

such as fluorescence microscopy, can be used to count the number of prokaryotes in a sample, providing us with N_{tot} . On the basis of these numbers, the authors make the following diversity estimates: oceans, 160 'species' ml^{-1} ; soil, 6400–38 000 g^{-1} ; sewage works, 70 ml^{-1} . This last number is a nice example of the paradox of enrichment, the phenomenon whereby fertilization (in the agricultural sense) reduces diversity. The low diversity estimated for the oceans and the high diversity for soils are consistent with results from studies using very different methodologies [8]. Understanding this difference is as important as understanding the latitudinal diversity gradients observed for macroscopic organisms. Also, the low diversity associated with high nutrient levels has also been observed in prokaryotic communities in sediments beneath fish farms [8].

My previous remark about the statistics of $N_{\text{max}}/N_{\text{tot}}$ in small samples requires elaboration. Suppose that, in the community, $N_{\text{max}}/N_{\text{tot}}$ is actually quite small. Then it is possible that our sample will not actually contain the most abundant species in the community. If we know the identity of this species, then we can restrict our analysis to samples that contain it, although this would require a modification of the estimation formula to accommodate this extra level of sampling. But if we have no means of restricting our analysis to such samples, then there appears to be no way to estimate diversity by this method, although this could simply be a failure of imagination on my part.

Finally, this new paper raises very starkly an extremely interesting question: just what, precisely, is the community being sampled? What does it mean to say that there are 70 species ml^{-1} in sewage works? It could mean anything from 'there are 70 species in sewage works, period' or 'there are billions of species in sewage works'.

Turning to macroscopic organisms, if we used this technique to estimate, say, tree diversity then we would need an understanding of β diversity (how diversity changes through space) to go beyond statements about numbers of trees ha^{-1} to numbers of trees in the world. In fact, it is entirely possible that the oceans of the world constitute a single prokaryotic community, just as aquatic microbial eukaryotes appear to be ubiquitous as a consequence of the abundance and dispersal abilities of the organisms concerned [9]. But it is not obvious that there is no β diversity in soils [8], and this needs further investigation.

References

- Floyd, R. *et al.* (2002) Molecular barcodes for soil nematode identification. *Mol. Ecol. Evol.* 11, 839–850
- Lopez-Garcia, P. *et al.* (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409, 603–607
- Moon-van der Staay, S.V. *et al.* (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* 409, 607–610
- Zettler, L.A.A. *et al.* (2002) Eukaryotic diversity in Spain's River of Fire – this ancient and hostile ecosystem hosts a surprising variety of microbial organisms. *Nature* 417, 137–137
- Moreira, D. and Lopez-Garcia, P. (2002) The molecular ecology of microbial eukaryotes unveils a hidden world. *Trends Microbiol.* 10, 31–38
- Kerr, R.A. (2001) Life – potential, slow, or long dead? *Science* 294, 1820–1821
- Curtis, T.P. *et al.* (2002) Estimating prokaryotic diversity and its limits. *Proc. Natl Acad. Sci. U. S. A.* 99, 10494–10499
- Torsvik, V. *et al.* (2002) Prokaryotic diversity – magnitude, dynamics and controlling factors. *Science* 296, 1064–1066
- Finlay, B.J. (2002) Global dispersal of free-living microbial eukaryote species. *Science* 296, 1061–1063

0169-5347/02/\$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0169-5347(02)00043-5

Letters

Shortcuts in systematics? A commentary on DNA-based taxonomy

Ole Seberg¹, Chris J. Humphries², Sandy Knapp², Dennis Wm. Stevenson³,
Gitte Petersen¹, Nikolaj Scharff⁴ and Nils Møller Andersen⁴

¹Department of Evolutionary Botany, Botanical Institute, University of Copenhagen, Gothersgade 140, DK-1123 Copenhagen K, Denmark

²The Natural History Museum, Cromwell Road, London SW7 5BD, UK

³The New York Botanical Garden, Bronx, New York, NY 10458, USA

⁴Entomology Department, Zoological Museum, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark

The primary aims of taxonomy are to name, circumscribe, describe and classify species. The first goal is convention but the remainder are science. The International Codes of Nomenclature are legislative instruments and nomenclature is simply a mechanism to ensure that a species

name is legitimately attached to a type specimen, regardless of scientific status. The type of a species does not serve, as Tautz *et al.* ([1,2], but see [3]) assert, as 'the central reference for comparisons'. The crucial link between names and scientific investigation is species circumscription followed by description. The Codes require Linnaean binomials: a genus name and a species epithet.

Corresponding author: Ole Seberg (oles@bot.ku.dk).

<http://tree.trends.com>

The rules are totally silent about what constitutes a species; rather this is a key goal of biological investigation. Circumscriptions of genera and species evolve as science progresses.

The Linnaean binomial system is not 'inherently unstable' but is used to interpret the underlying science. The problem that 'a name that has been used for a long time thus can suddenly disappear' (or reappear, for that matter) is a nuisance for everyone. However, if name changes are considered a serious problem, then the conservation and rejection criteria available in the Codes can be invoked.

Neither the Botanical nor the Zoological Code has a fixed authoritarian supervisory body; rather, they have committees chosen by a democratic process that oversee changes. The Codes work by consensus and are designed to be open and universally applicable [4]. Tautz *et al.* ([2], and see [3]) advocate universal, centralized, apparently obligatory registration, a concept emphatically rejected by the International Botanical Congress in 1999. Taxonomists from developing nations led the move against centralization, fearful that the wealthier nations were attempting to monopolize information. In our view, registration would eventually strangle systematics, as debate will be discouraged.

Most current taxonomy is pursued using low-cost technology. Mandatory introduction of DNA sequences into taxonomy seems to us a retrograde step. In most instances, a quick survey of morphology will serve the same purpose and, although morphology has its problems, DNA has as many pitfalls. A sufficiently different sequence might warrant the description of a new species, as will a sufficiently different morphology. An expensive and centralized DNA-based taxonomy would only add to the North–South divide in taxonomy, and might exclude the many taxonomists who have limited access to sequencing technology.

Acknowledging that there need not (or cannot) be universal agreement about which region of the genome to sequence, Tautz *et al.* suggest using house-keeping genes, especially the ribosomal genes (at least in animals). Although abundant, ribosomal sequences might be an inappropriate choice in the long run, because they have profound alignment problems and are subject to different degrees of concerted evolution.

Matching existing Linnaean names with DNA sequences is fraught with hazards. Although it is tempting to create a new starting date for priority for the Botanical and Zoological Codes [5], it seems futile to replace existing types (from which DNA in some cases cannot be extracted) with neotypes. As indicated by the authors [1], expert taxonomists are in short supply and many important groups are neglected. It takes little imagination to envisage the problems broad-scale designation of neotypes would cause.

Tautz *et al.* argue that existing DNA data bases represent bad taxonomy, as 'there is no guarantee that the correct species names were assigned by the submitter of the sequence, because there are no established taxonomic standards under which such submissions have to be done'. The circumscription of a species is an opinion

[6]. One might ask therefore who is going to decide on 'taxonomic standards'. Circumscription changes with increasing knowledge; that's the science. Many sequences are deposited in DNA data bases, but, if a specimen has been misidentified, only inspection of the voucher can solve this problem. It is naïve to think that 'phylogenetic analysis of query sequences, will readily place any sequences from new species', as if the problems of homoplasy, alignment, and even phylogenetic methods would not add to the unreliability of information in existing data bases.

The role of collections in systematics is vital. Collections represent a comparative model of diversity, and therefore, as the authors [1] point out, specimens should be retained as intact as possible. Destructive sampling of type specimens for any characters has always been a severe problem and such procedures are accepted only as a last resort. To destroy a sample to extract its DNA might be inadvisable in the light of future needs.

With regard to taxonomic and phylogenetic studies, current practise attempts to create higher level classifications that do not conflict directly with cladograms [7]. Today, phylogenetic methods applied to DNA data are a strong focus of systematics research, and it is difficult to believe that the pendulum might swing in another direction. However, methods do change and the fixed opinions of today might seem redundant in the future.

So what is the upshot? Indeed, what has the molecular revolution really achieved for taxonomy? The main advance is that it has provided access to copious data for cladistic analysis and, in our view, has provided new data sets and a new class of characters that can be extremely useful in broad-scale comparisons of everything from bacteria to mammals and plants. Often one is forced to use one, or a few, carefully selected specimens as representatives of a taxon, but experience has repeatedly shown that this can be a major mistake. Deliberately using a single specimen as a representative of the taxon will only create havoc in taxonomy, a fact long realized by taxonomists working with other types of data. Individual bases and DNA sequences are simply characters, tiny fragments of the lifecycle. It seems perverse to us to advocate using a DNA sequence as a mandatory identification tag for a species, even as a first approximation. Therefore, this plea for a DNA-based taxonomy is little more than a *cri de coeur* for bioinformatics and increased reliance on sequence data. We agree that there is a very strong need for efficient, although perhaps not centralized, repositories for DNA samples. How and where such samples should be curated, and how the curators should be funded, is an open question. We also agree that it would be a good idea to include DNA sequences in the diagnoses of taxa. However, we feel that a DNA-based taxonomy along the suggested lines would catastrophically retard taxonomic activity, and it would certainly not relieve the need for many more taxonomists, especially in the tropics, for the good of taxonomy's many dependent user groups.

References

- 1 Tautz, D. *et al.* (2002) DNA points the way ahead in taxonomy. *Nature* 418, 479
- 2 Tautz, D. *et al.* (2003) A plea for DNA taxonomy. *Trends Ecol. Evol.* 18 DOI: 10.1016/S0169-5347(02)00041-1
- 3 Minelli, A. (2003) The status of taxonomic literature. *Trends Ecol. Evol.* 18 DOI: 10.1016/S0169-5347(02)00051-4
- 4 Knapp, S. *et al.* (2002) Taxonomy needs evolution, not revolution. *Nature* 419, 559
- 5 Godfray, H.C.J. (2002) Challenges for taxonomy. *Nature* 417, 17–19
- 6 Mallet, J. (2003) Taxonomy: renaissance or Tower of Babel? DOI: 10.1016/S0169-5347(02)00061-7
- 7 Lipscomb, D. *et al.* (2003) The intellectual content of taxonomy: a comment on DNA taxonomy. DOI: 10.1016/S0169-5347(02)00060-5

0169-5347/02/\$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0169-5347(02)00059-9

The intellectual content of taxonomy: a comment on DNA taxonomy

Diana Lipscomb¹, N. Platnick² and Q. Wheeler³

¹Department of Biological Sciences, George Washington University, Washington DC 20052, USA

²Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA

³Department of Entomology, Comstock Hall, Cornell University, Ithaca, NY 14853, USA

Tautz *et al.* [1] propose to ‘solve’ the lack of adequate classifications and effective identification tools (the so-called ‘taxonomic impediment’) by replacing existing classifications with a system in which an infinitesimally tiny fraction of an organism’s genome is sequenced and used both to classify and identify the organism in question. The rationale for this suggested change, however, is specious and unlikely to produce a progressive research program. Such a system is already in use for unculturable prokaryotes, where the best we can do at present is collect sequence data from the environment, compile data bases of the results, and construct ‘classifications’ that reflect only the degree of similarity displayed by those sequences. This produces what is at best a caricature of real taxonomy, in which sequences that diverge by <5% are considered ‘conspecific’ (never mind that humans and chimpanzees might be far less divergent than that). Microbiologists would be the first to agree that, when the organisms can be cultured and their other attributes studied, we can do far better than this. Why, then, reduce the taxonomy of all other organisms to this impoverished state? The supposed advantages of DNA taxonomy do not stand up to rigorous scrutiny. For example, the claim that a sequence ‘is not influenced by subjective assessments’ ignores the difficulty of aligning sequences of different length, distinguishing paralogs from orthologs, or even selecting appropriate genes for any particular taxonomic study. Similarly, the supposition that DNA identification will lessen the confusion that sometimes results when taxonomic names change is unjustified. The only way that a DNA sequence identification tag could ameliorate confusion would be if the gene sequence used were constant among all members of the species but different in all other species. There is no evidence that most genes meet these criteria, and any

diagnostic character that meets these criteria would work – it need not be molecular.

The ‘taxonomic impediment’ is more effectively addressed by other means. Tautz *et al.* [1] are, in essence, suggesting that if we all become as ignorant of our organisms as are, perforce, those microbiologists working on currently unculturable taxa, the world would no longer suffer when knowledgeable specialists in a group die without having communicated all their knowledge to others. Programs such as the US–NSF’s PEET [Partnerships for Enhancing Expertise in Taxonomy (2003): <http://www.nsf.gov/pubsys/ods/getpub.cfm?nsf00140>], which provides funding to train a new generation of systematists, represent solutions far more effective than simply discarding everything we have learned to date about organismic diversity. Taxonomy might indeed be threatened, but the greatest threats might be from those who would usurp the resources that it needs to grow and thrive.

Tautz *et al.* [1] emphasis on the task of identification indicates a fundamental misunderstanding of the intellectual content and rigorous hypothesis testing that characterize contemporary taxonomy. Reducing taxonomy to identification alone makes it a technical task rather than a hypothesis-driven science. There is no credible reason to give DNA characters greater stature than any other character type. When species descriptions are based on a broad range of data, they become interesting scientific hypotheses making explicit predictions about the distribution of attributes among organisms. We reconstruct phylogenies to explain patterns of organismic diversity. Molecular data certainly contribute, but when nothing is known about organisms except their DNA, there are no evolutionarily interesting patterns to explain – just a tedious pattern of sequence similarity. The advocates of DNA taxonomy seem not to understand the peerless

Corresponding author: Diana Lipscomb (biold1@gwu.edu).