New Accomplishments and Approaches for Assessing Protistan Diversity and Ecology in Natural Ecosystems

DAVID A. CARON

Ecological research of microorganisms sensu latu *(archaea, bacteria, protists, viruses) has come of age within the last few decades. This newfound importance is a consequence of a greater appreciation for the enormous diversity present among these unseen entities, and an increasing recognition of the pivotal roles that these species play in food-web processes and geochemical cycles in aquatic and terrestrial ecosystems. These advances are due in large part to the incorporation of modern genetic and immunological approaches into ecological and physiological studies of natural assemblages and pure cultures of microorganisms. Molecular approaches have revolutionized bacterial and archaeal biology, and are beginning to transform our understanding of protistan ecology (unicellular eukaryotic algae and protozoa). Recent efforts have greatly improved our comprehension of the evolutionary relationships among protistan taxa; documented the existence of lineages of previously undetected protists; and catalyzed studies characterizing their diversity, nutritional modes, and trophic interactions. These extraordinary findings are only beginning to unfold as genetic databases for protists expand and as ecologists learn to interpret and exploit this wealth of genetic information.*

Keywords: protists, diversity, microbial ecology, biogeography, molecular ecology

Conventional wisdom dictates that life on Earth began as single-celled, microscopic forms nearly 4 billion years ago. These minute forms constituted all life on the planet for roughly half of Earth's biological history, and microbes have remained important determinants of organic matter production, trophic transfer, and degradation throughout Earth's entire history, although more attention and research tend to be focused on charismatic macrofauna. It therefore seems fitting, albeit overdue, that characterizing and understanding the importance of microbial diversity and function in natural ecosystems has become a focal point of ecological research in the 21st century.

Microbial ecology has risen to prominence in ecological research from its rather meager status as recently as the middle of the last century. Awareness of the diversity and importance of the larger phytoplankton (e.g., diatoms, dinoflagellates) increased rapidly in the early 20th century. However, even for this conspicuous component of aquatic food webs, studies throughout the latter half of that century significantly added to our knowledge of the standing stocks and diversity of these assemblages and initiated our awareness of the presence and importance of cyanobacteria and minute eukaryotic phototrophs (Malone 1971, Olson et al. 1990). Our knowledge of the ecological niches of aquatic bacteria and protozoa progressed more slowly, and with a few notable exceptions, these assemblages remained an ecological footnote relegated to vaguely defined decompositional processes until the last few decades of the 20th century (Pomeroy 1974, Sieburth 1979). Similarly, the ecological roles of soil bacteria and protozoa have been documented for more than a century, yet recognition of the central role that they play in organic matter degradation and nutrient uptake by plants did not improve dramatically until the latter half of the last century (Alexander 1961).

Today, microbes in the ocean and in freshwater ecosystems are widely recognized as essential participants in global biogeochemical cycles. These taxa constitute the bulk of the standing stock of biomass in most of the world's oceans (Caron et al. 1995), and primary production by cyanobacteria and eukaryotic phytoplankton is responsible for roughly half of the organic carbon and oxygen produced on Earth (and for removal of a commensurate amount of carbon dioxide).

BioScience 59: 287–299. ISSN 0006-3568, electronic ISSN 1525-3244. doi:10.1525/bio.2009.59.4.7

Microorganisms are also important agents for the trophic transfer of energy and carbon. Protists dominate herbivory and bacterivory in the ocean and many freshwater environments, and bacteria, phagotrophic protists (protozoa), and their viruses together process more than half of the total organic matter produced in the ocean, passing a significant fraction on to multicellular organisms and higher trophic levels. In addition, marine archaea, bacteria, protists, and viruses are collectively responsible for the remineralization of nonliving organic matter and the essential nutrients that fuel primary productivity (Suttle 2005, Karl 2007, Sherr et al. 2007). In soils, the activities of bacteria and heterotrophic protists exercise strong control over the decompositional rates of nonliving organic matter and nutrient availability to plants (Bonowski 2004). Disparate recent findings now support the emerging view that macroscopic species on Earth constitute only one aspect of what has been, and continues to be, a largely microbial world.

Protistan phylogeny

Depictions of the phylogeny and diversity of microscopic eukaryotes have changed dramatically and often during the past few decades; several issues remain volatile. These taxa are dominated by single-celled species formerly grouped within a single biological kingdom, the Protista, in the five-kingdom system of Whittaker (1969). Protists were separated from other kingdoms of eukaryotes in that scheme on the basis of their ability to exist as unicells. Within this kingdom, they were subdivided into two large collections of taxa, primarily in accordance with their mode of nutrition (phototrophy versus heterotrophy). This latter division possessed many artificialities, including numerous examples of the separation of morphologically similar taxa into different subkingdoms on the basis of the presence or absence of chloroplasts. This scheme also separated some multicellular forms from single-celled forms (e.g., some of the algal groups), even though they appeared to share a close evolutionary history.

The obvious evolutionary inconsistencies of the Whittaker scheme have motivated several recent reclassifications of protists, and the five-kingdom system has been replaced in recent years with a succession of hypotheses regarding eukaryote evolution and phylogenetic relationships (figure 1; Simpson and Roger 2004, Adl et al. 2005). Not surprisingly, genetic information (DNA sequence information) has played and continues to play an important role in these reorganizations. Protistan evolutionists are still debating some of the details of these new schemes, but there is general agreement that the emerging classification will emulate phylogeny more accurately than does the Whittaker scheme. Meanwhile, the term "protist" remains in common use for all eukaryotic taxa that are capable of existence as single cells and display phototrophic nutrition, heterotrophic nutrition, or some combination of these modes (Caron and Schnetzer 2007).

New approaches yield new views on protistan diversity

Wholesale reorganizations of the major groups of protistan taxa have captured considerable attention and stimulated animated discussions in the recent literature, but these activities have not overshadowed other significant advances in protistan ecology. Chief among these other breakthroughs have been (a) the discovery of several new lineages of protists that had previously gone undetected using traditional approaches of microscopy and culture, and (b) the detection of substantial cryptic diversity within presumably well-described lineages of minute protists.

The application of DNA sequencing to natural samples collected from a wide range of terrestrial and aquatic ecosystems has played a central role in these discoveries. The presence of DNA signatures representing novel microbial eukaryotes has been established primarily through the cloning and

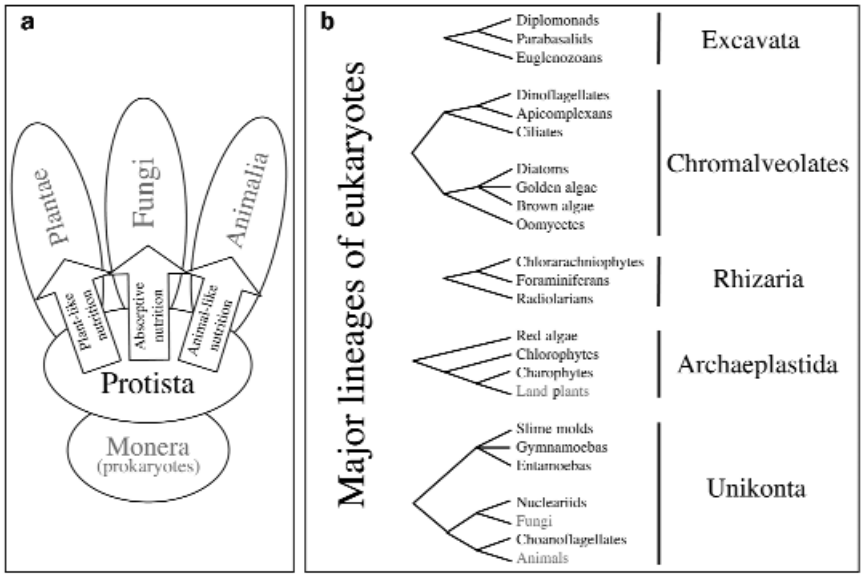

Figure 1. Changes in the generalized scheme of eukaryote phylogeny. (a) The scheme of Whittaker recognized five major kingdoms, with the protists (Protista) occupying one kingdom. The prominence of multicellular organisms was implied by the relative sizes of the balloons forming the Plantae, Fungi, and Animalia; redrawn from Whittaker (1969). (b) A modern hypothesis on the phylogeny of major eukaryote lineages as depicted in a basic biology textbook; redrawn from Campbell and Reece (2007). Note the placement of the plants, fungi, and animals (in gray) as relatively minor branches within the domain Eucarya. Details of the phylogenetic relationships among the major taxa have changed repeatedly in basic texts during the last decade. Many protistan taxa have not yet been placed with confidence within these schemes, and the evolutionary relationships of some taxa (even major lineages) are still debated.

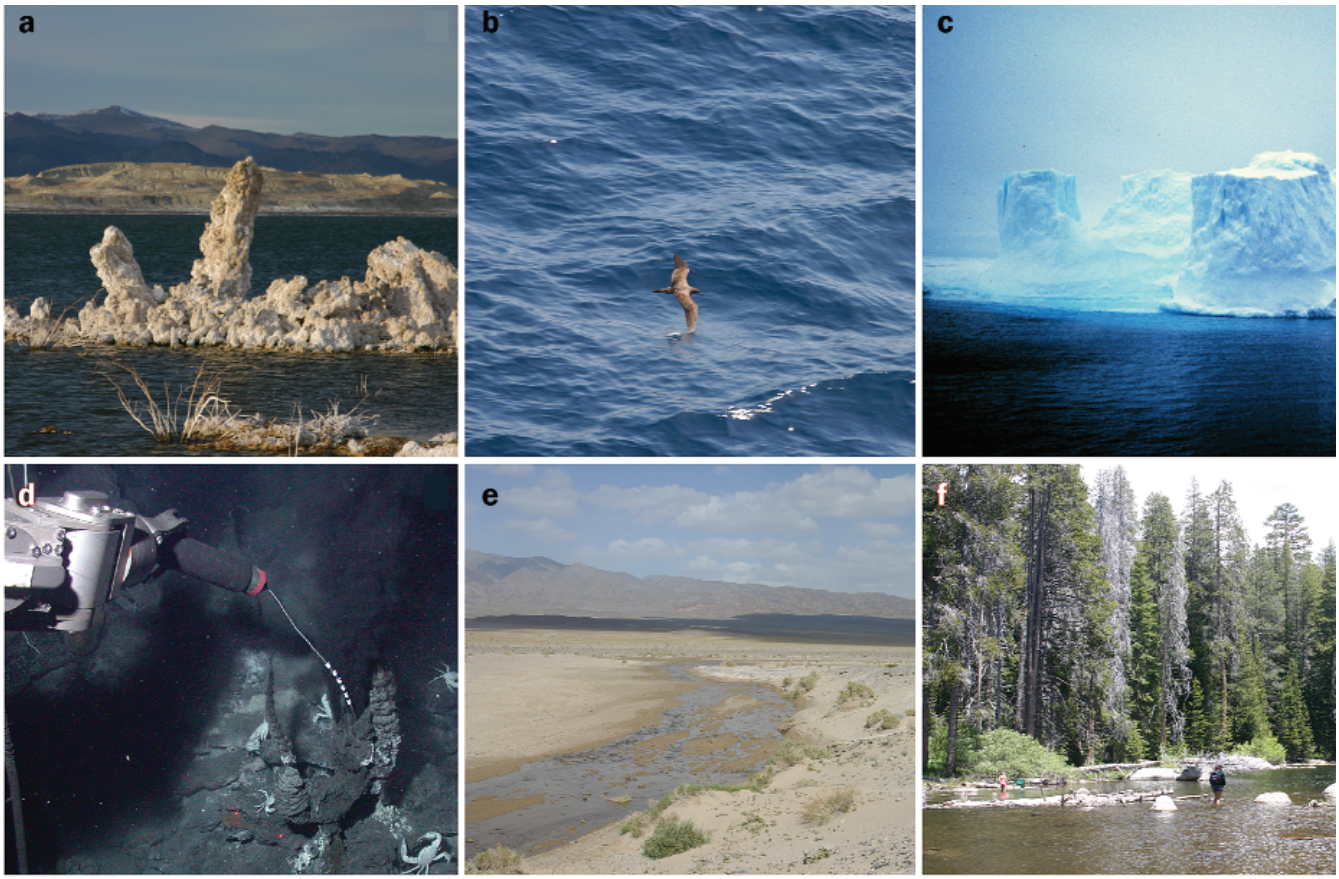

Figure 2. Remarkably diverse assemblages, many previously undescribed taxa, and even novel lineages of protists have recently been documented from a wide variety of aquatic and terrestrial ecosystems worldwide. Pictured are (a) the evaporative Mono Lake, California; (b) the open North Pacific Ocean; (c) the Ross Sea, Antarctica; (d) an East Pacific Rise hydrothermal vent; (e) a stream in Death Valley, California; and (f) Huntington Lake, California.

sequencing of small subunit (18S) ribosomal RNA (rRNA) genes extracted directly from environmental samples from a wide variety of geographical locations and depths (figure 2). These studies have indicated the presence of groups of sequences that have relatively low sequence similarity to any known, sequenced lineages of eukaryotes, although they clearly have their closest affinities with other unicellular eukaryotic (i.e., protistan) taxa.

Two such major clades of sequences representing previously undescribed protists have been documented. One group is composed of several small clades of sequences that show close affiliation to a number of lineages of Stramenopila within the Chromalveolata (Massana et al. 2004a). The stramenopiles are a diverse and abundant collection of taxa that include the diatoms (phototrophs bearing siliceous coverings), bicosoecids (small heterotrophic flagellates), chrysomonads (small phototrophic and heterotrophic flagellates), and a variety of other small phototrophic and heterotrophic forms. The novel environmental sequences that align with the stramenopiles were first observed in marine samples (and thus named MArine STramenopiles, or MAST, cells), and at least some of these taxa appear to be small heterotrophic forms (Massana et al. 2006). These cells have remained undetected until lately presumably because of their small sizes, nondistinctive morphologies, and inability to compete with other protists in enrichment cultures that are often used to isolate small heterotrophic flagellates.

A second major clade of novel 18S rRNA sequences obtained directly from environmental samples has phylogenetic affinity with the Alveolata of the Chromalveolata (López-García et al. 2001a, 2001b), but these sequences appear to be distinct from the ciliates, apicomplexans, and most dinoflagellates that make up the known lineages of alveolates. The novel alveolate sequences fall into two distinct groups (designated marine alveolates group I and group II). Recent evidence indicates that the group II alveolates may be related to a group of previously described but poorly characterized parasitic dinoflagellates (Groisillier et al. 2006), but the morphology and ecology of the protists in the group I alveolates are currently unknown. Indeed, DNA sequences are virtually the only form of information for these entities at the present time. Both the MAST sequences and the unknown alveolate sequences have been shown since their initial discoveries and descriptions to have widespread geographical distributions (Lovejoy et al. 2006, Stoeck et al. 2006, Countway et al. 2007).

The discovery of truly novel lineages of protists has been accompanied by the documentation of a tremendous breadth of diversity within lineages of protists previously thought to be well characterized through the traditional approaches of microscopy and culture. An excellent example is the extremely high diversity of minute chlorophytes documented within some freshwater ecosystems (Fawley et al. 2004), and the very large and diverse assemblages of minute protists from a wide variety of marine ecosystems (López-García et al. 2001b, Bass and Cavalier-Smith 2004, Massana et al. 2004b, Romari and Vaulot 2004, Countway et al. 2005, Lovejoy et al. 2007). In marine ecosystems, molecular phylogenetic studies based on sequences obtained from environmental samples have displayed sufficient sequence dissimilarity that several new algal classes have been erected to support the distinctions among these minute photosynthetic forms (Guillou et al. 1999, Kawachi et al. 2002). Subsequent ultrastructural, biochemical, and physiological information have generally supported these new classifications.

Why are we discovering so many previously undocumented phylotypes, and even novel protistan clades, through DNA sequencing campaigns? One contributing factor is that morphological characters have traditionally been the primary taxonomic criteria for defining protistan species, yet protists are extremely morphologically diverse. Their identification depends on a wide variety of methods for their collection, preservation, processing, and observation, and taxonomic criteria vary greatly among the different groups. It is therefore not surprising that much of the taxonomic breadth within some protistan groups has not yet been adequately defined. The existence of many small, morphologically amorphous species could easily explain why recent sequencing studies might reveal an enormous protistan diversity among these forms. While the presence of cryptic species has been known for many years, genetic methods are now allowing the rapid identification of physiologically distinct entities within morphospecies of protists. Fawley and colleagues (2004) recently documented extensive DNA sequence diversity among morphologically similar chlorophytes ("little green balls") within several lakes. Several studies revealing the different photosynthetic capabilities of these morphologically indistinguishable strains supported distinctions found through DNA sequencing.

In a similar sense, parasitic protists that have very limited or morphologically nondescript, free-living life stages complicate the identification of species and the assessment of protistan diversity using traditional approaches (microscopy and culture). Interestingly, it has been proposed that the unknown alveolate lineages may represent parasite taxa whose free-living stages have gone undetected by traditional methods (Moreira and López-García 2003). This hypothesis seems plausible, given the phylogenetic affinity of some of these alveolate sequences to known dinoflagellate parasites (Dolven et al. 2007). In addition, only relatively few of the myriad number of protistan species that exist in nature have ever been cultured and examined in detail, in part because of the highly selective nature of culture media and culture conditions. It is probable that a significant fraction of total protistan diversity remains to be documented, given the limitations of culture and direct microscopical observation to document the existence of these species.

Hurdles in estimating protistan diversity from genetic diversity

The assessment of protistan species diversity through DNA collection and sequencing is immune to the complications that assessment by morphological criteria imposes. The extraction, cloning, and sequencing of protistan genes are not believed to be taxon- or life-stage–dependent a priori. Therefore, the discovery of a diverse, and in many instances novel assemblage of protists might have been foreseen. However, genetic approaches may be influenced by a variety of problems such as gene copy number, extraction efficiency, efficiency of amplification of genes using the polymerase chain reaction (PCR), and the choice of cloning and sequencing primers. These issues may produce artifactual data, which has led some researchers to question whether many of the sequences obtained in environmental clone libraries actually represent the genetic signatures from real organisms. In addition, it is not yet clear how much of the genetic diversity observed within natural assemblages represents information that has morphological or physiological significance. Thus, at least two basic types of potential issues complicate the deduction of protistan diversity from DNA sequence information: the interpretation of species diversity from gene sequence diversity, and identification of methodological artifacts within molecular databases, such as chimeric sequences and sequencing artifacts (von Wintzingerode et al. 1997, Berney et al. 2004).

A fundamental assumption for applying genetic information to the study of protistan diversity is that the DNA sequences observed in environmental samples can be correlated directly to protistan species. This correlative task is not as trivial as many believe. Genetic dissimilarity exists between every two nonclonal organisms, and this dissimilarity confounds attempts to estimate protistan species diversity from sequence data. Even populations of protists in nature that might be expected to be homogeneous exhibit genetic diversity. For example, considerable genetic diversity has been documented within a single bloom population of the marine diatom *Ditylum brightwellii* (Rynearson and Armbrust 2004). Consequently, the use of DNA sequence information to document diversity must allow for an acceptable level of intraspecific sequence dissimilarity. Much of the resistance for accepting a molecular taxonomy for protists arises from disagreement over the definition of a reasonable boundary between intra- and interspecific variation for any given gene or genes.

Unfortunately, the problem of defining intra- and interspecies sequence variability is complicated by the confused species concept applied to protists. As noted above, protistan species traditionally have been defined on the basis of morphological characters, but reproductive and physiological criteria have also been used in species descriptions (Modeo

et al. 2003). For example, mating type compatibility, infectivity among opportunistically pathogenic protists, different feeding or nutrient-uptake kinetics, and the ability of certain phytoplankton taxa to produce toxins constitute ecologically important activities that have been employed as characters in defining species among strains that are sometimes morphologically indistinguishable, or nearly so. The tendency to employ multiple species concepts has increased in recent years, complicating the interpretation of protistan diversity in nature based on genetic information because these different species concepts present a moving target, so to speak, for ground-truthing a DNA-based taxonomy. The use of multiple species concepts of protists also affects our views regarding the geographical distributions of these species (see "Protistan distributions and biogeography," below).

The choice of a particular gene for establishing a molecular taxonomy useful for investigating protistan diversity is also not straightforward and it may not be universally applicable to all protistan taxa. Different protistan genes evolve (i.e., their sequences diverge) at different rates, and the rate for a given gene may not be the same for all protistan lineages (Sáez et al. 2003). The use of a rapidly evolving gene for molecular taxonomy may result in a single species being characterized as multiple entities, whereas the use of slowly evolving genes might group individuals that are generally accepted as different taxa. The potential for defining synonymous and cryptic species using sequence information is acute, given our limited state of available sequence information for protists and our meager knowledge of how that sequence information relates to traditional taxonomies. Nevertheless, although this state of our collective knowledge may limit the usefulness of molecular taxonomy for microbial ecologists now, it does not undermine the contribution that this approach will make in the future.

The use of DNA barcodes as a taxonomic scheme has detractors as well as strong proponents (Ebach and Holdrege 2005, Rubinoff et al. 2006); most of the detractors recognize the current limitations or complications with this approach, as noted above. Therefore, numerous refinements of a molecular taxonomy will assuredly take place (as they have for protistan molecular phylogeny) as more sequence information is amassed and combined with protistan species' descriptions based on traditional approaches. Ultimately, sequences of many genes may be employed to derive a robust DNA-based taxonomy (Blaxter 2004) in the way that multiple gene phylogenies are now used to provide multiple perspectives on the evolutionary history of protists (Harper et al. 2005). Progress is already being made in reconciling sequence-based phylogenies and taxonomies with more traditional species identifications for numerous taxa (Chantangsi et al. 2007, Evans et al. 2007, Hoef-Emden 2007). These refinements should greatly improve ecologists' ability to study protistan diversity in nature, because they will augment the tools presently available for conducting these studies. The incredible rate at which molecular analyses are being automated will also facilitate the processing of large numbers of samples, which are typical for ecological studies.

Among the methodological issues that challenge the use of DNA sequence information for estimating protistan diversity in natural microbial communities are the variable consistency and efficiency of DNA extraction, and the dependence on PCR amplification of DNA for cloning and sequencing studies. The latter artifacts include PCR primers and protocols that may fail to amplify a particular gene from all protistan species in a sample (Dawson and Pace 2002), and the generation of chimeric sequences during amplification. Chimeric sequences, or DNA strands produced from the fusion of two pieces of DNA from different species, are proposed as a possible explanation for the presence of unique sequences of microorganisms in environmental clone libraries (Berney et al. 2004). In addition, when using newly emerging DNA sequencing approaches, care must be taken to avoid errors that might generate spurious estimates of microbial diversity (Sogin et al. 2006).

Sequencing errors and the formation of chimeric sequences during sample processing create the possibility that at least some of the DNA sequences attributed to previously undetected protists are not valid sequences, but instead represent artifacts of the genetic approaches. However, it is highly unlikely that this problem could explain more than a small portion of the many unique sequences emerging from molecular ecological studies of microbial eukaryotes, because many of these novel sequences have now been recovered from numerous locales using slightly different cloning and sequencing approaches (López-García et al. 2001b, Fawley et al. 2004, Massana et al. 2004a, Groisillier et al. 2006, Countway et al. 2007, Lovejoy et al. 2007). Moreover, poor nucleic acid extraction efficiencies and PCR biases generally would tend to underestimate rather than overestimate the overall sequence diversity in a sample. It therefore seems clear that the use of genetic approaches to investigate the diversity of natural protistan assemblages is providing remarkable, believable new insights into the complexity and composition of microbial communities.

Estimating total protistan diversity

Protistan diversity is a very active research area, and molecular biological approaches play an important role in most of these studies. A rapidly growing number of protistan ecologists are involved in extensive cloning and sequencing campaigns whose overarching goal is the estimation and characterization of protistan diversity in natural microbial communities. As noted above, many previously undetected, undescribed protistan taxa have been discovered in the course of these studies. These studies are also beginning to alter our comprehension of the overall diversity, composition, and function of protistan assemblages, and they are generating new hypotheses on the relationship between diversity and the stability and resilience of microbial communities. The enormous diversity that is characteristic of natural protistan assemblages challenges even the considerable investigative power

afforded by molecular methods, but the constant and substantive advances in DNA sequencing and computational methods for exploiting sequence information are rapidly changing this situation.

Most molecular diversity studies have focused on the extraction, amplification, cloning, and sequencing of 18S rRNA genes because of the extensive public databases that exist for these genes, although other genes and intergenic spacers have also been employed. DNA sequences arising from these studies are compared in pairwise alignments to determine the number of operational taxonomic units (OTUs; i.e., the number of unique phylotypes, after a reasonable amount of intraspecific sequence dissimilarity has been determined) and the number of sequences that fall within each OTU. Sequences are routinely submitted to public databases to obtain as much taxonomic and phylogenetic information on OTUs as possible. The databases for protists are not yet as well developed as those for bacteria, and these databases contain many eukaryotic sequences that have not yet been related to specific taxa. For these reasons, submitted sequences are often identified as "unknown environmental" sequences. The capacity to obtain taxonomic and phylogenetic information on sequences is growing steadily and being refined constantly as databases expand and sequence information is linked to traditional taxonomic descriptions (Ludwig et al. 2004).

Cloning and sequencing DNA is a powerful approach for assessing the diversity of natural assemblages of protists, but it is costly and labor intensive. An alternative approach for large sequencing efforts is the extraction, amplification, and digestion of genes (usually 18S rRNA genes) into fragments, and the analysis of the number and size of these fragments as a means of examining changes in microbial community composition. This approach uses the digestion by endonucleases of the amplified genes to create a pattern of DNA fragments that provide information on the number of taxa in an assemblage. The underlying assumption is that the cleavage sites for an endonuclease are unique for each taxon in the assemblage, and thus the number and sizes of the fragments provide an indication of the species richness of the assemblage.

Many variations on this basic approach have been employed in DNA fragment analyses for protists, but the two most commonly used for protistan taxa are denaturing gradient gel electrophoresis (DGGE) and terminal random fragment length polymorphism (T-RFLP). These techniques provide a relatively rapid and inexpensive "snapshot" of community composition, and thus they have been applied for conducting comparative studies of multiple protistan assemblages from different sampling sites, temporally at a given location, or in response to environmental or experimental perturbations. Fragment patterns generated by these methods generally provide only fragment sizes and number, but it is possible to obtain some level of taxonomic information either by calibrating specific fragment sizes to microbial taxa (Kent et al. 2003) or by eluting and sequencing bands from a gel (Gast et al. 2004).

The use of fragment analysis for assessing the diversity of natural microbial communities has its detractors. Criticisms of the utility of these approaches have included the relatively low sensitivity of these methods (i.e., the ability to detect only the dominant taxa present in a sample), the potential for a single fragment to be composed of multiple taxa (i.e., the inability to discriminate between some taxa), and the inability to extrapolate from a particular fragment size to a named taxon (Bent et al. 2007, Danovaro et al. 2007). It is because of these criticisms, which are true in large part, that most genetic studies of protistan diversity have relied on gene sequencing. Nevertheless, fragment analysis is one of the easiest and least costly genetic ways to characterize the dominant taxa in an assemblage and to compare differences in the dominant taxa of the samples. Another common use of fragment analysis is the grouping of clones from a clone library on the basis of fragment lengths, and subsequent sequencing of one or a few clones from each group to reduce overall sequencing cost. This approach assumes that the different taxa in an assemblage all yield DNA fragments of unique length.

Emerging patterns in protistan diversity

The number of studies examining protistan diversity derived from gene sequences in samples collected across diverse natural communities is rapidly expanding (López-García et al. 2001a, 2001b, Dawson and Pace 2002, Moreira and López-García 2003, Fawley et al. 2004, Massana et al. 2004a, 2004b, Romari and Vaulot 2004, Countway et al. 2005, 2007, O'Brien et al. 2005, Groisillier et al. 2006, Lovejoy et al. 2007). Conclusions drawn from this research are predicated on the assumption that patterns of gene diversity obtained from molecular approaches can be extrapolated directly to species richness and the relative abundance of protistan species in nature. Although the validity of that assumption requires further testing, the results of these groundbreaking studies are providing new insights and stimulating novel hypotheses into protistan diversity, community composition, and community function. Through these studies it has been revealed that a relatively small number of taxa numerically dominate each assemblage, and that the dominant taxa appear to differ markedly among samples collected at different depths, locations, or sampling dates at a particular site. In addition, these assemblages are characterized by an extremely large (as yet undetermined) number of taxa at very low abundance. These features result in rank abundance curves for protistan communities whose overall shapes appear remarkably similar for numerous environmental samples (figure 3a), although the specific taxonomic composition of these curves differs greatly. In general, this pattern of protistan community structure is strongly analogous to the situation for bacterial assemblages in nature (Pedrós-Alió 2006, Sogin et al. 2006).

The tremendous diversity of phylotypes in natural ecosystems is easily visualized in a rarefaction curve, which graphically depicts the number of unique taxonomic units observed in a community as a function of individuals

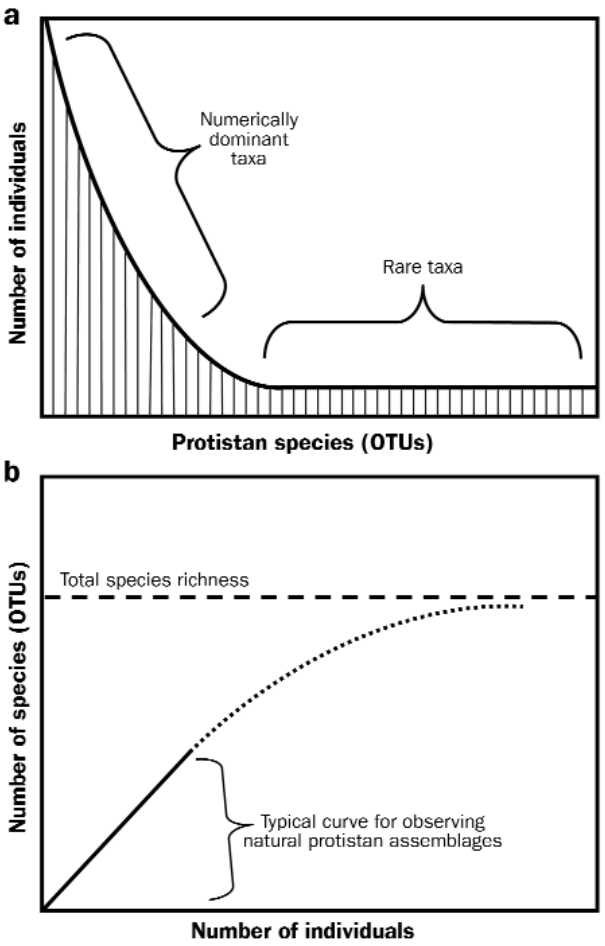

Figure 3. Idealized rank abundance curve (a) and rarefaction curve (b) of a natural protistan assemblage based on DNA sequence information. These assemblages typically are dominated numerically by relatively few taxa, but an extremely large number of rare taxa are also present (as shown on the right side of the rank abundance curve). These rare taxa exceed the ability of extant methods of microscopy, culture and genetics to assess total species richness. The inability of present methods to assess protistan diversity is indicated by the fact that rarefaction curves (b) often do not approach an asymptote, the total species richness of the assemblage (heavy dotted line). Abbreviation: OTUs, operational taxonomic units.

sampled (figure 3b). These curves approach a maximal value (total species richness) asymptotically as sampling effort reaches a point where most species have been observed at least once (Hughes and Hellmann 2005), but rarefaction curves constructed for protistan assemblages from clone libraries often show little, if any, inflection toward an asymptote (figure 3b; Countway et al. 2005, O'Brien et al. 2005, Stoeck et al. 2007). Our inability to exhaustively sample most natural assemblages of protists has led to the application of a number of diversity estimators to extrapolate from the available sequence information to the total diversity of the assemblage (Chao and Lee 1992, Hughes et al. 2001, Schloss and Handelsman 2004, Epstein and López-García 2007). These diversity indices use the shape of the rank abundance curve (or a portion of the curve) to estimate total species richness. These statistical estimators cannot account for PCR bias or poor extraction efficiencies, however, and thus diversity estimates based on these indices will probably increase as genetic methods provide a more accurate sampling of total protistan diversity and rank abundances of protists. The degree of uncertainty associated with these estimators is quite high at present, and the reality is that our knowledge of the overall diversity of most natural assemblages of protists is still rudimentary.

Protistan assemblage composition and reassembly

An interesting feature of the information now available regarding the rank abundance curves for natural assemblages of protists is the existence of very large numbers of rare phylotypes. This situation also typifies bacterial assemblages (Pedrós-Alió 2006, Sogin et al. 2006). What is the significance of this rare biosphere that appears to characterize microbial assemblages? Clearly, some of the sequences on which these claims are based could be the result of methodological artifact. The extraction and sequencing of genes from environmental samples does not preclude the inclusion of dead or moribund cells, or inactive (encysted) cells that may have been transported into a region but have no ecological function there. In addition, the complexities associated with defining species (i.e., OTUs) from sequence data may create "microdiversity" that is below the level that is ecologically relevant or taxonomically accepted, and true methodological issues such as the formation of chimeric sequences and sequencing artifacts may occur. Regardless, the sheer magnitude of rare taxa argues that this general pattern is real.

One hypothesis that was offered long ago as an explanation for the presence of rare microbial phylotypes is based on the assumption that small microbes are ubiquitous. Their small size and high abundances afford easy transportation, and thus one might expect global dissemination of these taxa. Baas Becking (1934) proposed the popular explanation that "everything is everywhere, but, the environment selects" to suggest that although microbes might be omnipresent (or constantly reintroduced) in all of Earth's ecosystems, local conditions provide strong selective pressure to maintain the vast majority of these species at very low abundances, which cannot be detected by most methods of observing microbes. This enduring paradigm has been recast many times and in slightly different ways to explain the apparent endemism of bacteria despite their global distribution (de Wit and Bouvier 2006). It is possible that we are now simply more aware of the huge "background" of rare microbial taxa in natural communities because molecular methods provide a means of detecting the presence of many taxa (i.e., phylotypes) that were previously below the limits of detection for microscopy and culture. Perhaps it is because we can now perceive these taxa that the issue of the rare biosphere has captured so much recent attention.

Many investigators have interpreted Baas Becking's suggestion to mean that rare taxa generally have little or no ecological purpose. It is possible that this supposition is true and that the vast abundances of protists in any given ecosystem make complete extinction of these populations difficult if not impossible (Fenchel 2005). If true, however, then it is not easy to explain why some species of small protists can be enriched to high abundances in cultures from undetectable abundances in natural samples collected from many regions of the globe (Finlay and Clarke 1999). Clearly, these species are viable and capable of growth even if they constitute part of the rare biosphere in most situations. It is also hard to explain the persistence of such a significant fraction of the total microbial biomass as rare taxa. Although each rare phylotype may appear only once or twice in a large clone library, collectively these taxa can constitute 20% to 30% of the total number of clones in the library (Caron et al. 2004, Countway et al. 2005, 2007). It is difficult to rationalize the existence of such a large fraction of total diversity as inactive microbes, given the strong predation pressure on microbes that is believed to typify most aquatic ecosystems (Sherr and Sherr 2002). At least two simple explanations seem plausible: either the transport of taxa over great distances is extremely common and extensive (and so rapidly replaces losses due to predation), or the rare taxa in ecosystems are occasionally active, and therefore autochthonous growth plays a role in maintaining their presence.

Do rare protistan taxa play a role in community function? Recent evidence suggests that the rare biosphere might play an important role by serving as a source of microbial taxa that can rapidly increase in abundance and take over important ecological activities when environmental conditions no longer favor the dominant species. As an example, a simple containment experiment conducted in the coastal North Atlantic in which ocean water was placed into containers maintained at ambient light and temperature, and sampled initially and after 24 and 72 hours of incubation, showed remarkable changes in the protistan assemblage composition (Countway et al. 2005). Sequencing of 18S rRNA genes revealed that only 18% of a total of 165 protistan OTUs observed in the study were present at all three time points, while 65% of the OTUs observed in the study were detected at only one of the three sampling times. These findings, although preliminary, have at least three potentially important implications for the diversity and composition of natural protistan assemblages. First, the composition of at least some protistan assemblages is extremely dynamic, much more dynamic than had been known or thought possible for microbial eukaryotes. Second, taxa that are below the limits of detection using extant methodology for assessing diversity and species composition are capable of becoming important components of the protistan assemblage in a short time. Third, very subtle changes in environmental conditions may result in major differences or rapid changes in the overall composition of a protistan assemblage.

It is also possible that rare taxa play important roles at low abundance (i.e., they need not be dominant to play pivotal ecological roles). This situation would be true if rare species fulfilled essential ecological roles that could affect overall system function (e.g., an important role has been proposed for nitrogen-fixing bacteria in many ecosystems, even though the bacteria may have low relative abundance). Similar situations exist for keystone predators among macroscopic communities, which exert considerable control over community structure and function at relatively low numerical abundance or biomass. The converse situation is also possible: high abundance of a taxon does not always indicate a pivotal ecological role. It can often indicate the relative nonparticipation of a species in ecosystem function (e.g., the accumulation of inedible algae in some aquatic ecosytems). Future studies must strive to look beyond the general patterns of community structure and interpret these patterns with knowledge of the ecological roles of the taxa present.

Notwithstanding these limitations of interpretation, initial studies to investigate protistan community structure make it clear that there is a critical need for methodologies that will enable researchers to sample much more deeply into the total species richness present in natural assemblages in order to better understand the dynamics of these assemblages. Early results indicate that many of the rare taxa can and do play significant ecological roles, and that relatively minor shifts in environmental conditions are sufficient to bring about dramatic changes in species composition. In addition, these changes occurred at rates that were much faster than might have been expected.

The implications of these results for biogeochemical cycles in natural microbial communities may be considerable. Many of the rate processes that are measured in aquatic ecosystems (e.g., rates of primary productivity, rates of herbivory or bacterivory) are obtained in experiments that involve the incubation of samples. These incubations may last from one to a few days and derive rates on the basis of changes in conditions in the bottles from the beginning to the end of the incubations (e.g., changes in chlorophyll concentration). One underlying assumption in these approaches has been that the microbial assemblages and their activities are not altered substantially by containment in bottles over the duration of the experiment. The results of Countway and colleagues (2005) raise questions as to whether this assumption is valid. On the other hand, it is possible that community functions—primary production, consumption, respiration—remain relatively unchanged during the incubations, even though the species conducting these processes change significantly. Therefore, characterizing and understanding both the changes occurring in microbial communities and the community activities during these experiments are fundamental to understanding the degree to which measurements of microbial activity in containers are representative of these processes in nature.

Protistan distributions and biogeography

Protistan biogeography is presently a contentious issue. If protistan species obey the line of reasoning asserted by many microbiologists that "everything is everywhere" (Baas Becking 1934), then many (perhaps most) protists should not demonstrate biogeography. At some point, however, size and other factors such as environmental tolerances present barriers to propagation that are sufficient to suppress global dissemination of organisms. There are, of course, many examples of endemism among macroscopic species, although human activities have changed the global distribution of many species during the past millennium.

Despite considerable rhetoric and highly divergent views on this topic in recent years, the available data would seem to argue that both the proponents for limited distributions of protists and those for ubiquitous distributions have valid cases. Some protistan species do appear to be globally distributed, judging from observations of these species from very disparate ecosystems around the world (Finlay and Clarke 1999). Nevertheless, many larger protists appear to have limited geographical ranges. For example, the utility of some diatoms, radiolaria, and foraminifera for studies of paleoclimatological reconstruction are predicated on the existence of geographically restricted distributions of these species as indicators of specific water masses in the world's oceans (Bé 1977). In an attempt to extend these observations, some researchers have argued recently that many protists display endemicity (Boenigk et al. 2006, Foissner 2006).

Both sides of this debate may have credible arguments because both claims may be true in part. Protistan species encompass a tremendous size range (< 1.0 to > 10,000 micrometers). It is probable that these species collectively span the critical size above which transport is restricted but below which global dissemination can take place easily. In addition, Fenchel (2005) estimated that beneath a 1-hectare surface area of shallow water there may be as many as 10^{16} protists, and as many as 10^{7} individuals that are undetectable to observers using extant methodology. The vast abundances of protists (even rare taxa) in a large volume of water or large expanse of soil might make the complete removal of competitive inferiors a very long process, much longer than the temporal scales for microbes to be introduced into these regions, or for environmental conditions to remain unfavorable for their growth. Therefore, although many species cannot be observed in a given environment, they may still be present. This argument seems particularly convincing for protists that produce cysts and for small protists that might go undetected because they possess unremarkable morphologies.

The morphological and genetic distinctions that have been documented within protistan morphospecies over relatively short geographical distances indicate that some degree of spatial heterogeneity or discontinuity exists for the distributions of many protistan taxa (Rynearson and Armbrust 2004, Foissner 2006). The degree to which these distinctions indicate endemism as opposed to intraspecific variability is not yet clear. Different researchers accept different levels of variability in the characters that are employed to describe protistan species, and therefore some nonsignificant component of this argument is probably semantic, a result of disagreement over what constitutes species-level distinctions. As a result, arguments regarding the ubiquity or endemism of protistan species have been intermingled with debates over the somewhat variable species concept employed for protists. Even so, not all species of protists can possibly be present in all habitats, especially unique, remote, or minute habitats, and not all species have sizes or life stages that would facilitate worldwide dispersal.

The debate regarding protistan biogeography has been pursued with vigor in the literature, and its outcome has important consequences for interpreting protistan diversity and protistan community structure. High levels of endemism among protistan species would imply that total protistan diversity might be extraordinarily high (Foissner 2006), while ubiquitous dispersal would imply a much lower overall diversity of protists (Finlay and Clarke 1999, Finlay and Fenchel 1999). Moreover, if protistan species are never truly driven to extinction in most ecosystems, then the rare biosphere may indeed play an important function as a source of species that can take on important ecological roles when environmental conditions no longer favor the growth of the dominant taxa. In that way, the ecological function of communities may be preserved, even though the species conducting these processes may change. The general tenets that species diversity and ecological redundancy among species have stabilizing effects on community function are common themes in both classical and protistan ecology (Naeem and Shibin 1997, Griffiths et al. 2004).

Autecological studies of protists

Although modern molecular biological methods and approaches are only beginning to modify our views of protistan species diversity, these techniques have already become immensely important for studying the autecology of some protistan species that have exceptional ecological importance or relevance to human health. Studies focused on species that are pathogens (e.g., *Giardia, Cryptosporidium, Perkinsus*) or toxic algae (e.g., *Alexandrium, Pseudo-nitzschia, Karenia*) have improved through the application of genetic and immunological methods that allow identification and enumeration of these species at relatively low abundances, thus enabling studies of the environmental factors that lead to their occurrence and activities in nature (Caron et al. 2004). These studies strive to understand the distributions and abundance of these species, and ultimately to develop a predictive understanding of their ecology. New tools that allow differentiation of target species from among a myriad of co-occurring protists are vital to improving our knowledge of these taxa.

Molecular approaches employed for these purposes have included immunological methods (immunofluorescence microscopy, immunofluorescence flow cytometry, enzyme-

linked immunosorbent assay), and genetic methods (fluorescence *in situ* hybridization, quantitative real-time PCR) to target individual protistan taxa in natural samples (Scholin et al. 1996, Vrieling and Anderson 1996, Popels et al. 2003, Countway and Caron 2006). These methods routinely provide much greater sensitivity than can the traditional approaches of microscopy and culture, and they typically are applicable over large ranges of abundance. For example, real-time PCR approaches can quantitatively measure the signal from a molecular target at concentrations ranging over several orders of magnitude. Additionally, the formats for some of these approaches allow the processing of multiple samples simultaneously or in a short period of time relative to traditional approaches, and thus enable ecological studies that typically involve large numbers of samples.

Emerging frontiers

Several research areas lie on the immediate scientific frontier of protistan biology and ecology, and molecular approaches will feature prominently in these ventures. Protistan phylogeny has been a focal point for this work for nearly two decades, and this work has changed our view of the evolution of eukaryotes dramatically in recent years (figure 1). Several higher-level reorganizations have taken place, but the validity of a number of these groupings, some of the evolutionary relationships among them, and most of the details regarding relationships of many smaller groups of protists to or within these higher-level groupings still require resolution. One recent opinion predicts that new, previously undetected clades of protists will continue to appear as sequencing campaigns continue, but that these clades will fall into well-established higher taxonomic groups (Richards and Bass 2005). Novel hypotheses on eukaryote evolution will certainly continue to emerge and be tested, and we can expect that protistan evolutionary schemes in basic biology texts will continue to change in the foreseeable future.

The application of molecular biological approaches to studies of protistan diversity now dominates the current discovery phase of ecological research on these species. Several dozen recent studies on this topic have focused on documenting protistan phylotype diversity in a wide assortment of environments. These studies have established the existence of enormous diversity, and many novel DNA sequences. This work must continue to garner attention as we delve deeply into natural protistan assemblages. New, high-throughput sequencing approaches will lead the way in this work.

In addition, these benchmark surveys of environmental DNA sequences are opening several new avenues of protistan research. Cloning and sequencing of DNA sequences has yielded information on putative protistan taxa, but inherent in this approach is the assumption that each phylotype represents a protist, and an active or potentially active microbe of a microbial community. That assumption is probably not completely valid. A concentrated effort is needed (a) to relate DNA sequences obtained from environmental samples to morphological entities, (b) to establish which of these taxa are active within a given sample and which are not, and (c) to employ this knowledge to ascertain the physiologies, ecological roles, and life histories of novel protistan taxa. The first of these goals will determine which of the many novel sequences that have been observed are actually protists, and which are the result of artifact (e.g., pseudogenes, chimeras, sequencing artifacts). The second will help establish which protistan taxa are contributing to community function at a given time and place. To date, functional contribution has largely been inferred from the most abundant taxa in rank abundance curves (figure 3a), and that is an unsatisfactory oversimplification of nature. Stoeck and colleagues (2007) reported substantial differences between sequences obtained from rDNA clone libraries and rRNA libraries obtained from the same samples. The latter clone libraries were constructed from ribosomal RNA, and therefore it was concluded that the rRNA libraries represented the metabolically active protistan taxa in the assemblage at the time of collection. The differences observed using these two methods of inferring protistan community composition imply that we still have much to learn regarding the active component of these assemblages in nature.

The third goal of extending the discovery of new phylotypes to assessments of the ecological roles of novel and uncultured taxa constitutes much of the present emphasis in bacterial and archaeal ecology. The discovery of novel DNA sequences representing new, previously undetected lineages of protists and the true breadth of species richness within known lineages has now begun to entrain studies utilizing traditional methods of culture and microscopy to establish the morphology and general biology of these taxa. These efforts will benefit from genomic and gene expression studies that are now beginning to focus on ecologically important protists (Wahlund et al. 2004, Lidie et al. 2005, La Claire 2006). These genetic campaigns will greatly expand our now limited knowledge of the physiological potential of protists in nature and enable unprecedented insights into their physiological response to environmental influences.

Researchers studying microbial eukaryotes are only now beginning to employ metagenomic and transcriptomic approaches to better understand protistan activities in nature. The scientific contribution of this work is still in its infancy, but its application in the future will doubtless be widespread. This research has gained momentum in ecological studies of bacteria and archaea because it provides a mechanism for exploring the physiology of microbial species that have not yet been brought into culture, thereby providing insight into the nutritional requirements and ecological roles of these species. Studies of bacteria and archaea have been facilitated by the relatively small size of their genomes relative to many eukaryotes. Technological advances of the future will undoubtedly sweep these methodological limitations aside, making our imaginations the only limiting step for application and exploitation of genomic information on protists.

Another main avenue for protistan ecological research will be employing new molecular methods to understand changes

in assemblage structure, and how these changes relate to overall protistan diversity and community function. Do changes in the composition of a protistan assemblage result in shifts in community function, or is community function relatively unchanged in the face of substantial reassembly of the dominant and rare taxa? An answer to this fundamental and long-standing question in ecology is now within the grasp of protistan ecologists as powerful new molecular approaches come on line.

Finally, microbial communities *sensu latu* represent excellent model systems for testing hypotheses concerning basic ecological principles. Microbes are eminently suited for studies of community function and response because they are easily contained and manipulated, although changes in community structure and function that might take place with containment will need to be characterized. Microbes have been employed for many years to test ecological theories, but their widespread use has been thwarted by the immense complexity of microbial assemblages and our limited ability to delineate community composition and response. Research under way at this time, and described in part in this article, will pave the way for a much better understanding of protistan community composition and dynamics, and the use of these assemblages as experimental model systems to understand evolution, ecological interactions among species, and biogeochemical processes.

Protists contribute meaningfully to the biomass and activities of microbial communities on a global scale. Together with bacteria, archaea, and viruses, they are the primary drivers of the production, utilization, and degradation of much of the organic matter on our planet and the cycling of many elements. Understanding their diversity of form and function is scientifically fascinating, and it is also fundamental to understanding how biological communities on our planet function today and how they have functioned throughout much of our planet's history.

Acknowledgments

The author is grateful to Rebecca J. Gast, David J. Patterson, Peter D. Countway, and Astrid Schnetzer for valuable discussions that contributed to the content of this manuscript, and to the officers of the International Society of Protistologists for proposing its inclusion in this journal. Preparation of the manuscript was supported in part from grants from the National Science Foundation (CCR-0120778, MCB-0084231, OPP-0125437, and OPP-0542456), the National Oceanic and Atmospheric Administration (NA160P2790), the Environmental Protection Agency (RD-83170501), and the Gordon and Betty Moore Foundation.

References cited

Adl SM, et al. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. Journal of Eukaryotic Microbiology 52: 399–451.

Alexander M. 1961. Introduction to Soil Microbiology. New York: Wiley.

Baas Becking LGM. 1934. Geobiologie of inleiding tot de milieukunde. The Hague: W. P. Van Stockkum and Zoon.

Bass D, Cavalier-Smith T. 2004. Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of Cercozoa (Protozoa). International Journal of Systematic and Evolutionary Microbiology 54: 2393–2404.

Bé AWH. 1977. An ecological, zoogeographic and taxonomic review of recent planktonic forminifera. Pages 1–100 in Ramsay ATS, ed. Oceanic Micropaleontology. New York: Academic Press.

Bent SJ, Peirson JD, Forney LJ. 2007. Measuring species richness based on microbial community fingerprints: The emperor has no clothes. Applied and Environmental Microbiology 73: 2399.

Berney C, Fahrni J, Pawlowski J. 2004. How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys. BMC Biology 4: 2–13.

Blaxter ML. 2004. The promise of a DNA taxonomy. Philosophical Transactions of the Royal Society of London B 359: 669–679.

Boenigk J, Pfandl K, Garstecki T, Harms H, Novarino G, Chatzinotas A. 2006. Evidence for geographic isolation and signs of endemism within a protistan morphospecies. Applied and Environmental Microbiology 72: 5159–5164.

Bonowski M. 2004. Protozoa and plant growth: The microbial loop in soil revisited. New Phytologist 162: 617–631.

Campbell NA, Reece JB. 2007. Biology. New York: Benjamin Cummings.

Caron DA, Schnetzer A. 2007. Protistan community structure. Pages 454–468 in Hurst CJ, Crawford RL, Garland JL, Lipson DA, Mills AL, Stetzenbach LD, eds. Manual of Environmental Microbiology. Washington (DC): ASM Press.

Caron DA, Dam HG, Kremer P, Lessard EJ, Madin LP, Malone TC, Napp JM, Peele ER, Roman MR, Youngbluth MJ. 1995. The contribution of microorganisms to particulate carbon and nitrogen in surface waters of the Sargasso Sea near Bermuda. Deep-Sea Research I 42: 943–972.

Caron DA, Countway PD, Brown MV. 2004. The growing contributions of molecular biology and immunology to protistan ecology: Molecular signatures as ecological tools. Journal of Eukaryotic Microbiology 51: 38–48.

Chantangsi C, Lynn DH, Brandl MT, Cole JC, Hetrick N, Ikonomi P. 2007. Barcoding ciliates: A comprehensive study of 75 isolates of the genus *Tetrahymena*. International Journal of Systematic and Evolutionary Microbiology 57: 2412–2423.

Chao A, Lee SM. 1992. Estimating the number of classes via sample coverage. Journal of the American Statistical Association 87: 210–217.

Countway PD, Caron DA. 2006. Abundance and distribution of *Ostreococcus* sp. in the San Pedro Channel, California (USA) revealed by qPCR. Applied and Environmental Microbiology 72: 2496–2506.

Countway PD, Gast RJ, Savai P, Caron DA. 2005. Protistan diversity estimates based on 18S rDNA from seawater incubations in the western North Atlantic. Journal of Eukaryotic Microbiology 52: 95–106.

Countway PD, Gast RJ, Dennett MR, Savai P, Rose JM, Caron DA. 2007. Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western N. Atlantic (Sargasso Sea and Gulf Stream). Environmental Microbiology 9: 1219–1232.

Danovaro R, Luna GM, Dell-Anno A, Pietrangeli B. 2007. Measuring species richness based on microbial community fingerprints: The emperor has no clothes: Authors' reply. Applied and Environmental Microbiology 73: 2399–2401.

Dawson SC, Pace NR. 2002. Novel kingdom-level eukaryotic diversity in anoxic environments. Proceedings of the National Academy of Sciences 99: 8324–8329.

de Wit R, Bouvier T. 2006. 'Everything is everywhere, but, the environment selects'; what did Baas Becking and Beijerinck really say? Environmental Microbiology 8: 755–758.

Dolven JK, Lindqvist C, Albert VA, Bjørklund KR, Yuasa T, Takahashi O, Mayama S. 2007. Molecular diversity of alveolates associated with neritic North Atlantic radiolarians. Protist 158: 65–76.

Ebach MC, Holdrege C. 2005. More taxonomy, not DNA barcoding. BioScience 55: 822–823.

Epstein S, López-García P. 2007. "Missing" protists: A molecular perspective. Biodiversity and Conservation 17: 261–276.

Evans KM, Wortley AH, Mann DG. 2007. An assessment of potential diatom "barcode" genes (*cox*1, *rbc*L, 18S and ITS rDNA) and their effectiveness in determining relationships in Sellaphora (Bacillariophyta). Protist 158: 349–364.

Fawley MJ, Fawley KP, Buchheim MA. 2004. Molecular diversity among communities of freshwater microchlorophytes. Microbial Ecology 48: 489–499.

Fenchel T. 2005. Cosmopolitan microbes and their 'cryptic' species. Aquatic Microbial Ecology 41: 49–54.

Finlay BJ, Clarke KJ. 1999. Ubiquitous dispersal of microbial species. Nature 400: 828.

Finlay BJ, Fenchel T. 1999. Divergent perspectives on protist species richness. Protist 150: 229–233.

Foissner W. 2006. Biogeography and dispersal of micro-organisms: A review emphasizing protists. Acta Protozoologica 45: 111–136.

Gast RJ, Dennett MR, Caron DA. 2004. Characterization of protistan assemblages in the Ross Sea, Antarctica by denaturing gradient gel electrophoresis. Applied and Environmental Microbiology 70: 2028–2037.

Griffiths BS, Kuan HL, Ritz K, Glover LA, McCaig AE, Fenwick C. 2004. The relationship between microbial community structure and functional stability, tested experimentally in an upland pasture soil. Microbial Ecology 47: 104–113.

Groisillier A, Massana R, Valentin K, Vaulot D, Guillou L. 2006. Genetic diversity and habitats of two enigmatic marine alveolate lineages. Aquatic Microbial Ecology 42: 277–291.

Guillou L, Chrétiennot-Dinet M-J, Medlin LK, Claustre H, Loiseaux-de Goer S, Vaulot D. 1999. *Bolidomonas*: A new genus with two species belonging to a new algal class, the Bolidophyceae (Heterokonta). Journal of Phycology 35: 368–381.

Harper JT, Waanders E, Keeling PJ. 2005. On the monophyly of chromalveolates using a six-protein phylogeny of eukaryotes. International Journal of Systematic and Evolutionary Microbiology 55: 487–496.

Hoef-Emden K. 2007. Revision of the genus *Cryptomonas* (Cryptophyceae) II: Incongruences between the classical morphospecies concept and molecular phylogeny in smaller pyrenoid-less cells. Phycologia 46: 402–428.

Hughes JB, Hellmann JJ. 2005. The application of rarefaction techniques to molecular inventories of microbial diversity. Methods in Enzymology 397: 292–308.

Hughes JB, Hellmann JJ, Ricketts TH, Bohannan BJM. 2001. Counting the uncountable: Statistical approaches to estimating microbial diversity. Applied and Environmental Microbiology 67: 4399–4406.

Karl DM. 2007. Microbial oceanography: Paradigms, processes and promise. Nature Reviews Microbiology 5: 759–769.

Kawachi M, Atsumi M, Ikemoto H, Miyachi S. 2002. *Pinguiochrysis pyriformis* gen. et sp. nov. (Pinguiophyceae), a new picoplanktonic alga isolated from the Pacific Ocean. Phycological Research 50: 49–56.

Kent AD, Smith DJ, Benson BJ, Triplett EW. 2003. Web-based phylogenetic assignment tool for the analysis of terminal restriction fragment length polymorphism profiles of microbial communities. Applied and Environmental Microbiology 69: 6768–6776.

La Claire JW II. 2006. Analysis of expressed sequence tags from the harmful alga, *Prymnesium parvum* (Prymnesiophyceae, Haptophyceae). Marine Biotechnology 8: 534–546.

Lidie KB, Ryan JC, Barbier M, Van Dolah FM. 2005. Gene expression in Florida red tide dinoflagellate *Karenia brevis*: Analysis of an expressed sequence tag library and development of a DNA microarray. Marine Biotechnology 7: 481–493.

López-García P, López-López A, Moreira D, Rodríguez-Valera F. 2001a. Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. FEMS Microbiology Ecology 36: 193–202.

López-García P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D. 2001b. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature 409: 603–607.

Lovejoy C, Massana R, Pedrós-Alío C. 2006. Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent seas. Applied and Environmental Microbiology 72: 3085–3095.

Lovejoy C, Vincent WF, Bonilla S, Roy S, Martineau M-J, Terrado R, Potvin M, Massana R, Pedrós-Alió C. 2007. Distribution, phylogeny and growth of cold-adapted picoprasinophytes in Arctic Seas. Journal of Phycology 43: 78–89.

Ludwig W, et al. 2004. ARB: A software environment for sequence data. Nucleic Acids Research 32: 1363–1371.

Malone TC. 1971. The relative importance of nannoplankton and netplankton as primary producers in tropical oceanic and neritic phytoplankton communities. Limnology and Oceanography 16: 633–639.

Massana R, Castresana J, Balagué V, Guillou L, Romari K, Groisillier A, Valentin K, Pedrós-Alió C. 2004a. Phylogenetic and ecological analysis of novel marine stramenopiles. Applied and Environmental Microbiology 70: 3528–3534.

Massana R, Balague V, Guillou L, Pedros-Alio C. 2004b. Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. FEMS Microbiology Ecology 50: 231–243.

Massana R, Guillou L, Terrado R, Forn I, Pedrós-Alió C. 2006. Growth of uncultured heterotrophic flagellates in unamended seawater incubations. Aquatic Microbial Ecology 45: 171–180.

Modeo L, Petroni G, Rosati G, Montagnes DJS. 2003. A multidisciplinary approach to describe protists: Redescriptions of *Novistrombidium testaceum* and *Strombidium inclinatum* Montagnes, Taylor and Lynn 1990 (Ciliophora, Oligotrichia). Journal of Eukaryotic Microbiology 50: 175–189.

Moreira D, López-García P. 2003. Are hydrothermal vents oases for parasitic protists? Trends in Parasitology 19: 556–558.

Naeem S, Shibin L. 1997. Biodiversity enhances ecosystem stability. Nature 390: 507–509.

O'Brien HE, Parrent JL, Jackson JA, Moncalvo J-M, Vilgalys R. 2005. Fungal community analysis by large-scale sequencing of environmental samples. Applied and Environmental Microbiology 71: 5544–5550.

Olson RJ, Chisholm SW, Zettler ER, Altabet MA, Dusenberry JA. 1990. Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. Deep-Sea Research 37: 1033–1051.

Pedrós-Alió C. 2006. Microbial diversity: Can it be determined? Trends in Microbiology 14: 257–263.

Pomeroy LR. 1974. The ocean's food web, a changing paradigm. BioScience 24: 499–504.

Popels LC, Cary SC, Hutchines DA, Forbes R, Pustizzi F, Gobler CJ, Coyne KJ. 2003. The use of quantitative polymerase chain reaction for the detection and enumeration of the harmful alga *Aureococcus anophagefferens* in environmental samples along the United States East Coast. Limnology and Oceanography 48: 92–102.

Richards TA, Bass D. 2005. Molecular screening of free-living microbial eukaryotes: Diversity and distribution using a meta-analysis. Current Opinion in Microbiology 8: 240–252.

Romari K, Vaulot D. 2004. Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. Limnology and Oceanography 49: 784–798.

Rubinoff D, Cameron S, Will K. 2006. Genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. Journal of Heredity 97: 581–594.

Rynearson TA, Armbrust EV. 2004. Genetic differentiation among populations of the planktonic marine diatom *Ditylum brightwellii* (Bacillariophyceae). Journal of Phycology 40: 34–43.

Sáez AG, Probert I, Geisen M, Quinn P, Young JR, Medlin LK. 2003. Pseudocryptic speciation in coccolithophores. Proceedings of the National Academy of Sciences 100: 7163–7168.

Schloss PD, Handelsman J. 2004. Status of the microbial census. Microbiology and Molecular Biology Reviews 68: 686–691.

Scholin CA, Buck KR, Britschgi T, Cangelosi G, Chavez FP. 1996. Identification of *Pseudo-nitzschia australis* (Bacillariophyceae) using rRNA-targeted probes in whole cell and sandwich hybridization formats. Phycologia 35: 190–197.

Sherr BF, Sherr EB, Caron DA, Vaulot D, Worden AZ. 2007. Oceanic protists. Oceanography 20: 102–106.

Sherr EB, Sherr BF. 2002. Significance of predation by protists in aquatic microbial food webs. Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology 81: 293–308.
Sieburth JM. 1979. Sea Microbes. New York: Oxford University Press.
Simpson AGB, Roger AJ. 2004. The real 'kingdoms' of eukaryotes. Current Biology 14: R693–696.
Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, Arrieta JM, Herndl GJ. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere." Proceedings of the National Academy of Sciences 103: 12115–12120.
Stoeck T, Hayward B, Taylor GT, Varela R, Epstein SS. 2006. A multiple PCR-primer approach to access the microeukaryotic diversity in environmental samples. Protist 157: 31–43.
Stoeck T, Zuendorf A, Behnke A, Breiner H-W. 2007. A molecular approach to identify active microbes in environmental eukaryote clone libraries. Microbial Ecology 53: 328–339.
Suttle CA. 2005. Viruses in the sea. Nature 437: 356–361.
von Wintzingerode F, Göbel UB, Stackebrandt E. 1997. Determination of microbial diversity in environmental samples: Pitfalls of PCR-based rRNA analysis. FEMS Microbiology Reviews 21: 213–229.
Vrieling EG, Anderson DM. 1996. Immunofluorescence in phytoplankton research: Applications and potential. Journal of Phycology 32: 1–16.
Wahlund TM, Hadaegh AR, Clark R, Nguyen B, Fanelli M, Read BA. 2004. Analysis of expressed sequence tags from calcifying cells of marine coccolithophorid (*Emiliania huxleyi*). Marine Biotechnology 6: 278–290.
Whittaker RH. 1969. New concepts of kingdoms of organisms. Science 163: 150–160.

David A. Caron (e-mail: dcaron@usc.edu) is with the Department of Biological Sciences at the University of Southern California, Los Angeles.