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

Journal of Eukaryotic Microbiology ISSN 1066-5234

SYMPOSIUM ARTICLE

# Termite Hindguts and the Ecology of Microbial Communities in the Sequencing Age

Vera Tai & Patrick J. Keeling

Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC, Canada

**Keywords**
Diversity; symbiosis; protist; oxymonad; parabasalia; single-cell; high-throughput.

**Correspondence**
V. Tai, Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC, Canada, V6T 1Z4
Telephone number: +604-822-4906;
FAX number: +604-822-6089;
e-mail: vera.tai@botany.ubc.ca

Received: 28 February 2013; accepted March 5, 2013.

doi:10.1111/jeu.12048

## ABSTRACT

Advances in high-throughput nucleic acid sequencing have improved our understanding of microbial communities in a number of ways. Deeper sequence coverage provides the means to assess diversity at the resolution necessary to recover ecological and biogeographic patterns, and at the same time single-cell genomics provides detailed information about the interactions between members of a microbial community. Given the vastness and complexity of microbial ecosystems, such analyses remain challenging for most environments, so greater insight can also be drawn from analysing less dynamic ecosystems. Here, we outline the advantages of one such environment, the wood-digesting hindgut communities of termites and cockroaches, and how it is a model to examine and compare both protist and bacterial communities. Beyond the analysis of diversity, our understanding of protist community ecology will depend on using statistically sound sampling regimes at biologically relevant scales, transitioning from discovery-based to experimental ecology, incorporating single-cell microbiology and other data sources, and continued development of analytical tools.

MICROBIAL ecosystems are poorly understood in large part because the diversity of microorganisms has been severely underestimated and inadequately described. A strong understanding of diversity is essential to address fundamental questions in community ecology, such as biogeography, community assembly, community stability, the links between taxonomic, genetic, and functional diversity, and the relationships between diversity and ecosystem function. Without accurate means to identify microbial diversity, elucidating the roles, functions, and interactions of microorganisms in natural environments is hampered and for these reasons microbial ecosystems have often been described as a "black box".

Descriptions of eukaryotic microbial diversity have classically depended on identifications using morphological, behavioural, or physiological criteria. These criteria, however, are inadequate to fully describe the diversity of microbes, in particular bacteria and archaea, where diversity often is manifested at the level of metabolism, but also in microbial eukaryotes (protists), where there is significant structural diversity but morphology alone is increasingly seen as insufficient for differentiating closely related species. Cryptic species abound in nature (see e.g. Gentekaki and Lynn 2010; Katz et al. 2012; Kosakyan et al. 2012; Lowe et al. 2010; Pfandl et al. 2009) and genetic diversity often dramatically outweighs the observed morphological diversity indicating the vast diversity of protists that have escaped discovery. Furthermore, obtaining morphological, behavioural, or physiological data is not possible for all microorganisms. Cultivation is typically required, and this is extremely challenging or near impossible for the vast majority of microorganisms. Microscopic examination is also very labour-intensive. Particularly from natural environments, obtaining morphological and other classical data for identifying protists is neither sufficient nor efficient for describing their diversity.

## PROTIST DIVERSITY THROUGH THE EYES OF HIGH-THROUGHPUT SEQUENCING

High-throughput sequencing provides a means to overcome the limitations of classical methods for describing microbial diversity. First, genetic data from molecular markers, such as the small subunit ribosomal RNA gene, more accurately describe diversity than morphological characters, so closely related species are more often differentiated and microbes that are not easily visualized are surveyed equally. Second, the level of sampling is higher and less biased. Thousands of sequences from a community of microorganisms can be collected from an environmental sample so a more complete survey of the diversity is obtained, which is particularly important in the identification

of the rarer species. Third, the process is not as labour-intensive; therefore the community of organisms from many samples can be handled efficiently. Microbial communities can also be sampled sufficiently to obtain meaningful and statistically significant comparisons. Finally, because genetic data are acquired, these data can be further analysed using phylogenetic and evolutionary models.

High-throughput sequencing unearths basic information about the diversity and composition of microbial communities that is essential for a deeper understanding of the interactions and processes that drive microbial ecosystems. It has been used extensively to study prokaryotic diversity, but only recently has high-throughput sequencing been applied to uncover eukaryotic diversity from natural environments in a systematic way. From primarily marine environments, using variable regions of the small subunit ribosomal RNA (SSU rRNA) gene as taxonomic markers, thousands of taxa are found; many more than could be documented by classical methods (Cheung et al. 2010; Edgcomb et al. 2011; Logares et al. 2012; Monchy et al. 2012; Orsi et al. 2012; Pawlowski et al. 2011; Stoeck et al. 2009, 2010). Existing classification schemes provided a broad taxonomic identity for the sequences that have been characterized, but at a finer scale many of the taxa were previously unknown. Sequencing has enabled the discovery of new lineages of prokaryotes and protists, but in the case of protists, it is important to remember that many taxa classically defined by morphological criteria have not been sampled at the molecular level; so many "novel molecular lineages" may turn out to correspond to known protist groups.

The first glimpses into the molecular diversity of marine protists using high-throughput sequencing demonstrated the extent to which diversity has been underestimated and also the impossibility of documenting this diversity using classical means. From these few studies, new ecological insights also emerged. For example, the composition of the protist community varied depending on the availability of oxygen in some marine environments (Logares et al. 2012; Orsi et al. 2012), while deep-sea sediment communities from arctic and antarctic polar regions were shown to be surprisingly similar (Pawlowski et al. 2011). Clearly, further sampling is needed to clarify these patterns, and doubtless this is only the beginning.

## THE GENETIC DIVERSITY YARD-STICK

As high-throughput sequencing has become more widely applied, new methodologies have also been developed to analyse the sheer quantity of data and provide diversity estimates (Caporaso et al. 2010; Schloss et al. 2009). At a more fundamental level, however, new methods also challenge our notions of microbial diversity: how is diversity described from purely molecular sequences?

Regardless of the methodology used, we must remain conscientious of the level of diversity that is measured. Different measures of diversity, whether based on morphology or genetic sequences, are not necessarily equivalent. Typically, DNA sequences are clustered at a defined level of similarity and a sequence chosen from this cluster, the operational taxonomic unit (OTU), represents that "taxon". Often, a similarity of 97% for SSU rRNA gene sequences is used to demarcate "species", but this is an assumption that should never be forgotten. In protists, multiple species concepts exist and there is no single concept that is adequate or universal. No measure of molecular diversity can be realistically and uniformly applied across such a large spectrum of biological diversity. Moreover, even for well-studied protists such as diatoms (Mann 2010) or ciliates (Hall and Katz 2011), where there has been some attempt to characterize a biological species concept (a species defined by reproductive isolation), it remains unclear which, if any, diversity measure corresponds to a biological species. For most protists, species boundaries are untested, their reproductive strategies are poorly known, and flexibility should be incorporated in our descriptions of their diversity.

The classification of marker sequences is useful to get an overview of the taxonomic composition of the community and to identify sequences that are similar to known taxa, but such classification is not always possible. Many environmental sequences are too different from sequences in existing databases to make clear identification possible. Moreover, the 18S rRNA gene is also a relatively conservative taxonomic marker so short sequences (typically ~250 bp) as obtained from current high-throughput sequencing technologies do not provide enough information to accurately classify the sequence beyond the genus level. In addition, classifications for many protists are uncertain because the organisms cannot be definitively placed in the tree of life, the classifications do not reflect phylogeny and are in flux, and the boundaries for a taxonomic level are difficult to demarcate, particularly when defining species. Because of these difficulties, the exact classification of an OTU may be misleading, but as we discover the range and extent of protist diversity, these molecular data can and should be used to clarify the classification and identification of protists.

These difficulties notwithstanding, the sequencing of molecular markers provides a measure of diversity that is vastly more ecologically and evolutionarily informative than classical descriptions. The finer scale of diversity they provide can reveal spatial and temporal patterns that were not previously possible using morphological descriptions. Even more detailed patterns will emerge from deep sampling of more divergent markers, for example, the internal transcribed spacer of the rRNA operon. These types of data are needed to gain a more sophisticated understanding of the relationship between diversity and ecosystem function, community assembly, and the response of the community to environmental changes.

## SAMPLING AND SCALING IN PROTIST ECOLOGY

Compared to multicellular organisms, fundamental information regarding protistan community ecology is lacking. Biogeographic patterns, for example, are completely

unknown for the vast majority of protists but are important for understanding the factors that may regulate their local or global distributions. The detection of biogeographical patterns depends on measuring diversity at an appropriate level and sampling the environment at a scale that matches the ecological effects driving the pattern, such as climate or nutrient/food availability. For multicellular organisms, the scaling of biogeographical patterns with diversity is a well-known phenomenon (Holt et al. 2013; McGill 2010). At a broad scale, a taxon could be ubiquitous (e.g. mammals), but at a finer scale of diversity, taxa comprising a broader taxon may show more restricted distributions (e.g. polar bears). For protists, the ease by which microbes can disperse or their large population sizes have often been cited as reasons for the observed ubiquity of some protist taxa, but there is generally little supporting evidence and the effects of cryptic species on these conclusions have not been thoroughly tested (Azovsky and Mazei 2012; Bass et al. 2007; Fenchel and Finlay 2004; Foissner 2007). More likely, greater sampling resolution with finer levels diversity will reveal more detailed patterns. High-throughput sequencing can facilitate the greater sampling effort that is needed to detect patterns of ubiquity and endemism of protist taxa, and reveal the diversity levels and the scale of ecological effects that dictate these biogeographical patterns.

Biogeographic patterns are also difficult to detect because the composition of natural communities is dynamic. Community composition can shift due to environmental changes, the influx/efflux of dispersing organisms, or historical effects such as the diversification and speciation of resident taxa. Marine environments, for example, are subject to numerous environmental factors that affect the composition and activities of the microbial community. A coastal marine environment is modified by currents, tides, terrestrial run-off, temperature, rainfall, nutrient influxes, and many others, which change the physical and chemical environment to varying degrees and influence the physiology, growth rates, interactions, and dispersal of microbes in the community. Randomized, replicated sampling regimes are relied upon to detect correlations between environmental variables and community composition. However, these environmental variables are often difficult to measure and in most cases their roles in regulating community composition or ecological functions are not clear.

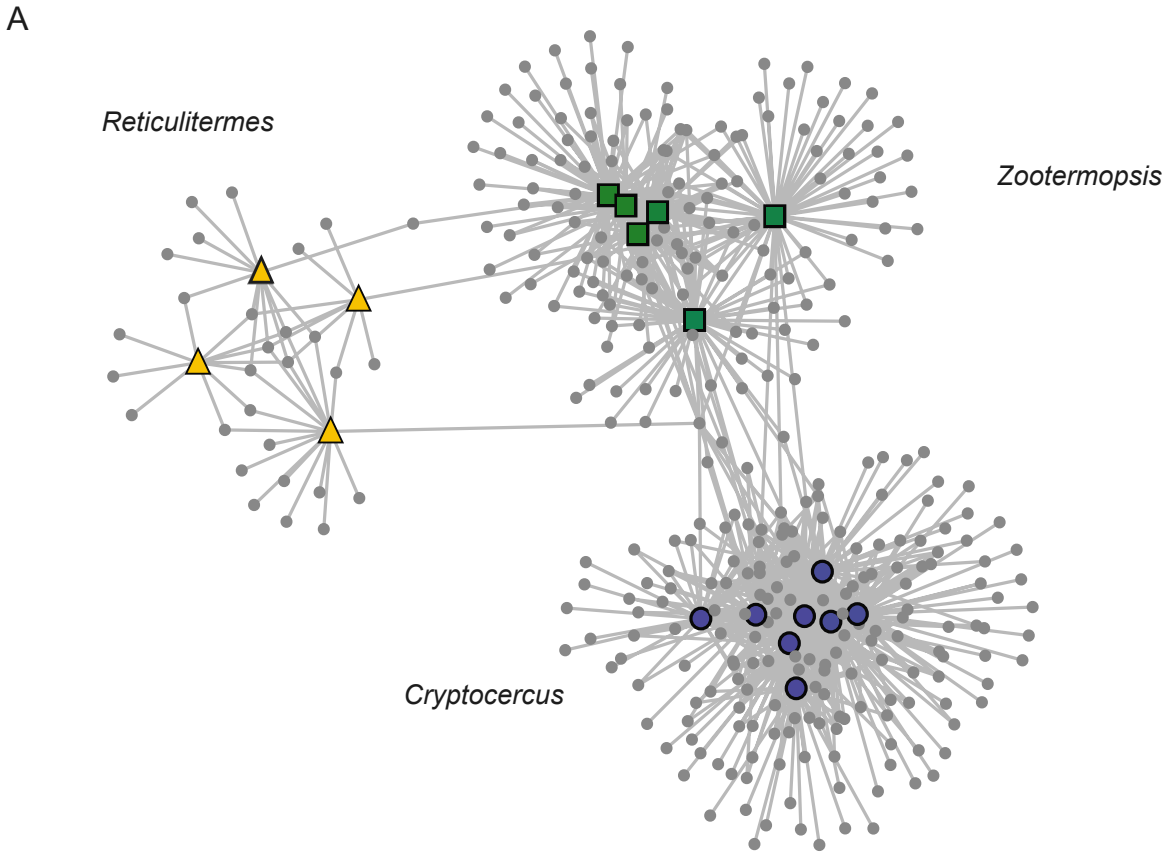
## PROTIST ECOLOGY OF TERMITE AND COCKROACH HINDGUTS

Marine environments are extremely important ecosystems to study because of their role in global biogeochemical cycles, climate, and productivity. But because the environment is dynamic, they are also extremely challenging to study and the results even more challenging to convert to general principles. Simpler and less dynamic ecosystems may not have the global impact of the oceans, but they nonetheless offer an important but overlooked contribution to our understanding of microbial ecology. Fundamental insights into the ecology and evolution of microorganisms and microbial communities will be more readily studied and interpreted in more controlled but still natural systems. In such systems, factors driving patterns of community composition can be determined with less sampling effort and fewer variables to account for.

The hindguts of lower termites and wood-eating *Cryptocercus* cockroaches are such environments. These insects' hindguts harbour microbial communities that are dominated by protists, but also include bacteria and archaea. The microbes are substantially isolated from external environmental fluctuations and community composition has largely been influenced by historical rather than ecological effects. The symbionts are vertically transmitted across generations (and among members of the colony) by the feeding of hindgut fluid (via proctodeal trophallaxis) to their young or newly moulted individuals (Nalepa et al. 2001), so dispersal of symbionts between host species is rare. This has resulted in the co-evolution of the insect hosts with their hindgut symbionts so that related hosts tend to share related symbionts (Kitade 2004; Noda et al. 2007). In closely related hosts (e.g. members of the same insect species), this association is very strong and stable over space and time: each host species of termite/cockroach generally harbours unique species of symbionts, so the microbial communities are highly endemic.

The protists inhabiting lower termite and cockroach hindguts are intriguing in their own right. Initially described as "parasites" (Leidy 1881), most of the protists are now considered beneficial symbionts that carry out the breakdown of lignocellulose to produce acetate, which is consumed by their hosts (Brune and Ohkuma 2011). These symbiont communities are often studied for biotechnological applications to convert lignocellulose to fuel (see e.g. Cairo et al. 2011; Scharf et al. 2011; Tartar et al. 2009; Todaka et al. 2010; Zhang et al. 2012). In terms of biodiversity, hindgut protist communities are predominantly composed of parabasalid and oxymonad taxa, many of which have been studied primarily for the discovery and characterization of novel and evolutionarily significant lineages (see e.g. Carpenter et al. 2011; Cepicka et al. 2010; Dacks and Redfield 1998; Heiss and Keeling 2006; Leander and Keeling 2004; Noda et al. 2009; Ohkuma et al. 2008; Saldarriaga et al. 2011).

The ecology of this environment is less well studied, but the hindgut communities of lower termites and *Cryptocercus* cockroaches offer a number of advantages over other environments to examine microbial ecology, especially for protists. As protist-dominated communities, the ecology and evolution of eukaryotic microbes can be addressed more readily. The historical imprint on community composition allows investigations into the diversification and adaptation of the symbionts within a host lineage. Ecological interactions between microbes and changes in community composition are also more tractable as the magnitude of diversity in the community is moderate and relatively consistent. Bacteria often occur as ecto- and endosymbionts of the protists and the nature



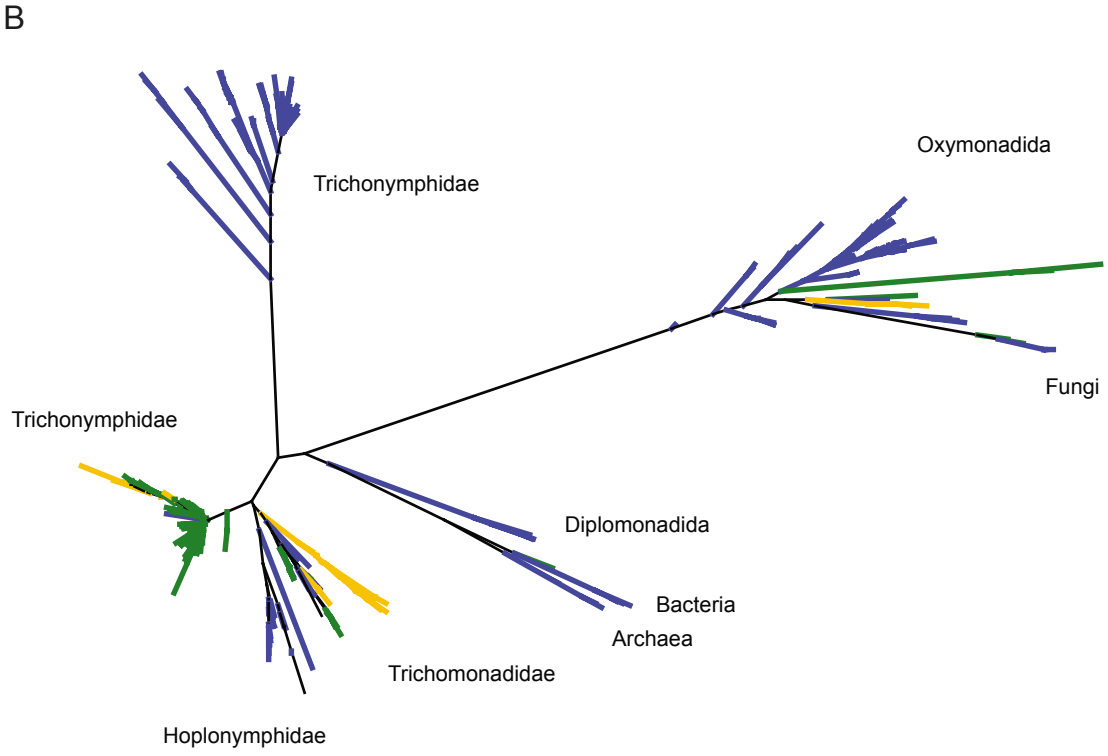



**Figure 1** Visualizations of protist diversity from the hindguts of *Cryptocercus* cockroaches, and *Reticulitermes* and *Zootermopsis* termites. These figures were generated from high-throughput sequencing data of the V4 region of the SSU rRNA gene (unpublished data) to give examples of what such data might look like when analysed and visualized differently. (**A**) Network analysis (force-directed layout) showing protist symbiont OTUs (grey dots) connected to their hosts (coloured circles, squares, or triangles) by a line. The relatively few symbiont OTUs shared between two hosts (indicated by grey dots with lines to more than one host), indicate a strong host-endemicity of symbionts. At the same time, the large number of symbiont OTUs that are shared between multiple individuals of a given host taxon indicates that there is a greater degree of symbiont variability, probably of closely related cryptic variants, within a host than previously thought, (**B**) Phylogenetic tree of all sequences obtained from the hindgut samples. Terminal and internal branches of the tree were coloured if the branch lead to descendant OTUs from a single host genus. The colouration reveals when in evolutionary time host-specific symbiont lineages evolved and diversified. Blue branches are *Cryptocercus*-specific lineages, yellow are *Reticulitermes*-specific, and green are *Zootermopsis*-specific.

of these protistan–bacterial interactions can be studied in a very directed, reproducible, and intensive way. Across hindguts from different host species, the composition of the communities may be different, but the general function of the community – to digest wood and provide fuel for their host – is common, so the relationship between diversity and ecosystem functioning can be investigated directly. In addition, some taxa may not have strictly co-evolved with their hosts, so the effect of dispersal on community composition and the ecology of the hindguts can also be examined.

As with most microbial communities, however, the microbial diversity in termite hindguts has been underestimated. Because the focus has been on understanding the evolution of particular taxa, the diversity of taxa within most hindguts are not fully described and biased towards the larger protists. In particular, the diversity of bacteria and archaea in most hosts is totally unknown. The protistan diversity is largely known through morphological descriptions, which are marked by the historical tendency to assume similar looking protists found in different hosts are most likely the same species. The use of molecular markers has shown this to be false in several examples and has generally expanded the known diversity (Gile et al. 2011; Harper et al. 2009; James et al. 2013; Ohkuma et al. 2005; Stingl and Brune 2003; Strassert et al. 2009). Remarkably, this trend is even true for large, easily identifiable symbionts from well-studied termites. For example, using even low-throughput molecular sequencing and DNA barcoding in the termite host whose hindgut has been the subject of arguably the most intensive study of any such environment, we recently showed that the species-level diversity of the largest and most easily identifiable symbionts has been underestimated by nearly twofold (Tai et al. 2013).

By using methods that increase sampling by many orders of magnitude, the diversity of the whole communities can be established, including smaller protists and bacteria that are more difficult to observe and identify. These data also inspire new ways to analyse and visualize the data from which different kinds of insights can emerge. For example, a network analysis demonstrates the richness of genetic diversity from the hindgut communities of *Cryptocercus* cockroaches, and *Reticulitermes* and *Zootermopsis* termites (Fig. 1A). The first observation from these networks is that diversity is much larger than observed morphologically, and larger even than has been observed using lower throughput molecular methods (e.g. Tai et al. 2013). Most of this diversity cannot be attributed to described species; some certainly represents population level diversity (or "strain level variation"), but there are also genetically distinct OTUs with species-level variation. The second observation, one that would have been more difficult to see from a phylogenetic representation, is that the protist taxa tend to be highly endemic to a host genus with few taxa occurring in more than one host. High-throughput sequencing data can also be analysed from an evolutionary point-of-view. For the symbionts from the hindguts of *Cryptocercus*, *Reticulitermes*, and *Zootermopsis*, many taxa originated from relatively recent radiations (Fig. 1B), and different patterns can be seen for *Cryptocercus* and the termites. These data can also be explored using multivariate statistics to identify specific trends to be tested more explicitly. From these initial findings, these new vantage points guide further investigations of the diversification, speciation, and co-evolutionary processes in protists.

## FROM COMMUNITIES TO SINGLE CELLS

Transitioning from the study of protist diversity patterns across communities to establishing ecological function will also be helped by high-throughput sequencing. On the one hand, supplementing microbial community assembly data with functional information such as meta-transcriptomics can establish links between diversity and ecosystem function. But probably more importantly, high-throughput sequencing can also be used to examine the biology of single cells. One of the primary limitations in studying natural microbial communities is the necessity of cultivation or collection of numerous cells to amass enough material for study. Cultivation is rarely readily possible and can induce major changes to the organism's physiology. Taking a meta-omic approach has been useful, but the entire community of microbes is analysed collectively. Members of the community cannot easily be sorted based on identity or function so ecosystem functions, such as primary productivity or nitrogen uptake rates, are measured with little knowledge of the contributing microorganisms. From a single cell, however, the analysis and interpretation steps are significantly simplified: the biology of the individual is recovered, individual variation can be assessed, and,

most importantly, a direct association can be made between identity and function.

Ecologically, single-cell sequencing can be used not only to explore the biology of that cell but also its interactions with other microorganisms. The ecological interactions between microorganisms are generally unknown in natural environments. For protists, single-cell sequencing has been successfully applied to discover interacting partners. These pioneering investigations have identified the bacterial prey, bacterial symbionts, and viruses interacting with specific protists (Martinez-Garcia et al. 2011; Thompson et al. 2012; Yoon et al. 2011).

In hindgut communities, many of the larger protists harbour bacterial ecto- and/or endosymbionts. High-throughput sequencing of symbiont bearing protists has enabled genome assemblies of the bacterial symbionts and provided insights into the biochemical nature of the interaction between protists and their symbionts. These interactions appear to be driven by a need for nitrogenous compounds that the bacteria can produce for their host (Hongoh et al. 2008a,b).

The analysis of single cells is a tremendous advance for microbial ecology. The biology of the individual can be resolved. Interactions between individual microorganisms can be determined. From single cells in a community, networks can be built of the interactions and functions that comprise the ecosystem.

## CONCLUSIONS

Sequencing technologies and other molecular tools are rapidly changing our knowledge of protist diversity, especially in natural environments. In the absence of cultivation, DNA sequencing has provided the means to more easily and fully assess environmental microbial diversity. Using high-throughput sequencing, the observed diversity is of a magnitude that was unfathomable just a decade ago.

The next challenge is to embrace this diversity and discover the ecological and evolutionary processes that explain it. This will entail the use of statistically sound sampling regimes at biologically relevant scales, transitioning from discovery-based to experimental ecology, incorporating single-cell microbiology, and continued development of analytical tools. Many natural communities, such as termite/cockroach hindguts, have features that simplify community analyses compared with more dynamic environments and can provide greater insight into the complexity of microbial ecosystems.

Although the most rapid advances in protist ecology are being made by applying high-throughput sequencing technologies, other data sources should be integrated to find the relationships between diversity and ecosystem functioning and the response of the microbial community to environmental changes. Ultimately, although we are in the sequencing age, this is but one tool to help us understand and manage the ecology of protists and their role as conduits between the micro- and "macro"-bial worlds.

## ACKNOWLEDGMENTS

Support for this manuscript was provided by a grant (227301) from the Natural Sciences and Engineering Research Council of Canada (NSERC). VT was funded as a Global Scholar of the Canadian Institute for Advanced Research (CIFAR). PJK is a CIFAR fellow.

## LITERATURE CITED

Azovsky, A. & Mazei, Y. 2012. Do microbes have macroecology? Large-scale patterns in the diversity and distribution of marine benthic ciliates. *Glob. Ecol. Biogeogr.*, 22:163–172.

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.

Brune, A. & Ohkuma, M. 2011. Role of the termite gut microbiota in symbiotic digestion. *In*: Bignell, D., Roisin, Y. & Lo, N. (ed.), Biology of Termites: A Modern Synthesis. Springer, New York. p. 439–475.

Cairo, J. P. L. F., Leonardo, F. C., Alvarez, T. M., Ribeiro, D. A., Büchli, F., Costa-Leonardo, A. M., Carazzolle, M. F., Costa, F. F., Leme, A. F. P., Pereira, G. A. & Squina, F. M. 2011. Functional characterization and target discovery of glycoside hydrolases from the digestome of the lower termite *Coptotermes gestroi*. *Biotechnol. Biofuels*, 4:50.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Meth.*, 7:335–336.

Carpenter, K. J., Horak, A., Chow, L. & Keeling, P. J. 2011. Symbiosis, morphology, and phylogeny of Hoplonymphidae (Parabasalia) of the wood-feeding roach *Cryptocercus punctulatus*. *J. Eukaryot. Microbiol.*, 58:426–436.

Cepicka, I., Hampl, V. & Kulda, J. 2010. Critical taxonomic revision of parabasalids with description of one new genus and three new species. *Protist*, 161:400–433.

Cheung, M. K., Au, C. H., Chu, K. H., Kwan, H. S. & Wong, C. K. 2010. Composition and genetic diversity of picoeukaryotes in subtropical coastal waters as revealed by 454 pyrosequencing. *ISME J.*, 4:1053–1059.

Dacks, J. B. & Redfield, R. J. 1998. Phylogenetic placement of *Trichonympha*. *J. Eukaryot. Microbiol.*, 45:445–447.

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.

Fenchel, T. & Finlay, B. 2004. The ubiquity of small species: patterns of local and global diversity. *Bioscience*, 54:777–784.

Foissner, W. 2007. Protist diversity and distribution: some basic considerations. *Biodivers. Conserv.*, 17:235–242.

Gentekaki, E. & Lynn, D. H. 2010. Evidence for cryptic speciation in *Carchesium polypinum* Linnaeus, 1758 (Ciliophora: Peritrichia) inferred from mitochondrial, nuclear, and, morphological markers. *J. Eukaryot. Microbiol.*, 57:508–519.

Gile, G. H., James, E. R., Scheffrahn, R. H., Carpenter, K. J., Harper, J. T. & Keeling, P. J. 2011. Molecular and morphological analysis of the family Calonymphidae with a description of

*Calonympha chia* sp. nov., *Snyderella kirbyi* sp. nov., *Snyderella swezyae* sp. nov. and *Snyderella yamini* sp. nov. *Int. J. Syst. Evol. Micr.*, 61:2547–2558.

Hall, M. S. & Katz, L. A. 2011. On the nature of species: insights from *Paramecium* and other ciliates. *Genetica*, 139:677–684.

Harper, J. T., Gile, G. H., James, E. R., Carpenter, K. J. & Keeling, P. J. 2009. The inadequacy of morphology for species and genus delineation in microbial eukaryotes: an example from the parabasalian termite symbiont *Coronympha*. *PLoS ONE*, 4:e6577.

Heiss, A. A. & Keeling, P. J. 2006. The phylogenetic position of the oxymonad *Saccinobaculus* based on SSU rRNA. *Protist*, 157:335–344.

Holt, B. G., Lessard, J. P., Borregaard, M. K., Fritz, S. A., Araujo, M. B., Dimitrov, D., Fabre, P. H., Graham, C. H., Graves, G. R., Jonsson, K. A., Nogues-Bravo, D., Wang, Z., Whittaker, R. J., Fjeldsa, J. & Rahbek, C. 2013. An update of Wallace's zoogeographic regions of the world. *Science*, 339:74–78.

Hongoh, Y., Sharma, V. K., Prakash, T., Noda, S., Taylor, T. D., Kudo, T., Sakaki, Y., Toyoda, A., Hattori, M. & Ohkuma, M. 2008a. Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell. *Proc. Natl Acad. Sci. USA*, 105:5555–5560.

Hongoh, Y., Sharma, V. K., Prakash, T., Noda, S., Toh, H., Taylor, T. D., Kudo, T., Sakaki, Y., Toyoda, A., Hattori, M. & Ohkuma, M. 2008b. Genome of an endosymbiont coupling $N_2$ fixation to cellulolysis within protist cells in termite gut. *Science*, 322:1108–1109.

James, E. R., Okamoto, N., Burki, F., Scheffrahn, R. H. & Keeling, P. J. 2013. *Cthulhu macrofasciculumque* n. g., n. sp. and *Cthylla microfasciculumque* n. g., n. sp., a newly identified lineage of parabasalian termite symbionts. *PLoS ONE*, 8:e58509.

Katz, L. A., DeBerardinis, J., Hall, M. S., Kovner, A. M., Dunthorn, M. & Muse, S. V. 2012. Heterogeneous rates of molecular evolution among cryptic species of the ciliate morphospecies *Chilodonella uncinata*. *J. Mol. Evol.*, 73:266–272.

Kitade, O. 2004. Comparison of symbiotic flagellate faunae between termites and a wood-feeding cockroach of the genus *Cryptocercus*. *Microbes Environ.*, 19:215–220.

Kosakyan, A., Heger, T. J., Leander, B. S., Todorov, M., Mitchell, E. A. D. & Lara, E. 2012. COI barcoding of nebelid testate amoebae (Amoebozoa: Arcellinida): extensive cryptic diversity and redefinition of the Hyalospheniidae Schultze. *Protist*, 163:415–434.

Leander, B. S. & Keeling, P. J. 2004. Symbiotic innovation in the oxymonad *Streblomastix strix*. *J. Eukaryot. Microbiol.*, 51:291–300.

Leidy, J. 1881. Parasites of the termites. *J. Acad. Nat. Sci. Phila.*, 8, Second Series:425–447.

Logares, R., Audic, S., Santini, S., Pernice, M. C., de Vargas, C. & Massana, R. 2012. Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. *ISME J.*, 6:1823–1833.

Lowe, C. D., Montagnes, D. J. S., Martin, L. E. & Watts, P. C. 2010. Patterns of genetic diversity in the marine heterotrophic flagellate *Oxyrrhis marina* (Alveolata Dinophyceae). *Protist*, 161:212–221.

Mann, D. G. 2010. Discovering diatom species: is a long history of disagreements about species-level taxonomy now at an end? *Pl. Ecol. Evol.*, 143:251–264.

Martinez-Garcia, M., Brazel, D., Poulton, N. J., Swan, B. K., Gomez, M. L., Masland, D., Sieracki, M. E. & Stepanauskas, R. 2011. Unveiling in situ interactions between marine protists and bacteria through single cell sequencing. *ISME J.*, 6:703–707.

McGill, B. J. 2010. Matters of scale. *Science*, 328:575–576.

Monchy, S., Grattepanche, J.-D., Breton, E., Meloni, D., Sanciu, G., Chabé, M., Delhaes, L., Viscogliosi, E., Sime-Ngando, T. & Christaki, U. 2012. Microplanktonic community structure in a coastal system relative to a *Phaeocystis* bloom inferred from morphological and tag pyrosequencing methods. *PLoS ONE*, 7: e39924.

Nalepa, C., Bignell, D. & Bandi, C. 2001. Detritivory, coprophagy, and the evolution of digestive mutualisms in Dictyoptera. *Insect. Soc.*, 48:194–201.

Noda, S., Kitade, O., Inoue, T., Kawai, M., Kanuka, M., Hiroshima, K., Hongoh, Y., Constantino, R., Uys, V., Zhong, J., Kudo, T. & Ohkuma, M. 2007. Cospeciation in the triplex symbiosis of termite gut protists (*Pseudotrichonympha* spp.), their hosts, and their bacterial endosymbionts. *Mol. Ecol.*, 16:1257–1266.

Noda, S., Mantini, C., Bordereau, C., Kitade, O., Dolan, M. F., Viscogliosi, E. & Ohkuma, M. 2009. Molecular phylogeny of parabasalids with emphasis on the order Cristamonadida and its complex morphological evolution. *Mol. Phylogenet. Evol.*, 52:217–224.

Ohkuma, M., Iida, T., Ohtoko, K., Yuzawa, H., Noda, S., Viscogliosi, E. & Kudo, T. 2005. Molecular phylogeny of parabasalids inferred from small subunit rRNA sequences, with emphasis on the Hypermastigea. *Mol. Phylogenet. Evol.*, 35:646–655.

Ohkuma, M., Noda, S., Hongoh, Y., Nalepa, C. A. & Inoue, T. 2008. Inheritance and diversification of symbiotic trichonymphid flagellates from a common ancestor of termites and the cockroach *Cryptocercus*. *P. Roy. Soc. B Bio.*, 276:239–245.

Orsi, W., Song, Y. C., Hallam, S. & Edgcomb, V. 2012. Effect of oxygen minimum zone formation on communities of marine protists. *ISME J.*, 6:1586–1601.

Pawlowski, J., Christen, R., Lecroq, B., Bachar, D., Shahbazkia, H. R., Amaral-Zettler, L. & Guillou, L. 2011. Eukaryotic richness in the abyss: insights from pyrotag sequencing. *PLoS ONE*, 6: e18169.

Pfandl, K., Chatzinotas, A., Dyal, P. & Boenigk, J. 2009. SSU rRNA gene variation resolves population heterogeneity and ecophysiological differentiation within a morphospecies (Stramenopiles, Chrysophyceae). *Limnol. Oceanogr.*, 54:171.

Saldarriaga, J. F., Gile, G. H., James, E. R., Horak, A., Scheffrahn, R. H. & Keeling, P. J. 2011. Morphology and molecular phylogeny of *Pseudotrichonympha hertwigi* and *Pseudotrichonympha paulistana* (Trichonymphea, Parabasalia) from neotropical rhinotermitids. *J. Eukaryot. Microbiol.*, 58:487–496.

Scharf, M. E., Karl, Z. J., Sethi, A. & Boucias, D. G. 2011. Multiple levels of synergistic collaboration in termite lignocellulose digestion. *PLoS ONE*, 6:e21709.

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.

Stingl, U. & Brune, A. 2003. Phylogenetic diversity and whole-cell hybridization of oxymonad flagellates from the hindgut of the wood-feeding lower termite *Reticulitermes flavipes*. *Protist*, 154:147–155.

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.

Stoeck, T., Behnke, A., Christen, R., Amaral-Zettler, L., Rodriguez-Mora, M. J., Chistoserdov, A., Orsi, W. & Edgcomb, V.

P. 2009. Massively parallel tag sequencing reveals the complexity of anaerobic marine protistan communities. *BMC Biol.*, 7:72.

Strassert, J. F. H., Desai, M. S., Brune, A. & Radek, R. 2009. The true diversity of devescovinid flagellates in the termite *Incisitermes marginipennis*. *Protist*, 160:522–535.

Tai, V., James, E. R., Perlman, S. J. & Keeling, P. J. 2013. Single-cell DNA barcoding using sequences from the small subunit rRNA and internal transcribed spacer region identifies new species of *Trichonympha* and *Trichomitopsis* from the hindgut of the termite *Zootermopsis angusticollis*. *PLoS ONE*, 8:e58728.

Tartar, A., Wheeler, M. M., Zhou, X., Coy, M. R., Boucias, D. G. & Scharf, M. E. 2009. Parallel metatranscriptome analyses of host and symbiont gene expression in the gut of the termite *Reticulitermes flavipes*. *Biotechnol. Biofuels*, 2:25.

Thompson, A. W., Foster, R. A., Krupke, A., Carter, B. J., Musat, N., Vaulot, D., Kuypers, M. M. M. & Zehr, J. P. 2012. Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science*, 337:1546–1550.

Todaka, N., Inoue, T., Saita, K., Ohkuma, M., Nalepa, C. A., Lenz, M., Kudo, T. & Moriya, S. 2010. Phylogenetic analysis of cellulolytic enzyme genes from representative lineages of termites and a related cockroach. *PLoS ONE*, 5:e8636.

Yoon, H. S., Price, D. C., Stepanauskas, R., Rajah, V. D., Sieracki, M. E., Wilson, W. H., Yang, E. C., Duffy, S. & Bhattacharya, D. 2011. Single-cell genomics reveals organismal interactions in uncultivated marine protists. *Science*, 332:714–717.

Zhang, D., Lax, A. R., Henrissat, B., Coutinho, P., Katiya, N., Nierman, W. C. & Fedorova, N. 2012. Carbohydrate-active enzymes revealed in *Coptotermes formosanus* (Isoptera: Rhinotermitidae) transcriptome. *Insect Mol. Biol.*, 21:235–245.