence/absence dissimilarities observed in our study are explained by the lack of a few key species. For example, *T. flocculosa* failed to be detected when using the 18S marker, despite the fact that this species is represented by two reference barcodes in the 18S-V4 database. The species is detected by the rbcL marker, so one could speculate that the Nordic *T. flocculosa* is not well amplified when using the 18S-V4 or that its 18S-V4 DNA fragment does not match the reference barcodes available.

While this study confirms the strong impact of the reference barcodes available, as much in quantity as in quality and some marker-specific difficulties, we cannot efficiently recommend one or the other marker. The 18S-V4 seems to have promising efficiency in ecological assessment and covered more of the Fennoscandian taxa morphologically identified at the time this study was undertaken, whereas the rbcL marker generated species lists more similar to the ones generated by the morphological approach. Indeed, the rbcL had more good-matches between species lists and better-correlated Shannon scores. However, its abundance discrepancies affected the IPS calculation more than the 18S ones.

Additionally, part of our results has facilitated the recent curation work on the rbcL marker reference barcodes, which greatly improved the quality of the database. No similar curation has yet been done for the 18S marker.

Conclusion

Overall, our findings that the metabarcoding method in general is also suitable for Northern European conditions are promising. However, based on our results, we are convinced that there is a need for further development of this method for the use for environmental assessment in Fennoscandia. The limitations of both techniques are multiple and correlated, making them difficult to isolate and properly quantify. Still, based on our results, we would recommend focusing first on the completion and maintenance of the reference databases, adding important missing species and carefully curating them to remove ambiguous barcodes and widen their use to broader ecosystems. Next, the abundance discrepancies were not the most common error source but clearly the one that mostly affected the ecological assessment. Thus, it would be interesting to adapt the Vasselon et al. (2018) biovolume correction factor to Fennoscandian diatoms' biovolumes and apply it to our dataset for the rbcL marker, to see if we could achieve an improvement for the ecological assessment.

Additionally, the great diversity of bioinformatics pipelines, currently available for diatoms' metabarcoding, poses another challenge. Which pipelines are currently being used by the different research groups dealing with diatom metabarcoding development and the way they affect the molecular data and ecological assessment need to be evaluated, perhaps as a first step toward a standardisation of the molecular process.

Finally, the current Fennoscandian way to set an ecological status class is based on the morphological method and the next step should be to focus on the integration of metabarcoding data into current ecological assessment methods, as recommended by Pawlowski et al. (2018). It has also been suggested that new indices should be developed directly on molecular data (Keck et al. 2017; Pawlowski et al. 2018) and our dataset of 180 samples, which include water chemistry data, could be used to test and possibly calibrate new indices currently being developed.

Acknowledgements

This project was funded by Stiftelsen Oscar och Lili Lamms Minne (http://www.stiftelsenlamm.a.se/) and by The Swedish Agency for Marine and Water management. We would like to thank Jón S. Ólafsson, from the Marine and Freshwater Resarch Institute (Iceland), for the collection of data from Iceland and for his helpful feedback. We also thank the Norwegian Institute for Water Research (NIVA), which provided the water chemistry data for all Norwegian samples. We would like to thank Teofana Chonova, Meline Corniquel and Sonia Lacroix who performed the preparation of DNA libraries for all samples at the molecular laboratory of INRA CARRTEL in Thonon (France) and The Plateforme Genome Transcriptome of INRA BIOGECO in Pierroton (France) for the HTS sequencing. Computer time for this study was provided by the computing facilities MCIA (Mésocentre de Calcul Intensif Aquitain) of the Université de Bordeaux and of the Université de Pau et des Pays de l'Adour and using the PlaFRIM experimental testbed, supported by Inria, CNRS (LABRI and IMB), Université de Bordeaux, Bordeaux INP and Conseil Régional d'Aquitaine (see https://www.plafrim.fr/). This work was supported by the European COST-Action DNAqua Net (CA15219).

References

Apothéloz-Perret-Gentil L, Cordonier A, Straub F, Iseli J, Esling P, Pawlowski J (2017) Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. Molecular Ecology Resources 17: 1231–1242. https://doi.org/10.1111/1755-0998.12668

Bauer DF (1972) Constructing Confidence Sets Using Rank Statistics. Journal of the American Statistical Association 67: 687–690. https://doi.org/10.2307/2284469

Cemagref (1982) Étude des méthodes biologiques quantitative d'appréciation de la qualité des eaux. Bassin Rhône-Méditerranée-Corse. Centre National du Machinisme Agricole du Génie rural des Eaux et des Forêts, Lyon, France.

CEN (2014) EN 13946:2014 Water quality – Guidance for the routine sampling and preparation of benthic diatoms from rivers and lakes CEN.

Chambers RG, Pope RD (1992) Engels Law and Linear-in-Moments Aggregation. American Journal of Agricultural Economics 74: 682–688. https://doi.org/10.2307/1242581

Chaumeil P, Fischer-Le Saux M, Frigerio J-M, Grenier E, Rimet F, Streito J-C, Laval V, Franc A (2018) R-Syst: a network providing curated molecular databases and data analysis tools for taxonomy and systematics (Prokaryotes and Eucaryotes). Portail Data Inra.

Clarke KR (1993) Nonparametric Multivariate Analyses of Changes in Community Structure. Australian Journal of Ecology 18: 117–143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

Elbrecht V, Leese F (2015) Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass-Sequence Relationships with an Innovative Metabarcoding Protocol. Plos One 10. https://doi.org/10.1371/journal.pone.0130324

Erlandsson M, Buffam I, Folster J, Laudon H, Temnerud J, Weyhenmeyer GA, Bishop K (2008) Thirty-five years of synchrony in the organic matter concentrations of Swedish rivers explained by variation in flow and sulphate. Global Change Biology 14: 1191–1198. https://doi.org/10.1111/j.1365-2486.2008.01551.x

Friberg N, Dybkjaer JB, Olafsson JS, Gislason GM, Larsen SE, Lauridsen TL (2009) Relationships between structure and function in streams contrasting in temperature. Freshwater Biology 54: 2051–2068. https://doi.org/10.1111/j.1365-2427.2009.02234.x

Frigerio JM, Rimet F, Bouchez A, Chancerel E, Chaumeil P, Salin F, Thérond S, Kahlert M, Franc A (2016) Diagno-syst: a tool for accurate inventories in metabarcoding. https://arxiv.org/abs/1611.09410.

Godhe A, Asplund ME, Harnstrom K, Saravanan V, Tyagi A, Karunasagar I (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

Gottschalk S, Kahlert M (2012) Shifts in taxonomical and guild composition of littoral diatom assemblages along environmental gradients. Hydrobiologia 694: 41–56. https://doi.org/10.1007/s10750-012-1128-7

Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos. Plos One 6. https://doi.org/10.1371/journal.pone.0017497

Hammer O, Harper D, Ryan P (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4: 1–9. https://palaeo-electronica.org/2001_1/past/past.pdf

Hollander M, Wolfe DA (1973) Nonparametric statistical methods. John Wiley & Sons, New York. https://doi.org/10.1002/9781119196037

Jarlman A, Kahlert M, Sundberg I, Herlitz E (2016) Påväxt i sjöar och vattendrag – kiselalgsanalys. Version 3:2: 2016-01-20. Handledning för miljöövervakning Undersökningstyp. Havs- och Vattenmyndigheten, Göteborg, 24 pp.

Johansson H, Persson G (2001) Svenska sjöar med höga fosforhalter. 790 naturligt eutrofa eller eutrofierade sjöar? Miljo¨analys Ifr Rapport 2001: 8.

Kahlert M, Albert R-L, Anttila E-L, Bengtsson R, Bigler C, Eskola T, Gälman V, Gottschalk S, Herlitz E, Jarlman A, Kasperoviciene J, Kokociński M, Luup H, Miettinen J, Paunksnyte I, Piirsoo K, Quintana I, Raunio J, Sandell B, Simola H, Sundberg I, Vilbaste S, Weckström J (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

Kahlert M, Kelly M, Albert R-L, Almeida SFP, Bešta T, Blanco S, Coste M, Denys L, Ector L, Fránková M, Hlúbiková D, Ivanov P, Kennedy B, Marvan P, Mertens A, Miettinen J, Picinska-Fałtynowicz

J, Rosebery J, Tornés E, Vilbaste S, Vogel A (2012) Identification versus counting protocols as sources of uncertainty in diatom-based ecological status assessments. Hydrobiologia 695: 109–124. https://doi.org/10.1007/s10750-012-1115-z

Keck F, Vasselon V, Rimet F, Bouchez A, Kahlert M (2018) Boosting DNA metabarcoding for biomonitoring with phylogenetic estimation of operational taxonomic units' ecological profiles. Molecular Ecology Resources 18: 1299–1309. https://doi.org/10.1111/1755-0998.12919

Keck F, Vasselon V, Tapolczai K, Rimet F, Bouchez A (2017) Freshwater biomonitoring in the Information Age. Frontiers in Ecology and the Environment 15: 266–274. https://doi.org/10.1002/fee.1490

Kelly M, Bennett C, Coste M, Delgado C, Delmas F, Denys L, Ector L, Fauville C, Ferreol M, Golub M, Jarlman A, Kahlert M, Lucey J, Ni Chathain B, Pardo I, Pfister P, Picinska-Faltynowicz J, Rosebery J, Schranz C, Schaumburg J, van Dam H, Vilbaste S (2009) A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. Hydrobiologia 621: 169–182. https://doi.org/10.1007/s10750-008-9641-4

Kelly M, Boonham N, Juggins S, Kille P, Mann D, Pass D, Sapp M, Sato S, Glover R (2018) A DNA based diatom metabarcoding approach for Water Framework Directive classification of rivers. https://www.researchgate.net/profile/David_Mann6/publication/323960290_A_DNA_based_diatom_metabarcoding_approach_for_Water_Framework_Directive_classification_of_rivers/links/5ab4ec7e45851515f59965be/A-DNA-based-diatom-metabarcoding-approach-for-Water-Framework-Directive-classification-of-rivers.pdf?origin=publication_detail

Kelly M, Urbanic G, Acs E, Bennion H, Bertrin V, Burgess A, Denys L, Gottschalk S, Kahlert M, Karjalainen SM, Kennedy B, Kosi G, Marchetto A, Morin S, Picinska-Faltynowicz J, Poikane S, Rosebery J, Schoenfelder I, Schoenfelder J, Varbiro G (2014) Comparing aspirations: intercalibration of ecological status concepts across European lakes for littoral diatoms. Hydrobiologia 734: 125–141. https://doi.org/10.1007/s10750-014-1874-9

Kermarrec L, Franc A, Rimet F, Chaumeil P, Frigerio JM, Humbert JF, Bouchez A (2014) A next-generation sequencing approach to river biomonitoring using benthic diatoms. Freshwater Science 33: 349–363. https://doi.org/10.1086/675079

Kermarrec L, Franc A, Rimet F, Chaumeil P, Humbert JF, Bouchez A (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

Lecointe C, Coste M, Prygiel J (1993) Omnidia – Software for Taxonomy, Calculation of Diatom Indexes and Inventories Management. Hydrobiologia 269: 509–513. https://doi.org/10.1007/BF00028048

Ólafsson JS, Ingimundardóttir GV, Hansen I, Sigurðardóttir SG (2010) Smádýralíf í afrennslisvatni frá háhitasvæðunum við Kröflu, Ölkelduháls og í Miðdal í Henglinum. Veiðimálastofnun, Náttúrufræðistofnun Íslands, Náttúrustofa Norðausturlands og Líffræðistofnun Háskólans, Reykjavík. https://orkustofnun.is/gogn/Orkusjodur/Orkusjodur-08-Smadyralif-i-afrennslisvani-fra-hahitasvaedum.pdf

Pawlowski J, Kelly-Quinn M, Altermatt F, Apotheloz-Perret-Gentil L, Beja P, Boggero A, Borja A, Bouchez A, Cordier T, Domaizon I, Feio MJ, Filipe AF, Fornaroli R, Graf W, Herder J, van der Hoorn B, Jones JI, Sagova-Mareckova M, Moritz C, Barquin J, Piggott JJ, Pinna M, Rimet F, Rinkevich B, Sousa-Santos C, Specchia V, Trobajo R, Vasselon V, Vitecek S, Zimmerman J, Weigand A, Leese F, Kahlert M (2018) The future of biotic indices in the ecogenomic era: Integrating (e) DNA metabarcoding in biological assessment of aquatic ecosystems. Science of the Total Environment 637: 1295–1310. https://doi.org/10.1016/j.scitotenv.2018.05.002

R-CoreTeam (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Ramsay W (1898) Über die Geologische Entwicklung der Halbinsel Kola in der Quartärzeit, 151 pp.

Rimet F, Bouchez A (2012) Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. Knowl Managt Aquatic Ecosyst: 01. https://doi.org/10.1051/kmae/2012018

Rimet F, Chaumeil P, Keck F, Kermarrec L, Vasselon V, Kahlert M, Franc A, Bouchez A (2016) R-Syst::diatom: an open-access and curated barcode database for diatoms and freshwater monitoring. Database-the Journal of Biological Databases and Curation. https://doi.org/10.1093/database/baw016

Rimet F, Vasselon V, A-Keszte B, Bouchez A (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 SF, Vasselon V, Jacquet S, Bouchez A, Ariztegui D, Rimet F (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

Smol JP, Stoermer EF (2010) The diatoms: applications for the environmental and earth sciences. Cambridge University Press, 667 pp. https://doi.org/10.1017/CBO9780511763175

Stein ED, Martinez MC, Stiles S, Miller PE, Zakharov EV (2014) Is DNA Barcoding Actually Cheaper and Faster than Traditional Morphological Methods: Results from a Survey of Freshwater Bioassessment Efforts in the United States? Plos One 9. https://doi.org/10.1371/journal.pone.0095525

Stein ED, White BP, Mazor RD, Miller PE, Pilgrim EM (2013) Evaluating Ethanol-based Sample Preservation to Facilitate Use of DNA Barcoding in Routine Freshwater Biomonitoring Programs Using Benthic Macroinvertebrates. Plos One 8. https://doi.org/10.1371/journal.pone.0051273

Student (1908) The Probable Error of a Mean. Biometrika 6: 1–25. https://doi.org/10.1093/biomet/6.1.1

Swedish Environmental Protection Agency (2007) Bedömningsgrunder för sjöar och vattendrag. Bilaga Bilaga A Till Handbok 2007:4: Handbok för miljöövervakning: Programområde: Sötvatten. Retrieved from Stockholm. http://www.naturvardsverket.se/Documents/publikationer/620-0148-3.pdf

Taxalista (2018) kiselalger i svenska sötvatten. http://miljodata.slu.se/mvm/Content/Static/Current/Kiselalger%20i%20svenska%20sötvatten.xlsx [accessed 18 July.2018]

Vasselon V, Bouchez A, Rimet F, Jacquet S, Trobajo R, Corniquel M, Tapolczai K, Domaizon I (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, Domaizon I, Rimet F, Kahlert M, Bouchez A (2017a) 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

Vasselon V, Rimet F, Tapolczai K, Bouchez A (2017b) Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). Ecological Indicators 82: 1–12. https://doi.org/10.1016/j.ecolind.2017.06.024

Visco JA, Apotheloz-Perret-Gentil L, Cordonier A, Esling P, Pillet L, Pawlowski J (2015) Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. Environmental Science & Technology 49: 7597–7605. https://doi.org/10.1021/es506158m

Zabell SL (2008) On Student's 1908 article – "The Probable Error of a Mean". Journal of the American Statistical Association 103: 1–7. https://doi.org/10.1198/016214508000000030

Zhan AB, Hulak M, Sylvester F, Huang X, Adebayo AA, Abbott CL, Adamowicz SJ, Heath DD, Cristescu ME, MacIsaac HJ (2013) High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods in Ecology and Evolution 4: 558–565. https://doi.org/10.1111/2041-210X.12037

Zimmermann J, Abarca N, Enk N, Skibbe O, Kusber WH, Jahn R (2014) Taxonomic Reference Libraries for Environmental Barcoding: A Best Practice Example from Diatom Research. Plos One 9. https://doi.org/10.1371/journal.pone.0108793

Zimmermann J, Glockner G, Jahn R, Enke N, Gemeinholzer B (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, Jahn R, Gemeinholzer B (2011) Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. Organisms Diversity & Evolution 11: 173–192. https://doi.org/10.1007/s13127-011-0050-6