The Journal of Eukaryotic Microbiology
Published by the International Society of Protistologists

Journal of Eukaryotic Microbiology ISSN 1066-5234

SYMPOSIUM ARTICLE

Towards a Molecular Taxonomy for Protists: Benefits, Risks, and Applications in Plankton Ecology

David A. Caron

Department of Biological Sciences, University of Southern California, Los Angeles, California

Keywords
microbial ecology; molecular ecology.

Correspondence
D.A. Caron, Department of Biological Sciences, University of Southern California, 3616 Trousdale Parkway, AHF 301, Los Angeles, CA 90089-0371, USA
Telephone number: +213-740-0203;
FAX number: +213-740-6720;
e-mail: dcaron@usc.edu

Received: 6 March 2013; accepted March 11, 2013.

doi:10.1111/jeu.12044

ABSTRACT

The increasing use of genetic information for the development of methods to study the diversity, distributions, and activities of protists in nature has spawned a new generation of powerful tools. For ecologists, one lure of these approaches lies in the potential for DNA sequences to provide the only immediately obvious means of normalizing the diverse criteria that presently exist for identifying and counting protists. A single, molecular taxonomy would allow studies of diversity across a broad range of species, as well as the detection and quantification of particular species of interest within complex, natural assemblages; goals that are not feasible using traditional methods. However, these advantages are not without their potential pitfalls and problems. Conflicts involving the species concept, disagreements over the true (physiological/ecological) meaning of genetic diversity, and a perceived threat by some that sequence information will displace knowledge regarding the morphologies, functions and physiologies of protistan taxa, have created debate and doubt regarding the efficacy and appropriateness of some genetic approaches. These concerns need continued discussion and eventual resolution as we move toward the irresistible attraction, and potentially enormous benefits, of the application of genetic approaches to protistan ecology.

PROTISTAN taxonomy provides a rich and complex field of study for scientists attempting to organize and classify the incredible diversity that exists among single-celled, eukaryotic taxa. For ecologists, taxonomy provides a tool with which to group individuals into ecological and hopefully evolutionary entities that represent the basic units of physiological ability and genetic inheritance within protistan lineages. Ecological studies are complicated by the complexity of taxonomies that have been and continue to be applied to protists, a historical consequence of the adoption of morphological characters as the "gold standard" for identifying and cataloging these species. The huge breadth of morphological diversification present among protistan taxa has resulted in multiple procedures and protocols for collecting, preserving, staining, observing, and enumerating protists to distinguish the different morphological characters used to characterize these species (Adl et al. 2005). This morphology-based species concept has expanded in recent decades to include ultrastructural features enabled by the use of electron microscopy, physiological and ecological criteria, and most recently DNA sequence information for defining and identifying protists.

The last of these criteria, gene sequencing, has begun to spawn a variety of important approaches and tools for the study of protistan ecology. Numerous applications are now emerging for assessing the presence and abundance of particular species, examining community structure, and for estimating the total species richness of natural protistan assemblages (see below). These successes have encouraged the development of an informal molecular taxonomy for protists as a means of normalizing collection, processing and analysis of samples across all taxa, thereby reducing or eventually eliminating the difficulties of applying multiple traditional taxonomies when examining complex, natural assemblages of protists (Caron 2009b). This movement towards a molecular taxonomy is not unique to protistan researchers, but has also been proposed for bacteria, plants, and animals (Blaxter 2004; Hebert et al. 2003; Tautz et al. 2003). The development of a molecular taxonomy has been welcomed by many ecologists eager to obtain new tools for their studies, but it also has raised objections among many biologists (Ebach and Holdrege 2005).

Resistance to a molecular taxonomy has been based largely on the argument that traditional taxonomies for

microbes might not be adequately or accurately captured by DNA sequence analysis. This argument has been particularly strong regarding a molecular taxonomy based on a single gene, the most studied of which has been the small subunit ribosomal RNA gene (ssu rDNA, the 16S and 18S genes for bacteria and eukaryotes, respectively). This gene is considered inappropriate by some researchers as the basis for establishing a bacterial molecular taxonomy because bacterial taxonomy has historically been based primarily on a complex array of physiological abilities, and 16S sequences do not necessarily map well or consistently to these abilities due to multiple occurrences of lateral gene transfer in evolutionarily time. A DNA taxonomy for protists has begun to attract attention by a number of researchers, but the development and application of such a scheme is still in its infancy (De Jonckheere 2004).

Why bother with a molecular taxonomy for protists? Ecological studies typically require the analysis of large numbers of samples. This expectation, coupled with the extremely high species diversity of most natural microbial assemblages and the complexity of methods employed for identifying and enumerating these species (and the taxonomic expertise required), make it difficult to characterize even a few of the protistan groups in a sample. A single approach that could be employed across all taxonomic groups of protists would be a tremendous boon for ecological studies.

BENEFITS AND APPLICATIONS

Sequence data have been used extensively in recent decades to aid in testing hypotheses regarding the phylogenetic relationships among eukaryotes. The evolutionary insight resulting from these endeavors has been nothing short of profound. The emergence of the large amounts of sequence information that have supported studies of phylogeny subsequently enabled the development of approaches and methods that have begun to dramatically improve the performance of ecological observations and experimental studies of protistan assemblages.

Facilitating "discovery"

Sequencing of environmental samples has been used extensively within the last decade to investigate species richness of natural protistan assemblages. These studies have documented the presence of extraordinarily large numbers of species comprising natural protistan assemblages, particularly in the marine pelagic realm (see reviews by Vaulot et al. 2008; Caron and Gast 2008; Caron et al. 2012). Large numbers of previously unknown or uncultured taxa or whole lineages have been identified within a number of protistan groups (Guillou et al. 2008; Massana et al. 2006; Not et al. 2009). Molecular approaches have been particularly effective for demonstrating the presence of morphologically nondescript, but potentially ecologically important taxa such as a number of previously unknown species of small, phototrophic, and heterotrophic flagellates (Massana et al. 2006; Shi et al. 2009). In addition, DNA sequencing coupled with physiological studies has revealed the presence of apparent cryptic species within various morphospecies of small protists (Boenigk et al. 2006; Fawley et al. 2006; Stoeck et al. 2008; Yubuki et al. 2010).

High specificity of species identification

The rapid expansion of databases resulting from the sequencing of cultured protists and environmental samples has provided the fodder for the development of tools that can differentiate and identify specific protistan taxa within mixed, natural assemblages. These tools have been applied to monitor the presence and abundance of species of ecological interest, greatly facilitating autecological and biogeographical studies of these taxa. DNA probes adapted to various fluorescent in situ hybridization approaches are now used extensively for the detection of various noxious and toxic algal taxa (Greenfield et al. 2008; Medlin et al. 2010). These methods have been particularly useful for protistan species that lack distinctive morphological features such as minute eukaryotic algae and small heterotrophic flagellates, and thus cannot be easily differentiated from co-existing taxa (Fuller et al. 2006; Thaler and Lovejoy 2012). Species-specific oligonucleotides have also been adapted for use in microarrays, expanding the number of species that can be characterized at one time from a given sample, although microarray approaches have not been as commonly applied to eukaryotes as they have to bacteria (Ahn et al. 2006; Gentry et al. 2006; Medlin et al. 2006; Metfies and Medlin 2007).

High sensitivity of detection

Other methodologies, developed in conjunction with species-specific probes and primers, have leveraged species-specific DNA signatures to dramatically improve the sensitivity with which targeted protistan taxa can be identified and counted. A number of quantitative polymerase chain reaction methods have been developed that allow detection of protistan species at extremely low abundances. Detecting protistan taxa with very high sensitivity in complex, natural microbial assemblages is particularly useful for some toxic algae that produce powerful neurotoxins, or parasitic taxa that might be difficult to detect using microscopy or culture methods (Andree et al. 2011; Audemard et al. 2004; Bowers et al. 2006; Garneau et al. 2011; Hosoi-Tanabe and Sako 2005; Howard et al. 2012; Zamor et al. 2012). The ability to follow the population dynamics of targeted species at sub-bloom abundances greatly facilitates our ability to investigate the environmental conditions leading to the success of these species in microbial communities.

Community-level analyses

One of the most significant contributions of genetic information to protistan ecology is the potential for the

development of new approaches that will enable community-level studies of these species. This level of analysis is presently unattainable, given the vast diversity of species (and their attendant taxonomies) in most ecosystems. The full potential of this approach will not be realized until a method for assessing *all* protistan taxa in a natural sample is fully validated (i.e. addressing issues of PCR bias, extraction, amplification efficiency, etc.), but the potential to eventually be able to examine the response of whole assemblages of protists to environmental forcing factors is exciting. These approaches are enabling studies that examine the biogeographic patterns of protists (Bass et al. 2007; Countway et al. 2007; Salani et al. 2012), and they are providing a means of testing hypotheses regarding protistan community structure and reassembly in response to environmental change or alterations in trophic structure (Caron and Countway 2009). Examining the relationship between diversity and ecosystem function (Cardinale 2011) will enhance our understanding of the role of protistan species diversity in energy production and flow.

High sample throughput

The potential to rapidly analyze large numbers of samples, coupled with whole-community analysis, make the lure of a DNA taxonomy almost irresistible to many protistan ecologists. To be sure, these methods are still emerging and rapidly evolving. This work was initiated with cloning and sequencing surveys of environmental samples (generally targeting 18S genes) approximately a decade ago, but these studies were rather limited due to the prohibitive cost of the approach. More cost-efficient approaches such as DNA fragment analysis, and lower costs associated with high-throughput DNA sequencing (e.g. 454 pyrosequencing or Illumina sequencing) have begun to greatly expand the depth of sequencing that can be obtained for natural samples. The latter approaches will undoubtedly increase, improve, and become less expensive in the future.

RISKS AND PROBLEMS OF A MOLECULAR TAXONOMY

The potential benefits of employing genetic signatures in protistan ecology are not without significant caveats, especially at the present time. Some of these concerns and/or shortcomings will wane as sequence databases expand and methodological improvements and innovations progress, but some issues (e.g. the species concept for protists) present fundamental stumbling blocks that thwart widespread acceptance and application of a molecular taxonomy and its application to protistan ecology.

Overestimating or underestimating species richness

The use of DNA sequences as proxies for the identification of protistan species presently suffers from the potential for significantly underestimating *or* overestimating the total species richness of natural communities of microbial eukaryotes. This problem stems from the present approaches for grouping DNA sequences into "operational taxonomic units" (OTUs). On one hand, underestimation of species richness can occur if OTUs are constructed in such a way that multiple species (defined by the classical standards of morphology or physiology) are condensed into a single OTU. Underestimation of species richness can also result if sample coverage is simply too sparse to allow reasonable estimates of total species richness. In this regard, the inadequacies of some approaches, particularly DNA fragment analysis, to adequately describe community diversity have been clearly stated (Bent and Forney 2008; Bent et al. 2007). Fragment analysis approaches can provide useful assessments of the dominant taxa within an assemblage, and it remains a viable and cost effective approach for examining changes in the dominant taxa of assemblages when large numbers of samples must be processed. However, fragment analysis is inappropriate for estimating total species richness due to a limited ability of these approaches to detect sub-dominant taxa in an assemblage.

Cloning and sequencing of genes from environmental samples, and more recently high-throughput sequencing approaches, have provided more robust assessments of species richness, but even these methods may be unable to plumb the full depths of many natural microbial assemblages. Therefore, a number of species richness estimators have been developed and applied to extrapolate from limited sample coverage to the total estimated species richness in a sample (Bunge 2011; Schloss et al. 2009).

The potential for overestimating total species richness of protistan assemblages has clearly created more controversy in microbial ecology than the possibility of underestimating this feature. Overestimating richness using sequence information results from the separation of slightly dissimilar DNA sequences arising from a single species into more than one OTU. The potential for overestimation of richness in this manner is significant, given that individuals within any population will harbor some degree of DNA sequence dissimilarity. What constitutes meaningful sequence variability (i.e. is indicative of different species) and what is neutral mutation among individuals within a single species? The answer to that question is not clear at this time, and it has been argued that much of the sequence variability observed in a protistan morphospecies might be intraspecific, neutral mutation, and of no ecological importance (Fenchel 2005). Disagreement over the significance of sequence variability as an indicator of cryptic species of protists within morphospecies is at the heart of a significant debate in protistan ecology relating to the biogeography of these species (Caron 2009a).

Algorithms for calling OTUs are still relatively unconstrained because the relationship between sequence variability and traditional protistan species identifications are not yet well defined (Caron et al. 2009; Jones et al. 2011; Powell et al. 2011; Schloss and Westcott 2011), and the efficacy of different genes or gene fragments is still being examined (Dunthorn et al. 2012). Converting short DNA sequences presently provided by high throughput

sequencing methods into OTUs is particularly problematic. Approaches for how to interpret short DNA sequences vis-à-vis microbial species richness are just appearing (Amaral-Zettler et al. 2009; Huse et al. 2008), as are cautions on the limitations and improvements in these approaches (Huse et al. 2007; Kunin et al. 2009; Quince et al. 2009).

Loss of information on the activities of protists

Concern over a molecular taxonomy among some protistologists derives from an apprehension that DNA sequence information will displace knowledge of protistan morphology and physiology. This concern seems unfounded for protistan ecologists because ecologists are innately interested in the morphology, behavior, and physiology of these species. For ecologists, DNA sequences simply represent additional approaches (and potentially more rapid and powerful ones) for characterizing the presence and abundance of species of protists. Sequences, by themselves, have little meaning unless they are linked to traditional taxonomies, behaviors, and physiologies of these species. In short, there is little fear that protistan ecologists will forget what protists look like or what they do simply because sequences might be employed to tell which species are present in a sample.

Fear of unfulfilled expectations

Could a molecular taxonomy clarify the species concept for protists? Molecular approaches and methods are becoming so integrated into modern biology that there is an expectation among some investigators that sequence information will eventually solve all extant issues relating to the species concept for microbes. To be sure, the use of sequence data has helped resolve taxonomic impasses for some protistan taxa, and our understanding of many aspects of protistan ecology have been significantly expanded through the exploitation of sequence data. However, it may be unrealistic to expect that sequence information alone will resolve the present debate over the species concept, and it would be unfortunate to find that this unfulfilled expectation thwarts enthusiasm for the integration of sequence information into the species concept, and consequently delays the development of new genetic methods that can be applied in ecological studies.

MOVING FORWARD WITH A MOLECULAR TAXONOMY FOR PROTISTS

The application of genetic approaches to the study of protistan ecology is still an emerging field, but many techniques for identifying and quantifying protistan taxa, or whole assemblages, have emerged within the last several years. Many limitations and caveats involving these approaches remain to be resolved, but these advancements are quite astounding given that DNA sequence information for protists has become commonplace only within the last decade while the taxonomy of these species began approximately 350 yr ago. A few guiding principles that might hasten the development of a protistan molecular taxonomy, and application to studies of protistan ecology, are given below.

Choosing the best gene(s)

Small subunit ribosomal RNA genes (18S) have received the most attention with respect to developing a molecular taxonomy for protists, although cytochrome oxidase I and other genes have also received attention, particularly for barcoding animals and plants (Hebert et al. 2003). The choice of a gene to serve as a basis for a taxonomic scheme must fundamentally account for the rate of mutation of the gene relative to the rates of speciation among the taxa. Sequence dissimilarity between species must be sufficient to allow separation into different OTUs but not so great as to cause ambiguity in defining OTUs. However, the rate of mutation of a given gene is not necessarily constant for different groups of protists (Pawlowski et al. 1997). For that reason, genes may differ in their appropriateness for establishing OTUs for different protistan groups, and an appropriate molecular taxonomy will have to take this fact into account (De Jonckheere et al. 2012). Genes other than 18S have been employed for barcoding diatoms (Hamsher et al. 2011), amoebae (Nassonova et al. 2010) and heterotrophic flagellates (Wylezich et al. 2010) with success. These studies indicate that we are still early in the development of a DNA taxonomy, that there are many useful candidate genes or genomic regions that might be more informative as taxonomic tools for specific groups of protists, and that as a community we are probably moving toward the adoption of a multigene taxonomy, as has been the case with phylogenetic studies of protists (Riisberg et al. 2009; Santoferrara et al. 2012a).

It is also important to recognize that the present state of high-throughput sequencing technology still yields relatively short sequences that can be problematic when attempting to demarcate OTUs. There is a need to map gene fragments onto whole genes to improve taxonomic/phylogenetic informational content (Dawson and Hagen 2009). These new technologies can yield enormous numbers of sequences, allowing the detection of many, many OTUs. However, more work is needed to ensure that the sequences are grouped into OTUs that have ecological meaning and are not simply the result of intraspecific sequence variation.

Standardizing, bench-marking, cross-checking

The most expeditious way to ensure that OTUs derived from DNA sequences constitute ecologically relevant information is to directly compare and map sequence data against traditional morphological and physiological characters and features. The application of a DNA taxonomy to protists is useful to ecologists only if the genetic signatures can be interpreted in a manner that links those signatures to behavior, ecology, and biogeochemistry.

A number of comparisons of different genetic approaches or genes have been undertaken recently in part to identify useful genes or gene fragments for use as taxonomic characters (Edgcomb et al. 2011; Marie et al. 2006; Stoeck et al. 2010). Work has also begun in earnest to incorporate sequence information into taxonomic descriptions of a number of protistan groups including ciliates (Xu et al. 2012), amoebae (De Jonckheere 2004; Lahr et al. 2008), diatoms (Moniz and Kaczmarska 2010) and other taxa. These studies are essential for integrating DNA sequence information into the fabric of protistan taxonomy. However, there are still relatively few direct comparisons of protistan species richness employing both morphological and molecular approaches (Medinger et al. 2010; Santoferrara et al. 2012b). Such comparisons are necessary to establish the nature and magnitude of the mismatch between these approaches and to eventually merge them into a single (and hopefully more effective) tool for protistan ecology.

CONCLUSION

Protistan ecologists are faced with a daunting, multi-faceted species concept in conducting their research: gross morphological features of the cell augmented by ultrastructure and physiology. More recently, DNA sequencing has been employed as a means of resolving conflicts or inadequacies relating to these latter features, with the expectation that ultimately these approaches will merge to provide a more robust taxonomy (Schlegel and Meisterfeld 2003). Concurrently, DNA sequence information has become extremely useful in protistan ecology because it provides a huge set of characters for developing methods that can specifically detect and count protistan taxa in natural samples. These methods offer the potential to be universally applicable across protistan lineages, obviating the need for different procedures and taxonomic schemes that are presently required for the many disparate groups of protists. Molecular approaches will enable much more holistic assessments of natural protistan assemblages, enhancing observational and experimental studies to understand microbial community structure and its relationship to ecosystem function. The time is right for consensus building to arrive at a species concept that integrates DNA sequence information with morphological and physiological features.

ACKNOWLEDGMENTS

Support for this manuscript was provided by National Science Foundation grants MCB-0,703,159 and OCE-1,136,818 and a grant from the Gordon and Betty Moore Foundation.

LITERATURE CITED

Adl, S. M., Simpson, A. G. B., Farmer, M. A., Andersen, R. A., Anderson, O. R., Barta, J. R., Bowser, S. S., Brugerolle, G., Fensome, R. A., Fredericq, S., James, T. Y., Karpov, S., Kugrens, P., Krug, J., Lane, C. E., Lewis, L. A., Lodge, J., Lynn, D. H., Mann, D. G., McCourt, R. M., Mendoza, L., Moestrup, O., Mozley-Standridge, S. E., Nerad, T., Shearer, C. A., Smirnov, A. V., Spiegel, F. W. & Taylor, M. F. J. R. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Euk. Microbiol.*, 52:399–451.

Ahn, S., Kulis, D. M., Erdner, D. L., Anderson, D. M. & Walt, D. R. 2006. Fiber-optic microarray for simultaneous detection of multiple harmful algal bloom species. *Appl. Environ. Microbiol.*, 72:5742–5749.

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W. & Huse, S. M. 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS ONE*, 4:e6372.

Andree, K. B., Fernandez-Tejedor, M., Elandaloussi, L. M., Quijano-Scheggia, S., Sampedro, N., Garces, E., Camp, J. & Diogene, J. 2011. Quantitative PCR coupled with melt curve analysis for detection of selected *Pseudo-nitzschia* spp. (Bacillariophyceae) from the northwestern Mediterranean sea. *Appl. Environ. Microbiol.*, 77:1651–1659.

Audemard, C., Reece, K. S. & Burreson, E. M. 2004. Real-time PCR for detection and quantification of the protistan parasite *Perkinsus marinus* in environmental waters. *Appl. Environ. Microbiol.*, 70:6611–6618.

Bass, D., Richards, T. A., Matthai, L., Marsh, V. & Cavalier-Smith, T. 2007. DNA evidence for global dispersal and probable endemicity of protozoa. *BMC Evol. Biol.*, 7:162.

Bent, S. J. & Forney, L. J. 2008. The tragedy of the uncommon: understanding limitations in the anlaysis of microbial diversity. *ISME J.*, 2:689–695.

Bent, S. J., Peirson, J. D. & Forney, L. J. 2007. Measuring species richness based on microbial community fingerprints: the emperor has no clothes. *Appl. Environ. Microbiol.*, 73:2399.

Blaxter, M. L. 2004. The promise of a DNA taxonomy. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.*, 359:669–679.

Boenigk, J., Jost, S., Stoeck, T. & Garstecki, T. 2006. Differential thermal adaptation of clonal strains of a protist morphospecies originating from different climatic zones. *Environ. Microbiol.*, 9:593–602.

Bowers, H. A., Tomas, C., Tengs, T., Kempton, J. W., Lewitus, A. J. & Oldach, D. W. 2006. Raphidophyceae (Chadefaud ex Silva) systematics and rapid identification: sequence analyses and real-time PCR assays. *J. Phycol.*, 42:1333–1348.

Bunge, J. (2011) Estimating the number of species with Catchall. Proceedings of Pacific Symposium on Biocomputing' 2011. Kohala Coast, Hawaii.

Cardinale, B. J. 2011. Biodiversity improves water quality through niche partitioning. *Nature*, 472:86–89.

Caron, D. A. 2009a. Protistan biogeography: why all the fuss? *J. Euk. Microbiol.*, 56:105–112.

Caron, D. A. 2009b. New accomplishments and approaches for assessing protistan diversity and ecology in natural ecosystems. *Bioscience*, 59:287–299.

Caron, D. A. & Gast, R. J. 2008. The diversity of free-living protists: seen and unseen, cultured and uncultured. *In*: Zengler, K. (ed.) Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back. ASM Press, Washington, DC. p. 67–93.

Caron, D. A. & Countway, P. D. 2009. Hypotheses on the role of the protistan rare biosphere in a changing world. *Aq. Microb. Ecol.*, 57:227–238.

Caron, D. A., Countway, P. D., Jones, A. C., Kim, D. Y. & Schnetzer, A. 2012. Marine protistan diversity. *Ann. Rev. Mar. Sci.*, 4:467–493.

Caron, D. A., Countway, P. D., Savai, P., Gast, R. J., Schnetzer, A., Moorthi, S. D., Dennett, M. R., Moran, D. M. & Jones, A. C. 2009. Defining DNA-based operational taxonomic units for microbial eukaryote ecology. *Appl. Environ. Microbiol.*, 75:5797–5808.

Countway, P. D., Gast, R. J., Dennett, M. R., Savai, P., Rose, J. M. & Caron, D. A. 2007. Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). *Environ. Microbiol.*, 9:1219–1232.

Dawson, S. C. & Hagen, K. D. 2009. Mapping the protistan 'rare biosphere'. *J. Biol.*, 8:105.

De Jonckheere, J. F. 2004. Molecular definition and the ubiquity of species of the genus *Naegleria*. *Protist*, 155:89–103.

De Jonckheere, J. F., Gryseels, S. & Eddyani, M. 2012. Knowledge of morphology is still required when identifying new amoeba isolates by molecular techniques. *Eur. J. Protistol.*, 48:178–184.

Dunthorn, M., Klier, J., Bunge, J. & Stoeck, T. 2012. Comparing the hyper-variable V4 and V9 regions of the small subunit rDNA for assessment of ciliate environmental diversity. *J. Euk. Microbiol.*, 59:185–187.

Ebach, M. C. & Holdrege, C. 2005. More taxonomy, not DNA barcoding. *Bioscience*, 55:822–823.

Edgcomb, V., Orsi, W., Bunge, J., Jeon, S., Christen, R., Leslin, C., Holder, M., Taylor, G. T., Suarez, P., Varela, R. & Epstein, S. 2011. Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. *ISME J.*, 5:1344–1356.

Fawley, M. W., Dean, M. L., Dimmer, S. K. & Fawley, K. P. 2006. Evaluating the morphospecies concept in the Selenastraceae (Chlorophyceae, Chlorophyta). *J. Phycol.*, 42:142–154.

Fenchel, T. 2005. Cosmopolitan microbes and their 'cryptic' species. *Aq. Microb. Ecol.*, 41:49–54.

Fuller, N. J., Tarran, G. A., Cummings, D. G., Woodward, E. M. S., Orcutt, K. M., Yallop, M., Le Gall, F. & Scanlan, D. J. 2006. Molecular analysis of photosynthetic picoeukaryote community structure along an Arabian Sea transect. *Limnol. Oceanogr.*, 51:2502–2514.

Garneau, M.-E., Schnetzer, A., Countway, P. D., Jones, A. C., Seubert, E. L. & Caron, D. A. 2011. Examination of the seasonal dynamics of the toxic dinoflagellate *Alexandrium catenella* at Redondo Beach, California, by quantitative PCR. *Appl. Environ. Microbiol.*, 77:7669–7680.

Gentry, T. J., Wickham, G. S., Schadt, C. W., He, Z. & Zhou, J. Z. 2006. Microarray applications in microbial ecology research. *Microb. Ecol.*, 52:159–175.

Greenfield, D., Marin, R., Doucette, G., Mikulski, C., Jones, K., Jensen, S., Roman, B., Alvarado, N. & JFeldman, J. & Scholin, C. , 2008. Field applications of the second-generation Environmental Sample Processor (ESP) for remote detection of harmful algae: 2006–2007. *Limnol. Oceanogr. Methods*, 6:667–679.

Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R., Scanlan, D. J. & Worden, A. Z. 2008. Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environ. Microbiol.*, 10:3349–3365.

Hamsher, S. E., Evans, K. M., Mann, D. G., Poulícková, A. & Saunders, G. W. 2011. Barcoding diatoms: exploring alternatives to COI-5P. *Protist*, 162:405–422.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. London Ser. B*, 270:313–321.

Hosoi-Tanabe, S. & Sako, Y. 2005. Species-specific detection and quantification of toxic marine dinoflagellates *Alexandrium tamarense* and *A. catenella* by real-time PCR assay. *Mar. Biotechnol.*, 7:506–514.

Howard, M. D. A., Jones, A. C., Schnetzer, A., Countway, P. D., Tomas, C. R., Kudela, R. M., Hayashi, K., Chia, P. & Caron, D. A. 2012. Quantitative real-time PCR for *Cochlodinium fulvescens* (Dinophyceae), a potentially harmful dinoflagellate from California coastal waters. *J. Phycol.*, 48:384–393.

Huse, S., Huber, J., Morrison, H., Sogin, M. & Welch, D. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.*, 8:R143.

Huse, S., Dethlefsen, L., Huber, J. A., Welch, D. M., Relman, D. A. & Sogin, M. L. 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.*, 4:e1000255.

Jones, M., Ghoorah, A. & Blaxter, M. 2011. jMOTU and taxonerator: turning DNA barcode sequences into annotated operational taxonomic units. *PLoS ONE*, 6:e19259.

Kunin, V., Engelbrektson, A., Ochman, H. & Hugenholtz, P. 2009. Wrinkles in the rare biosphere: pyrosequencing errors lead to artificial inflation of diversity estimates. *Environ. Microbiol.*, 12:118–123.

Lahr, D. J. G., Bergmann, P. J. & Lopes, S. G. B. C. 2008. Taxonomic identity in microbial eukaryotes: a practical approach using the testate amoeba *Centropyxis* to resolve conflicts between old and new taxonomic descriptions. *J. Euk. Microbiol.*, 55:409–416.

Marie, D., Zhu, F., Balagué, V., Ras, J. & Vaulot, D. 2006. Eukaryotic picoplankton communities of the Mediterranean Sea in summer assessed by molecular approaches (DGGE, TTGE, QPCR). *FEMS Microbiol. Ecol.*, 55:403–415.

Massana, R., Terrado, R., Form, I., Lovejoy, C. & Pedrós-Alió, C. 2006. Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. *Environ. Microbiol.*, 8:1515–1522.

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. *Mol. Ecol.*, 19:32–40.

Medlin, L., Metfies, K., Mehl, H., Wiltshire, K. & Valentin, K. 2006. Picoeukaryotic plankton diversity at the Helgoland time series site as assessed by three molecular methods. *Microb. Ecol.*, 52:53–71.

Medlin, L. K., Diercks, S. & Beszteri, S. 2010. Mini review: probes for detecting *Prymnesium parvum* and preliminary results from gene expression studies. *J. Am. Water Resources Assoc.*, 46:144–152.

Metfies, K. & Medlin, L. K. 2007. Refining cryptophyte identification with DNA-microarrays. *J. Plankton Res.*, 29:1071–1075.

Moniz, M. B. J. & Kaczmarska, I. 2010. Barcoding of diatoms: nuclear encoded ITS revisited. *Protist*, 161:7–34.

Nassonova, E., Smirnov, A., Fahrni, J. & Pawlowski, J. 2010. Barcoding amoebae: comparison of SSU, ITS and COI genes as tools for molecular identification of naked lobose amoebae. *Protist*, 161:102–115.

Not, F., del Campo, J., Balagué, V., de Vargas, C. & Massana, R. 2009. New insights into the diversity of marine picoeukaryotes. *PLoS ONE*, 4:e7143.

Pawlowski, J., Bolivar, I., Fahrni, J. F., de Vargas, C., Gouy, M. & Zaninetti, L. 1997. Extreme differences in rates of molecular evolution of Foraminifera revealed by comparison of ribosomal DNA sequences and the fossil record. *Mol. Biol. Evol.*, 14:498–505.

Powell, J. R., Monaghan, M. T., ÖPik, M. & Rillig, M. C. 2011. Evolutionary criteria outperform operational approaches in

producing ecologically relevant fungal species inventories. *Mol. Ecol.*, 20:655–666.

Quince, C., Lanzen, A., Curtis, T. P., Davenport, R. J., Hall, N., Head, I. M., Read, L. F. & Sloan, W. T. 2009. Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat Meth*, 6:639–641.

Riisberg, I., Orr, R. J. S., Kluge, R., Shalchian-Tabrizid, K., Bower, H. A., Patil, V., Edvardsen, B. & Jakobsen, K. S. 2009. Seven gene phylogeny of heterokonts. *Protist*, 160:191–204.

Salani, F. S., Arndt, H., Hausmann, K., Nitsche, F. & Scheckenbach, F. 2012. Analysis of the community structure of abyssal kinetoplastids revealed similar communities at larger spatial scales. *ISME J.*, 6:713–723.

Santoferrara, L. F., McManus, G. B. & Alder, V. A. 2012a. Phylogeny of the order Tintinnida (Ciliophora, Spirotrichea) inferred from small- and large-subunit rRNA genes. *J. Euk. Microbiol.*, 59:423–426.

Santoferrara, L. F., McManus, G. B. & Alder, V. A. 2012b. Utility of genetic markers and morphology for species discrimination within the order Tintinnida (Ciliophora, Spirotrichea). *Protist*, 164:24–36.

Schlegel, M. & Meisterfeld, R. 2003. The species problem in protozoa revisited. *Eur. J. Protistol.*, 39:349–355.

Schloss, P. D. & Westcott, S. L. 2011. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Appl. Environ. Microbiol.*, 77:3219–3226.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J. & Weber, C. F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.*, 75:7537–7541.

Shi, X. L., Marie, D., Jardillier, L., Scanlan, D. J. & Vaulot, D. 2009. Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic South East Pacific Ocean. *PLoS ONE*, 4:e7657.

Stoeck, T., Jost, S. & Boenigk, J. 2008. Multigene phylogenies of clonal *Spumella*-like strains, a cryptic heterotrophic nanoflagellate, isolated from different geographical regions. *Int. J. Syst. Evol. Microbiol.*, 58:716–724.

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. *Mol. Ecol.*, 19:21–31.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2003. A plea for DNA taxonomy. *Trends Ecol. Evol.*, 18:70–74.

Thaler, M. & Lovejoy, C. 2012. Distribution and diversity of a protist predator *Cryothecomonas* (Cercozoa) in Arctic marine waters. *J. Euk. Microbiol.*, 59:291–299.

Vaulot, D., Eikrem, W., Viprey, M. & Moreau, H. 2008. The diversity of small eukaryotic phytoplankton (≤ 3 μm) in marine ecosystems. *FEMS Microbiol. Rev.*, 32:792–820.

Wylezich, C., Nies, G., Mylinikov, A. P., Tautz, D. & Arndt, H. 2010. An evaluation of the use of the LSU rRNA D1-D5 domain for DNA-based taxonomy of eukaryotic protists. *Protist* 161:342–352

Xu, D., Sun, P., Shin, M. K. & Kim, Y. O. K. 2012. Species boundaries in tintinnid ciliates: a case study – morphometric variability, molecular characterization, and temporal distribution of *Helicostomella* species (Ciliophora, Tintinnina). *J. Euk. Microbiol.*, 59:351–358.

Yubuki, N., CÉZa, V., Cepicka, I., Yabuki, A., Inagaki, Y., Nakayama, T., Inouye, I. & Leander, B. S. 2010. Cryptic diversity of free-living parabasalids, *Pseudotrichomonas keilini* and *Lacusteria cypriaca* n. g., n. sp., as inferred from small subunit rDNA Sequences. *J. Euk. Microbiol.*, 57:554–561.

Zamor, R. M., Glenn, K. L. & Hambright, K. D. 2012. Incorporating molecular tools into routine HAB monitoring programs: using qPCR to track invasive Prymnesium. *Harmful Algae*, 15:1–7.