Received: 28 November 2018 | Revised: 19 July 2019 | Accepted: 4 September 2019

DOI: 10.1111/fwb.13413

ORIGINAL ARTICLE

Freshwater Biology WILEY

Phylogenetic placement of environmental sequences using taxonomically reliable databases helps to rigorously assess dinophyte biodiversity in Bavarian lakes (Germany)

Marc Gottschling[1] | **Juliana Chacón**[1] | **Anže Žerdoner Čalasan**[1,2] | **Stefan Neuhaus**[3] | **Juliane Kretschmann**[1] | **Herwig Stibor**[4] | **Uwe John**[3,5]

[1]Department Biologie, Systematische Botanik und Mykologie, GeoBio-Center, Ludwig-Maximilians-Universität München, München, Germany

[2]Department Biologie/Chemie, Botanik, Universität Osnabrück, Osnabrück, Germany

[3]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

[4]Department Biologie, Aquatic Ecology, Ludwig-Maximilians-Universität München, Martinsried-Planegg, Germany

[5]Helmholtz-Institut für Funktionelle Marine Biodiversität an der Universität Oldenburg, Oldenburg, Germany

Correspondence
Uwe John, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12, D – 275 70 Bremerhaven, Germany.
Email: uwe.john@awi.de

Abstract

1. Reliable determination of organisms is a prerequisite to explore their spatial and temporal occurrence and to study their evolution, ecology, and dispersal. In Europe, Bavaria (Germany) provides an excellent study system for research on the origin and diversification of freshwater organisms including dinophytes, due to the presence of extensive lake districts and ice age river valleys. Bavarian freshwater environments are ecologically diverse and range from deep nutrient-poor mountain lakes to shallow nutrient-rich lakes and ponds.
2. We obtained amplicon sequence data (V4 region of small subunit-rRNA, *c.* 410 bp long) from environmental samples collected at 11 sites in Upper Bavaria. We found 186 operational taxonomic units (OTUs) associated with Dinophyceae that were further classified by means of a phylogenetic placement approach.
3. The maximum likelihood tree inferred from a well-curated reference alignment comprised a systematically representative set of 251 dinophytes, covering the currently known molecular diversity and OTUs linked to type material if possible. Environmental OTUs were scattered across the reference tree, but accumulated mostly in freshwater lineages, with 79% of OTUs placed in either *Apocalathium*, *Ceratium*, or *Peridinium*, the most frequently encountered taxa in Bavaria based on morphology.
4. Twenty-one Bavarian OTUs showed identical sequences to already known and vouchered accessions, two of which are linked to type material, namely *Palatinus apiculatus* and *Theleodinium calcisporum*. Particularly within Peridiniaceae, delimitation of *Peridinium* species was based on the intraspecific sequence variation.
5. Our approach indicates that high-throughput sequencing of environmental samples is effective for reliable determination of dinophyte species in Bavarian lakes. We further discuss the importance of well-curated reference databases that remain to be developed in the future.

KEYWORDS
amplicon sequencing, dinoflagellates, distribution, molecular phylogenetics, operational taxonomic unit

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors. *Freshwater Biology* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Solid knowledge of ecosystem functioning and community dynamics during seasonal or longer periods, as well as conservation strategies and the impact of invasive species, essentially relies on precise original data about the spatial and temporal occurrence of the inhabiting organisms. Recent advances in the application of high-throughput sequencing (HTS) to analyse the molecular diversity in aquatic environments have enabled a better understanding of the community composition and species distribution in such ecosystems. Within the last decade, our knowledge of the microbial biodiversity has increased faster than ever, especially after Edwards et al. (2006) published the first metagenome analysis of environmental samples using *next-generation* sequencing technologies. An unexpectedly high protist diversity has been described since then in various aquatic environments (Cuvelier et al., 2010; Keeling et al., 2014; Kohli, Neilan, Brown, Hoppenrath, & Murray, 2014; Massana et al., 2015; Stoeck et al., 2010), revealing also the existence of seasonal variation in many phytoplankton taxa (Tillmann, Salas, Jauffrais, Hess, & Silke, 2014; Toebe, Joshi, et al., 2013) and demonstrating the power of these methods in discovering the world's hidden microbial diversity (Lindeque, Parry, Harmer, Somerfield, & Atkinson, 2013; Medinger et al., 2010). Unknown sequences derived from environmental samples may additionally assist discovering new species and even new lineages (Seenivasan, Sausen, Medlin, & Melkonian, 2013).

One of the most basic and important questions in evolutionary biology of microorganisms refers to mechanisms that have shaped their current distribution (O'Dwyer, Kembel, & Sharpton, 2015). The relative importance of principal processes such as divergence, dispersal, and selection by ecological filtering (Vellend, 2010) is a key point that needs to be worked out rigorously. There is an ongoing debate (Bass & Boenigk, 2011; Caron, 2009; Foissner, 2011) into whether microbes are all cosmopolitan, and lack distinct distributions (Fenchel & Finlay, 2004; Finlay, 2002; Read et al., 2013), or rather follow a moderate endemism model (Bass, Richards, Matthai, Marsh, & Cavalier-Smith, 2007; Bates et al., 2013; Coleman, 2001). Under a less dogmatic perspective, plankton communities may actually consist of both widespread species and those with a smaller range (Bik et al., 2012; Coleman, 2001; Foissner, 2008; Žerdoner Čalasan, Kretschmann, Filipowicz, et al., 2019), and occurrences for such species remain to be identified individually.

A limitation that still precludes a comprehensive knowledge of protist distribution is the taxonomic confusion that exists due to complex determination procedures (Wheeler, 2008). To take full advantage of environmental sequencing, curated reference collections are necessary, linking the molecular data with scientific binominals. Much progress has been made in the past years to build such *DNA-barcoding* databases, not only for animals and plants (Hollingsworth, 2011; Vernooy et al., 2010), but also for various microorganisms (Quast et al., 2013) such as fungi, diatoms, and dinophytes (Del Campo et al., 2018; Fourtanier & Kociolek, 2009; Mordret et al., 2018; Peršoh, 2015; Schoch et al., 2012; Stern et al., 2012). These databases are continuously updated and curated (Rigden, Fernandez-Suarez, & Galperin, 2016), preferably based on information from type material (Pawlowski et al., 2012). However, the availability of reference data for environmental sequencing studies is still highly unbalanced across various taxonomic protist lineages.

A prerequisite to any assessment of the microbial biodiversity found in a given ecosystem is the confident placement of resulting molecular operational taxonomic units (OTUs) within particular lineages in the Tree of Life. Due to a huge number of reads, and the diversity of organisms captured by HTS of environmental DNA, initial OTU annotation using reference databases usually allows for confident taxonomic classifications only up to coarser taxonomic ranks (such as phylum, class, or order: Quast et al., 2013). A complementary approach, which enables finer levels of taxonomic assignments (i.e. family, genus, or species), is the phylogenetic placement of reads in reference trees, which also offers the advantage of estimating statistical support values for annotations, regardless of the read's length (Dunthorn et al., 2014). This method has been recently used in barcoding and biomonitoring projects of protists (Elferink et al., 2017; Keck, Vasselon, Rimet, Bouchez, & Kahlert, 2018; Medinger et al., 2010; Vergin et al., 2013).

In comparison to those living in marine environments, the freshwater dinophytes are relatively well understood. Currently, 350 species are listed based on morphology (Mertens, Rengefors, Moestrup, & Ellegaard, 2012; Moestrup & Calado, 2018), and continuous efforts to revise them taxonomically have been made in the past years (Moestrup, Lindberg, & Daugbjerg, 2009; Craveiro, Pandeirada, Daugbjerg, Moestrup, & Calado, 2013; Takano, Yamaguchi, Inouye, Moestrup, & Horiguchi, 2014; Kretschmann, Filipowicz, Owsianny, Zinßmeister, & Gottschling, 2015; Kretschmann, Owsianny, Žerdoner Čalasan, & Gottschling, 2018, to mention only a few studies). This also includes the generation of reference DNA sequence information being available for a considerable portion of the species. Nevertheless, there is still a significant portion of freshwater dinophyte flora that is neither taxonomically revised nor molecularly investigated. As such, it hinders the transparent research in this field and calls all newly obtained data into question. Moreover, the phenomenon of cryptic speciation (i.e. molecular diversification without morphological differentiation), which is taxonomically challenging in a number of marine dinophytes (Hariganeya et al., 2013; Söhner, Zinßmeister, Kirsch, & Gottschling, 2012; Toebe, Alpermann, et al., 2013), appears rare (or understudied) in the freshwater habitat. From the phylogenetic perspective, freshwater dinophytes are a heterogeneous group and have colonised their habitats from the marine environment several times independently (Logares et al., 2007; Žerdoner Čalasan, Kretschmann, & Gottschling, 2019). Dinophyte development is complex and frequently includes at least two ecologically differentiated stages of life-history, namely a motile theca (*vegetative cell*) and an immotile, coccoid cell (colloquially termed *cyst*: von Stosch, 1973; Pfiester & Anderson, 1987).

In Germany, a number of dinophyte species have been originally described from the Bavarian region (Baumeister, 1957; Lindemann, 1920; Schrank, 1802) also being the focus of the present study.

However, most algae inhabiting Bavarian waters are presumably wider-ranging, with type localities outside of that region. Mauch, Schmedtje, Maetze, and Fischer (2003) reports 62 dinophyte species (online table available at www.gewaesser-bewertung.de/files/taxaliste.pdf), but this checklist does not provide any source references, is taxonomically not rigorous and also includes species from the marine environment (e.g. *Heterocapsa rotundata*, which is unlikely to be present in Bavarian freshwater habitats). Morphology based records of Bavarian dinophytes are otherwise rather sporadic, and only a few species have been reported so far including common *Ceratium hirundinella*, *Gyrodinium* (≡ *Gymnodinium*) *helveticum*, and *Peridinium willei* (Raeder, 1990; Schaumburg, 1996; Siebeck, 1982), as well as rarer species such as *Cystodinium cornifax* and *Gloeodinium montanum* (Höll, 1928). The true number of dinophyte species in Bavaria, and their spatial occurrences, is thus unknown at present, and more thorough biodiversity assessments of these key protists will benefit from the application of contemporary, high-throughput molecular methods.

In the present study, we provide an initial attempt to apply environmental amplicon sequences with the principal aim to uncover the yet unexplored dinophyte diversity in Bavarian lakes as an exemplary system. We determine their species by placing the ribotypes in phylogenetic reference trees (i.e. rRNA sequence comparison of multiple sequences including GenBank vouchers collected all around the world). Bavarian freshwater environments range from deep nutrient-poor mountain lakes to shallow nutrient-rich lakes and ponds (basic information is available at www.lfu.bayern.de/). These are characterised by a broad range of environmental conditions and resource levels and represent potential habitats for phytoplankton communities. We expect the dinophyte communities in those lakes to be composed of species with different ecological requirements and potentials for dispersal (Žerdoner Čalasan, Kretschmann, Filipowicz, et al., 2019), resulting in different effects on their distribution. Our results will lay the basis for a better knowledge of ecosystem functioning and evolutionary dynamics of protists such as freshwater dinophytes.

2 | METHODS

Surface plankton tow samples were collected from piers at 13 localities in Upper Bavaria (Germany) in April 2017 using a plankton net (mesh size 20 μm). The localities included 10 lakes (two lakes were sampled at two sites) and one subsidiary river, to cover standing and flowing bodies of water as well (Table 1, Figure 1). Geographic coordinates were recorded for all sites using a standard GPS Garmin Ltd device. Cells were observed, documented, and measured under a CKX41 inverted microscope (Olympus; Hamburg, Germany) equipped with a phase-contrast option and a DP73 digital camera (Olympus).

Environmental DNA was extracted using the Genomic DNA from Soil kit (Machery-Nagel; Düren, Germany) following the manufacturer's protocol. The small subunit (SSU or 18S) of the ribosomal RNA (rRNA) operon V4 region (*c.* 410 bp) was the amplification target. Due to polymerase chain reaction (PCR) biases or PCR errors that may artificially increase diversity, each PCR was performed in triplicates (Acinas, Sarma-Rupavtarm, Klepac-Ceraj, & Polz, 2005). Forward and reverse primers were those used by Xiao, Wu, Liu, Xu, and Chi (2017). Amplification of DNA (PCR) for subsequent amplicon sequencing (Illumina) was carried out using 5 ng/μl template DNA, 1 μM of each primer and 2× KAPA Hifi HotStart Ready Mix (Roche; Penzberg, Germany). Resulting PCR products were visualised in 1% agarose gels and purified using AMPure XP Beads (Beckman Coulter). Dual indices and Illumina sequence adapters were attached by means of an Index PCR using the Nextera XT Index Kit (Illumina), and final PCR products were again purified using AMPure XP Beads. The library was validated using an Agilent 2100 Bioanalyzer Software and a DNA 1000 Chip (Agilent Technologies) to verify the size of the resulting fragments. The final DNA libraries were equimolarly pooled and run in a MiSeq System (Illumina) after combining the denatured PhiX control library (15%) and the denatured amplicon library. Some 6.5 million 2 × 300 bp paired-end reads were produced and demultiplexed into 13 samples.

Using Trimmomatic (v0.38; Bolger, Lohse, & Usadel, 2014), 3′-ends of the reads were trimmed based on read quality information. PEAR (v0.9.10; Zhang, Kobert, Flouri, & Stamatakis, 2014) with default settings was used to merge the paired-end reads. Sequences that could not be merged were discarded. Primer-matching sequence segments were truncated from the amplicons by cutadapt (v1.9; Martin, 2011) and amplicons were only kept in the sequence pool if both the segments of the forward and of the reverse primer could be found. Remaining sequences were filtered for further quality features by vsearch (v2.3.0; Rognes, Flouri, Nichols, Quince, & Mahé, 2016). Sequences were discarded if they were outside a 50 bp radius above or below the median length of the primer-truncated amplicon (*c.* 387 bp), if they carried any ambiguity, or if the expected number of miscalled bases of a sequence (sum of all base error probabilities of a sequence) was >1. Chimera were predicted also by vsearch utilising the UCHIME algorithm (Edgar, Haas, Clemente, Quince, & Knight, 2011) with default settings in de novo mode for each sample separately and removed from the sample files. About 4 million sequences passed all filtering steps and were used as input for the OTU-clustering, which was done using the tool Swarm (v2.1.8; Mahé, Rognes, Quince, De Vargas, & Dunthorn, 2015) with default settings. The most abundant amplicon of each OTU-cluster was used as an OTU representative. These sequences were annotated by the RDP classifier (Wang, Garrity, Tiedje, & Cole, 2007) implemented in mothur (v1.38.1; Schloss et al., 2009) using the Ref_NR99 version of release 128 of the SILVA SSU sequence set (Quast et al., 2013) using a reference with a confidence cut-off of 90. The annotation of each representative sequence was used as annotation of the OTU-cluster as well and added to the corresponding line of the OTU table.

Operational taxonomic units classified as Dinoflagellata as search strings and with an abundance of ≥10 were classified more accurately by phylogenetic placement onto a reference tree based on concatenated rRNA alignments. Full voucher information of this

TABLE 1 Geographic origin of the samples used in the molecular analyses. To characterise the localities in terms of their ecology, we provide the corresponding trophic state index (Nürnberg, 1996; OECD, 1982)

Collectors	Collection number	Locality (trophic state index)	Geographic coordinates (latitude, longitude)	Area (km^2)	Max. depth (m)	Mean depth (m)	Altitude (m)
J. Kretschmann & J. Schmid	D075	Lake Kochel (mesotrophic peri-alpine lake)	47°39.200′N 11°19.001′E	3.95	66	31	601
J. Kretschmann & J. Schmid	D076	Riegsee (mesotrophic peri-alpine lake)	47°41.395′N 11°13.849′E	1.97	15.4	6.8	661
J. Kretschmann & J. Schmid	D077	Staffelsee (oligotrophic peri-alpine lake)	47°41.406′N 11°10.631′E	7.66	39.4	9.8	648
A. Žerdoner Čalasan & M. Gottschling	D078	Starnberger See (mesotrophic peri-alpine lake)	47°41.032′N 11°11.342′E	56.36	127.8	53.2	533
A. Žerdoner Čalasan & M. Gottschling	D079	Eitzenberger Weiher (eutrophic moor pond)	47°47.122′N 11°21.757′E	0.0534	10	Unknown	602
A. Žerdoner Čalasan & M. Gottschling	D080	Oster See (mesotrophic peri-alpine lake)	47°47.687′N 11°18.357′E	0.235	29.7	9.38	589
A. Žerdoner Čalasan & M. Gottschling	D081	Hardtbach near Haarsee (mesotrophic water reservoir)	47°49.044′N 11°13.206′E	0.0647	8.8	Unknown	643
A. Žerdoner Čalasan & M. Gottschling	D082	Zellsee (mesotrophic water reservoir)	47°52.335′N 11°03.848′E	1	<5	Unknown	583
A. Žerdoner Čalasan & M. Gottschling	D083	Lech at Apfeldorf (peri-alpine river bay)	47°54.120′N 10°55.900′E	–	–	–	640
A. Žerdoner Čalasan & M. Gottschling	D084	Ammersee (Diessen) (mesotrophic peri-alpine lake)	47°57.017′N 11°06.571′E	46.6	81.1	37.5	530
A. Žerdoner Čalasan & M. Gottschling	D085	Ammersee (Herrsching) (mesotrophic peri-alpine lake)	47°59.570′N 11°10.188′E	46.6	81.1	37.5	531
A. Žerdoner Čalasan & M. Gottschling	D086	Pilsensee (mesotrophic peri-alpine lake)	48°01.421′N 11°11.764′E	1.95	17.1	9.3	525
A. Žerdoner Čalasan & M. Gottschling	D087	Lake Wörth (mesotrophic peri-alpine lake)	48°04.110′N 11°12.013′E	4.34	34	14.7	553

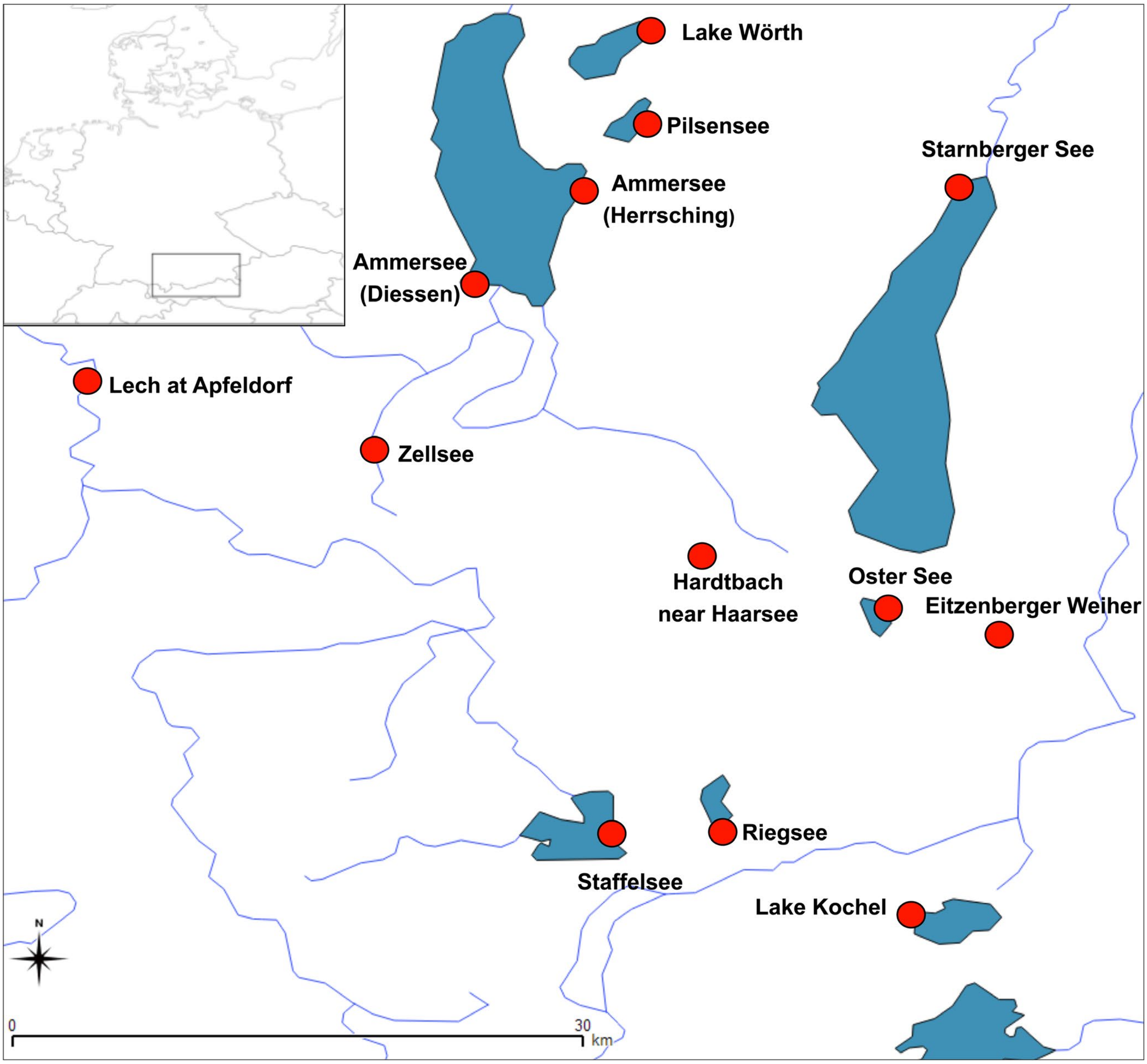

FIGURE 1 Selected Bavarian lakes under investigation. Collection sites are indicated by red dots, GPS coordinates are given in Table 1. The map was made using geographic data for the inland waters in Germany, which are freely available at the DIVA-GIS website (http://www.diva-gis.org/gdata, accessed 12 April, 2019) and the free software QGIS, Free Open Source Geographic Information System (QGIS Development Team, 2019)

systematically representative set comprising 241 dinophytes (plus 10 outgroup accessions) are provided in Table S1. To further explore dinophyte identities based on DNA sequences, we performed BLAST searches (Altschul, Gish, Miller, Myers, & Lipman, 1990).

For alignment constitution, separate matrices of the rRNA operon (i.e. SSU, ITS, LSU) were constructed, aligned using "MAFFT" v6.502a (Katoh & Standley, 2013) and concatenated afterwards. The aligned matrices are available as *.nex files upon request. Dinophyte phylogenetic analyses were carried out using maximum likelihood (ML) and Bayesian approaches, as described in detail previously (Gottschling et al., 2012) using the resources available from the CIPRES Science Gateway (Miller, Pfeiffer, & Schwartz, 2010). The Bayesian analysis was performed using "MrBayes" v3.2.6 (Ronquist et al., 2012, freely available at http://mrbayes.sourceforge.net/download.php) under the GTR+Γ substitution model and the random-addition-sequence method with 10 replicates. We ran two independent analyses of four chains (one cold and three heated) with 20,000,000 generations, sampled every 1,000th cycle, with an appropriate burn-in (10%) as inferred from the evaluation of the trace files using Tracer v1.5 (http://tree.bio.ed.ac.uk/software/ tracer/). For the ML calculation, the MPI version of "RAxML" v8.2.4 (Stamatakis, 2014, freely available at http://www.exelixis-lab.org/) was applied using the GTR+Γ substitution model under the CAT approximation. We determined the best-scoring ML tree and performed 1,000 non-parametric bootstrap replicates (rapid analysis) in a single step. Statistical support values (LBS: ML bootstrap support, BPP: Bayesian posterior probabilities) were drawn on the resulting, best-scoring tree. To assess whether certain amplicons are nested in other freshwater clades, or belong to yet unidentified lineages deriving from marine clades, we focussed on phylogenetic context of reads with <95% similarity to known species.

3 | RESULTS

In the plankton tow samples analysed using light microscopy, we observed a considerable morphological diversity of dinophytes (Figures 2 and 3), comprising about 10 different species and 2–3 easily distinguishable species per locality. It included photoautotroph taxa such as the relatively easy recognisable *Apocalathium* (Figure 2f,g), *Ceratium* (Figure 3), and *Peridinium* (Figure 2d,h–n), but also heterotroph taxa such as a *Pfiesteria*-like dinophytes (Figure 2e). Occasionally, coccoid developmental stages of dinophytes were also observed (Figure 2b–d).

The rRNA reference alignment of dinophytes was 1,857 + 1,670 + 3,745 bp long and was composed of 965 + 839 + 1,963 parsimony-informative sites (51.8%, mean of 15.0 per terminal taxon) and 5,308 distinct RAxML alignment patterns. Figure 4 shows the best-scoring ML tree (−ln = 232,502.267554), with the internal topology not fully resolved. However, Dinophyceae were monophyletic (99LBS), many nodes were statistically well if not maximally supported, and a number of lineages at high taxonomic

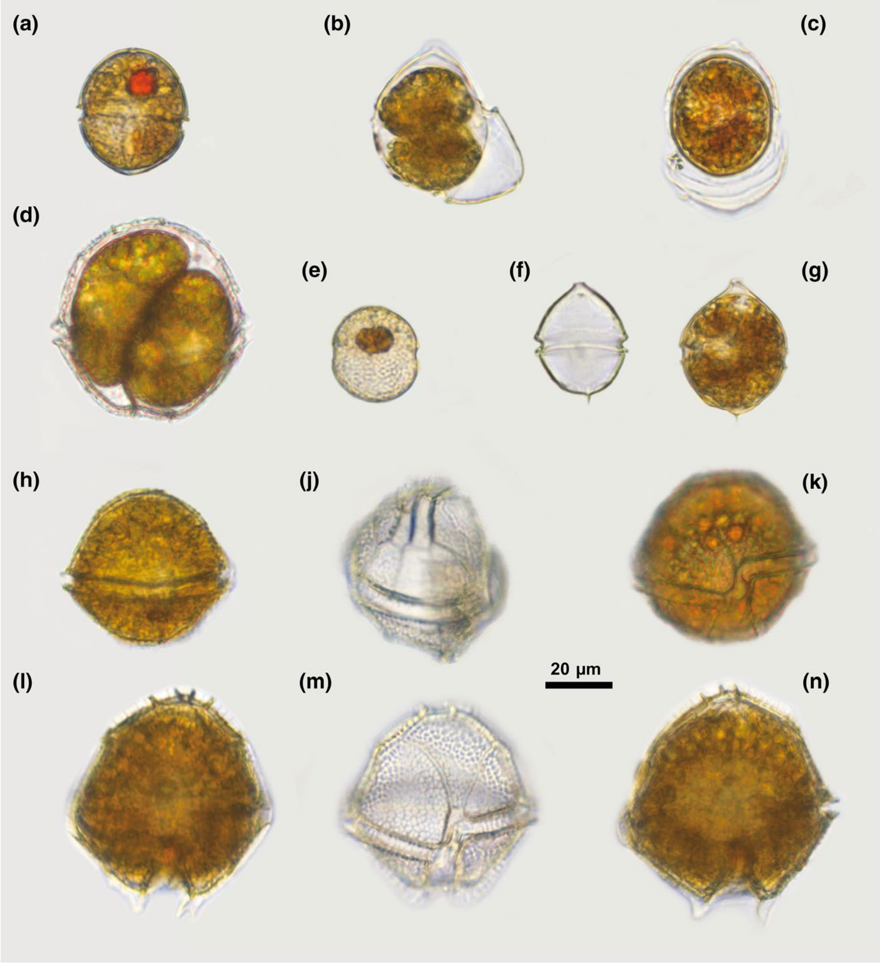

FIGURE 2 Morphological diversity of dinophytes from Bavarian lakes. (a) Theca of unidentified dinophyte (Ammersee [Herrsching]). (b) Developmental stage of unidentified dinophyte (Lech at Apfeldorf). (c) Developmental stage of unidentified dinophyte (Lech at Apfeldorf). (d) Developmental stage of *Peridinium cinctum* (Lake Kochel). (e) Theca of *Pfiesteria*-like dinophyte (Riegsee). (f) Empty theca of *Apocalathium aciculiferum*. (g) Theca of *Apocalathium aciculiferum*. (h) Theca of *Peridinium cinctum* (dorsal view; Starnberg). (j) Empty theca of *Peridinium willei* showing epitheca conformation and areolate surface (Lake Wörth, where high abundance of the species is confirmed by amplicon sequences). (k) Theca of *Peridinium cinctum* (interior ventral view; Lake Kochel). (l) Theca of *Peridinium bipes* (Oster See, where high abundance of the species is confirmed by amplicon sequences). (m) Empty theca of *Peridinium bipes* showing areolate surface (interior ventral view; Hardtbach, where high abundance of the species is confirmed by amplicon sequences). (n) Theca of *Peridinium bipes* (Riegsee). Image adjustments (such as scaling, cropping, white-balancing, colour management; l–n mirrored) were carried out in Photoshop® and Illustrator® (Adobe Systems; Munich, Germany) and images were arranged in QuarkXPress® (Quark Software; Hamburg, Germany)

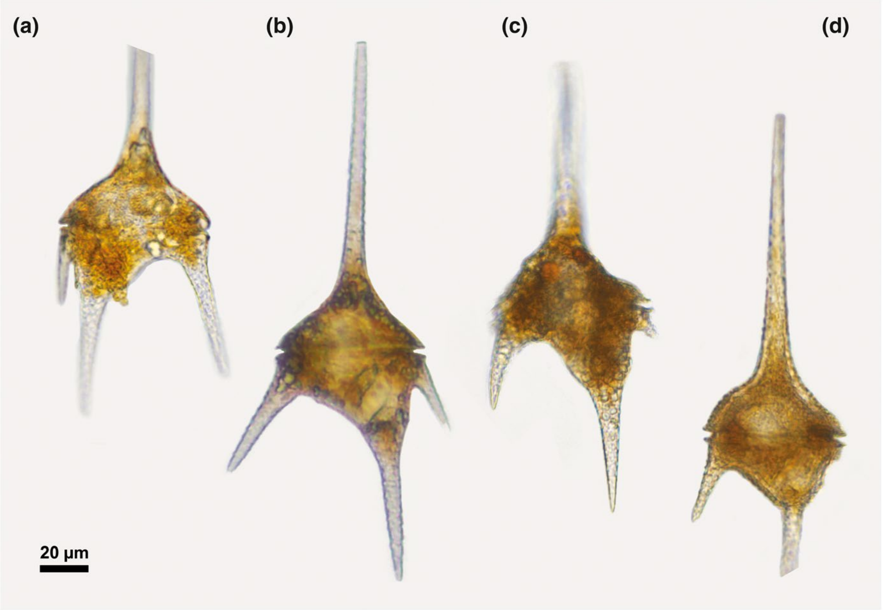

FIGURE 3 Morphological diversity of *Ceratium* from Bavarian lakes. (a) Theca (Ammersee [Diessen], where high abundance of *Ceratium* sp. is confirmed by amplicon sequences). (b) Theca (Starnberger See, where high abundance of *Ceratium* sp. is confirmed by amplicon sequences). (c) Theca (Lake Kochel). (d) Theca (Riegsee, where high abundance of *Ceratium* sp. is confirmed by amplicon sequences)

level such as Dinophysales (100LBS, 1.00BPP), Gonyaulacales, Gymnodiniales (75LBS, 1.00BPP), Peridiniales, Prorocentrales and Suessiales (91LBS, 1.00BPP), but also Amphidomataceae (95LBS, 1.00BPP), Brachidiniaceae (97LBS), and Tovelliaceae (82LBS, 1.00BPP) were recognised. Only 13 of 241 dinophytes (5.4%) were not assigned to any of such lineages. At least 21 distantly related lineages of freshwater dinophytes were scattered over the reference tree, showing a polyphyletic pattern.

In total, we obtained 1,155 OTUs classified as dinophytes, most of which were singletons, doubletons, or occurred in frequencies of <10 (data available upon request). All subsequent work was performed with those 186 dinophyte OTUs with an abundance of ≥10. The vast majority of environmental OTUs were assigned to the photoautotrophic dinophytes and were placed mostly in the freshwater lineages (Figure 5). Large amounts of OTUs were determined as *Apocalathium* (16.1%), *Ceratium* (45.2%), or *Peridinium* (17.7%). However, 18 OTUs (9.7%) were assigned to so far marine taxa including, for example, *Blastodinium*, †*Calcicarpinum* (both Peridiniales), *Dinophysis* (Dinophysales), and *Tripos* (Gonyaulacales). A number of OTUs exhibited considerable phylogenetic distance to known vouchered sequences. These were found among Blastodiniaceae, Borghiellaceae, Gonyaulacaceae and particularly *Pfiesteria*-like dinophytes (Thoracosphaeraceae; not shown).

Figure 6 shows a phylogeny, as a cut-off from the reference tree plus the environmental OTUs, including *Peridinium* with considerable sequence variation. All respective OTUs were placed in one of the established species *Peridinium bipes*, *Peridinium cinctum*, *Peridinium gatunense*, or *Peridinium willei*, each based on statistical support values ≥ 75LBS and .92BPP, respectively. Twenty-one OTUs (11.3%) were (almost) identical to known (reference) sequences, and were therefore reliably assigned to, for example, *Naiadinium polonicum* and *Theleodinium calcisporum* from the Thoracosphaeraceae as well as *Palatinus apiculatus* from the Peridiniopsidaceae (Figure 6) and *Biecheleria brevisulcata* from the Suessiaceae (the latter three species are linked to type material: Table S1). Five OTUs identical to vouchered sequences were assigned to the marine species "*Alexandrium*" *hiranoi*, *Lingulodinium polyedra* (both Gonyaulacales), *Phalacroma rotundatum* (Dinophysales), *Pelagodinium bei*, and *Polarella glacialis* (both Suessiales). Reliable species determinations of our study are summarised in Table 2.

4 | DISCUSSION

Reliable determination of organisms is a necessary prerequisite to explore their spatial and temporal occurrence and to rigorously test hypotheses on their diversification, ecology, and dispersal. Flowering plants, insects, and larger animals are well represented in extensive collections (Krupnick & Kress, 2005; Mayer et al., 2013; Rocha et al., 2014; Steinicke, 2014). In numerous cases, these have also been digitised over the course of the past decade, providing enduring and exact publicly available occurrence data (e.g. GBIF, GBOL, JSTOR, Tropicos®). Such powerful and continuously curated databases are scarce for protists, which are too small for direct observation and need microscopic expertise for examination. However, the problem is recognised, and considerable efforts have been made to build curated sequence databases and reference phylogenetic trees for dinophytes (Del Campo et al., 2018; Elferink et al., 2017; Mordret et al., 2018; Quast et al., 2013) and other microbial taxa.

Our reference tree comprising the known dinophycean sequence diversity is largely in agreement with previous rRNA approaches (Gu et al., 2013) as well as those based on excessive transcriptome sequence data (although using a much smaller taxon sample: Janouškovec et al., 2017; Price & Bhattacharya, 2017). The tree reflects traditional taxonomic concepts based on morphology better than previous (frequently single-locus) approaches, and we are able to recognise major monophyletic lineages of dinophytes. Trees such as that obtained in this study serve as a taxonomic scaffold for the systematic placement of newly generated dinophyte sequences, even if they are short as it is in the case of next-generation

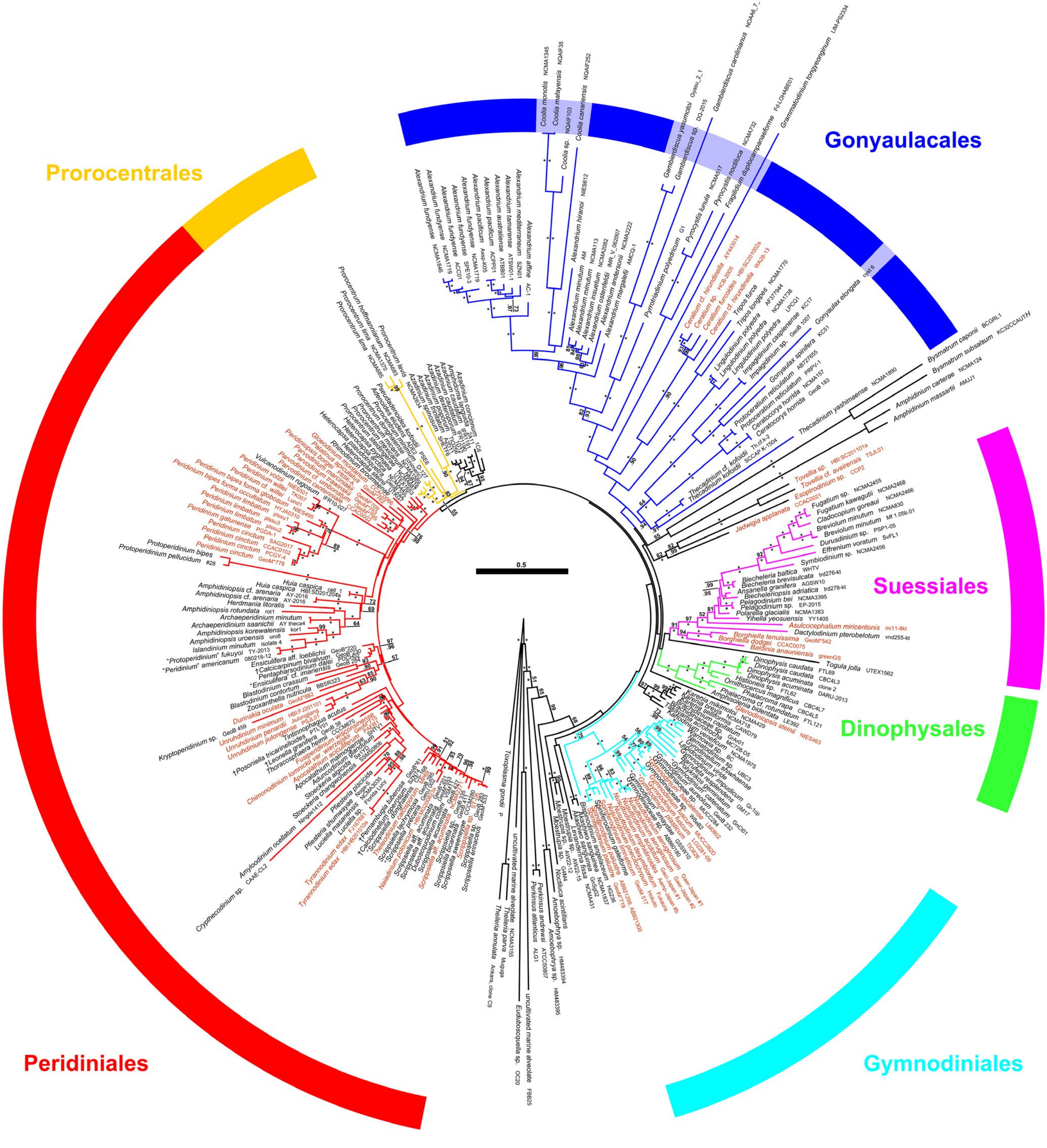

FIGURE 4 A molecular reference tree recognising major groups of dinophytes. Maximum likelihood (ML) tree of 241 systematically representative dinophyte sequences (with strain number information) as inferred from a rRNA nucleotide alignment (3,767 parsimony-informative positions). Numbers on branches are ML bootstrap (above) and Bayesian support values (below) for the clusters (asterisks indicate maximal support values, values under 50 and 0.90, respectively, are not shown). Colour coding: black, marine OTU; red, freshwater OTU

sequencing amplicons. Our approach to determine dinophyte species using reference trees as inferred from multi-locus rRNA alignments is proven successful to a certain degree, at least for samples from the freshwater environment (providing also some new dinophyte records for Bavaria: Table 2), but also from the marine realm (Elferink et al., 2017; Wohlrab et al., 2018). Only if taxonomically curated and kept up to date can the rRNA alignments also serve as a tool for placing other unknown protist environmental sequences into the Tree of Life. However, such rRNA-based reference databases are still in their infancy for most other protist lineages.

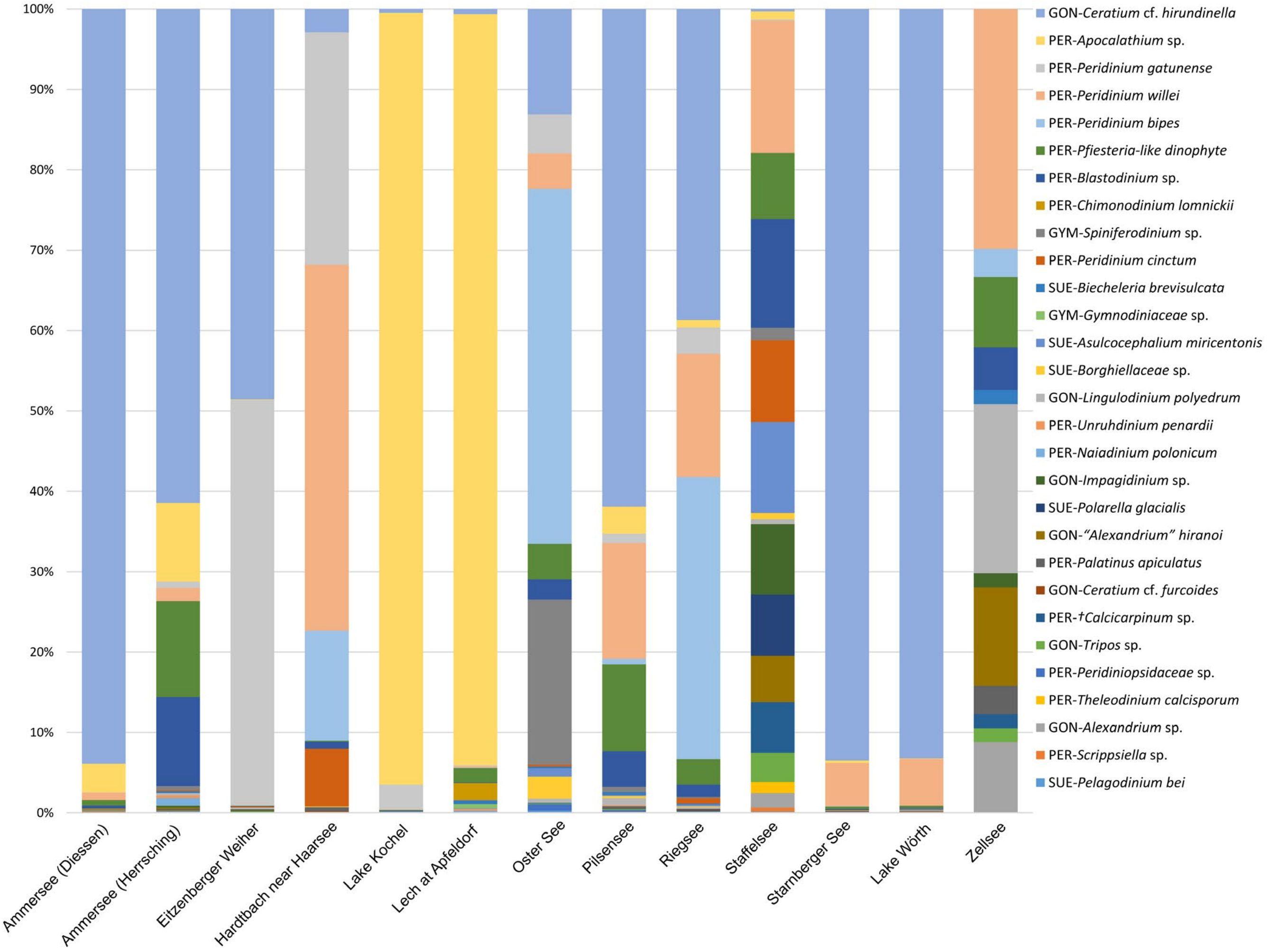

FIGURE 5 Variation in dinophyte abundances in Bavarian lakes. Abundance of environmental OTUs in the different collection sites, with names based on phylogenetic placement. GON, Gonyaulacales; GYM, Gymnodiniales; PER, Peridiniales; SUE, Suessiales

The most frequently encountered taxa correspond to those that are known from morphological surveys in Bavaria, which comprise *Apocalathium* (Mauch et al., 2003; Mischke, Riedmüller, Hoehn, Deneke, & Nixdorf, 2015; Schaumburg, 1996), *Ceratium* (Mauch et al., 2003; Raeder, 1990; Schaumburg, 1996; Schaumburg & Hehl, 2001), and *Peridinium* (Fröbrich, Mangelsdorf, Schauer, Streil, & Wachter, 1977; Mauch et al., 2003; Mischke et al., 2015; Raeder, 1990; Schaumburg, 1996), summing-up to almost 80% of all identified OTUs. Our sequence-based findings seem to support our own morphological confirmations in the samples that we have investigated, and such combinatorial approaches are needed (Medinger et al., 2010; Rimet, Vasselon, A-Keszte, & Bouchez, 2018; Mora et al., 2019), as long as the taxonomic impediment and uncertainty with taxon determination continue to exist. However, when comparing morphological and genetic species diversity assessments, one has to take into account the morphological variation and the cryptic speciation. On a broader spectrum, these might namely either seemingly corroborate or contradict each other. Thus, having a broad knowledge on the biology of investigated microbiota is of great importance, as for the example given below.

The Peridiniaceae are one of the most important groups of freshwater dinophytes and may comprise about a dozen species (Gottschling, Kretschmann, & Žerdoner Čalasan, 2017; Moestrup & Calado, 2018), half of which are already known from molecular DNA sequences (Table S1). In rRNA sequences, they show a considerable variation even within species (Izquierdo López, Kretschmann, Žerdoner Čalasan, & Gottschling, 2018) leading to long branches in phylogenetic trees (Gu et al., 2013). This intraspecific variability, in combination with a reliable taxonomy at the species level, makes the Peridiniaceae a good example of how species determination can be effective using environmental amplicon sequences. We are able to assign all peridiniacean OTUs gained in this study to an established species of *Peridinium* (at least as long as no cryptic speciation has been documented in this lineage). *Peridinium bipes* exhibits a distinct morphology, but our approach allows for efficient differentiation between, for example, *P. cinctum* and *P. willei* that are challenging to tell apart using just light microscopy in monitoring studies. However, sequence variation might be lower (and closely related species may exhibit identical SSU sequences) in other groups such as *Scrippsiella*, and it is therefore not always possible to differentiate between

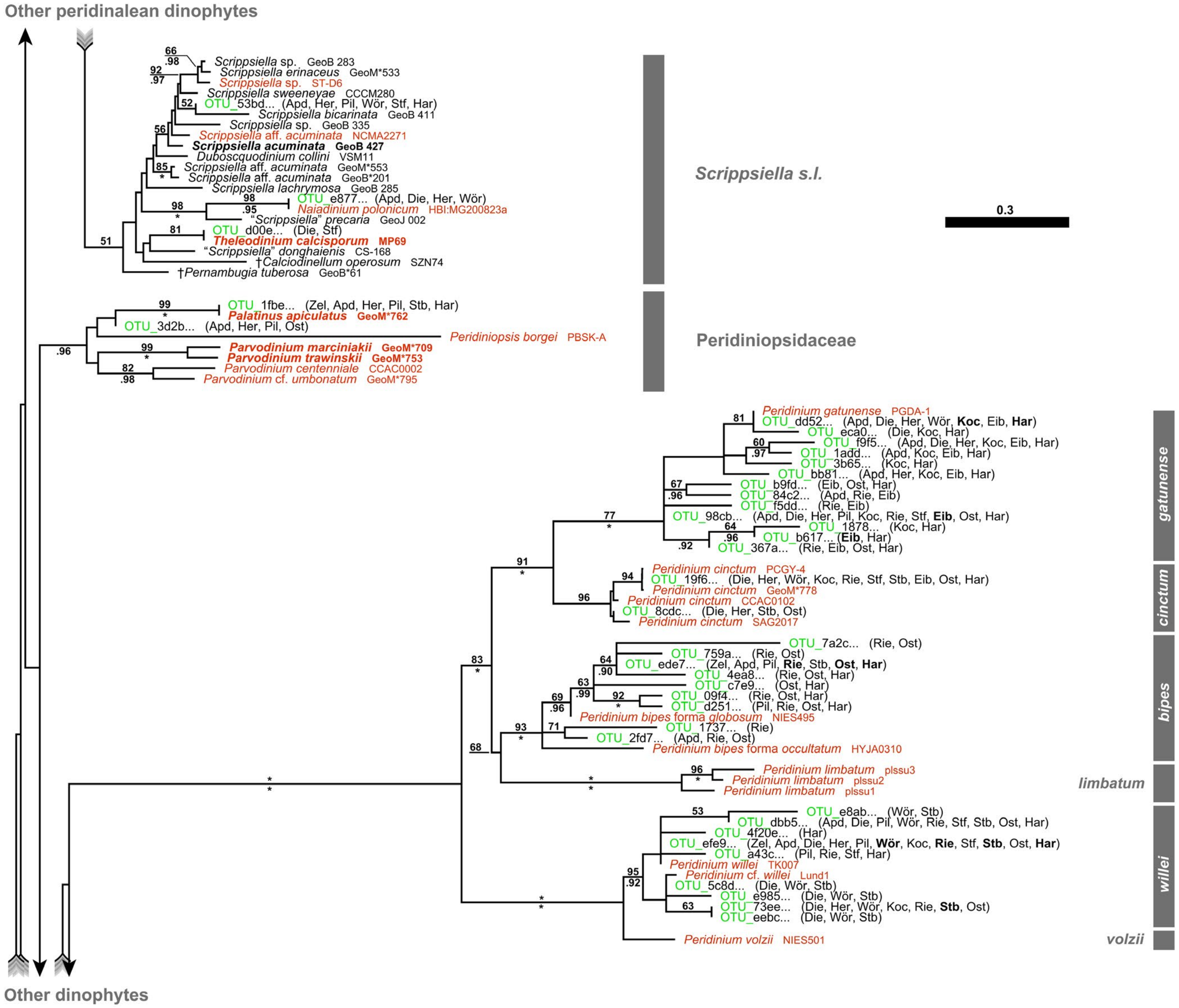

FIGURE 6 Reliable placement of selected Bavarian dinophycean OTUs on a phylogenetic reference tree. Maximum likelihood (ML) tree of 241 systematically representative dinophyte sequences plus 186 environmental OTU sequences (cut out of freshwater Peridiniaceae as well as *Scrippsiella s.l.* and Peridiniopsidaceae; black: marine, blue: freshwater). Numbers on branches are ML bootstrap (above) and Bayesian support values (below) for the clusters (asterisks indicate maximal support values, values under 50 and .90, respectively, are not shown). Colour/font coding: black, marine OTU; red, freshwater OTU; green, environmental OTU; bold, OTU represented by type material. Localities of environmental OTUs are indicated: Apd, Lech at Apfeldorf; Die, Ammersee (Diessen); Eib, Eitzenberger Weiher; Har, Hardtbach near Haarsee; Her, Ammersee (Herrsching); Koc, Lake Kochel; Ost, Oster See; Pil, Pilsensee; Rie, Riegsee; Stb, Starnberger See; Stf, Staffelsee; Wör, Lake Wörth; Zel, Zellsee. Note that species of *Peridinium* show intraspecific variability regarding rRNA sequences (Izquierdo López et al., 2018) and that all species, to which OTUs are assigned (i.e. *Peridinium bipes*, *Peridinium cinctum*, *Peridinium gatunense*, *Peridinium willei*), show high LBS statistical support ≥ 80. Further note that we found OTUs being identical to known sequences of, for example, *Naiadinium polonicum*, *Palatinus apiculatus*, and *Theleodinium calcisporum* (the latter two represented by type material)

species based on one particular molecular locus (Söhner et al., 2012; Žerdoner Čalasan, Kretschmann, Filipowicz, et al., 2019). Because of the large heterogeneity of evolutionary rates in the rRNA operon (Gu et al., 2013; Saldarriaga Echavarría, Taylor, Cavalier-Smith, Menden-Deuerd, & Keeling, 2004; Stern et al., 2012; Žerdoner Čalasan, Kretschmann, & Gottschling, 2019), there is no such thing as an universal *DNA barcode* for all constituting lineages, and other loci including ITS or LSU may prove better suited for reliable species determination of such organism sets using HTS.

The potential identification of few marine taxa in the freshwater environment based on sequences comes as a surprise. Contamination due to the applied methods is very unlikely, as these species are not cultivated in our laboratories, but the incomplete assessment of freshwater dinophytes in DNA reference databases may account for the ostensible detection of marine taxa in our samples. However, some dinophytes have colonised freshwater from marine habitats relatively recently (Žerdoner Čalasan, Kretschmann, & Gottschling, 2019). Furthermore, different rates of the rRNA operon evolution

TABLE 2 Occurrence of selected dinophytes that could be reliably determined to the species level based on environmental rRNA sequence data. DIN, Dinophysales; GON, Gonyaulacales; GYM, Gymnodiniales; PER, Peridiniales; SUE, Suessiales. Taxa not listed in Mauch et al. (2003) are indicated with an asterisk

Species (taxonomic assignment)	Ammersee (Diessen)	Ammersee (Herrsching)	Eitzenberger Weiher	Hardtbach near Haarsee	Lake Kochel	Lech at Apfeldorf	Oster See	Pilsensee	Riegsee	Staffelsee	Starnberger See	Lake Wörth	Zellsee
"*Alexandrium*" *hiranoi** (GON)										+			
*Asulcocephalium miricentonis** (SUE)					+		+			+			
Biecheleria brevisulcata (SUE)	+					+							
Chimonodinium lomnickii (PER)	+					+							
*Lingulodinium polyedra** (GON)													+
Naiadinium polonicum (PER)		+				+							
Palatinus apiculatus (PER)											+		
Peridinium bipes (PER)				+			+		+				
Peridinium cinctum (PER)				+					+	+			
Peridinium gatunense (PER)			+	+	+	+	+		+				
Peridinium willei (PER)	+	+		+	+	+	+	+	+	+	+	+	+
*Phalacroma rotundatum** (DIN)										+			
*Polarella glacialis** (SUE)										+			
Spiniferodinium palustre (GYM)							+	+					
*Theleodinium calcisporum** (PER)										+			
Unruhdinium penardii (PER)						+							

may impact the phylogenetic position of certain reads. For instance, *Gonyaulax clevei* has been reported from German lakes (Hickel & Pollhinger, 1986), but has never been sequenced. It is possible that those sequences once gained will group with marine taxa such as species of *Lingulodinium*. Moreover, a few organisms are physiologically able to successfully overcome the physiological barrier between the oceans and freshwater habitats (Pokorný, 2009) and establish new populations. As living models, species such as *Huia caspica* and *Kolkwitziella acuta* might be a key in this respect, as they are found in both marine and freshwater habitats (Gu, Mertens, & Liu, 2016; Mertens et al., 2015). By any means, the precise biological role, and the overall biological activity, has to be worked out for marine taxa in freshwater environments. Future research should rigorously use physical specimens and preferably living strains. Moreover, dinophyte species richness in Bavarian lakes may be greater than that previously reported in the literature based on our genetic analysis. Particularly, the considerable number of so far unknown *Pfiesteria*-like sequences (Burkholder & Marshall, 2012; Calado, Craveiro, Daugbjerg, & Moestrup, 2009; Litaker et al., 2005) is impressive. A targeted search in future will assess this diversity and address questions, such as whether these are already accepted species without hitherto known DNA sequence information or even new species.

In the microbial world, the importance of DNA sequences linked to type material cannot be overestimated. In this respect, our approach to place OTUs on a reference tree using curated and vouchered representatives has been proven successful with the documentation of sequences identical to *Biecheleria brevisulcata* (Suessiaceae), *Palatinus apiculatus* (Peridiniopsidaceae), and *Theleodinium calcisporum* (Thoracosphaeraceae). The first and latter species have been described only a few years ago (Craveiro et al., 2013; Takahashi, Sarai, & Iwataki, 2014), but *Palatinus apiculatus* refers to a historical name from the 19th century (Ehrenberg, 1838). Usually, such names are not linked to DNA sequence information, but the application of epitypification has made the determination of such species unambiguous (Kretschmann, Žerdoner Čalasan, Kusber, & Gottschling, 2018). This strategy has not—to the best of our knowledge—been applied for the species of *Ceratium*, which makes the determination of numerous OTUs gained in this study at the species level vague. Once identified, strategic taxonomic clarifications of target organisms are possible and may refer to names such as *Ceratium macroceras* and *Ceratium tetraceros*, both being described from Bavaria (Schrank, 1793, 1802). However, as the potential intraspecific and interspecific variability together with possible cryptic speciation and morphological variability of *Ceratium* are currently unknown, the taxonomic clarification must be performed with a great caution, taking all such aspects into account. The taxonomic tool of epitypification provided by the International Code of Nomenclature for algae, fungi, and plants (Turland et al., 2018) is still not regularly used in phycology, despite being of great importance for stable taxonomy. Such approaches must be carried out, because the resolution power of all available HTS methods, as exemplified in this study, always comes down to the basic necessity of a reliable application of names.

Curated contemporary reference databases leave further room for improvements. For example, a number of the OTU sequences obtained in this study were assigned to *Scrippsiella* by the SILVA (Quast et al., 2013) reference, which is a predominantly marine dinophyte lineage. Such entries misleadingly refer to it as "*Scrippsiella*" *hangoei*, which is a name being classified today under *Apocalathium*. Therefore, the correct species name for the OTUs is (probably) freshwater *Apocalathium aciculiferum*. It might be only a matter of time until this particular taxonomic confusion is corrected for future releases of the SILVA databases, as this error has been already corrected in the DINOREF database (Mordret et al., 2018). The latter, in turn, relies on SSU reference sequence data only and is therefore unable to place environmental OTUs of studies using LSU (Elferink et al., 2017) and/or ITS sequences (Lutz, McCutcheon, McQuaid, & Benning, 2018). Concomitantly, taxa such as *Peridiniella*, *Sphaerodinium* and *Tyrannodinium*, of which only LSU sequences are known at present, are subsequently undetectable using DINOREF (however, †*Leonella* and †*Posoniella* are also missing, although SSU reference sequences are already available). The variety of sequencing approaches thus requires a database that assures both, provision of extensive rRNA sequence information and taxonomic reliability. In this respect, the indication of sequences that have been gained from type material is also important (Pawlowski et al., 2012) and needs to be added in DINOREF. Moreover, curation of the DINOREF database has not prevented the use of names such as 'Goniodomataceae', which have been nomenclaturally rejected (Elbrächter & Gottschling, 2015; Prud'homme Van Reine, 2017).

Our approach to detect dinophytes in Bavarian lakes is powerful and will lay the basis for solid information on which species are widely distributed and abundant, and which species are rarer and represent rather endemic entities with narrower distributions. If occurrences of dinophyte species correlate with environmental traits, then improved species circumscriptions also taking their ecological niche into account are possible. With our project, we may start to understand not only *that* a certain species occurs in a given freshwater habitat, but also *why*. Our example of field mapping, and the pursued predictability of freshwater dinophyte occurrences, has thus a great potential to serve as a model for other taxonomic groups and / or the investigation of similar and alternative environments in other parts of the world.

ACKNOWLEDGMENTS

We are thankful to Susanne S. Renner (Munich) for funding the Postdoc position of JC, Nancy Kühne (Bremerhaven) and Martina Silber (Munich) for technical assistance in the molecular labs, Stephanie Elferink (Bremerhaven) for support with the sequence analysis and Mia Stockenreiter (Munich) for helpful discussions. Sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB34585 (https://www.ebi.ac.uk/ena/data/view/PRJEB34585) and all other data will be available in ZENEDO (https://zenodo.org) DOI: 10.5281/zenodo.3462171. The authors declare that the

research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORCID

Marc Gottschling https://orcid.org/0000-0002-4381-8051

Uwe John https://orcid.org/0000-0002-1297-4086

REFERENCES

Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., & Polz, M. F. (2005). PCR-induced sequence artifacts and bias: Insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and Environmental Microbiology*, *71*, 8966–8969. https://doi.org/10.1128/aem.71.12.8966-8969.2005

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403–410. https://doi.org/10.1016/s0022-2836(05)80360-2

Bass, D., & Boenigk, J. (2011). Everything is everywhere: A twenty-first century de-/reconstruction with respect to protists. In D. Fontaneto (Ed.), *Biogeography of microscopic organisms. Is everything small everywhere?* (pp. 88–110). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9780511974878

Bass, D., Richards, T. A., Matthai, L., Marsh, V., & Cavalier-Smith, T. (2007). DNA evidence for global dispersal and probable endemicity of Protozoa. *BMC Evolutionary Biology*, *7*, 162. https://doi.org/10.1186/1471-2148-7-162

Bates, S. T., Clemente, J. C., Flores, G. E., Walters, W. A., Parfrey, L. W., Knight, R., & Fierer, N. (2013). Global biogeography of highly diverse protistan communities in soil. *ISME Journal*, *7*, 652–659. https://doi.org/10.1038/ismej.2012.147

Baumeister, W. (1957). Neue Dinococcalen aus dem niederbayerischen Hügelland zwischen Isar und Inn (I). *Archiv für Protiskenkunde*, *102*, 21–43.

Bik, H. M., Sung, W., De Ley, P., Baldwin, J. G., Sharma, J., Rocha-Olivares, A., & Thomas, W. K. (2012). Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. *Molecular Ecology*, *21*, 1048–1059. https://doi.org/10.1111/j.1365-294x.2011.05297.x

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Burkholder, J. M., & Marshall, H. G. (2012). Toxigenic *Pfiesteria* species—Updates on biology, ecology, toxins, and impacts. *Harmful Algae*, *14*, 196–230. https://doi.org/10.1016/j.hal.2011.10.022

Calado, A. J., Craveiro, S. C., Daugbjerg, N., & Moestrup, Ø. (2009). Description of *Tyrannodinium* gen. nov., a freshwater dinoflagellate closely related to the marine *Pfiesteria*-like species. *Journal of Phycology*, *45*, 1195–1205. https://doi.org/10.1111/j.1529-8817.2009.00735.x

Caron, D. A. (2009). Past president's address: Protistan biogeography: Why all the fuss? *Journal of Eukaryotic Microbiology*, *56*, 105–112. https://doi.org/10.1111/j.1550-7408.2008.00381.x

Coleman, A. W. (2001). Biogeography and speciation in the *Pandorina*/*Volvulina* (Chlorophyta) superclade. *Journal of Phycology*, *37*, 836–851. https://doi.org/10.1046/j.1529-8817.2001.01043.x

Craveiro, S. C., Pandeirada, M. S., Daugbjerg, N., Moestrup, Ø., & Calado, A. J. (2013). Ultrastructure and phylogeny of *Theleodinium calcisporum* gen. et sp. nov., a freshwater dinoflagellate that produces calcareous cysts. *Phycologia*, *52*, 488–507. https://doi.org/10.2216/13-152.1

Cuvelier, M. L., Allen, A. E., Monier, A., McCrow, J. P., Messié, M., Tringe, S. G., ... Worden, A. Z. (2010). Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 14679–14684. https://doi.org/10.1073/pnas.1001665107

Del Campo, J., Kolisko, M., Boscaro, V., Santoferrara, L. F., Nenarokov, S., Massana, R., ... Parfrey, L. W. (2018). EukRef: Phylogenetic curation of ribosomal RNA to enhance understanding of eukaryotic diversity and distribution. *PLoS Biology*, *16*, e2005849. https://doi.org/10.1371/journal.pbio.2005849

Dunthorn, M., Otto, J., Berger, S. A., Stamatakis, A., Mahé, F., Romac, S., ... Biomarks, C. (2014). Placing environmental next-generation sequencing amplicons from microbial eukaryotes into a phylogenetic context. *Molecular Biology and Evolution*, *31*, 993–1009. https://doi.org/10.1093/molbev/msu055

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, *27*, 2194–2200. https://doi.org/10.1093/bioinformatics/btr381

Edwards, R. A., Rodriguez-Brito, B., Wegley, L., Haynes, M., Breitbart, M., Peterson, D. M., ... Rohwer, F. (2006). Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genomics*, *7*, 57. https://doi.org/10.1186/1471-2164-7-57

Ehrenberg, C. G. (1838). Zwölfte Familie: Kranzthierchen. In *Die Infusionsthierchen als vollko mmene Organismen. Ein Blick in das Tiefere organische Leben der Natur* (pp. 249–259, pl. XXII). Leipzig, Germany: Leopold Voss.

Elbrächter, M., & Gottschling, M. (2015). 2383) Proposal to reject the name Goniodomataceae (Dinophyceae). *Taxon*, *64*, 1052–1053. https://doi.org/10.12705/645.19

Elferink, S., Neuhaus, S., Wohlrab, S., Toebe, K., Voss, D., Gottschling, M., ... John, U. (2017). Molecular diversity patterns among various phytoplankton size-fractions in West Greenland in late summer. *Deep-Sea Research Part I-Oceanographic Research Papers*, *121*, 54–69. https://doi.org/10.1016/j.dsr.2016.11.002

Fenchel, T., & Finlay, B. J. (2004). The ubiquity of small species: Patterns of local and global diversity. *BioScience*, *54*, 777–784. https://doi.org/10.1641/0006-3568(2004)054[0777:tuossp]2.0.co;2

Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species. *Science*, *296*, 1061–1063. https://doi.org/10.1126/science.1070710

Foissner, W. (2008). Protist diversity and distribution: Some basic considerations. *Biodiversity and Conservation*, *17*, 235–242. https://doi.org/10.1007/s10531-007-9248-5

Foissner, W. (2011). Dispersal of protists: The role of cysts and human introductions. In D. Fontaneto (Ed.), *Biogeography of microscopic organisms. Is everything small everywhere?* (pp. 61–87). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9780511974878

Fourtanier, E., & Kociolek, J. P. (2009). Catalogue of diatom names. Part 1. Introduction and Bibliography. *Occasional Papers of the California Academy of Sciences*, *156*, 1–168.

Fröbrich, G., Mangelsdorf, J., Schauer, T., Streil, J., & Wachter, H. (1977). Gewässerkundliche Studie über sechs Seen bei Füssen im Allgäu. *Schriftenreihe des Bayerischen Landesamtes für Wasserwirtschaft*, *3*, 1–176.

Gottschling, M., Kretschmann, J., & Žerdoner Čalasan, A. (2017). Description of Peridiniopsidaceae, fam. nov (Peridiniales, Dinophyceae). *Phytotaxa*, *299*, 293–296. https://doi.org/10.11646/phytotaxa.299.2.16

Gottschling, M., Söhner, S., Zinßmeister, C., John, U., Plötner, J., Schweikert, M., ... Elbrächter, M. (2012). Delimitation of the Thoracosphaeraceae (Dinophyceae), including the calcareous dinoflagellates, based on large amounts of ribosomal RNA sequence data. *Protist*, *163*, 15–24. https://doi.org/10.1016/j.protis.2011.06.003

Gu, H., Kirsch, M., Zinßmeister, C., Söhner, S., Meier, K. J. S., Liu, T., & Gottschling, M. (2013). Waking the dead: Morphological and molecular characterization of extant †*Posoniella tricarinelloides* (Thoracosphaeraceae, Dinophyceae). *Protist*, *164*, 583–597. https://doi.org/10.1016/j.protis.2013.06.001

Gu, H., Mertens, K. N., & Liu, T. (2016). *Huia caspica* gen. & comb. nov., a dinoflagellate species that recently crossed the marine-freshwater boundary. *Phycological Research*, *64*, 251–258. https://doi.org/10.1111/pre.12146

Hariganeya, N., Tanimoto, Y., Yamaguchi, H., Nishimura, T., Tawong, W., Sakanari, H., … Adachi, M. (2013). Quantitative PCR method for enumeration of cells of cryptic species of the toxic marine dinoflagellate *Ostreopsis* spp. in coastal waters of Japan. *PLoS ONE*, *8*, e57627.

Hickel, B., & Pollhinger, U. (1986). On the morphology and ecology of *Gonyaulax apiculata* (Penard) Entz from the Selenter See (West Germany). *Archiv für Hydrobiologie Supplement*, *73*, 227–232.

Höll, K. (1928). *Oekologie der Peridineen. Studien über den Einfluß chemischer und physikalischer Faktoren auf die Verbreitung der Dinoflagellaten im Süßwasser.* Jena: Fischer.

Hollingsworth, P. M. (2011). Refining the DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 19451–19452. https://doi.org/10.1073/pnas.1116812108

Izquierdo López, A., Kretschmann, J., Žerdoner Čalasan, A., & Gottschling, M. (2018). The many faces of *Peridinium cinctum*: Morphological and molecular variability in a common dinophyte. *European Journal of Phycology*, *53*, 156–165. https://doi.org/10.1080/09670262.2017.1397198

Janouškovec, J., Gavelis, G. S., Burki, F., Dinh, D., Bachvaroff, T. R., Gornik, S. G., … Saldarriaga Echavarría, J. F. (2017). Major transitions in dinoflagellate evolution unveiled by phylotranscriptomics. *Proceedings of the National Academy of Science of the United States of America*, *114*, E171–E180. https://doi.org/10.1073/pnas.1614842114

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

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

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., … Worden, A. Z. (2014). The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biology*, *12*, e1001889. https://doi.org/10.1371/journal.pbio.1001889

Kohli, G. S., Neilan, B. A., Brown, M. V., Hoppenrath, M., & Murray, S. A. (2014). Cob gene pyrosequencing enables characterization of benthic dinoflagellate diversity and biogeography. *Environmental Microbiology*, *16*, 467–485. https://doi.org/10.1111/1462-2920.12275

Kretschmann, J., Filipowicz, N. H., Owsianny, P. M., Zinßmeister, C., & Gottschling, M. (2015). Taxonomic clarification of the unusual dinophyte *Gymnodinium limneticum* Wołosz. (Gymnodiniaceae) from the Tatra Mountains. *Protist*, *166*, 621–637. https://doi.org/10.1016/j.protis.2015.09.002

Kretschmann, J., Owsianny, P. M., Žerdoner Čalasan, A., & Gottschling, M. (2018). The hot spot in a cold environment: Puzzling *Parvodinium* (Peridiniopsidaceae, Peridiniales) from the Polish Tatra Mountains. *Protist*, *169*, 206–230. https://doi.org/10.1016/j.protis.2018.02.004

Kretschmann, J., Žerdoner Čalasan, A., Kusber, W.-H., & Gottschling, M. (2018). Still curling after all these years: *Glenodinium apiculatum* Ehrenb. (Peridiniales, Dinophyceae) repeatedly found at its type locality in Berlin (Germany). *Systematics and Biodiversity*, *16*, 200–209. https://doi.org/10.1080/14772000.2017.1375045

Krupnick, G. A., & Kress, W. J. (2005). *Plant conservation: A natural history approach*. Chicago, IL: University of Chicago Press.

Lindemann, E. B. L. W. (1920). Untersuchungen über Süßwasserperidineen und ihre Variationsformen II. *Archiv für Naturgeschichte*, *84*, 121–194.

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., & Atkinson, A. (2013). Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PLoS ONE*, *8*, e81327.

Litaker, R. W., Steidinger, K. A., Mason, P. L., Landsberg, J. H., Shields, J. D., Reece, K. S., … Tester, P. A. (2005). The reclassification of *Pfiesteria shumwayae* (Dinophyceae): *Pseudopfiesteria*, gen. nov. *Journal of Phycology*, *41*, 643–651. https://doi.org/10.1111/(issn)1529-8817

Logares, R., Rengefors, K., Kremp, A., Shalchian-Tabrizi, K., Boltovskoy, A., Tengs, T., … Klaveness, D. (2007). Phenotypically different microalgal morphospecies with identical ribosomal DNA: A case of rapid adaptive evolution? *Microbial Ecology*, *53*, 549–561. https://doi.org/10.1007/s00248-006-9088-y

Lutz, S., McCutcheon, J., McQuaid, J. B., & Benning, L. G. (2018). The diversity of ice algal communities on the Greenland Ice Sheet as revealed by oligotyping. *Microbial Genomics*, *4*, 159.

Mahé, F., Rognes, T., Quince, C., De Vargas, C., & Dunthorn, M. (2015). Swarm v2: Highly-scalable and high-resolution amplicon clustering. *PeerJ*, *3*, e1420.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*, 10–12.

Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., … De Vargas, C. (2015). Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environmental Microbiology*, *17*, 4035–4049. https://doi.org/10.1111/1462-2920.12955

Mauch, E., Schmedtje, U., Maetze, A., & Fischer, F. (2003). Taxaliste der Gewässerorganismen Deutschlands zur Kodierung biologischer Befunde. *Informationsberichte des Bayerischen Landesamtes für Wasserwirtschaft*, 1–367.

Mayer, G. C., Coyne, J. A., Losos, J. B., Foufopoulos, J., Shubin, N., Futuyma, D. J., … Edwards, S. V. (2013). Museums' role: Increasing knowledge. *Science*, *339*, 1148–1149. https://doi.org/10.1126/science.339.6124.1148-b

Medinger, R., Nolte, V., Pandey, R. V., Jost, S., Ottenwälder, B., Schlötterer, C., & Boenigk, J. (2010). Diversity in a hidden world: Potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. *Molecular Ecology*, *19*, 32–40. https://doi.org/10.1111/j.1365-294x.2009.04478.x

Mertens, K. N., Rengefors, K., Moestrup, Ø., & Ellegaard, M. (2012). A review of recent freshwater dinoflagellate cysts: Taxonomy, phylogeny, ecology and palaeocology. *Phycologia*, *51*, 612–619. https://doi.org/10.2216/11-89.1

Mertens, K. N., Takano, Y., Yamaguchi, A., Gu, H., Bogus, K., Kremp, A., … Matsuoka, K. (2015). The molecular characterization of the enigmatic dinoflagellate *Kolkwitziella acuta* reveals an affinity to the *Excentrica* section of the genus *Protoperidinium*. *Systematics and Biodiversity*, *13*, 829–844.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, US–LA, pp. 1–8. https://doi.org/10.1109/GCE.2010.5676129

Mischke, U., Riedmüller, U., Hoehn, E., Deneke, R., & Nixdorf, B. (2015). Handbuch für die Seenbewertung mittels Plankton – Phyto-See-Index (Teil A) und PhytoLoss-Modul Zooplankton (Teil B), 1–143.

Moestrup, Ø., & Calado, A. J. (2018). *Dinophyceae*. Berlin: Springer.

Moestrup, Ø., Lindberg, K., & Daugbjerg, N. (2009). Studies on woloszynskioid dinoflagellates IV: The genus *Biecheleria* gen. nov. *Phycological Research*, *57*, 203–220. https://doi.org/10.1111/j.1440-1835.2009.00540.x

Mora, D., Abarca, N., Proft, S., Grau, J. H., Enke, N., Carmona, J., … Zimmermann, J. (2019). Morphology and metabarcoding: A test with stream diatoms from Mexico highlights the complementarity of identification methods. *Freshwater Science*, *38*, 448–464.

Mordret, S., Piredda, R., Vaulot, D., Montresor, M., Kooistra, W. H. C. F., & Sarno, D. (2018). DINOREF: A curated dinoflagellate (Dinophyceae) reference database for the 18S rRNA gene. *Molecular Ecology Resources*, *18*, 974–987. https://doi.org/10.1111/1755-0998.12781

Nürnberg, G. K. (1996). Trophic state of clear and colored, soft- and hardwater lakes with special consideration of nutrients, anoxia, phytoplankton and fish. *Journal of Lake and Reservoir Management*, *12*, 432–447. https://doi.org/10.1080/07438149609354283

O'Dwyer, J. P., Kembel, S. W., & Sharpton, T. J. (2015). Backbones of evolutionary history test biodiversity theory for microbes. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 8356–8361. https://doi.org/10.1073/pnas.1419341112

OECD (1982). *Eutrophication of waters. Monitoring, assessment and control*. Paris: OECD.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., … De Vargas, C. (2012). CBOL Protist Working Group: Barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS Biology*, *10*, e1001419. https://doi.org/10.1371/journal.pbio.1001419

Peršoh, D. (2015). Plant-associated fungal communities in the light of meta'omics. *Fungal Diversity*, *75*, 1–25. https://doi.org/10.1007/s13225-015-0334-9

Pfiester, L. A., & Anderson, D. M. (1987). Dinoflagellate reproduction. In F. J. R. Taylor (Ed.), *The biology of dinoflagellates* (pp. 611–648). Oxford: Blackwell.

Pokorný, P. (2009). Evolution and function of freshwater ecosystems. In V. Cílek (Ed.), *Earth system: History and natural variability* (pp. 118–137). Oxford: EOLSS.

Price, D. C., & Bhattacharya, D. (2017). Robust Dinoflagellata phylogeny inferred from public transcriptome databases. *Journal of Phycology*, *53*, 725–729. https://doi.org/10.1111/jpy.12529

Prud'homme Van Reine, W. F. (2017). Report of the Nomenclature Committee for Algae: 15. *Taxon*, *66*, 191–192. https://doi.org/10.12705/661.16

QGIS Development Team (2019). QGIS Geographic Information System. Open Source Geospatial Foundation Project. Retrieved from http://qgis.osgeo.org

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*, D590–D596.

Raeder, U. B. C. (1990). Vergleichende Untersuchungen der Phytoplanktonsukzession in 15 Seen des Osterseengebiets. PhD thesis, Technische Universität München, München.

Read, B. A., Kegel, J., Klute, M. J., Kuo, A., Lefebvre, S. C., Maumus, F., … Grigoriev, I. V. (2013). Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature*, *499*, 209–213. https://doi.org/10.1038/nature12221

Rigden, D. J., Fernandez-Suarez, X. M., & Galperin, M. Y. (2016). The 2016 database issue of Nucleic Acids Research and an updated molecular biology database collection. *Nucleic Acids Research*, *44*, D1–D6. https://doi.org/10.1093/nar/gkv1356

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 and Evolution*, *18*, 51–62. https://doi.org/10.1007/s13127-018-0359-5

Rocha, L. A., Aleixo, A., Allen, G., Almeda, F., Baldwin, C. C., Barclay, M. V. L., … Witt, C. C. (2014). Specimen collection: An essential tool. *Science*, *344*, 814–815. https://doi.org/10.1126/science.344.6186.814

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., … Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, *61*, 539–542. https://doi.org/10.1093/sysbio/sys029

Saldarriaga Echavarría, J. F., Taylor, F. J. R., Cavalier-Smith, T., Menden-Deuerd, S., & Keeling, P. J. (2004). Molecular data and the evolutionary history of dinoflagellates. *European Journal of Protistology*, *40*, 85–111. https://doi.org/10.1016/j.ejop.2003.11.003

Schaumburg, J. (1996). Seen in Bayern. Limnologische Entwicklungen von 1980 bis 1996. *Informationsberichte des Bayerischen Landesamtes für Wasserwirtschaft*, *1*, 1–211.

Schaumburg, J., & Hehl, I. (2001). Kleinseen in Bayern – Ökologische Bewertung von Freiwasser, Sediment, Ufer und Einzugsgebiet. *Informationsberichte des Bayerischen Landesamtes für Wasserwirtschaft*, *1*, 1–281.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., … Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, *75*, 7537–7541. https://doi.org/10.1128/aem.01541-09

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., … Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 6241–6246. https://doi.org/10.1073/pnas.1117018109

Schrank, F. V. P. (1793). III. Mikroskopische Wahrnehmungen. *Der Naturforscher*, *27*, 26–37, pl. III 37.

Schrank, F. V. P. (1802). *Briefe naturhistorischen, physikalischen und ökonomischen Inhaltes*. Erlangen: Schubart.

Seenivasan, R., Sausen, N., Medlin, L. K., & Melkonian, M. (2013). *Picomonas judraskeda* gen. et sp. nov.: The first identified member of the Picozoa phylum nov., a widespread group of picoeukaryotes, formerly known as 'picobiliphytes'. *PLoS ONE*, *8*, e59565. https://doi.org/10.1371/journal.pone.0059565

Siebeck, O. (1982). Der Königssee. Eine limnologische Projektstudie. *Nationalpark Berchtesgaden : Forschungsberichte*, *5*, 1–131.

Söhner, S., Zinßmeister, C., Kirsch, M., & Gottschling, M. (2012). Who am I – and if so, how many? Species diversity of calcareous dinophytes (Thoracosphaeraceae, Peridiniales) in the Mediterranean Sea. *Organisms Diversity and Evolution*, *12*, 339–348. https://doi.org/10.1007/s13127-012-0109-z

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Steinicke, H. (2014). *Challenges and opportunities of integrative taxonomy for research and society. Taxonomic research in the era of OMICS technologies*. Halle/Saale: Leopoldina.

Stern, R. F., Andersen, R. A., Jameson, I., Küpper, F. C., Coffroth, M.-A., Vaulot, D., … Keeling, P. J. (2012). Evaluating the ribosomal Internal Transcribed Spacer (ITS) as a candidate dinoflagellate barcode marker. *PLoS ONE*, *7*, e42780. https://doi.org/10.1371/journal.pone.0042780

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H. W., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology*, *19*, 21–31. https://doi.org/10.1111/j.1365-294x.2009.04480.x

Takahashi, K., Sarai, C., & Iwataki, M. (2014). Morphology of two marine woloszynskioid dinoflagellates, *Biecheleria brevisulcata* sp. nov. and *Biecheleriopsis adriatica* (Suessiaceae, Dinophyceae), from Japanese coasts. *Phycologia*, *53*, 52–65. https://doi.org/10.2216/13-192.1

Takano, Y., Yamaguchi, H., Inouye, I., Moestrup, Ø., & Horiguchi, T. (2014). Phylogeny of five species of *Nusuttodinium* gen. nov. (Dinophyceae), a genus of unarmoured kleptoplastidic dinoflagellates. *Protist*, *165*, 759–778. https://doi.org/10.1016/j.protis.2014.09.001

Tillmann, U., Salas, R., Jauffrais, T., Hess, P., & Silke, J. (2014). AZA: The producing organisms—biology and trophic transfer. In L. M. Botana (Ed.), *Seafood and freshwater toxins. Pharmacology, physiology, and detection* (pp. 773–798). Boca Raton, FL: CRC Press.

Toebe, K., Alpermann, T. J., Tillmann, U., Krock, B., Cembella, A., & John, U. (2013). Molecular discrimination of toxic and non-toxic *Alexandrium* species (Dinophyta) in natural phytoplankton assemblages from the Scottish coast of the North Sea. *European Journal of Phycology*, *48*, 12–26. https://doi.org/10.1080/09670262.2012.752870

Toebe, K., Joshi, A. R., Messtorff, P., Tillmann, U., Cembella, A., & John, U. (2013). Molecular discrimination of taxa within the dinoflagellate genus *Azadinium*, the source of azaspiracid toxins. *Journal of Plankton Research*, *35*, 225–230. https://doi.org/10.1093/plankt/fbs077

Turland, N. J., Wiersema, J. H., Barrie, F. R., Greuter, W., Hawksworth, D. L., Herendeen, P. S., … Smith, G. F. (2018). *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*. Glashütten: Koeltz.

Vellend, M. (2010). Conceptual synthesis in community ecology. *Quarterly Review of Biology*, *85*, 183–206. https://doi.org/10.1086/652373

Vergin, K. L., Beszteri, B., Monier, A., Thrash, J. C., Temperton, B., Treusch, A. H., … Giovannoni, S. J. (2013). High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME Journal*, *7*, 1322–1332. https://doi.org/10.1038/ismej.2013.32

Vernooy, R., Haribabu, E., Muller, M. R., Vogel, J. H., Hebert, P. D., Schindel, D. E., … Singer, G. A. (2010). Barcoding life to conserve biological diversity: Beyond the taxonomic imperative. *PLoS Biology*, *8*, e1000417.

von Stosch, H. A. (1973). Observations on vegetative reproduction and sexual life cycles of two freshwater dinoflagellates, *Gymnodinium pseudopalustre* Schiller and *Woloszynskia apiculata* sp. nov. *British Phycological Journal*, *8*, 105–134. https://doi.org/10.1080/00071617300650141

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, *73*, 5261–5267. https://doi.org/10.1128/aem.00062-07

Wheeler, Q. D. (2008). Taxonomic shock and awe. In Q. D. Wheeler (Ed.), *The new taxonomy* (pp. 211–226). Boca Raton: CRC Press. https://doi.org/10.1201/crcsysassspe

Wohlrab, S., Falcke, J. M., Lin, S., Zhang, H., Neuhaus, S., Elferink, S., Voss, D., Zielinski, O., & John, U. (2018). Metatranscriptome profiling indicates size-dependent differentiation in plastic and conserved community traits and functional diversification in dinoflagellate communities. *Frontiers in Marine Science*, *5*, UNSP 358. https://doi.org/10.3389/fmars.2018.00358

Xiao, C., Wu, X., Liu, T., Xu, G., & Chi, R. (2017). Microbial community structure of activated sludge for biosolubilization of two different rock phosphates. *Applied Biochemistry and Biotechnology*, *182*, 742–754. https://doi.org/10.1007/s12010-016-2358-3

Žerdoner Čalasan, A., Kretschmann, J., Filipowicz, N. H., Irimia, R.-E., Kirsch, M., & Gottschling, M. (2019). Towards global distribution maps of unicellular organisms such as calcareous dinophytes based on DNA sequence information. *Marine Biodiversity*, *49*, 749–758.

Žerdoner Čalasan, A., Kretschmann, J., & Gottschling, M. (2019). They are young, and they are many: Dating freshwater lineages in unicellular dinophytes. *Environmental Microbiology*. https://doi.org/10.1111/1462-2920.14766

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, *30*, 614–620. https://doi.org/10.1093/bioinformatics/btt593

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Gottschling M, Chacón J, Žerdoner Čalasan A, et al. Phylogenetic placement of environmental sequences using taxonomically reliable databases helps to rigorously assess dinophyte biodiversity in Bavarian lakes (Germany). *Freshw Biol*. 2019;00:1–16. https://doi.org/10.1111/fwb.13413