

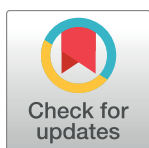
RESEARCH ARTICLE

The best of both worlds: A combined approach for analyzing microalgal diversity via metabarcoding and morphology-based methods

Sophie Groendahl^{1*}, Maria Kahlert², Patrick Fink^{1,3}

1 University of Cologne, Zoological Institute, Aquatic Chemical Ecology, Cologne, Germany, **2** Swedish University of Agricultural Sciences, Department of Aquatic Sciences and Assessment, Uppsala, Sweden, **3** University of Duesseldorf, Institute for Cell Biology and Zoology, Duesseldorf, Germany

* sgroenda@uni-koeln.de



Abstract

An increasing number of studies use next generation sequencing (NGS) to analyze complex communities, but is the method sensitive enough when it comes to identification and quantification of species? We compared NGS with morphology-based identification methods in an analysis of microalgal (periphyton) communities. We conducted a mesocosm experiment in which we allowed two benthic grazer species to feed upon benthic biofilms, which resulted in altered periphyton communities. Morphology-based identification and 454 (Roche) pyrosequencing of the V4 region in the small ribosomal unit (18S) rDNA gene were used to investigate the community change caused by grazing. Both the NGS-based data and the morphology-based method detected a marked shift in the biofilm composition, though the two methods varied strongly in their abilities to detect and quantify specific taxa, and neither method was able to detect all species in the biofilms. For quantitative analysis, we therefore recommend using both metabarcoding and microscopic identification when assessing the community composition of eukaryotic microorganisms.

OPEN ACCESS

Citation: Groendahl S, Kahlert M, Fink P (2017) The best of both worlds: A combined approach for analyzing microalgal diversity via metabarcoding and morphology-based methods. PLoS ONE 12(2): e0172808. doi:10.1371/journal.pone.0172808

Editor: Hideyuki Doi, University of Hyogo, JAPAN

Received: November 10, 2016

Accepted: February 9, 2017

Published: February 24, 2017

Copyright: © 2017 Groendahl et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data are available from the Dryad database (DOI: [10.5061/dryad.tb923](https://doi.org/10.5061/dryad.tb923)).

Funding: This study was funded by the Deutsche Forschungsgemeinschaft (www.dfg.de), grant FI 1548/5-1 to PF. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

During the last century biodiversity has declined due to anthropogenic influences [1]. This results in reduced ecosystem stability, functioning and provision of ecosystem services [2–4]. While loss of biodiversity in larger organisms is well studied [1], we know far less about the effects of biodiversity loss for microorganisms such as algae. As microalgae are the base of aquatic food webs [5, 6], a reduction in algal diversity may have great repercussions for higher trophic levels. During the last century, most studies on the diversity, distribution, and abundance of algae were based on morphological characteristics e.g. [7–9]. However, this is a challenging and time-consuming method, since microalgal samples typically contain many cryptic, small, and rare species. Therefore considerable taxonomic expertise are required, which are unfortunately becoming increasingly rare [10]. The global number of algal species is estimated to be over a million species [11]. A conservative approach anticipated 72,500 algal

species, of which only 60% have been described to date [11]. In order to quantify and to reduce the decline in biodiversity, we need more precise assessments of the current biodiversity. A newly developed method, called DNA metabarcoding, via next-generation sequencing (NGS), enables the rapid identification of species in environmental samples. Metabarcoding uses short gene sequences in order to identify species. It proven to be fast and cost effective [12, 13], making it possible to track and measure biodiversity over vast areas and time spans [14, 15]. One of the great advantages of metabarcoding is that rare species may be detected with greater sensitivity [16]. This is one of the reasons why NGS-based methods are being increasingly used to detect invasive species [17]. Moreover, DNA metabarcoding allows for identification of species that cannot be discriminated based on morphological features. However, we currently know little about the reliability of this method, and as yet there are no clear methodological guidelines regarding how it should be used. Therefore, in order to evaluate how accurate NGS is when it comes to estimating microorganism diversity, comparisons with morphology-based analysis need to be made. Many studies use pre-defined communities in order to verify their data e.g. [18–20]. The usage of mock communities may indeed improve the metabarcoding data validation process, but it has some shortcomings. Most mock communities often contain a lower number of species than field samples, and the quality of the DNA is often higher. Only a few rigorous studies have compared molecular versus morphology-based methods using samples from the same sampling sites, particularly with regard to microalgal diversity [21, 22]. Furthermore, the results from these studies are inconclusive. In a paper by Abdad et al. [21] similar spatial and temporal trends of taxonomic diversity were found for metabarcoding and microscopic studies of zooplankton, but not for phytoplankton. Moreover, they found that DNA metabarcoding could have the potential for semi-quantitative assessment of organisms' abundances as well. Conversely, Zimmermann et al. [22] found no correlation between the number of reads obtained by NGS sequencing and the number of cells counted in a light microscope.

We here describe a mesocosm experiment to investigate the potential of using DNA metabarcoding to gain primary producer diversity and composition estimates. Previous studies have demonstrated that consumer species richness may increase the number of primary producer species due to complementary effects [23, 24]. We therefore designed the mesocosm experiment to test how consumer species richness influences primary producer species richness. A clear alteration of primary-producer community composition due to consumer species grazing is expected to be reflected by both morphological and molecular methods.

We allowed two species of benthic grazers (a gastropod and a mayfly larva) belonging to two different functional feeding groups, to feed upon natural, benthic microalgae-dominated biofilms (periphyton) over a period of 50 days. We sampled the biofilms for molecular (Roche 454 sequencing) and morphology-based analyses, and compared the abundances and diversity estimates generated from the two methods. We find great discrepancies between the methods; the NGS approach provided us with valuable information regarding species richness, whereas the morphology-based method provided species abundance and biodiversity measurements.

Material and methods

The experiment was conducted in freshwater mesocosms on campus at the University of Cologne, NW Germany. Algal biofilm was pre-grown on tiles (4.7 x 4.7 cm) for two months in tanks with WC medium [25] based on an inoculum from a nearby pond. Two months later, when a thick algal biofilm had developed, the tiles were divided between 12 ten-liter buckets (two tiles in each bucket), each filled with 5 L of water. A net was placed over the buckets to prevent other potential consumer species from feeding upon the algae. The treatments

consisted of a control treatment without consumer species, a treatment with juvenile *Lymnaea stagnalis* gastropods, a treatment with *Cloeon dipterum* mayfly larvae, and a treatment with both consumer species; each replicated threefold (S1 Table). The consumer species were collected from the same pond from which the water was taken. The consumer species were not endangered nor were they legally protected. All conditions for animal maintenance and experiments were carefully optimized to ensure that they met the animals' needs. A specific ethical approval by the university's Institutional Animal Care and Use Committee is not required for work with invertebrates according to German law. Regardless, we undertook all necessary measures to minimize animal suffering and followed guidelines for the use of animals in research and teaching activities [26]. We estimated the fresh weight of the consumer species to ensure that it did not differ between treatments. The experimental units were checked daily for emerging *C. dipterum*. In case of emergence, *C. dipterum* larvae were replaced. The animals were weighed every three days to ensure the same fresh weight (± 5 mg) in all treatments. If the fresh mass ratios deviated, animals were either removed or added to ensure the same fresh mass of consumers in each enclosure. After 50 days, the experiment was terminated. Samples were taken for dry mass, as well as for morphological and molecular determination of algal community composition.

DNA extraction, PCR and sequencing

DNA was isolated using DNeasy Blood and Tissue Kits (Qiagen, Netherlands, S1 Fig). 200 μ l of buffer ATL and 20 μ l of proteinase K were added to the samples. The samples were then homogenized with Minilys (Bertin instruments, France) using Precellys ceramic kit 1.4/2.8 mm (Bertin instruments, France) for 2 x 20 sec at 5000 rpm. The samples were incubated for 10 min at 56°C in a thermomixer and then centrifuged for 3 min at 17000 x g. The subsequent steps were conducted in accordance with the instructions of the manufacturer of the kit. The sample was incubated for 1 min at room temperature and then centrifuged for 1 min at 6000 x g. The DNA was stored at -20°C until further processing.

The DNA concentration was estimated and purity was verified in a NanoDrop spectrophotometer (Implen, Germany) at 260 and 280 nm. DNA purity and quantity were also verified by electrophoresis in 1.2% agarose gels on 1 x TAE buffer stained with 1 x GelRedTM (Biotium Inc., USA). The samples' DNA content was adjusted to 25 ± 3 ng/ μ l prior to PCR amplification. The hypervariable V4 region of the 18S of the rRNA genes was amplified using the forward primer 5' -AATTCCAGCTCCAATAGCGTATAT-3' and the reverse primer 5' -TTTCA GCCTTGCGACCATAC-3'. The primer pair was designed to ensure amplification of both green algae and diatoms. 500 random algae 18S rDNA sequences were aligned in Geneious [27], after which Primer3 [28] was used to find the best primer sites. NetPrimer (PREMIER Biosoft International, Palo Alto, CA) was used to further optimize the design of the primer pair. To ensure sample recognition in downstream analyses, the samples were amplified with tagged primers—the forward primer was tagged specifically for each sample using unique 10 bp tags (MID sequences). For all primers, the GS FLX Shotgun adaptors were attached in order to make them compatible with pyrosequencing procedures. The 25 μ l PCR assay comprised of 25 ng of DNA, 0.02 U/ μ l Phusion Green Hot Start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific, USA), 0.5 μ M of each primer, 5 μ l of 5x Phusion HF Buffer (Thermo Fisher Scientific, USA), 200 mM dNTP mix (VWR International, USA) and 15.75 μ l of ultrapure water. PCR amplifications were carried out using FlexCycler (Analytik Jena, Germany). PCR cycling parameters consisted of an initial denaturation step at 98°C for 30 s, followed by 30 amplification cycles of 98°C for 10 s, 59°C for 30 s, 72°C for 30 s and 10 min final extension at 72°C. In order to reduce the effect of PCR biases that may have occurred in any given reaction,

each of the samples were PCR-amplified three times. The PCR products were purified using the GenElute™ PCR Clean-Up Kit (Sigma-Aldrich, USA) following the manufacturer's protocol. The DNA concentration of the samples was again estimated using a NanoDrop spectrophotometer, adjusted to 25 ± 3 ng/μl and pooled. The ready-to-load library was then sent to GATC Biotech (Germany) for 454 GS FLX paired-end sequencing (Roche Applied Science).

Sequence analysis

We amplified approximately 500–600 bp within the 18S gene, including the hypervariable V4 region [29, 30] (S1 Fig). Mothur v. 1.34.4 [31] was used to filter the raw data. First the FASTQ files were converted to FASTA and QUAL files, and the sequences were trimmed using the trim.flows command in Mothur (pdiffs = 0, bdiffs = 0). After this step, all sequences shorter than 200 bp were omitted (mintlength = 200), together with sequences with homopolymers longer than 8 bp (maxhomop = 8). No ambiguous base calls were allowed (maxambig = 0). Approximately 70% of all sequences remained after these steps. The resulting reads were dereplicated (collapsed to unique sequences) using the unique.seqs command and aligned to the SILVA reference alignment, v. 123, [32] with the align.seqs command. We then ran the screen.seqs command in Mothur to ensure that all sequences overlapped in the same alignment space. Approximately 20% of all sequences remained after this step. Thereafter, all sequences that are within 2 bp of a more abundant sequence were clustered. Chimeras were detected using UCHIME [33] with the chimera.uchime command. OTUs (operational taxonomic units) were built using the dist.seqs command in Mothur (cutoff = 0.15). Reads were clustered with sequence divergence threshold at 1% with the cluster command and singletons were discarded. For taxonomic annotation, the resulting sequences were imported into Geneious [27], where we performed a local MegaBLAST [34] search of each OTU versus a local DNA reference library with gap costs set to linear and with a match score of 1 and mismatch score of -2. The local reference databases were constructed based on the PR² database [35]. The best MegaBLAST hit against our local database was used to classify each sequence, and a positive identification was defined as a hit with at least 90% identity and 100% query coverage. The OTUs were not assigned to the closest hit in the database; we instead based the final taxonomic determination upon the pairwise identity to the reference sequence to avoid erroneous identification. Reads with a 99, 97 and 90% pairwise identity were assigned to species, genus, or family, respectively [36]. The rarefaction curves were constructed using the rarefaction.single command (calc = sobs, freq = 100). To construct the phylogenetic tree, we used the dist.seqs command in Mothur. With the dist.seqs command the uncorrected pairwise distances between aligned DNA sequences are calculated. By default, a gap is penalized once and terminal gaps are penalized. We then ran the clearcut program (<http://bioinformatics.hungry.com/clearcut/>) within Mothur, using the clearcut command. Clearcut required the distance matrix created by the dist.seqs command. The clearcut program uses relaxed neighbor joining (RNJ) algorithm [37] when constructing phylogenetic trees. By running the clearcut command a file called abrecovery.tre is generated. Finally, we uploaded the abrecovery.tre file to the Interactive Tree of Life (iTOL) tool [38] in order to illustrate the taxonomic community compositions via a phylogenetic tree.

Morphology-based analysis

Benthic algae were counted in three subsamples of each Lugol-preserved sample (diluted 1:32 with tap water) using an inverted microscope at a magnification of 400–1000x. In order to obtain the total biovolume of the species in each sample, the average number of cells per taxon was multiplied with the biovolume estimates, based on the typical cell morphology [39]. The

taxonomic identity was determined to the lowest possible level; some taxa were grouped into non-taxonomical groups where the method did not allow for higher taxonomic resolution.

Statistics

Before the statistical tests, all data were checked for homoscedasticity using Levene's test. All statistical tests were conducted in SigmaPlot (v.11, SysStat). A one-way ANOVA was conducted to test for differences in algal dry mass between the grazing treatments, with algae dry mass as dependent variable and the grazing treatments as independent variable. A two-way ANOVA tested for the grazing effect upon single taxa followed by *post-hoc* comparisons with Tukey's HSD. Here the treatment and algae species were independent variables and the number of cells was dependent variable. The data were ranked in MS Excel prior to the two-way ANOVA due to heteroscedasticity. Linear regressions were conducted to test for correlations between abundances of single taxa as determined by both methods.

Results

The morphology-based approach

18 primary producer taxa were identified microscopically (Table 1): 2 were identified to species level and 11 to genus level. We found one euglenoid, two cyanobacteria, two diatom, and twelve green algal taxa. The biovolume per cell differed strongly between taxa, ranging from 5 to 11,665 μm^3 per cell.

Based on cell numbers, *Pseudanabena* sp. was the dominant primary producer, while *Cosmarium* sp. was dominant in terms of biovolume due to their large cell size (S2 Fig; Table 1).

The molecular approach

We obtained 139,702 sequences, of which 15,662 sequences remained after quality filtering. Clustering at 99% identity yielded 195 OTUs (Fig 1). The number of reads per sample after

Table 1. Primary producers found in the biofilm communities determined under the light microscope, given together with the geometric shapes of their cells and the cell-specific biovolume as calculated on basis of geometric figures according to [39].

Algae groups	Taxa	Geometric shape approximation	Average biovolume (μm^3)
Cyanobacteria	<i>Anabaena</i> sp.	Cylinder, two half spheres	85
Cyanobacteria	<i>Pseudanabena</i> , cf	Cylinder, two half spheres	10
Diatoms	Diatoms, ribbon colonies	Box	260
Diatoms	Diatoms, solitary	Prism on elliptic base	150
Green algae	<i>Chlamydomonas</i> , cf	Prolate spheroid	675
Green algae	<i>Closterium</i> sp.	Double cone	7070
Green algae	Coccoid green algae	Sphere	60
Green algae	<i>Cosmarium</i> sp.	Double spheroid	11665
Green algae	<i>Monoraphidium</i> sp.	Double cone	75
Green algae	<i>Oedogonium</i> sp.	Cylinder	435
Green algae	<i>Ooystis</i> sp.	Prolate spheroid	250
Green algae	<i>Pediastrum</i> sp.	Box	310
Green algae	<i>Scenedesmus</i> sp., two celled colonies	Prolate spheroid	35
Green algae	<i>Scenedesmus</i> sp., four celled colonies	Prolate spheroid	9
Green algae	<i>Selenastrum</i> sp.	Double cone	5
Green algae	<i>Tetraedon caudatum</i>	Box	65
Green algae	<i>Tetraedon minimum</i>	Box	185
Euglenoids	<i>Trachelomonas</i> sp.	Sphere	2145

doi:10.1371/journal.pone.0172808.t001

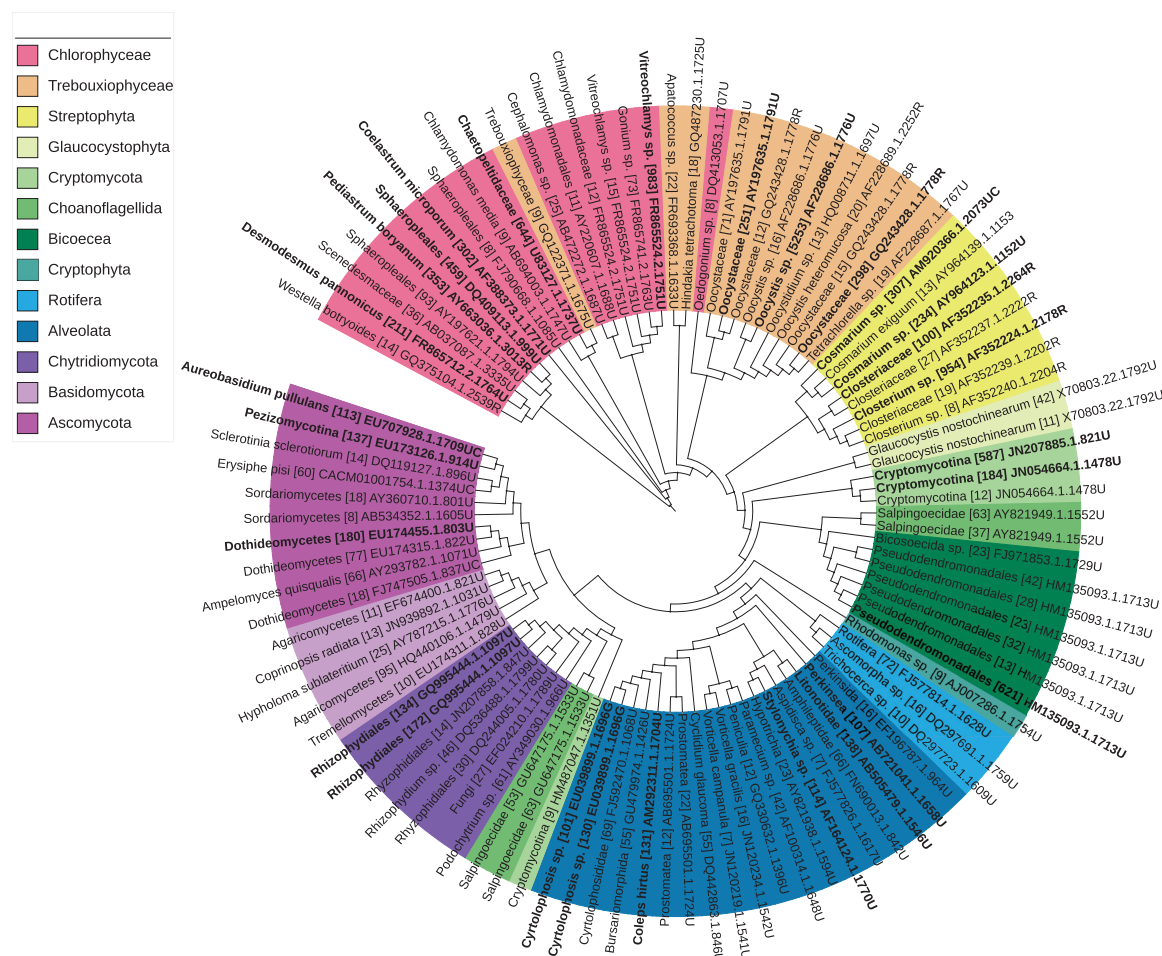


Fig 1. Phylogenetic tree based upon the 100 most abundant OTUs. OTUs with more than 100 sequences are marked in bold. The number of OTUs obtained is displayed in parenthesis. The OTUs are clustered at 99% identity and blasted against the PR² reference database for taxonomic identification. The phylogenetic tree was constructed by the relaxed neighbor joining algorithm in the program Clearcut and visualized using iTOL.

doi:10.1371/journal.pone.0172808.g001

quality filtering varied from 325–2,393 (S3 Fig). The majority of the sequences were classified as fungi (36%), followed by algae (33%), ciliophora (15%), bicoecia (5%) and choanoflagellida (4%). 64 OTUs were identified as algae. The algal OTUs were identified at different taxonomic levels: 13 to species (20%), 25 to genus (39%), 20 to family (31%) level. The remaining six OTUs could only be determined at the level of order due to a lack of higher resolution in the reference database.

Comparison of the morphology-based and the molecular approach

Eight algal taxa were detected using both the molecular and the morphology-based approach (S3 Table). *Monorhaphidium* sp., *Selenastrum* sp., *Trachelomonas* sp. and the cyanobacteria were not identified by the metabarcoding approach (S3 Table). Moreover, the algal community composition varied strongly between the molecular and the morphology-based approach (Fig 2). Based upon the morphological data (excluding the cyanobacteria), coccoid green algae were the most abundant group of primary producers (58%), followed by *Scenedesmus* sp. (22%). In comparison, the molecular approach yielded *Oocystaceae* as the most abundant group of primary producers (45%) followed by *Closteriaceae* (18%).

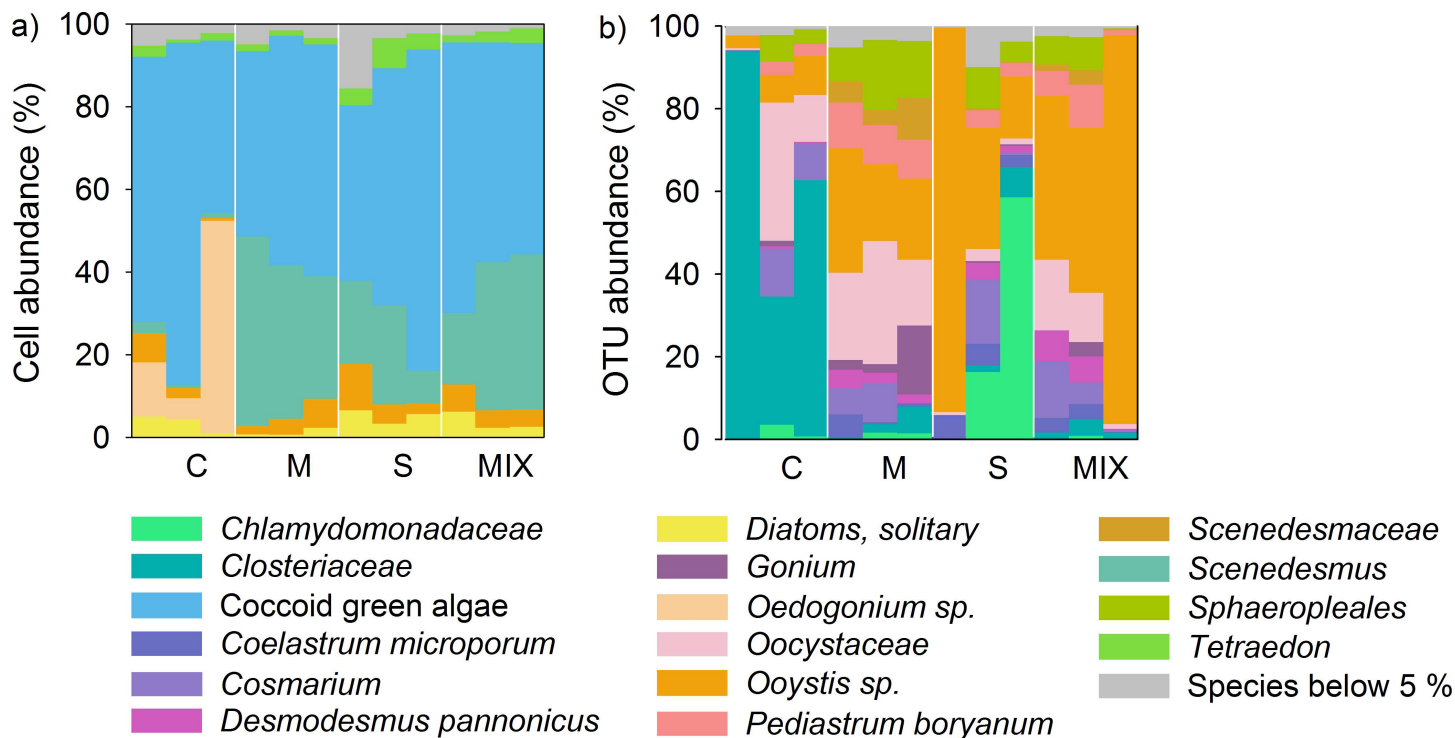


Fig 2. The species composition from morphology-based (a) and molecular (b) data of all twelve samples. The two species of *Tetraedon*, identified with the morphology-based method, were grouped and cyanobacteria were removed (as they cannot be detected with the 18S primer set for eukaryotes) for a better comparison. The biofilms were subjected to grazing by *C. dipterum*, *L. stagnalis* or by both consumer species over a period of 50 days. No consumer species were present in the control treatment. The four treatments are labeled C = Grazer-free control, M = *C. dipterum* (Mayfly), S = *L. stagnalis* (Snail), MIX = *C. dipterum* and *L. stagnalis*.

doi:10.1371/journal.pone.0172808.g002

We conducted linear regressions between the abundances of taxa across all treatments identified using both the molecular and the morphology-based methods (Fig 3, Table 2), and found significant correlations between the morphology-based and molecular abundance data for *Closterium* sp. (Fig 3B and 3F), *Cosmarium* sp. (Fig 3C and 3G), and *Oedogonium* sp. (Fig 3D and 3H, Table 2).

Grazing effects

The consumer species varied significantly in their feeding preferences, resulting in distinct algal community composition changes (S4 Fig; S2 Table). We found a highly significant interaction between the grazing treatments and the cell number of particular primary producer species (Table 3).

For 13 of 18 primary producer species, significant grazing effects were found upon cell number abundances (S4 Fig). For example, the morphology-based approach revealed a higher abundance of *Closterium* sp. and *Oedogonium* sp. in the control treatment when compared to the grazer treatments (Fig 2; S4 Fig). Similarly, the DNA metabarcoding approach found a high abundance of *Closteriaceae* in the control treatment in comparison to the grazer treatments (Fig 2), however very few *Oedogonium* sp. sequences were detected. Still, the DNA metabarcoding displayed a greater variation between the OTU richness and abundances between the treatments when compared to the morphology-based approach (Fig 2). The biological replicates analyzed using the both the DNA metabarcoding approach and the morphology-based approach were highly consistent, indicating that both methods provide stable and

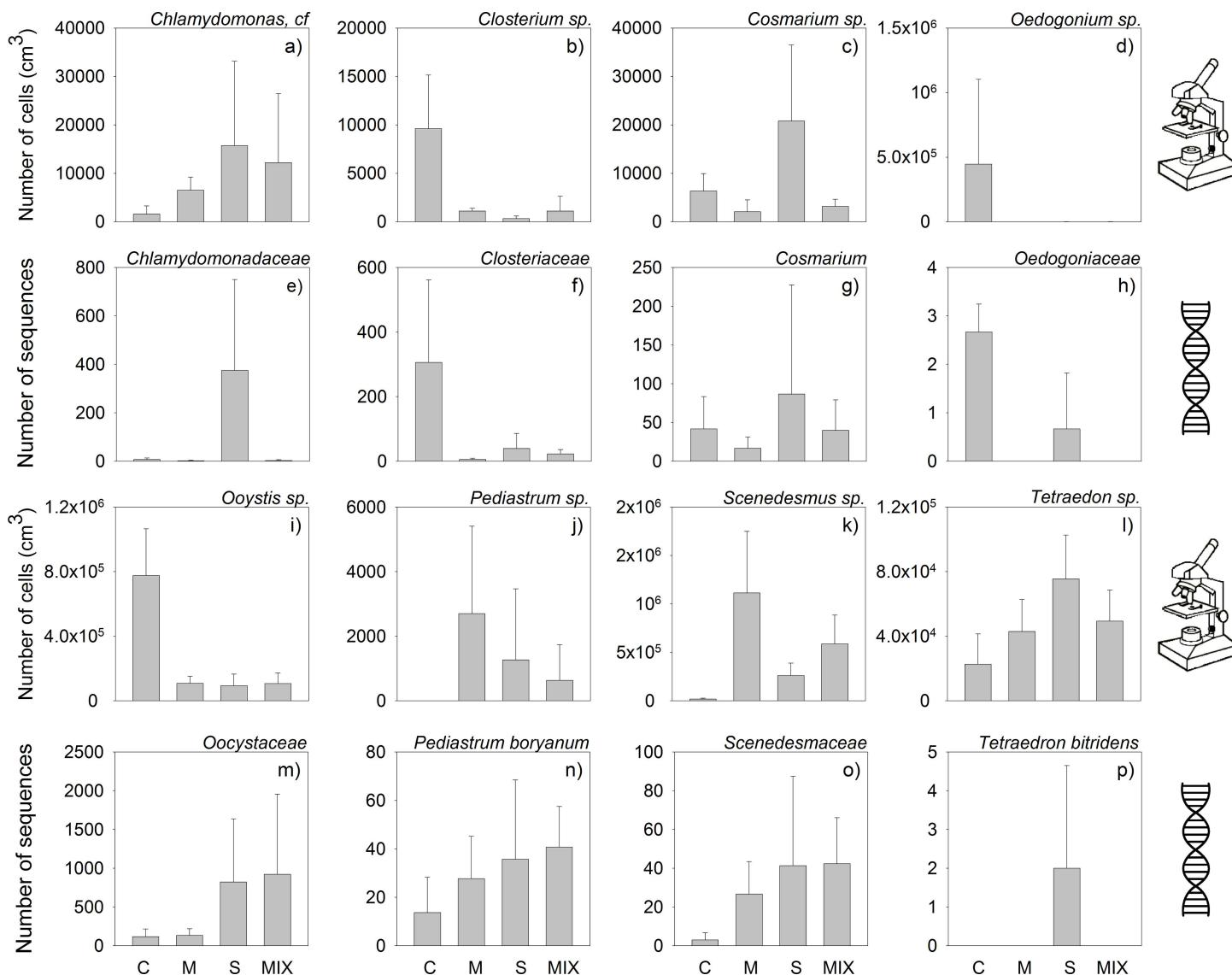


Fig 3. Comparisons of grazing impact on single algal taxa based upon microscopic identification (a—d, i—l) and sequence data (e—h, m—p). The biofilms were subjected to grazing by *C. dipterum*, *L. stagnalis*, or by both consumer species over a period of 50 days. No consumer species were present in the control treatment. The four treatments are labeled C = Grazer-free control, M = *C. dipterum* (Mayfly), S = *L. stagnalis* (Snail), MIX = *C. dipterum* and *L. stagnalis*. Bars represent algal cell (a—d, i—l) and sequence numbers (e—h, m—p, mean \pm SD, N = 3).

doi:10.1371/journal.pone.0172808.g003

reliable results (Fig 2). Yet, we could not find a significant correlation between consumer species richness and primary producer using neither the morphology-based ($y = 14.500 - (0.167x)$, $R^2 = 0.02$, $P = 0.67$; S6A Fig) nor the DNA metabarcoding approach ($y = 20.500 + (2.333x)$, $R^2 = 0.04$, $P = 0.51$; S6B Fig).

Discussion

This study represents one of the first comparative analyses of algal community structure employing both morphology-based and molecular methods to investigate environmental samples originating from the same sampling areas. Metabarcoding uncovered a vast taxonomic diversity, largely exceeding the 18 taxa yielded by the morphology-based survey. Without any

Table 2. Results of linear regressions (F-statistics and P-values; d.f. = degrees of freedom) between the abundances of specific taxa obtained via the morphology-based and the molecular method.

Algae taxa	Equation	df	F	P
<i>Chlamydomonas sp.</i>	$y = 1.06 + (0.03x)$	3	2.017	0.291
<i>Closterium sp.</i>	$y = -3.64 + (0.032x)$	3	62.349	0.016
<i>Cosmarium sp.</i>	$y = 19.815 + (0.00326x)$	3	25.082	0.038
<i>Oedogonium sp.</i>	$y = 0.220 + (0.00000548x)$	3	30.249	0.032
<i>Ooystis sp.</i>	$y = 704.959 - (0.000763x)$	3	1.086	0.407
<i>Pediastrum sp.</i>	$y = 26.437 + (0.00259x)$	3	0.138	0.746
<i>Scenedesmus sp.</i>	$y = 21.458 + (0.0000139x)$	3	0.295	0.642
<i>Tetradron sp.</i>	$y = -1.365 + (0.0000391x)$	3	5.348	0.147

doi:10.1371/journal.pone.0172808.t002

additional efforts (or identification expertise), the molecular approach identified additional eukaryotic taxa from the fungal and animal kingdoms. In total, 195 eukaryotic OTUs were detected in the DNA-based data set. Similarly, other studies investigating phytoplankton communities revealed a far greater species richness when using metabarcoding [21]. However, many metabarcoding studies have been criticized for misclassification of erroneous OTUs to new species (due to PCR and sequencing errors), which thereby may cause an overestimation of species diversity [40–42]. Therefore, any observations of low-abundance taxa need to be evaluated critically. Moreover, the intra- and interspecific diversity may greatly vary between species [43]. Consequently, the usage of a clustering divergence threshold may lead to an over and/or underestimation of the species richness. Here, the DNA metabarcoding approach covered most of the species found with the morphology-based approach, but not all. *Monoraphidium sp.*, *Selenastrum sp.* and *Trachelomonas sp.* were not identified by the metabarcoding, despite the fact that all three genera are included in the reference database. This may be explained by erroneous morphological classification, or misclassification of OTUs due to low variability within the metabarcoding marker.

We were only able to identify 59% of our OTUs to genus level. This is within the range of previous studies performed on unicellular eukaryotes [15, 21]. The lack of complete and curated DNA reference libraries is one of the limiting factors for the identification of species in metabarcoding studies [44, 45]. If DNA metabarcoding is to become a standard alternative or supplement to morphology-based approaches, we need to fill gaps in the DNA reference databases, which in turn is impossible without comprehensive taxonomic expertise. To increase the level of taxonomic identification of our sequences, we choose to use the PR² database [35]. This is a database specializing in 18S rDNA sequences of protists, with the benefit of being curated.

Metabarcoding is emerging as a promising method in biodiversity research. Yet, the lack of rigorous frameworks for analyzing NGS data makes comparisons between studies difficult. Among other factors, the taxonomic resolution in DNA metabarcoding studies depends heavily on the selection of the primer pairs. We designed a primer pair to ensure amplification

Table 3. Two-way ANOVA on the effect of consumer species combination (treatment) on the cell number of single algal taxa. F-statistics and P-values are shown. d.f. = degrees of freedom.

Effect	df	F	P
Algal taxa	16	0.85	0.63
Treatment	3	0.91	0.44
Algal taxa x Treatment	48	4.37	0.000

doi:10.1371/journal.pone.0172808.t003

of green algae and diatoms, targeting the v4 region in the 18S rDNA gene. As the 18S rDNA gene is a highly conserved gene found in all eukaryotes, this allowed us to analyze a broad taxonomic spectrum of the eukaryotic diversity. Universal primer pairs were shown to amplify only half of the OTUs revealed when using more selective primer pairs [46]. The usage of a specific primer set may have resulted in an increased OTU richness. However, universal primer pairs (e.g. the 18S rDNA gene) are more often included in reference databases [32, 35], the use of a more selective primer pair may therefore have reduced the number of identified OTUs.

Due to the coarse scale of the taxonomic identification of algae species provided by the morphology-based approach, we were not able to compare some of the common groups identified with the DNA metabarcoding results (e.g. coccoid green algae). Nevertheless, we found that the abundance pattern of the NGS data clearly differed from that of the morphology-based analysis. Based upon the molecular data, the second most common algal family (after *Oocystaceae*) was *Closteriaceae*, which comprised less than two percent of the primary producer community according to the morphology-based analysis. These deviations in the abundance data between the morphology-based and molecular method may have been caused by copy number variation [47, 48], pseudogenes [49] and/or number of nuclei per individual. Copy number of the rDNA gene can vary dramatically across taxa [47] and even within species [50]. While copy number variation is found in archaea and bacteria [51], it is far more extensive for eukaryotes, where up to tens of thousands copies have been found [47]. Moreover, DNA extraction [52] and PCR amplification [53] are also known to influence abundance estimates.

Even if there is a bias in the PCR procedure, DNA extraction and interspecific copy number variation, we may still be able to compare the abundance of single species between treatments. With respect to *Closterium* sp. and *Oedogonium* sp., the abundance of the algal species was significantly reduced in the consumer species treatments when compared to the control in both morphology-based and molecular-based approaches. Nevertheless, the molecular and morphology-based abundance data only correlated in 37.5% of the cases. This means that in the majority of the cases, the outcome of the abundance data differed between methods. Similar results have been found regarding pollen [54], plants [55], nematodes [56], macroinvertebrates [57], diatoms [18] and phytoplankton [21].

Grazing effects

It was obvious from both methodological approaches that the algal communities were strongly impacted by invertebrate grazing activity. With regards to the morphology-based data, we found that the abundance of *Oedogonium* sp. (a filamentous green algae) and diatoms (ribbon colonies) decreased in the presence of grazers. Both *Radix peregra* (a close relative to *L. stagnalis*) and *C. dipterum* were shown to prefer filamentous algae by previous studies [58, 59]. The cell abundances of *Closterium* sp., the second largest algae taxa identified, decreased in the presence of grazers. This results confirms previous findings, where snails [60–62] and mayfly larvae [59] were shown to prefer larger algal cells. We also identified grazer-specific effects. *C. dipterum* seemed to prefer *Trachelomonas* sp., whereas *L. stagnalis* ingested a higher proportion of *Scenedesmus* sp. Conversely, *Monoraphidium* sp., *Oocystis* sp., *Pediastrum* sp., and *Selenastrum* sp. decreased with grazer absence, likely caused by an increased competition between the algae species.

We expected a positive linear correlation between consumer species richness and primary producer richness, but there was no statistically significant result confirming this. However, with the DNA metabarcoding approach, we found a non-significant increase of primary producer richness with increasing grazer richness as expected, suggesting that the DNA metabarcoding approach may be more sensitive in detecting environmental changes.

Conclusions

Although considerable progress has recently been made in DNA metabarcoding, many challenges remain. Additional steps need to be taken to promote clearer methodological guidelines for analyzing NGS data. The methods and results reported herein contribute to the development of a fast, reliable and cost effective way to analyze algae communities. In conclusion, using DNA metabarcoding together with morphology-based traditional methods when assessing algal biodiversity increases the reliability of the outcomes. Species richness estimates can be made if careful measures are taken to avoid the overestimation of taxa from sequencing errors. DNA metabarcoding is a useful tool when it comes to detecting rare taxa and overall changes in community compositions, however, without comprehensive curated reference libraries, DNA metabarcoding lacks the power to contribute to algal species richness estimates.

Supporting information

S1 Fig. Workflow of experimental pipeline. The major bioinformatics steps of the experimental pipeline.

(TIF)

S2 Fig. The average primary producer community composition (including cyanobacteria) based on cell number (a) and biovolume (b), determined microscopically in relation to the consumer treatment: C = Grazer-free control, M = *C. dipterum* (Mayfly), S = *L. stagnalis* (Snail), MIX = *C. dipterum* and *L. stagnalis*.

(TIF)

S3 Fig. Rarefaction curve of all OTUs clustered at 99% identity in the four grazer treatments: C = Grazer-free control, M = *C. dipterum* (Mayfly), S = *L. stagnalis* (Snail), MIX = *C. dipterum* and *L. stagnalis*.

(TIF)

S4 Fig. Impact of grazing on single algal species based on microscope counts. Bars represent mean (\pm SD) algal cell numbers in the different treatments: C = Grazer-free control, M = *C. dipterum* (Mayfly), S = *L. stagnalis* (Snail), MIX = *C. dipterum* and *L. stagnalis*. Means that were found to be significantly different after *post-hoc* comparisons are labeled with superscript letters. Only taxa where significant grazing effects were found are displayed.

(TIF)

S5 Fig. Mean (\pm SD of $n = 3$) algal dry mass dependent on the consumer treatment: C = grazer-free control, M = *C. dipterum* (Mayfly), S = *L. stagnalis* (Snail), MIX = *C. dipterum* and *L. stagnalis*. One-way ANOVA (d.f. = 11, F = 3.31, P = 0.08, N = 3).

(TIF)

S6 Fig. The effect of consumer species richness on primary producer richness. Primary producer species richness a) and OTU richness b) (N = 3) after grazing by 0–2 consumer species over a period of 50 days. The results of the linear regression are represented as a solid line. The four treatments are labeled C = Grazer-free control, M = *C. dipterum* (Mayfly), S = *L. stagnalis* (Snail), MIX = *C. dipterum* and *L. stagnalis*.

(TIF)

S1 Table. Experimental setup. To ensure equal grazing pressure in all units, equal fresh weight of the consumer species were added.

(DOCX)

S2 Table. P-values derived from Tukey post-hoc comparisons regarding the effect of grazing on the cell number of single algal taxa in the treatments: C = Grazer-free control, M = *C. dipterum* (Mayfly), S = *L. stagnalis* (Snail), MIX = *C. dipterum* and *L. stagnalis*).
(DOCX)

S3 Table. The primary producer species detected using the morphology-based method and the molecular method, and taxa detected using both methods. If taxa were detected at higher taxonomic rank using the alternative method, the higher taxonomic rank is noted up to family level.
(DOCX)

Author Contributions

Conceptualization: SG PF.

Data curation: SG.

Formal analysis: SG.

Funding acquisition: PF.

Investigation: SG MK PF.

Methodology: SG MK PF.

Project administration: SG PF.

Resources: PF MK.

Software: SG.

Supervision: MK PF.

Validation: SG MK PF.

Visualization: SG.

Writing – original draft: SG PF.

Writing – review & editing: SG MK PF.

References

1. Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM. Accelerated modern human-induced species losses: entering the sixth mass extinction. *Sci Adv*. 2015.
2. Balvanera P, Pfisterer AB, Buchmann N, He JS, Nakashizuka T, Raffaelli D, et al. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol Lett*. 2006; 9(10):1146–56. doi: [10.1111/j.1461-0248.2006.00963.x](https://doi.org/10.1111/j.1461-0248.2006.00963.x) PMID: [16972878](https://pubmed.ncbi.nlm.nih.gov/16972878/)
3. Loreau M, Naeem S, Inchausti P, Bengtsson J, Grime J, Hector A, et al. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*. 2001; 294(5543):804–8. doi: [10.1126/science.1064088](https://doi.org/10.1126/science.1064088) PMID: [11679658](https://pubmed.ncbi.nlm.nih.gov/11679658/)
4. McCann KS. The diversity–stability debate. *Nature*. 2000; 405(6783):228–33. doi: [10.1038/35012234](https://doi.org/10.1038/35012234) PMID: [10821283](https://pubmed.ncbi.nlm.nih.gov/10821283/)
5. Arrigo KR. Marine microorganisms and global nutrient cycles. *Nature*. 2005; 437(7057):349–55. doi: [10.1038/nature04159](https://doi.org/10.1038/nature04159) PMID: [16163345](https://pubmed.ncbi.nlm.nih.gov/16163345/)
6. Sommer U. Comparison between steady state and non-steady state competition: experiments with natural phytoplankton. *Limnol Oceanogr*. 1985; 30(2):335–46.
7. Lehmalz JT, Sandgren CD. Species-specific rates of growth and grazing loss among freshwater algae'. *Limnol Oceanogr*. 1985; 30:34–46.

8. Power ME, Matthews WJ. Algae-grazing minnows (*Camptostoma anomalum*), piscivorous bass (*Micropterus* spp.), and the distribution of attached algae in a small prairie-margin stream. *Oecologia*. 1983; 60(3):328–32.
9. Porter KG. Selective grazing and differential digestion of algae by zooplankton. *Nature*. 1973; 244:179–80.
10. Tang CQ, Leasi F, Obertegger U, Kieneke A, Barraclough TG, Fontaneto D. The widely used small sub-unit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *Proc Natl Acad Sci U S A*. 2012; 109(40):16208–12. doi: [10.1073/pnas.1209160109](https://doi.org/10.1073/pnas.1209160109) PMID: [22988084](https://pubmed.ncbi.nlm.nih.gov/22988084/)
11. Guiry MD. How many species of algae are there? *J Phycol*. 2012; 48(5):1057–63. doi: [10.1111/j.1529-8817.2012.01222.x](https://doi.org/10.1111/j.1529-8817.2012.01222.x) PMID: [27011267](https://pubmed.ncbi.nlm.nih.gov/27011267/)
12. Bourlat SJ, Borja A, Gilbert J, Taylor MI, Davies N, Weisberg SB, et al. Genomics in marine monitoring: new opportunities for assessing marine health status. *Mar Pollut Bull*. 2013; 74(1):19–31. doi: [10.1016/j.marpolbul.2013.05.042](https://doi.org/10.1016/j.marpolbul.2013.05.042) PMID: [23806673](https://pubmed.ncbi.nlm.nih.gov/23806673/)
13. Ji Y, Ashton L, Pedley SM, Edwards DP, Tang Y, Nakamura A, et al. Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecol Lett*. 2013; 16(10):1245–57. doi: [10.1111/ele.12162](https://doi.org/10.1111/ele.12162) PMID: [23910579](https://pubmed.ncbi.nlm.nih.gov/23910579/)
14. Janzen DH, Hajibabaei M, Burns JM, Hallwachs W, Remigio E, Hebert PD. Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philos Trans R Soc Lond B Biol Sci*. 2005; 360(1462):1835–45. doi: [10.1098/rstb.2005.1715](https://doi.org/10.1098/rstb.2005.1715) PMID: [16214742](https://pubmed.ncbi.nlm.nih.gov/16214742/)
15. De Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, et al. Eukaryotic plankton diversity in the sunlit ocean. *Science*. 2015; 348(6237):1261605. doi: [10.1126/science.1261605](https://doi.org/10.1126/science.1261605) PMID: [25999516](https://pubmed.ncbi.nlm.nih.gov/25999516/)
16. Zhan A, Hulak M, Sylvester F, Huang X, Adebayo AA, Abbott CL, et al. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. *Methods Ecol Evol*. 2013; 4(6):558–65.
17. Comtet T, Sandionigi A, Viard F, Casiraghi M. DNA (meta) barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. *Biological Invasions*. 2015; 17(3):905–22.
18. Keramarrec L, Franc A, Rimet F, Chaumeil P, Humbert J-F, Bouchez A. Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms. *Mol Ecol Resour*. 2013; 13(4):607–19. doi: [10.1111/1755-0998.12105](https://doi.org/10.1111/1755-0998.12105) PMID: [23590277](https://pubmed.ncbi.nlm.nih.gov/23590277/)
19. Port JA, O'Donnell JL, Romero-Maraccini OC, Leary PR, Litvin SY, Nickols KJ, et al. Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Mol Ecol*. 2016; 25(2):527–41. doi: [10.1111/mec.13481](https://doi.org/10.1111/mec.13481) PMID: [26586544](https://pubmed.ncbi.nlm.nih.gov/26586544/)
20. Elbrecht V, Taberlet P, Dejean T, Valentini A, Usseglio-Polatera P, Beisel J-N, et al. Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ*. 2016.
21. Abad D, Albaina A, Aguirre M, Laza-Martínez A, Uriarte I, Iriarte A, et al. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. *Mar Biol*. 2016; 163(7):1–13.
22. Zimmermann J, Glöckner G, Jahn R, Enke N, Gemeinholzer B. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Mol Ecol Resour*. 2015; 15(3):526–42. doi: [10.1111/1755-0998.12336](https://doi.org/10.1111/1755-0998.12336) PMID: [25270047](https://pubmed.ncbi.nlm.nih.gov/25270047/)
23. Burkpile DE, Hay ME. Herbivore species richness and feeding complementarity affect community structure and function on a coral reef. *Proc Natl Acad Sci U S A*. 2008; 105(42):16201–6. doi: [10.1073/pnas.0801946105](https://doi.org/10.1073/pnas.0801946105) PMID: [18845686](https://pubmed.ncbi.nlm.nih.gov/18845686/)
24. Jaschinski S, Aberle N, Gohse-Reimann S, Brendelberger H, Wiltshire KH, Sommer U. Grazer diversity effects in an eelgrass–epiphyte–microphytobenthos system. *Oecologia*. 2009; 159(3):607–15. doi: [10.1007/s00442-008-1236-2](https://doi.org/10.1007/s00442-008-1236-2) PMID: [19082631](https://pubmed.ncbi.nlm.nih.gov/19082631/)
25. Guillard RR. Culture of phytoplankton for feeding marine invertebrates. *Culture of marine invertebrate animals*: Springer; 1975. p. 29–60.
26. ASAB. Guidelines for the treatment of animals in behavioural research and teaching. *Anim Behav*. 2012; 83:301–9.
27. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012; 28(12):1647–9. doi: [10.1093/bioinformatics/bts199](https://doi.org/10.1093/bioinformatics/bts199) PMID: [22543367](https://pubmed.ncbi.nlm.nih.gov/22543367/)
28. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 2012; 40(15):e115–e. doi: [10.1093/nar/gks596](https://doi.org/10.1093/nar/gks596) PMID: [22730293](https://pubmed.ncbi.nlm.nih.gov/22730293/)
29. Hancock JM. The contribution of DNA slippage to eukaryotic nuclear 18S rRNA evolution. *J Mol Evol*. 1995; 40(6):629–39. PMID: [7643414](https://pubmed.ncbi.nlm.nih.gov/7643414/)

30. Hwang UW, Ree HI, Kim W. Evolution of hypervariable regions, V4 and V7, of insect 18S rRNA and their phylogenetic implications. *Zoolog Sci.* 2000; 17(1):111–21. doi: [10.2108/zsj.17.111](https://doi.org/10.2108/zsj.17.111) PMID: [18494566](https://pubmed.ncbi.nlm.nih.gov/18494566/)
31. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol.* 2009; 75(23):7537–41. doi: [10.1128/AEM.01541-09](https://doi.org/10.1128/AEM.01541-09) PMID: [19801464](https://pubmed.ncbi.nlm.nih.gov/19801464/)
32. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013.
33. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics.* 2011; 27(16):2194–200. doi: [10.1093/bioinformatics/btr381](https://doi.org/10.1093/bioinformatics/btr381) PMID: [21700674](https://pubmed.ncbi.nlm.nih.gov/21700674/)
34. Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequences. *Journal of Computational biology.* 2000; 7(1–2):203–14. doi: [10.1089/10665270050081478](https://doi.org/10.1089/10665270050081478) PMID: [10890397](https://pubmed.ncbi.nlm.nih.gov/10890397/)
35. Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, et al. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 2012.
36. Flynn JM, Brown EA, Chain FJ, Maclsaac HJ, Cristescu ME. Toward accurate molecular identification of species in complex environmental samples: testing the performance of sequence filtering and clustering methods. *Ecol Evol.* 2015; 5(11):2252–66. doi: [10.1002/ece3.1497](https://doi.org/10.1002/ece3.1497) PMID: [26078860](https://pubmed.ncbi.nlm.nih.gov/26078860/)
37. Evans J, Sheneman L, Foster J. Relaxed neighbor joining: a fast distance-based phylogenetic tree construction method. *J Mol Evol.* 2006; 62(6):785–92. doi: [10.1007/s00239-005-0176-2](https://doi.org/10.1007/s00239-005-0176-2) PMID: [16752216](https://pubmed.ncbi.nlm.nih.gov/16752216/)
38. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 2016.
39. Hillebrand H, Dürselen CD, Kirschtel D, Pollinger U, Zohary T. Biovolume calculation for pelagic and benthic microalgae. *J Phycol.* 1999; 35(2):403–24.
40. Kunin V, Engelbrektson A, Ochman H, Hugenholtz P. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol.* 2010; 12(1):118–23. doi: [10.1111/j.1462-2920.2009.02051.x](https://doi.org/10.1111/j.1462-2920.2009.02051.x) PMID: [19725865](https://pubmed.ncbi.nlm.nih.gov/19725865/)
41. Bachy C, Dolan JR, López-García P, Deschamps P, Moreira D. Accuracy of protist diversity assessments: morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. *ISME J.* 2013; 7(2):244–55. doi: [10.1038/ismej.2012.106](https://doi.org/10.1038/ismej.2012.106) PMID: [23038176](https://pubmed.ncbi.nlm.nih.gov/23038176/)
42. Behnke A, Engel M, Christen R, Nebel M, Klein RR, Stoeck T. Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ Microbiol.* 2011; 13(2):340–9. doi: [10.1111/j.1462-2920.2010.02332.x](https://doi.org/10.1111/j.1462-2920.2010.02332.x) PMID: [21281421](https://pubmed.ncbi.nlm.nih.gov/21281421/)
43. Brown EA, Chain FJ, Crease TJ, Maclsaac HJ, Cristescu ME. Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably describe zooplankton communities? *Ecol Evol.* 2015; 5(11):2234–51. doi: [10.1002/ece3.1485](https://doi.org/10.1002/ece3.1485) PMID: [26078859](https://pubmed.ncbi.nlm.nih.gov/26078859/)
44. Leray M, Knowlton N. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proc Natl Acad Sci U S A.* 2015; 112(7):2076–81. doi: [10.1073/pnas.1424997112](https://doi.org/10.1073/pnas.1424997112) PMID: [25646458](https://pubmed.ncbi.nlm.nih.gov/25646458/)
45. Cowart DA, Pinheiro M, Mouchel O, Maguer M, Grall J, Miné J, et al. Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. *PLoS One.* 2015; 10(2):e0117562. doi: [10.1371/journal.pone.0117562](https://doi.org/10.1371/journal.pone.0117562) PMID: [25668035](https://pubmed.ncbi.nlm.nih.gov/25668035/)
46. Lentendu G, Wubet T, Chatzinotas A, Wilhelm C, Buscot F, Schlegel M. Effects of long-term differential fertilization on eukaryotic microbial communities in an arable soil: a multiple barcoding approach. *Mol Ecol.* 2014; 23(13):3341–55. doi: [10.1111/mec.12819](https://doi.org/10.1111/mec.12819) PMID: [24888892](https://pubmed.ncbi.nlm.nih.gov/24888892/)
47. Prokopyowich CD, Gregory TR, Crease TJ. The correlation between rDNA copy number and genome size in eukaryotes. *Genome.* 2003; 46(1):48–50. doi: [10.1139/g02-103](https://doi.org/10.1139/g02-103) PMID: [12669795](https://pubmed.ncbi.nlm.nih.gov/12669795/)
48. Zhu F, Massana R, Not F, Marie D, Vault D. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol.* 2005; 52(1):79–92. doi: [10.1016/j.femsec.2004.10.006](https://doi.org/10.1016/j.femsec.2004.10.006) PMID: [16329895](https://pubmed.ncbi.nlm.nih.gov/16329895/)
49. Santos SR, Kinzie RA, Sakai K, Coffroth MA. Molecular characterization of nuclear small subunit (ISS)-rDNA pseudogenes in a symbiotic Dinoflagellate (*Symbiodinium*, Dinophyta). *J Eukaryot Microbiol.* 2003; 50(6):417–21. PMID: [14733432](https://pubmed.ncbi.nlm.nih.gov/14733432/)
50. Lyckegaard EM, Clark AG. Evolution of ribosomal RNA gene copy number on the sex chromosomes of *Drosophila melanogaster*. *Mol Biol Evol.* 1991; 8:458–74. PMID: [1921706](https://pubmed.ncbi.nlm.nih.gov/1921706/)

51. Pei AY, Oberdorf WE, Nossa CW, Agarwal A, Chokshi P, Gerz EA, et al. Diversity of 16S rRNA genes within individual prokaryotic genomes. *Appl Environ Microbiol.* 2010; 76(12):3886–97. doi: [10.1128/AEM.02953-09](https://doi.org/10.1128/AEM.02953-09) PMID: [20418441](https://pubmed.ncbi.nlm.nih.gov/20418441/)
52. Roh C, Villatte F, Kim B-G, Schmid RD. Comparative study of methods for extraction and purification of environmental DNA from soil and sludge samples. *Appl Biochem Biotechnol.* 2006; 134(2):97–112. PMID: [16943632](https://pubmed.ncbi.nlm.nih.gov/16943632/)
53. Gonzalez JM, Portillo MC, Belda-Ferre P, Mira A. Amplification by PCR artificially reduces the proportion of the rare biosphere in microbial communities. *PLoS One.* 2012.
54. Richardson RT, Lin C-H, Sponsler DB, Quijia JO, Goodell K, Johnson RM. Application of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem. *Appl Plant Sci.* 2015; 3(1):1400066.
55. Hiiesalu I, Oepik M, Metsis M, Lilje L, Davison J, Vasar M, et al. Plant species richness belowground: higher richness and new patterns revealed by next-generation sequencing. *Mol Ecol.* 2012; 21(8):2004–16. doi: [10.1111/j.1365-294X.2011.05390.x](https://doi.org/10.1111/j.1365-294X.2011.05390.x) PMID: [22168247](https://pubmed.ncbi.nlm.nih.gov/22168247/)
56. Porazinska DL, GIBLIN-DAVIS RM, Faller L, Farmerie W, Kanzaki N, Morris K, et al. Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Mol Ecol Resour.* 2009; 9(6):1439–50. doi: [10.1111/j.1755-0998.2009.02611.x](https://doi.org/10.1111/j.1755-0998.2009.02611.x) PMID: [21564930](https://pubmed.ncbi.nlm.nih.gov/21564930/)
57. Hajibabaei M, Spall JL, Shokralla S, van Konynenburg S. Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecol.* 2012; 12(1):1.
58. Calow P. Studies on the natural diet of *Lymnaea pereger obtusa* (Kobelt) and its possible ecological implications. *J Mollus Stud.* 1970; 39(2–3):203–15.
59. Brown DS. The ingestion and digestion of algae by *Chloëon dipterum* L. (*Ephemeroptera*). *Hydrobiologia.* 1960; 16(1):81–96.
60. Lopez GR, Levinton JS. The availability of microorganisms attached to sediment particles as food for *Hydrobia ventrosa* Montagu (Gastropoda: Prosobranchia). *Oecologia.* 1978; 32(3):263–75.
61. Groendahl S, Fink P. The effect of diet mixing on a nonselective herbivore. *PLoS One.* 2016.
62. Fenchel T, Kofoed L, Lappalainen A. Particle size-selection of two deposit feeders: the amphipod *Corophium volutator* and the prosobranch *Hydrobia ulvae*. *Mar Biol.* 1975; 30(2):119–28.