Cosmopolitan microbes and their ‘cryptic' species

Tom Fenchel*

Marine Biological Laboratory, University of Copenhagen, Strandpromenaden 5, 3000 Helsingør, Denmark

ABSTRACT: The notion that microbial species have cosmopolitan distribution is discussed for protists in light of the considerable genetic variation within nominal species. It is concluded that there is no compelling evidence to indicate that different genotypes (‘cryptic species’) have biogeographies or that they originated through allopatric speciation. It is suggested that variation in rRNA gene sequences reflects accumulated neutral mutations which do not necessarily reflect functional diversity. The fundamental value of the species concept is that it provides information on the organism in terms of phenotypic properties.

KEY WORDS: Cosmopolitan distribution · Protists · Genetic variation · Functional diversity

INTRODUCTION

Microbial species appear to have cosmopolitan distribution. With respect to eukaryotic microbes, this observation has been made repeatedly over the last ~150 yr by Ehrenberg, Darwin, Schewiakoff, Kahl, and others. About a century ago, Beijerinck made a similar claim for the bacteria; he based this on the use of selective media that made it possible to isolate different functional types of bacteria from almost anywhere. The phrase ‘Everything is everywhere — the environment selects’ (Baas Becking 1934) expresses the idea somewhat more succinctly. Any given microbial species will occur anywhere on the surface of the Earth provided that particular habitat requirements are met. The distribution of microbes does not depend on contingencies occurring over evolutionary or ecological time scales — as in animals and plants — but only on the properties of the habitat (for historical references, see Finlay 2002).

Of course, not all protists are cosmopolitan. Some species (e.g. certain ciliates, foraminifera, and dinoflagellates) are restricted to particular climatic zones. However, in such cases, warm-water species appear to be pantropical and cold-water species seem to be bipolar (e.g. Dragesco 1968, Montresor et al. 2003). While accepting that most protists have cosmopolitan distribution, some authors also maintain that certain species, especially among (terrestrial) testate amoebae and ciliates, do have real biogeography in the sense of being confined to certain continents or to a particular hemisphere, with the implication that these exemplify allopatric speciation (e.g. Foissner 1999, Wilkinson 2001). This cannot, of course, be ruled out, but in many cases putative endemisms have later proved to be the result of undersampling. It stands to reason that the distribution of symbiotic, host-specific microbes is limited by the distribution of the host species.

The relatively recent introduction of gene sequencing (especially SSU rRNA genes) has disclosed that many nominal species represent large genetic distances between different isolates. This has again led to the suggestion that a ‘cryptic’ functional diversity exists that reflects particular adaptations among morphologically indistinguishable members of nominal species, and also that such cryptic species may result from allopatric speciation and display limited geographic distributions.

The present study discusses the apparent absence of biogeography of protists in the light of recent genetic data, and in this context also considers what constitutes a microbial species. The conclusion tends to be that the large genetic distances found within many microbial species reflect huge absolute population sizes in combination with a substantial geological age of most microbial phenotypes. Genetic variation in SSU rRNA genes reflects the accumulation of neutral mutations and may or may not correlate with phenotypic differentiation. So far, there is no compelling evidence to indicate that genotypic variation within nominal species shows any geographic patterns, or that the large genetic variation within nominal species necessarily correlates with a corresponding functional diversity.

*Email: tfenchel@bi.ku.dk

© Inter-Research 2005 · www.int-res.com
THE COSMOPOLITAN DISTRIBUTION OF NOMINAL SPECIES

The majority of nominal protist species occur worldwide wherever appropriate habitat requirements are met. One corollary of a global distribution of microbial species is that their global species number is relatively low in comparison with macroscopic organisms. The huge number of macroscopic animal and plant species is a consequence of endemisms and allopatric speciation. Compilations of the global number of species are uncertain, and taxonomic experts may disagree on the delimitation of species. Furthermore, although many species may remain undiscovered, synonyms abound. Nevertheless, it is a fairly robust conclusion that most nominal species measure around 1 cm in length and that the number of nominal species below this size range decreases with decreasing body size. This pattern applies to aquatic and terrestrial organisms alike (May 1988, Fenchel & Finlay 2004). It has been estimated that there are about 11,000 named protozoa (Finlay 2001) corresponding to about 6% of all named aquatic animals. Local species size spectra are therefore different from those at the global scale. Locally, small organisms dominate in terms of number of species. In an attempt to identify all eukaryotes in a 1 ha pond and a 2 ha part of a marine shallow water bay (Fenchel & Finlay 2004, Finlay & Fenchel 2004), it was found that the dominating size class of species is within the range 10 to 100 μm in body length. Protists constitute about 2/3 of all eukaryotic species recorded at these sites (a total of ca. 1000 species at each site). While the registration of the macrofauna and -flora is probably almost complete, it will undoubtedly be possible to recover an additional number of protists at these sites.

It was further found that the identified organisms represented an increasing fraction of the global species pool with decreasing size. Thus, almost 50% of all known marine heterotrophic flagellates and about 20% of all known marine ciliates were recovered at the marine site. Finally, it can be shown that the fraction of ‘cosmopolitan’ species (defined somewhat arbitrarily as occurring in at least 2 oceans and both hemispheres, or in 2 biogeographic regions and both hemispheres) increases with decreasing body size. This pattern applies not only to the comparison between protists and macroscopic organisms, but also to the continuous relationship, so that small metazoans also tend to have a wider geographical range than larger ones. The fact that the meiofauna (usually defined as benthic metazoans that can pass through a 1 mm sieve) tend to have cosmopolitan distribution has previously been observed (e.g. Giere 1993).

Undersampling, multiple levels of taxonomic resolution, synonyms, and diverse practices among different taxonomic experts remain important problems, and the categorical claim that all protists can be recovered anywhere on Earth given the appropriate habitat is premature. However, there is little doubt that this applies to the majority of protists. Additional evidence that protist species are essentially cosmopolitan can be found in Fenchel et al. (1997), Finlay & Clarke (1999), Esteban et al. (2000), Azovsky (2002), and Finlay et al. (2002).

The explanation for these observations is essentially numbers: population sizes of particular organisms are roughly inversely proportional to their body size, and microbes therefore have huge population sizes. Beneath a water surface of 1 ha, there will be around 10^{19} bacterial cells and perhaps 10^{16} protists. Although these numbers must be divided among many species, the number of individuals of easily detectable species will still be enormous and ‘rare’ species with perhaps 10^7 cells below the 1 ha water surface will, in practice, be undetectable.

Large population sizes create 2 effects: (1) efficient dispersal, because — everything else being equal — the probability that a viable individual is transported to another site is proportional to absolute population size; and (2) the probability of local extinction becomes very small. Extinction is caused by environmental and demographic stochasticity; the expected time for extinction, however, grows quickly with increasing population size. Any habitat therefore includes (in addition to a relatively limited number of thriving species) a large number of relatively small species populations. These may grow only occasionally under particular conditions or not at all, but they are constantly supplemented with immigrants from elsewhere in the form of spores, cysts, or vegetative cells.

This mechanism has been established in plant and animal ecology — empirically as well as theoretically. Species with high local abundances tend to have a wide regional occurrence and local species populations may be sustained through immigration in habitats where they would otherwise eventually become extinct (Hanski & Gyllenberg 1997). The observation that for microbes ‘everything is everywhere’ — the environment selects’ is simply an extrapolation of this principle to organisms with huge population sizes.

‘CRYPTIC’ SPECIES

It has long been known that some nominal ciliate species, like many metazoans, represent complexes of sibling (or ‘cryptic’) species. These are morphologically similar or almost similar, but are completely or partially genetically isolated by sterility barriers. Some of these complexes also include asexual strains that can be characterised by SSU rRNA gene sequences. The genetic distances between these sibling species may...
be slight or large (Nanney & McCoy 1976, Nanney et al. 1998). In some cases, rather subtle functional differences have been suggested to characterise individual cryptic species.

In adherence to the biological species concept, these different strains have been given binominal species names. The wisdom of this is, perhaps, debatable. Some ciliate species apparently do not form such complexes (and isolates from different parts of the world always mate readily), but it is likely that complexes of sibling species are common and most are unlikely ever to be revealed.

It has recently been suggested that the *Tetrahymena*-complex (so far including 39 types) represents examples of limited geographical distribution (see Nanney 2004 at http://www.life.uiuc.edu/nanney/biogeography.html). However, several of the sibling species do occur worldwide. Furthermore, sampling sites are strongly biased in favour of North America: among 25 types isolated in other continents, 20 have been found in North America. Finally, in addition to undersampling in most parts of the world, the absence of any *a priori* idea about how many more such undiscovered strains may exist renders premature any conclusion about limited distribution patterns of sibling species.

WHAT IS A NAME?

Discussions about microbial diversity often imply that a theoretically based species concept exists. Following Mallet (2001), a species simply represents the basic rank of classification as delimited by taxonomic experts and the reality of species is often overestimated. A species does not constitute an evolutionary unit. The evolutionary unit is a clone in the case of asexual organisms or a population (a common gene pool) in the case of sexual outbreeders. The latter corresponds to the biological species concept, but this has meaning only for a given interbreeding population at a given time. For asexual organisms (all prokaryotes and a large fraction of eukaryotic microbes as well as many plants and some animals), a species is simply a collection of clones with similar phenotypic traits maintained through stabilising selection. Although each clone evolves independently of other clones, natural selection in most cases tends to preserve particular discrete combinations of phenotypic traits that taxonomists recognise as ‘species’ (Hutchinson 1968).

In sexual outbreeders, the taxonomist’s delimitation of species may often agree with the biological species concept in the sense that the members are interfertile and evolutionarily isolated from other species by sterility barriers. However, there is rarely a possibility to test this; sterility barriers are not a simple function of genetic differentiation, and in practice as in many animals and plants there are often many other complications such as clines, complex patterns of sterility barriers, and apomictic plants.

The fundamental value of the species concept is that it provides information on the organism in terms of phenotypic properties. Organisms tend to come in discrete combinations of correlated phenotypic properties. These are the units of biodiversity in terms of functional properties, and ecology is concerned with the interaction between such units (species) and their biotic and abiotic environment.

ENVIRONMENTAL DNA — WHERE ARE ALL THE SPECIES?

Recently, attempts have been made to estimate microbial diversity from DNA extracted from natural environments, with subsequent sequencing of rRNA genes (for eukaryotic microorganisms, e.g. López-Garcia et al. 2001, Dawson & Pace 2002, Moreira & López-Garcia 2002).

The result of such attempts is always a number of sequences that are more or less similar (but very rarely completely identical) to sequences already recorded in databases, and also some sequences that are different from anything previously recorded, so that the sequence can only be assigned to some higher taxonomic level. The conclusion is invariably that a hidden and undetected microbial diversity exists.

This approach poses many difficulties: the provenance of the DNA is unknown and may, for example, derive from parasites; in addition, no robust estimate of the frequency of the sequenced organisms is available, and replication errors during polymerase chain reaction (PCR) is not usually evaluated. While the claim is always for a ‘surprisingly high diversity’ (e.g. López-Garcia et al. 2001), the number of sequences sampled in some particular habitat is, in fact, not very high in comparison with the number of nominal protist species that can be obtained from a water sample by means of microscopy and enrichment cultures. A major problem is that in the literature a ‘new species’ is simply a sequence that is not already in a database. However, the rRNA-genes of most nominal protists have still not been sequenced, and the ‘ghost species’ may in fact have been well known (even in culture) for a long time.

In contrast to prokaryotes, it is probably always possible to identify eukaryotic microorganisms microscopically to an appropriate taxonomic level (or perhaps to recognise a species as being new and undescribed). It is surprising that the studies on environmental DNA from eukaryotes have not combined this approach with microscopy, or have done so rather superficially. It is difficult to believe, and con-
trary to common experience, that a plethora of undescribed and phenotypically distinct species exist in well-studied habitats such as the marine water column. At this stage, I believe that studies on environmental DNA are misleading, at least in terms of functional diversity.

**GENETIC DISTANCES WITHIN NOMINAL SPECIES**

It has been recognised for some time that the genetic distances within nominal species of microbes are often substantial. It is also often believed that such variation reflects adaptive specialisations. This may be the case, but it is not necessary to assume this in order to explain the observations. It is useful to discuss sexual and asexual organisms separately.

**Asexual organisms**

Sex is absent in a number of major groups of protists including several flagellate groups and naked amoebae, and also in representatives of many other groups. Two recent studies present SSU rRNA gene sequences of some bodonid flagellates collected from a wide variety of sites around the world (Von der Heyden et al. 2004a, Koch & Ekelund 2005). It is characteristic that (1) no 2 isolates are genetically identical and (2) the phylogenetic trees of nominal species (e.g. *Bodo desig-nis* and *B. saltans*) have deep roots. While no calibration of the molecular clock is available in the absence of fossils, it is obvious that the last common ancestors of these phenotypes are very ancient in geological terms.

It may be tempting to assume that such genetic branching somehow reflects adaptive radiation, but this is not necessarily so. The phenotypes represent a local adaptive peak in fitness space and have been maintained there over long geological periods by stabilising selection. However, the numerous clonal lineages have each independently accumulated neutral mutations over millions of years. The large genetic variation and distances within the nominal species can be explained in terms of the neutral theory of molecular evolution (Kimura 1983) in combination with huge absolute population sizes, leading to low extinction rates of individual clonal lineages, and stabilising selection that has maintained the particular phenotypes over a long geological period. This must be the null hypothesis and, in the absence of experimental evidence, anything else is speculation.

The fact that all isolated strains of nominal flagellate species were genetically distinct in the studies of Von den Heyden et al. (2004a) and Koch & Ekelund (2005) suggests that the actual number of existing genotypes is huge. Neither these nor other similar studies (e.g. Atkins et al. 2000) indicate any geographic patterns in the sense that strains isolated close together tend to be genetically more similar than strains isolated from sites separated by great distance.

**Sexual outbreeders**

Many groups of protists include sexual outbreeders (e.g. ciliates, foraminifera, dinoflagellates, diatoms). There is substantial variation in the frequency of sex. In some ciliates, conjugation is necessary after a number of mitotic divisions; in other species sex appears to occur rarely and sporadically, or is sometimes induced by degrading environmental conditions. Representatives of some other protists produce gametes seasonally, or alternate between sexual and asexual generations.

Genetic divergence is constrained within populations of interbreeding individuals, and the interfertile *Tetrahymena*-strains discussed above can accordingly be characterised by identical rRNA gene sequences (Nanney et al. 1998).

Striking examples of identical SSU rRNA genes of several planktonic protists (foraminifera and a dinoflagellate) collected from subantarctic and arctic waters, respectively, have recently been published (Darling et al. 2000, Montresor et al. 2003). Another planktonic foraminiferan with bipolar distribution provided evidence of some genetic differentiation (Darling et al. 2004). This may be an example of the initiation of allopatric speciation in these large protists separated by the longest possible distance on Earth. Under all circumstances, complete similarity of SSU rRNA genes does not, of course, preclude genetic isolation or differentiation in genes other than rRNA genes, but it does constrain the time for the latest exchange of genes between these antipodal populations to the later part of the Pleistocene.

Notwithstanding this example of recent genetic differentiation in foraminifera, allopatric speciation is unlikely to play any role in microbes. In organisms with clonal evolution, there are no constraints on sympatric genetic differentiation. In sexual species, clado genesis may occur as a result of chromosomal mutations; a process that in most protists is facilitated by the fact that they are capable of reproducing by mitosis for many generations (similar to sympatric speciation in plants due to hybridisation or chromosomal mutations). Sympatric phenotypic differentiation is a possibility following sterility barriers, but as in clonal organisms in which each cell division creates 2 new 'biological species' this does not necessarily take place due to stabilising selection.
GENOTYPES AND PHENOTYPIC DIFFERENTIATION

Of course, directional selection does occur, and may even trace environmental changes over short periods of time. However, the direction of selection changes over time and will often tend to average out over long periods so that the phenotypes are largely conserved. Nevertheless, it is to be expected that nominal species may represent some phenotypic differentiation among different habitats or climatic zones. Differences between rRNA gene sequences are selectively neutral and not per se involved in any particular adaptive phenotypic trait. The question is whether there is any correlation between genetic distance (in terms of rRNA-sequences) and possible functional diversity.

There are relatively few studies on phenotypic diversification within nominal species of protists, and not all are convincing in terms of experimental reproducibility (e.g. growth rate constants measured in populations that did not display balanced growth).

Koch & Ekelund (2005) found that isolates of Bodo desiginis from freshwater could not grow at salinities exceeding 10 to 20 ppt, while marine isolates seemed to thrive at all salinities (from freshwater to full strength seawater). The corresponding data on gene sequences data are, however, too limited to determine whether the marine and freshwater strains represent separate clusters in terms of rRNA gene sequences. Such clusters of marine and limnic isolates of goniomonad flagellates were, however, demonstrated by Von der Heyden et al. (2004b). Similarly, B. J. Finlay et al. (unpubl.) found that freshwater isolates of the ciliate Cyclidium glaucoma represent a separate branch on the genealogical tree for this species, and there is evidence that isolates from hyperhaline habitats (natural salt lakes and salterns) constitute another branch that is somewhat closer to that including marine and brackish water isolates. A marine isolate of C. glaucoma grew well at all salinities from freshwater to hyperhaline water (90 ppt), but 2 freshwater isolates failed to grow at salinities exceeding 20 to 25 ppt. In this case, it seems that C. glaucoma must have split into 2 clades at some point in its geological past; one of these clades adapted to life in freshwater and subsequently lost the ability to grow in full strength seawater.

In contrast, a recent study (Lowe et al. 2005) of the euryhaline dinoflagellate Oxyrrhis marina showed no such correlation between tolerance to different salinities and genotype. As in other such studies, there were substantial genetic distances between some of the strains, again indicating a substantial geological age of the last common ancestor of the isolates. No geographic structuring of the genotypes was evident, e.g. strains from the Baltic Sea and from Bahrain were identical, but substantially different from isolates from the Isle of Man and North America. While all strains were quite euryhaline, the Bahrain strain failed to grow at salinities below ~20 ppt while the Baltic strain grew at salinities ranging from very dilute brackish water to a salinity of ~40 ppt. Like the Bahrain strain, some of the North American strains could not grow at low salinities. The differential adaptation to different salinities is therefore not correlated with the rRNA genealogical tree, but must have evolved independently on different occasions.

Comparing different protist nominal species, it is obvious that niche width varies. For example, many marine species are considerably less tolerant of brackish water than the above-mentioned creatures. Some species are also more selective with respect to food particles than others. A wider niche, e.g. in terms of being able to grow over a wide span of temperatures and salinities, has energetic costs. As we have seen, such specialisations (mainly in terms of loss of some capability) can be observed for particular strains within nominal species. However, it is equally evident that the niches are often much wider than required in the habitat from which a particular strain has been isolated. For example, the Oxyrrhis marina discussed above will never be exposed to salinities exceeding 10 ppt in the Baltic, but it nevertheless thrives at a Mediterranean salinity. The characterisation of O. marina as a euryhaline species still applies irrespective of the fact that it is composed of many distantly related genotypes. Similarly, some nominal protist species tolerate only a limited temperature range (e.g. species of plankton foraminifera that occur only in certain climatic zones). In contrast, isolates of the ciliates Uronema marinum and U. nigricans from Greenland and Denmark, respectively, showed exactly the same temperature span for growth (~1.0 to 38°C), although seawater temperatures in Greenland have probably not exceeded 5°C since the Pliocene (T. Fenchel unpubl.).

The ‘within genotype niche width’ does differ among species — some are more specialised than others. However, niches cannot become infinitely narrow for any species. All environments are, to some extent, stochastic with respect to the biotic and physical-chemical factors, and conditions also vary from place to place and seasonally within any one place. It is therefore unlikely that the plethora of genotypes displayed by microbes reflects some subtle and undetected ‘fine-tuning’ to different environments. If so, this must be demonstrated experimentally, or it will remain pure speculation.

CONCLUSION

For microbes, ‘everything is everywhere — the environment selects’, and this also seems to apply at the level of different rRNA genotypes within nominal spe-
cies. Other genetic variation caused by selection in response to, for example, different climates or salinities may evolve rapidly and may show spatial patterns — just as is found in animal and plant species with a wide distribution range. However, microbes do not display biogeography in the sense that there are penguins in the Antarctic, but puffins and auks in the Arctic.

Techniques of molecular genetics have greatly increased our understanding of microbial phylogeny. It has also supported classical taxonomy in that gene sequencing usually shows that nominal species are represented by monophyletic sequence clusters. Its application for understanding microbial diversity in natural habitats has, however, resulted in a misleading picture of functional diversity. The measured genotypic variation must largely be selectively neutral and above all reflect the molecular clock. Discovering and understanding functional diversity and microbial interactions must primarily depend on direct observation of phenotypes and on experimental work.

Acknowledgements. I am grateful to Bland J. Finlay for valuable comments on an earlier draft of the manuscript.

LITERATURE CITED


Baas Becking LGM (1934) Geobiologie of inleidung tot de milieukunde. WP van Stockum & Zoon, Den Haag


Submitted: July 26, 2005; Accepted: September 5, 2005

Proofs received from author(s): October 14, 2005