



## Concepts in protistology: Species definitions and boundaries

Jens Boenigk<sup>a,\*</sup>, Marc Ereshefsky<sup>b</sup>, Kerstin Hoef-Emden<sup>c</sup>, James Mallet<sup>d</sup>, David Bass<sup>e</sup>

<sup>a</sup>General Botany, University Duisburg-Essen, Universitätsstr. 5, 45117 Essen, Germany

<sup>b</sup>Department of Philosophy, University of Calgary, 618 Campus Place N.W., Calgary, Canada

<sup>c</sup>Botanical Institute, Cologne Biocenter, University of Cologne, Zùlpicher Str. 47b, 50674 Cologne, Germany

<sup>d</sup>Galton Laboratory, Department of Genetics, Evolution and Environment (GEE), University College London, Gower Street, London WC1E 6BT, United Kingdom

<sup>e</sup>Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom

Received 24 August 2011; received in revised form 28 November 2011; accepted 28 November 2011

### Abstract

This paper summarises the Symposium ‘Concepts in Protistology’, during the VI European Congress of Protistology, Berlin, 25–29 July 2011. There is an increasing focus on cataloguing the number of species on earth, species barcoding initiatives, and the increasing need to reconcile molecular with morphological data in protists within a taxonomic framework. We identify several obstructions to defining species in protists, including the high incidence of asexuality, high levels of both morphological conservation and evolutionary convergence, high levels of genetic diversity that cannot so far be correlated with phenotypic characters, conflicting signals between both genetic and phenotypic taxonomic markers, and different requirements and challenges of species definition in different protist groups. We assert that there is no species ‘category’ for protists, and recommend that a working definition of species is clarified on a case-by-case basis. Thus, a consensus approach may emerge within protist groups, but any one approach is unlikely to encompass a wide phylogenetic range. However, as long as clarity of intent and method is maintained, the utility of the term ‘species’ in protists will also be maintained as a reproducible and convenient (if artificial) way of referring to particular lineages within a tightly defined context.

© 2011 Published by Elsevier GmbH.

**Keywords:** Species concept; Protist; SSU rDNA; ITS rDNA; COI; OTU

### Introduction

Although the high-level taxonomy of protists is a well-grounded (although not resolved) area of research, the opposite end of the taxonomic spectrum, alpha-taxonomy, and what constitutes protist species are issues fraught with uncertainty and disagreement. The fact that there is no generally accepted basis for delimiting species in protists has many unfortunate consequences, prime among them being (1) a

lack of basic communicability about fundamental biological units (with obvious negative implications for barcoding), (2) lack of clarity regarding their evolutionary and ecological significance, and (3) a drastic underestimation of protist diversity and importance in more general biodiversity papers. A good example of the latter is shown by Mora et al. (2011), in which (by no fault of theirs) estimates of species numbers of protists (in particular) are unrealistically depressed, in part because of the problem with defining ‘species’ and also because of the rapidly changing and relatively unstructured nature of protist taxonomy overall.

The ECOP workshop did not set out to ‘decide’ on the ‘best’ species concept to apply to protists, but rather to assess

\*Corresponding author.

E-mail address: [jens.boenigk@uni-due.de](mailto:jens.boenigk@uni-due.de) (J. Boenigk).

to what extent such a concept is possible, or even desirable, and what the best working bases for protist alpha-taxonomy should be. The main body of this paper summarises the issues and outcomes of the interactive process culminating in the oral presentations. The individual contributions of the authors are reproduced in the three supplementary files.

### The historical context: shifts in the conception of megasystematics and species

The era of protistology has seen major methodological advances starting with the invention of simple microscopes, which allowed the first visualization of individual microbes by Leeuwenhoek in 1674 (cf. Dobell 1932). Early observers lumped the small microorganisms into broad categories, for instance as “*Vermis punctiformis*” (cf. the genus *Monas* Müller 1773), as “Punktthierchen”, “Kugelthierchen” and “Ovalthierchen” (von Gleichen and Freiherr 1778; compare also *Monas punctum* Ehrenberg 1838), or in the genus *Chaos* (Linnaeus 1758). The general system achieved by the end of the 19th century (cf. Bütschli 1880–89; Doflein 1916; Kent 1880–81; Pascher and Lemmermann 1914) remained in place for most of the 20th century. The general quality of the taxon diagnoses of protists improved little during this time and was largely based on light microscopy, even though the invention of electron microscopy in the 1930s (cf. Agar 1996) further increased taxonomic resolution.

The advent of molecular methods in the 1990s provided a very different perspective on microbial eukaryotic diversity. Many new genetic lineages were detected, a substantial number of which were highly distinct from those previously known, suggesting novel elements of biodiversity at high taxonomic ranks. At the other end of the scale, very high levels of genetic diversity were found within and around already known lineages, suggesting an abundance of cryptic species and sister taxa. Another revelation was that protist morphology is highly and often surprisingly convergent – for example ‘amoebae’ are found over most of the tree of life, classical Heliozoa and Radiolaria are several fold polyphyletic, and many lineages that were thought to be fungal were shown to branch elsewhere on the tree when placed by molecular phylogenetics (fungal analogues such as oomycetes, plasmodiophorids, labyrinthulids, etc.). Thus another layer of poorly resolved complexity was added to the already muddled story emerging from earlier morphological studies.

A recent approach for overcoming the drawbacks of insufficient taxonomic coverage of the studied diversity is the use of operational taxonomic units (OTUs) (e.g. Green et al. 2004). Especially in sequence-based diversity studies, such OTUs are defined by using sequence similarity or distance thresholds (e.g., Schloss and Handelsman 2005). In many cases these OTUs were treated synonymous to species, and were used, for instance, for the estimation of species richness data. The taxonomic power of lumping together organisms in OTUs based on sequence similarities was never evaluated

for eukaryotic microorganisms. However, in the case of prokaryotes the underestimation of real species numbers by approaches defining OTUs based on 16S rRNA sequences is obvious (Stackebrandt and Ebers 2006).

### A consistent species concept for protists?

Biodiversity research and ecology rely on safe identifications of species and on reproducible species counts, but often this requirement is not met. Various problems result from partially inconsistent species concepts, insufficient taxonomic coverage (undersampling), and uncertainty about which characters should best be used as bases for deciding species boundaries. Recent methodological progress has highlighted severe inconsistencies between the conceptual and the practical historic approaches to species and biodiversity. The dispute is currently stirred up by inconsistencies between molecular phylogenies on the one hand and morphological species denominations and traditional classification concepts on the other.

The conceptual conflict embraces the differences between zoological, botanical, and microbiological concepts of species. Due to ambiguous, contradictory and/or inconsistent species descriptions, species numbers obtained from biodiversity surveys using the traditional morphospecies concept are not comparable to those obtained from environmental DNA. As different methodological approaches are – in part – linked to different concepts of species and of diversity the dispute on protist species, protist diversity, and protist systematics often fails to differentiate differences in the conceptual basis from methodological limitations and real variation. For instance, the existence of newly revealed cryptic species obviously results in increasing biodiversity estimates. By contrast, taxon-independent (OTU-based) diversity studies as often applied for microorganisms tend to fail to resolve species and thus tend to underestimate biodiversity. This basic problem has many consequences, and represents a serious obstacle to understanding key aspects of protist biology and ecology, for example the ‘everything is everywhere’ debate and the perception of protist biodiversity.

In this paper we review the difficulties and challenges of alpha-taxonomy and species delineation. We do not attempt to review each protist group in an attempt to decide how the basic taxonomic units in each group should be defined; that is up to the experts working on them. However, examples are given from some of the groups in which the authors do have expertise, and the inclusion of some very well-studied metazoa provides a phylogenetically distant but conceptually relevant and informative perspective on issues which are often thought to be particularly problematic in protists. A significant element of our discussion is not associated with any organismal group, but considers the ‘problem’ from a philosophical/logical standpoint, which we feel offers perhaps the strongest direction and encouragement for alpha-taxonomy. We see the key issues in protist alpha-taxonomy to be:

- (1) confusion arising from high levels of evolutionary convergence and morphological conservation in protists,
- (2) morphological plasticity in many protist groups can lead to unreliable morphological diagnoses,
- (3) different relative rates of phenotypic and genotypic divergence among protist groups, meaning that a single genetic marker will provide very inconsistent taxonomic signals across protists as a whole,
- (4) uncertainty about which characters are best used to distinguish species,
- (5) high levels of diversity revealed by molecular techniques that are not associated with known cells, and genotypic diversity being much higher than (easily measured) morphological diversity,
- (6) limited genetic information about many protist lineages.

## Materials and Methods

Boenigk and Bass invited three contributors to their session, whose work is concerned with different aspects of species delineation, definitions, and concepts. Overall, we did not attempt to be taxonomically comprehensive; rather to provide an intellectual framework for the future consideration of ‘the species problem’ in protists. The invited speakers were

- (1) James Mallet, Professor of Biological Diversity at University College London (Galton Laboratory, Department of Genetics, Evolution and Environment (GEE)). One of Mallet’s core research questions is how new species arise; his group focuses on the evolution and ecology of butterflies and moths (Lepidoptera), particularly near the species boundary.
- (2) Marc Ereshefsky, Professor of Philosophy at the University of Calgary, Canada, who has written many papers and books on species and taxonomy, and lists his research interests as: Species; Natural Kinds; Kinds and Individuals in Biology; Homology; Historicity.
- (3) Kerstin Hoef-Emden from the Botanical Institute, University of Cologne, Germany; an expert in the taxonomy and biology of cryptophytes, and recipient of the Tyge Christensen Prize from the International Phycological Society in recognition of the best algal paper published in *Phycologia* during 2007.

Each contributor provided a written document for this publication as well as an oral presentation at the Symposium. These (edited) documents are provided as supplementary materials (Suppl. texts 1–3), and are referred to in the Results and Discussion, which attempts to synthesise and summarise their content with that of the oral presentations and the group discussion that ended the Symposium.

## Results and Discussion

Many of the problems in defining species in protists have been attributed to differences between protists and multicellular organisms (multicells), in which species boundaries have been thought to be in many cases clearer and confirmable. Multicells are often morphologically distinctive, and it was based on discrete character sets between sets of populations that the concept of ‘species’ was generated. Protists are small, difficult to observe, and morphologically relatively homogeneous (certainly within groups, but also between many groups that we now know to be only distantly related), providing relatively few characters with which to separate them. In addition, they can be phenotypically very variable even within a clonal line. Combined with the difficulty of observing a single specimen over a period long enough to properly characterise it, this means that many light microscopy-based species descriptions are at best incomplete and frequently effectively useless or misleading. The advent of molecular methods and the use of marker genes to count and measure diversity has in some respects made the situation even more complicated. This showed that most protist morphospecies harbour high levels of genetic diversity that often cannot easily be associated with phenotypic differences, either because the latter are too small or difficult to detect, or because the direction and scale of one kind of distance has a complicated or contradictory relationship with the other (i.e. very high levels of morphological conservation and convergent evolution in protists).

However, molecular findings have also complicated species delineation in multicells. Mallet (Suppl. text 1) explains that analyses of sequence data have tended to increase the number of species in animal groups, by ‘re-recognising’ varieties and races that had been lumped together as polytypic species by the biological species concept (BSC). One of the intentions of the BSC was partly to minimise confusion caused by the multitude of geographical races/varieties within biological species by not recognising such varieties as separate species unless they were reproductively isolated in sympatry. Mallet explains that at least 10% of all (metazoan) species hybridise with at least one other species (and not just sister species), and that it is a fallacy that species delimitation in macro-organisms is more robust than in microbes. There is often a continuum of intermediate types, which means that a consistent species concept will be elusive. The genomic consequences of introgression and horizontal gene transfer mean that the use of single genetic markers (or barcodes) is fundamentally flawed. Different ‘good’ species can share the same barcode, and hybrid introgression can spread barcodes among species. There is a lack of correspondence between phenotypic characters, genome-level divergence, and individual marker genes. Therefore, it is necessary to integrate multiple lines of evidence to resolve species boundaries. The broader implications of these findings lead Mallet to argue that the idea of a ‘true species reality’ is unsound, partly because different scientists believe that

different processes are important as bases for delimiting species. Therefore, attempting to select from one of the c. 24 existing species concepts is futile and ignores the biological realities. Species delimitation should be based on the practicalities of each case; i.e. the meaning of 'species' is in their use, not as absolute entities.

These conclusions are strongly concordant with Ereshefsky's (Suppl. text 2) arguments. Both authors agree that there is no species category. Scientists have been trying in vain for hundreds of years to find the correct definition of 'species', but the simplest conclusion is that one does not exist. The important thing to notice about the variety of species concepts of which Ereshefsky gives a partial but illuminating overview (Suppl. text 2) is that they pick out different types of lineages, by defining their collection in different ways. What this should tell us is not that we need to look harder for the correct definition of 'species', but that there are many different types of fundamental lineages in the world: 'species' comprise a multitude of different types of taxa. Therefore, if there is no species category, then we should stop trying to find the correct definition of 'species'. However, this does not mean that species do not exist – they are real lineages in nature. The key requirement is that scientists are explicit about which species delineation approach is being taken. Ereshefsky cites Darwin: "I mean by species, those collections of individuals, which have commonly been so designated by naturalists". This applies to all life, but at the beginning of the 21st Century perhaps with particular and valid force to protists, which are in the dawn of their true elucidation at the alpha-taxonomic level, as so many multicellular taxa were when Darwin wrote those words. Furthermore, the advent of molecular data resets the taxonomic clock in many important respects, and compels us to reconsider how species may be most effectively and informatively delineated.

These considerations have already been adopted and embraced to an extent by many protistologists. The contribution of Hoef-Emden and Bass (Suppl. text 3) gives examples of how alpha-taxonomic decisions have been made in cryptophytes, and cercozoan and glissomonad Cercozoa, where the phenotypic characters discernable by light microscopy can be difficult to distinguish and/or highly plastic. A consensus approach is emerging that combines and reconciles phenotypic with genotypic data to delineate species, in which the importance and influence of sequence-based characters are particularly emphasised in cases where phenotypic separation is difficult. Hoef-Emden and Bass advocate high sampling within and between closely related lineages to maximise consistency and robustness of inference, and a good phylogenetic model on which to map phenotypic characters.

Unlike in many multicellular groups, where mitochondrial (mtDNA) genes are commonly used as barcodes/markers, in protists the genes and spacer regions on the ribosomal RNA (rRNA) gene cistron are currently most commonly used. There are several reasons for this: this region (principally 18S, the small subunit rRNA gene) has been routinely sequenced from culture collection and new strains, as it is relatively

easily PCR-amplified and sequenced, primer sites are well developed, and SSU rDNA can be used to generate relatively good single-gene phylogenies. However, the rDNA regions also have disadvantages, principally that they are multicopy with variable levels of intragenomic variation among groups, and the SSU is generally considered too conserved to distinguish among species. Therefore, in the better-known groups the faster evolving ITS rDNA is used for species delimitation, taking the cue from Annette Coleman's investigations into the utility of compensatory base changes in particular ITS2 regions as good markers of biological species boundaries. The mitochondrial COI gene is espoused by some as a good 'species' marker in some protist groups. However, its widespread use is currently not feasible as many groups lack information about it. For example, a rough assessment of GenBank entries for Cercozoa show ~3700 entries for SSU, 110 for LSU, 500 for the ITS regions, and 10–20 for COI. The high incidence of derived mitochondria and anaerobic groups in protists relative to other eukaryotes also preclude the universal use of mitochondrial markers as protist barcodes.

It is important to realise that any 'species level' marker currently under consideration can only ever be exactly that – a marker for taxonomic boundaries, rather than a directly linked indicator. None of these is associated with sexual compatibility or genetic isolation. However, we do not consider this to be a fatal problem as the benefits of using such highly sampled and widely accessible (i.e. relatively easily sequenced regions from a wide phylogenetic range of protists) markers outweighs the theoretical desirability of a 'functionally relevant' marker. If (when) the latter are found they are almost certainly going to (and should) be taxon group-specific. Research on the vast majority of protist groups is not yet so advanced. Importantly, the sexuality of most protist groups is unknown – many are assumed to be long-term asexuals. These are discussed in more detail by Mallet (Suppl. text 1); however we note here that their presumed asexuality does not preclude the use of any of the commonly used marker regions (including the 'biological species marker' ITS2 rDNA) being used as one of the species delimitation criteria. As Mallet states, "It seems most sensible to call them [asexual protist lineages] separate species if they display major disease-causing differences or have other ecological or biological differences that seem worth recognising as species".

Mallet also points out that the very large effective population sizes of protists should theoretically harbour very high levels of neutral or nearly neutral polymorphism in sexual species. In these cases, coalescence times will be very large, which could lead to much sharing of ancestral polymorphism between recent species. Consequently, the use of single markers to delimit species becomes very risky, as this could create many apparent species that bear no relation with patterns elsewhere in the genome. Asexual species with very large populations can become highly genetically structured (high levels of disequilibrium), also risking overestimating the number of units that can reasonably be treated as species.

These features of protist populations may result in an even smoother continuum of differences among what might be considered species, and between what previously (with even more incomplete knowledge) may have been recognised as 'good' species (Fenchel and Finlay 2006). However, as Hoef-Emden and Bass show, it is the case in at least some protist groups that very closely related lineages as measured by the most commonly used markers (SSU and ITS rDNA) can show significant phenotypic and ecological differences that reasonably justify them being referred to as different species, as it is likely that they have different functional profiles. These and other studies refute the idea that 'morphospecies' are insufficiently resolving as a basis for protist alpha-taxonomy. However, we believe that a 'good' species should be genotypically and phenotypically distinct from all others; 'phenotype' is a much more inclusive category than morphology alone. Blindly creating species on the basis of equivalent genetic (marker) differences along can lead to a massive proliferation of 'species' for which there is no known phenotypic basis. As sequence data are often much easier to obtain quickly than the more labour-intensive job of phenotypic/ecological characterisation this is a real temptation, pending further investigation. We feel that in these cases the best approach is to designate strain names/codes, linking an actual organism to its genetic data, and only to create new species when some other line of evidence reinforces that decision.

The advent of next generation sequencing (NGS; principally 454 Sequencing and Illumina platforms at the time of writing) arguably constitutes a second 'molecular revolution' in the way that biological diversity is measured and modelled. These technologies have allowed us to sequence much more deeply into genomes and transcriptomes of organisms across the tree of life and into the lineage diversity of (particularly) microbial communities using culture/specimen-independent environmental molecular probing. Genome projects for individual protist lineages are increasingly growing in number, facilitated by NGS technologies and associated development of analytical methods. These are revealing fundamental truths about many aspects of the biology and evolution of the sequenced organisms. However, multilocus genome-level data can be generated in less intensive than reconstructing full genomes, and these offer a powerful way of delineating fundamental taxonomic units, avoiding many of the pitfalls outlined above in this paper. Statistical packages known as 'assignment tests' identify genetically distinguishable populations using multilocus genotypic data, and assign individuals to those populations (Pritchard et al. 2000; Huelsenbeck and Andolfatto 2007; Falush et al. 2003, 2007; Corander et al. 2003, 2004). All of these implementations use inhomogeneities in the data, particularly linkage disequilibrium, to detect groups; one does not expect to see heterozygote deficits or linkage disequilibrium within populations (see Pritchard et al. 2000). The results are usually displayed as a bar graph, with the bars for each individual coloured to represent Bayesian posterior probabilities, the

assignment probabilities, of belonging to each group (see Figure in Suppl. text 1). If the samples are all taken from a single area, species delimitation via genotypic clusters in sympatry consists of detecting inhomogeneities of population structure in multilocus genetic data; the resulting assignment probabilities represent probabilities of belonging to a particular species. Mallet regards assignment tests as the method of choice when implementing the genotypic cluster idea of species as a delimitation procedure. Their use in protists, despite the caveats described above about large populations sizes and large numbers of asexual taxa, is untested but could be very informative. The small amount of genome fingerprinting work done on protist populations suggests that populations are much more highly structured than the prevalent marker-based assessments show, and that this diversity is ecologically informative (e.g. Logares et al. 2009), and therefore taxonomically relevant. Therefore, NGS can open up a genomic perspective in taxonomic studies, allowing more inclusive and adaptively/functionally informative differences to be measured among strains and inform taxon delineation studies.

Another application of NGS technology to protistology is environmental marker tag sequencing to measure microbial assemblage structure and response to environmental conditions (e.g. Creer et al. 2010; Medinger et al. 2010; Stoeck et al. 2010), when environmental variables are co-measured with nucleic acid sample collection. Although this approach does not give much information about each lineage detected, it does provide a very detailed perspective on community structure and diversity. Methods are currently being developed to integrate such sequence diversity data with environmental parameters relating to the same samples, so that of 'ecotype' genotypes can be defined by their multi-dimensional ecological cluster profiles. Ecotypes are genetically consistent functionally informative units below the level of morphospecies (Finlay 2004; Finlay et al. 2006), and can form the basis of future genomic/transcriptomic investigation once they are isolated from the environment. This approach, an ecological analogue to the genomic fingerprinting/assignment tests described above, could provide the basis for delineating ecologically defined 'species', even when no other information about the organisms involved is available (Suppl. text 3).

## Conclusions

The overriding outcome of the Symposium was that although a species category does not exist in nature (and there is no reason why it should), and that species are not fixed biological points in time or space, the concept of species is not an outmoded or unhelpful one. However, what scientists refer to as species will differ between biological situations, and therefore the methods used to delineate species should be made clear in each study. Sometimes this will require more clarification than in others. An increasing battery of

sampling and analytical tools and approaches offer many opportunities for achieving this. As the application of the term 'species' to a group of organisms always implies some kind of hypothesis (as we are imposing these artificial or transient boundaries on nature), an informed and flexible use of the term should actually enhance and inform the scientific process.

The central proposal therefore is to skip the search for a correct general concept but in each case firstly to clearly state the concept (basis for taxonomic subdivision) that is being used. Secondly, for biodiversity studies based on molecular data we need to accept that such data target different levels of resolution in different groups of organisms. Thirdly, it is generally unlikely that current genetic marker-based diversity studies are serendipitously 'counting' species, although the marker chosen may be a close approximation to a sensible species-level distinction. The extent to which marker vs. phenotypic vs. genomic differences vary *within* apparently well-defined and homogeneous groups is also largely unknown, and must be accounted for as much as possible. The idea of the 'species' as a basic unit of diversity is in many cases problematic and alternative, complementary views and measures of biodiversity, functional diversity, and evolutionary complexity should be brought into play.

To address the issues outlined at the end of the Introduction, we recommend the following guidelines for the future of protist alpha-taxonomy:

- (A) Species may be delineated in a variety of ways. The delineation method can be determined by the case in point and may be based entirely on phenotype or genotype, if unavoidable. In such cases, it might be prudent to wait until complementary data are available before formally describing species to avoid later changes.
- (B) Ideally, species will be delineated based on more than one line of evidence, and their distinguishing features made explicit in a diagnosis. Such species need not correspond to any existing species 'definition', but the basis for delineation should be made clear by the authors.
- (C) Genome-level profiling (genomic fingerprinting, sequencing, multi-gene datasets) has great potential for describing and quantifying differences between lineages, particularly as single-cell genomics is now possible in many cases. This is not currently an approach accessible to many protistologists, but should be actively developed as a taxonomic tool. In addition, generating genomic/transcriptomic studies of multiple protist lineages from all branches of the tree of life should be an ongoing objective of high priority for the research community.
- (D) A robust phylogenetic framework is essential for correct interpretation of phenotypic characters and divergence. This should be a base standard for the establishment of new species.
- (E) Environmental sequence data offers great potential for assessing species diversity in nature. However, this

approach requires knowledge of how genotypic and phenotypic divergence is related in the groups detected in order to be reliable and informative. Methods should be developed to integrate ecological and genotypic data from environmental studies to enable robust delineation of ecotypes and (surrogate) species.

## Acknowledgments

We would like to thank Prof. Dr. Klaus Hausmann, President of VI ECOP, and his organizing team, for funding and assisting us to organize and present this symposium at VI ECOP, Berlin, July 2011.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.ejop.2011.11.004](https://doi.org/10.1016/j.ejop.2011.11.004).

## References

- Agar, A.W., 1996. The story of European commercial electron microscopes. In: Mulvey, T. (Ed.), *The growth of electron microscopy. Advances in Imaging and Electron Physics* 96. Academic Press, London, pp. 415–584.
- Bütschli, O., 1880–89. *Protozoa, Bronns Klassen und Ordnungen im Tierreich*. CF Winter, Leipzig.
- Corander, J., Waldmann, P., Sillanpää, M.J., 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* 163, 367–374.
- Corander, J., Waldmann, P., Marttinen, P., Sillanpää, M.J., 2004. BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* 20, 2363–2369.
- Creer, S., Fonseca, V.G., Porazinska, D.L., Giblin-Davis, R.M., Sung, W., Power, D.M., Packer, M., Carvalho, G.R., Blaxter, M.L., Lamshead, P.J.F., Thomas, W.K., 2010. Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. *Mol. Ecol.* 19 (s1), 4–20.
- Dobell, C., 1932. *Antony van Leeuwenhoek and his 'little animals'*. Bale and Danielsson, London.
- Doflein, F., 1916. *Lehrbuch der Protozoenkunde*. Fischer, Jena.
- Ehrenberg, C.G., 1838. *Die Infusionstierchen als Vollkommene Organismen*. Voss, Leipzig.
- Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587.
- Falush, D., Stephens, M., Pritchard, J.K., 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* 7, 574–578.
- Fenchel, T., Finlay, B.J., 2006. The diversity of microbes: resurgence of the phenotype. *Philos. Trans. R. Soc. B* 361, 1965–1973.
- Finlay, B.J., 2004. Protist taxonomy: an ecological perspective. *Philos. Trans. R. Soc. B* 359, 599–610.

- Finlay, B.J., Esteban, G.F., Brown, S., Fenchel, T., Hoef-Emden, K., 2006. Multiple cosmopolitan ecotypes within a microbial eukaryote morphospecies. *Protist* 157, 377–390.
- Green, J.L., Holmes, A.J., Westoby, M., et al., 2004. Spatial scaling of microbial eukaryote diversity. *Nature* 432, 747–750.
- Huelsenbeck, J.P., Andolfatto, P., 2007. Inference of population structure under a Dirichlet process model. *Genetics* 175, 1787–1802.
- Kent, W.S., 1880–81. *A manual of the Infusoria*, Vol. 1–3. David Bogue, London.
- Linnaeus, C., 1758. *Systema Naturae*, 10th ed. Laurentii Salvii, Stockholm.
- Logares, R., Boltovskoy, A., Bensch, S., Laybourn-Parry, J., Rengefors, K., 2009. Genetic diversity patterns in five protist species occurring in lakes. *Protist* 160, 301–317.
- Medinger, R., Nolte, V., Vinay Pandey, R., Jost, S., Ottenwalder, B., Schlotterer, C., Boenigk, J., 2010. Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. *Mol. Ecol.* 19 (s1), 32–40.
- Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., Worm, B., 2011. How many species are there on earth and in the ocean? *PLoS Biol.* 9, e1001127.
- Muller, O.F., 1773. *Vermium Terrestrium et Fluviatilium*. Hauniae et Lipsiae.
- Pascher, A., Lemmermann, E., 1914. *Die Susswasserflora von Deutschland, osterreich und der Schweiz*. Band 1: Flagellatae. Gustav Fischer, Jena.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Schloss, P.D., Handelsman, J., 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* 71, 1501–1506.
- Stackebrandt, E., Ebers, J., 2006. Taxonomic parameters revisited: tarnished gold standards. *Microbiol. Today* 33, 152–155.
- 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 (s1), 21–31.
- von Gleichen, Freiherr, W.F., 1778. *Abhandlung uber die Saamen- und Infusionsthierchen- und uber die Erzeugung*. Winterschmidt, Nurnberg.