Cladistics (1992) 8:67–72

INDIVIDUAL ORGANISMS AS TERMINAL ENTITIES: LAYING THE SPECIES PROBLEM TO REST

Paul Vrana[1,2] and Ward Wheeler[1]

[1] *Department of Invertebrates, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024-5192, U.S.A. and* [2] *Department of Biological Sciences, Columbia University, New York, NY 10027, U.S.A.*

Received for publication 12 April 1990; accepted 2 December 1991

It has become increasingly clear that the role and definition of species in a phylogenetic systematics context must be clarified both in order to ascertain the nature of the entities used in cladistic analysis and to define the limits, if any, of that analysis. We found the recent papers on this subject by both de Queiroz and Donoghue (1990a,b) and Wheeler and Nixon (1990 and Nixon and Wheeler, 1990) to be lacking. Both place great weight on the importance of species (or "interbreeding populations" in the case of de Queiroz and Donoghue) as fundamental units in phylogenetic analysis. We believe this concept itself is faulty, and that individual organisms rather than any interbreeding "group" in which they are placed should be used as the terminal entities in phylogenetic analysis. There are two primary reasons for doing so: avoidance of both non-monophyletic groups and the generalization of character states over large numbers of organisms. In a phylogenetic context, we would espouse a species definition based on *demonstrated* monophyly. Yet, as Nelson (1989a) has pointed out, establishing the monophyly of a "species" may prove quite difficult. Any other definition, however, leaves one open to greater problems. We see nothing special about taxa based on the unpolarizable attribute of ability to interbreed (e.g. see Nelson, 1989b) or on the attempt (successful or not) to interbreed.

It should be noted that we follow Nelson's (1989b) view on monophyly and the nature of taxa: that is taxa are seen as relationships of organisms, rather than groups of them (Fig. 1). In other words taxa belong to the characters, rather than characters to taxa. Thus "descent" plays no necessary role and the arguments of de Queiroz and Donoghue (1988, 1990a) on paraphyly (e.g. that some descendents of a given ancestor going extinct render a group paraphyletic, or the general proposition that descent has a role in a definition of monophyly) do not apply. We feel this view of monophyly is a key point not only here but for cladistic analysis in general, particularly in its relationship to process theories.

Both Wheeler and Nixon and de Queiroz and Donoghue feel that there is some inherent taxonomic "line of death" below which systematic analyses should not be attempted. This strikes us as being quite similar to the arguments of "too much homoplasy" (e.g. Kirsch and Archer, 1982) invoked by evolutionary taxonomists to show how cladistics might fail under particular circumstances. One cannot know this *a priori*; it is a conclusion to be made after analysis, not an observation. Indeed, the idea that there is no structure below the level of interbreeding populations is merely the untested null hypothesis (G. Nelson, pers. comm.). As de Queiroz and Donoghue note,

0748-3007/92/010067+06$03.00/0

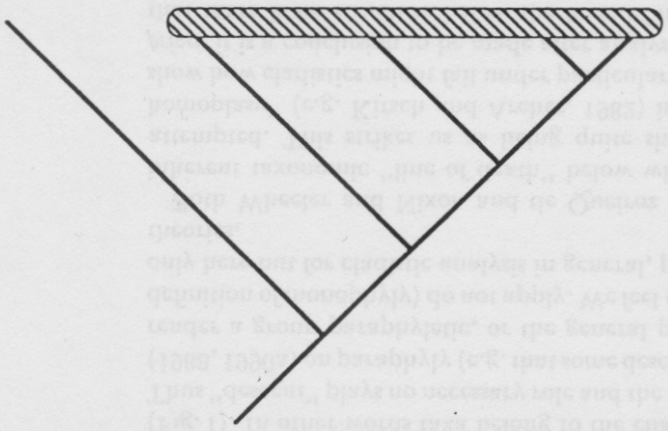

Fig. 1. Our interpretation of Nelson's (1989b) view of monophyly. A taxon is seen as a relationship among terminal units based solely on pattern, here indicated by the hatched area. Note that no descent is implied, and that nodes are not included in the indicated monophyletic group. The nature of these terminal entities is unspecified. We feel they should either be individual organisms, or a demonstratably monophyletic taxon.

such pattern is frequently found at this level with the use of molecular techniques. Even Avise et al. (1987: 518) have pointed out: "Phylogenetic differences within species are qualitatively of the same kind as ... those normally pictured in higher order phylogeny reconstructions".

Certainly every organism is likely to have apomorphies at the DNA level, even if they are uniquely acquired mutations. That is, given that genetic variation has been shown to be present between any two given organisms (e.g. the hypervariable regions used in DNA "fingerprinting"), one cannot argue that there are no varying characteristics (be they "traits" or "characters" *sensu* Nixon and Wheeler, 1990) upon which to test the existence of structure within any group of organisms. Indeed, it is precisely whether these characteristics are traits or characters that is being tested. It cannot be emphasized too strongly that one cannot know this *a priori*. For example, much has been made of Hennig's figure illustrating the difference between tokogeny and phylogeny (see Nixon and Wheeler, 1990, fig. 1; Hennig, 1966, fig. 6). However, this figure is heavily assumption-laden: it presumes that some inexplicable process at the "Species" level renders an entirely clean break between reticulation and divergence, that this break is obvious, and the process is constant across all organisms. In other words, this figure rather than empirical fact, *is an unfounded process statement* (see Nelson, 1989b, for a treatment of process assumptions and their inapplicability to species and taxa definitions). The level at which reticulation occurs and pattern is not discernable can only be determined empirically.

Many recent studies have implicitly attempted to ascertain this level, usually through the use of mitochondrial DNA RFLP or sequence data (due to its high levels of variation), though some have been based on allozyme data. Many of these studies have found discernable structure below that level termed the "species". Nixon and Wheeler (1990) correctly point out that mtDNA's uniparental inheritance may yield a pattern not directly relevant to the question of hierarchy among the organisms—that is a bi-parentally inherited gene could reveal reticulation, indicating that the mtDNA data was a "gene tree" rather than "species" tree. Similar "gene tree" arguments can be made for multi-copy nuclear genes which appear to be have undergone concerted evolution.

Indeed, one can invoke molecular evolution scenarios to explain why *any* locus might not show the "true" organismal phylogeny. While we agree that nuclear loci are preferred, we feel in that the absence of other factors, mtDNA or any molecular data should be taken at face value—namely as the best existing hypothesis of structure. Theory cannot be preferred over data.

To illustrate this view, we refer to Allard and Honeycutts's (1991) recent nuclear ribosomal DNA restriction site phylogeny of the rodent genus *Onychomys*. In this study each of three *Onychomys* species was represented by several individuals, and in the case of *O. leucogaster*, by 15 individuals from three different U.S. states (New Mexico, Texas and Oregon). While 1500 most parsimonious trees were produced by "exact" solution algorithms (Hennig 86's ie* and PAUP's Branch and Bound), a good deal of structure was retained in the strict consensus cladogram presented. In this cladogram, *O. leucogaster* individuals from each state formed a monophyletic group, though there was no resolution below this level (implying reticulation since individual organisms were unique). Furthermore, the Texas and New Mexico clades formed a larger group to which the Oregon clade was sister, thus preserving monophyly of the species. The genus *Onychomys* was also found to be monophyletic with respect to rodents of three other genera (four other species), each represented by a single individual. Thus, *Onychomys leucogaster* was merely a nodal designation below which there was discernable hierarchy.

Does this mean that these state clades of *O. leucogaster* should be considered different species, as Wheeler and Nixon would no doubt suggest? Our answer is simply that we don't know (nor, in this instance do we feel it should be a concern). Certainly, it might be a consistent and useful definition if the term "species" always applied to that level below which reticulation occurred. The caveat is that there is no way to know this prior to analysis, thus in many cases applying the term by this definition must be a statement of blind faith.

While we agree that sexually reproducing organisms with the potential plesiomorphy of interfertility would be more likely to show reticulate patterns than cladistic resolution, this seems a poor excuse for a lack of investigation [Ability to interbreed is always the known character. Panmixis is a possible conclusion to be reached after (probably impossibly) exhaustive study, rather than an assumption to be made.]. Indeed, given that we cannot recognize ancestors and that many species may be paraphyletic in the sense that individuals in that group are in fact more closely related to members of another species, many genera might be equally likely to show lack of resolution(Nelson, 1989a). Analysis of such situations must be based purely on pattern, rather than on how we theorize that any tokogeny/phylogeny interface works.

De Queiroz and Donoghue also raise the problem of individuals being occasionally non-spatio/temporally bounded. This seems a non-argument to us, as populations (which they favor as terminal taxa) exemplify this problem to an exponentially larger degree. Our advice is simply not to use individuals that have received organ transplants (in the case of slime molds, one can simply grow monoclonal colonies). More importantly we feel that cladistic analysis is necessary to establish these groups in the first place. A "population" (let alone "species") not defined by synapomorphy does not seem useful to us.

Donoghue (1985: 177) has been much closer to agreement with our view in the past: "If one wishes to resolve phylogenetic relationships as far as possible—to find the smallest monophyletic groups of organisms—then it seems reasonable that individual organisms should be used as terminal taxa in analyzing relationships ... Organisms

should be placed into more and less inclusive monophyletic groups using shared derived characters as evidence, just as species, genera, or families are united on this basis". However, he has changed his view largely based on the premise that "sex matters" (Donoghue, pers. comm.). This seems to be the character that unites de Queiroz and Donoghue and Nixon and Wheeler to our exclusion. While this proposition may often be true, there is no *a priori* way to know or any established procedure to determine when and how sex will "matter". There is certainly no reason to believe that sex will have equal import across all groups.

It is therefore obvious (by the emphasis on sex within definitions) that there is a bias on the part of species definitions by authors towards multi-cellular sexually reproducing groups. We believe this is undesireable, and that any such theory should extend to all organisms. The notion that asexual organisms constitute an insignificant minority unworthy of consideration in a general theory of biology is based on lack of information, given that few systemic studies focus on protozoan, bacterial or viral groups. Wheeler (pers. comm. and Nixon and Wheeler, 1990) as well as de Queiroz and Donoghue (1988) have agreed that any number of unicellular asexual organisms which show pattern may be termed the equivalent of species. If one uses this not-unreasonable definition (given that this taxon is not seen as a group of organisms or tokogeny = phylogeny) and the realization that each of the "sexual species" hosts at least several of the "asexual"—protozoan, fungal, bacterial and viral—then it must be the former that seem insignificant and aberrant in general biological mechanisms. Some may attempt to portray sex as a more general mechanism, extending to many of the "asexual species". It should not be glibly assumed that the occasional (often one-way) exchange of genetic material in these organisms is equivalent to sexual reproduction in metazoans. This is tantamount to stating that retroviral infection of a human host is that person having sex with the virus. Espouse such a definition if you must, but be aware of the ramifications. Certainly such processes should play no part in the delimitation of the smallest taxa possible.

In practice, every systematist examines individual organisms. Yet, the unit used in the systematic analysis is the species. When there is variation among individuals this is either ignored (common = primitive) or coded as ambiguous—an artificial suite of characters not necessarily observed in any real organism. These seem to us to be the best reasons for using individuals as terminal units in phylogenetic analysis. If there is polymorphism within species in characters that one would like to compare among species, perhaps one should consider establishing the monophyly of the group first. If a number of individuals are identical for all characters considered, they could then be considered one taxon for the purposes of phylogenetic analysis (this may lessen practical objections to our view, as it renders the most parsimonious cladogram more obtainable).

We realize that such a species definition, that is as arbitrary as any other level, has disturbing implications for biodiversity. Casting doubt on the validity of the units being counted renders the conclusions of these studies somewhat dubious. Our view does, however, make cladistic analysis relevant to population genetics, a field riddled with untestable assumptions (and responsible for many of the molecular evolution paradigms), as well as recent biogeographic events (see Allard and Honeycutt, 1991, as well as Riddle and Honeycutt, 1990, for the implications of *Onychomys* phylogeny to U.S. desert biogeography) and conservation biology. As an example of the latter, one may apply cladistic analysis to a number of organisms belonging to a taxon known to be endangered ("species", "genus" or whatever) and obtain a cladogram showing

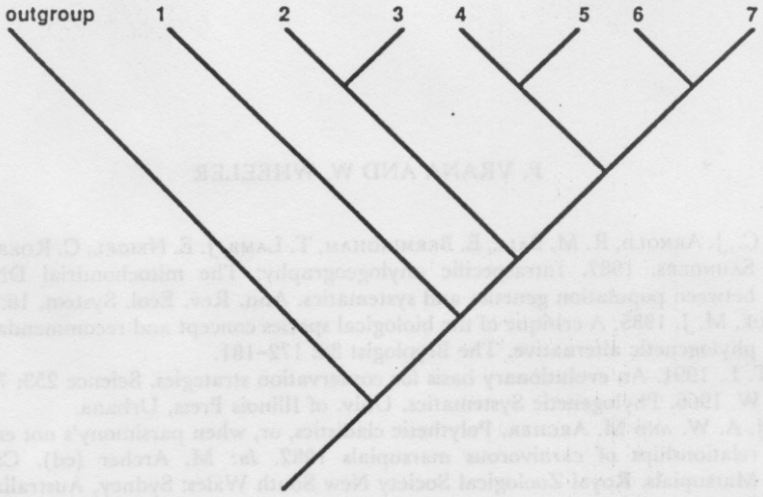

Fig. 2. Hypothetical scheme of cladistic analysis used as a decision criterion in conservation biology. Entities 1–7 represent populations of an endangered species each representing a different biogeographic locality and identified by a unique genetic marker. The outgroup represents the non-endangered sister taxon. Given this cladogram and the ability (e.g. funds) to preserve two of these groups, we would choose group 1, and any from groups 2–7. Given the ability to preserve three of the groups, we would choose 1, either 2 or 3, and one from 4–7. Clearly other factors may influence this decision process, but this procedure is the most logically defensible in preserving the greatest cladistic diversity, given the pattern analysis.

structure between less inclusive taxa of these organisms. A defensible strategy of conservation can then be devised, based on the principle of maximizing cladistic diversity (Fig. 2; Vane-Wright et al., 1991, as well as Erwin, 1991, have also proposed cladogram-based conservation strategies). Thus, in our view cladistics is a more generally applicable method of analysis.

The cladistically determined hierarchy of individuals implies a picture of genealogy—the categorical rank of species is irrelevant. Which, if any, of the groups revealed by geneaology is a species is an entirely different question. We suggest no change in terminology, however; the term species may still be applied to groups on the basis of interfertility. Indeed, this may be appropriate for those explicitly pursuing questions of process; we simply ask that such groups not be used in phylogenetic analysis. As Nelson (1989a) has said: "there seem to be no basic taxonomic unit and no particular unit of evolution" and (species) "problems are insoluble, for they stem from a false assumption: that there is an empirical difference between species and other taxa". In keeping with this, we feel that use of individual organisms as terminal entities is the only logical choice.

Acknowledgements

We thank G. Nelson as well as M. De Pinna for commenting on early versions of this manuscript. M. Donoghue, K. Nixon and Q. Wheeler for their "helpful" and "constructive" criticisms during the review process. We thank J. Gatesy and especially R. DeSalle for commenting on later drafts.

REFERENCES

Allard, M. W. and R. L. Honeycutt. 1991. Ribosomal DNA variation within and between species of rodents, with emphasis on the genus *Onychomys*. Mol. Biol. Evol. 8: 71–84.

AVISE, J. C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB, J. E. NEIGEL, C. ROEB AND N. C. SAUNDERS. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Ann. Rev. Ecol. System. 18: 489–522.

DONOGHUE, M. J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. The Bryologist 88: 172–181.

ERWIN, T. L. 1991. An evolutionary basis for conservation strategies. Science 253: 750–752.

HENNIG, W. 1966. Phylogenetic Systematics. Univ. of Illinois Press, Urbana.

KIRSCH, J. A. W. AND M. ARCHER. Polythetic cladistics, or, when parsimony's not enough: the relationships of carnivorous marsupials 1982. *In:* M. Archer (ed). Carnivorous Marsupials. Royal Zoological Society New South Wales: Sydney, Australia.

NELSON, G. 1989a. Species and taxa: systematics and evolution. *In:* D. Otte and J. A. Endler (eds). Speciation and its Consequences. Sinauer Associates, Sunderland, Massachusetts.

NELSON, G. 1989b. Cladistics and evolutionary models. Cladistics 5: 275–289.

NIXON, K. C. AND Q. D. WHEELER. 1990. An amplification of the phylogenetic species concept. Cladistics 6: 211–223.

DE QUEIROZ, K. AND M. J. DONOGHUE. 1988. Phylogenetic systematics and the species problem. Cladistics 4: 317–338.

DE QUEIROZ, K. AND M. J. DONOGHUE. 1990a. Phylogenetic systematics or Nelson's version of cladistics? Cladistics 6: 61–75.

DE QUEIROZ, K. AND M. J. DONOGHUE. 1990b. Phylogenetic systematics and species revisited. Cladistics 6: 83–90.

RIDDLE, B. R. AND R. L. HONEYCUTT. 1990. Historical biogeography in north american arid regions: an approach using mitochondrial DNA phylogeny in grasshopper mice (genus Onychomys). Evolution. 44: 1–15.

VANE-WRIGHT, R. I., C. J. HUMPHRIES AND P. H. WILLIAMS. 1991. Biol. Conserv. 55: 235.

WHEELER, Q. D. AND K. C. NIXON. 1990. Another way of *looking at* the species problem: a reply to de Querioz and Donoghue. Cladistics 6: 77–81.