

Species realities and numbers in sexual vertebrates: Perspectives from an asexually transmitted genome

(speciation/phylogeny/sister species/comparative molecular evolution/phylogeography)

JOHN C. AVISE* AND DEETTE WALKER

Department of Genetics, University of Georgia, Athens, GA 30602-7223

Contributed by John C. Avise, December 11, 1998

ABSTRACT A literature review is conducted on the phylogenetic discontinuities in mtDNA sequences of 252 taxonomic species of vertebrates. About 140 of these species (56%) were subdivided clearly into two or more highly distinctive matrilineal phylogroups, the vast majority of which were localized geographically. However, only a small number (two to six) of salient phylogeographic subdivisions (those that stand out against mean within-group divergences) characterized individual species. A previous literature summary showed that vertebrate sister species and other congeners also usually have pronounced phylogenetic distinctions in mtDNA sequence. These observations, taken together, suggest that current taxonomic species often agree reasonably well in number (certainly within an order-of-magnitude) and composition with biotic entities registered in mtDNA genealogies alone. In other words, mtDNA data and traditional taxonomic assignments tend to converge on what therefore may be “real” biotic units in nature. All branches in mtDNA phylogenies are nonanastomose, connected strictly via historical genealogy. Thus, patterns of historical phylogenetic connection may be at least as important as contemporary reproductive relationships *per se* in accounting for microevolutionary unities and discontinuities in sexually reproducing vertebrates. Findings are discussed in the context of the biological and phylogenetic species concepts.

More than 60 years ago, Dobzhansky (1) wrote that “biological classification is simultaneously a man-made system of pigeon-holes devised for the pragmatic purpose of recording observations in a convenient manner and an acknowledgment of the fact of organic discontinuity.” How well do perceived compartments at the species level (the taxonomic assignments in present use) record “authentic” biological discontinuities in nature? One traditional approach in assessing species realities has been to ask whether different human societies perceive biotic units similarly. For example, Mayr (ref. 2; see also ref. 3) found that preliterate peoples of New Guinea had vernacular names for 136 of the 137 native birds recognized as separate species by academically trained Western zoologists. Similar conclusions were drawn regarding particular Amazonian tree species as recognized by native peoples and academic botanists (4). Such outcomes suggest that species perceptions in these taxonomic groups are culture-independent (5) and hence, perhaps, that the biotic units are “real.”

In this same spirit, we ask here whether biotic discontinuities as seen through the eyes of laboratory-based mitochondrial geneticists tend to bear resemblance in number and composition to the biological units currently recognized as taxonomic species. There are additional reasons for interest in the outcome. First, discontinuities might be evident in local biotas

(the nondimensional species perception) but may blur when geographic variation is taken into account. Molecular phylogeographic studies address this issue, because they explicitly analyze spatial variation (6, 7). Second, under the biological species concept (BSC), a sexual species usually is perceived as a reproductive community whose gene pool retains coherency primarily via the bonds of interbreeding and genetic exchange (1, 8); however, mtDNA molecules are transmitted asexually, and matrilines are nonreticulate. Thus, any genuine unities within (and discontinuities between) groups of organisms in mtDNA genotype cannot be attributed to “horizontal” patterns of contemporary lineage anastomosis via mating *per se*. Instead, they must be caused by “vertical” connections (and partitions) in matrilineal phylogenies. However, vertical connections themselves are functions of the demographic histories of population units demarcated by temporally extended patterns of interbreeding and gene flow.

MATERIALS AND METHODS

This report is based on recent summaries of the primary mtDNA literature for vertebrates dealing with intraspecific phylogeographic patterns (9, 10) and interspecific genetic distances (11). These reviews, which should be consulted for details, drew from primary studies that assayed more than 150 bp of mtDNA sequence per individual and, in the case of phylogeographic analyses, dealt typically with multiple samples from widely spaced localities across significant portions of a species’ range. Reptiles and amphibians were poorly represented in the literature relative to mammals, birds, and fishes; as such, these two vertebrate classes are pooled in the presentations that follow.

The original mtDNA phylogeographic studies at the intraspecific level employed a variety of assay procedures, including direct sequencing of particular genes and restriction fragment length polymorphism analyses of the whole genome or particular loci. Estimates of sequence divergence (where presented) and phylogenetic appraisals of mtDNA haplotypes (as summarized by a wide variety of tree-building algorithms) were taken directly from the primary papers. To avoid a focus on unduly “shallow” mtDNA clades that are of little interest in the current context, studies that monitored the rapidly evolving control region (12–14) were disregarded, except where ancient and explicit divergence dates from control-region sequences were proposed in the original publications.

A literature review on interspecific distances (11) summarized levels of mtDNA sequence divergence between sister species, other congeners, and confamilial genera across the vertebrates. Estimates of mtDNA sequence divergence, calculated with Kimura’s two-parameter method (15), were derived

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: BSC, biological species concept; PSC, phylogenetic species concept.

*To whom reprint requests should be addressed. e-mail: avise@arches.uga.edu.

from comparisons of a total of 1,832 cytochrome b (*cytb*) gene sequences (1 per species) retrieved from GenBank.

RESULTS

This review encompasses intraspecific phylogeographic reports on a total of 252 vertebrate species that met the assay criteria listed above. Of these species, 140 (56%) were subdivided clearly into two or more highly distinctive matrilineal phylogroups, as gauged, typically, by relatively large genetic gaps between respective branches that received strong bootstrap support in an estimated mtDNA gene tree. Typical examples involving a mammal, bird, turtle, and two species of fish are illustrated in Fig. 1, and numerous other such cases are pictured or described in refs. 7 and 16. The remaining 112 species (44%) displayed either relatively shallow or no significant phylogeographic structure in the available mtDNA assays.

Within the 140 species that were strongly sundered genealogically, the principal intraspecific phylogroups nearly always displayed a strong geographic orientation. Indeed, 93% of these 140 species conformed to phylogeographic category I (6), in which distinct mtDNA genealogical assemblages are grouped geographically. The species that were subdivided strongly in mtDNA phylogeny usually showed only two principal phylogroups apiece, although some were separated into as many as six deep matrilineal assemblages (Fig. 2). An example of a species with four major phylogeographic units is depicted in Fig. 3.

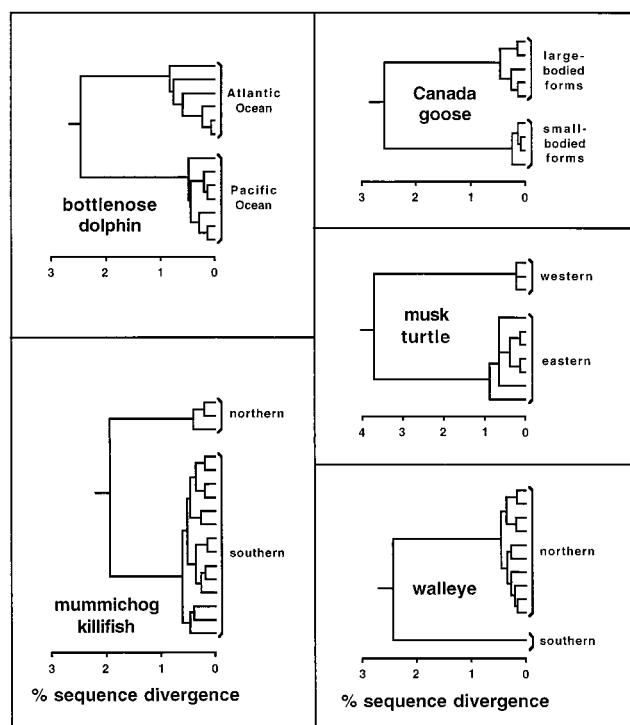


FIG. 1. Examples of a common situation (“phylogeographic category I”; ref. 6) in which two distinctive, geographically oriented mtDNA phylogroups have been reported within a vertebrate species. Shown are cluster phenograms for the bottlenose dolphin (*Tursiops truncatus*; patterned after data presented in ref. 37), the Canada goose (*Branta canadensis*; patterned after data presented in ref. 38), the musk turtle (*Sternotherus minor*; patterned after data presented in refs. 39 and 40), the killifish (*Fundulus heteroclitus*; patterned after data presented in ref. 41), and the walleye fish (*Stizostedion vitreum*; patterned after data presented in ref. 42; see also ref. 44 for a recent update).

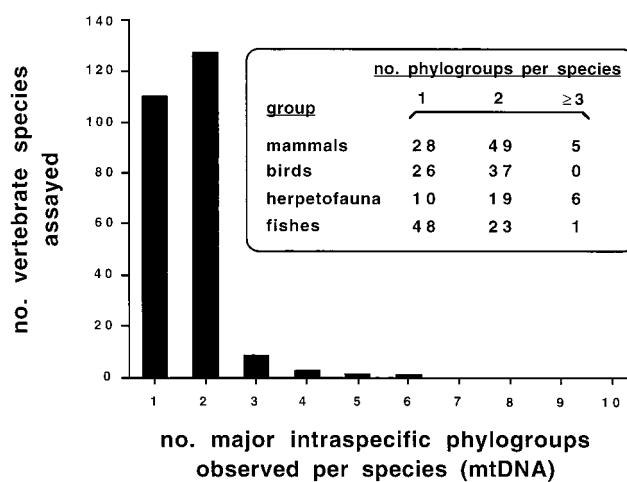


FIG. 2. Histogram of the number of distinctive phylogeographic units per species in mtDNA surveys conducted on more than 250 vertebrates. (Inset) Data are broken down by taxonomic group.

In most cases (as exemplified in Fig. 1), the identification of deep mtDNA phylogroups and the assignment of a species to phylogeographic category I were rather unambiguous. However, because of a great heterogeneity in the methods of data presentation and/or phylogenetic analysis in the original papers, definitive universal criteria for phylogroup recognition are difficult to specify. Typically, salient clades interpreted as intraspecific phylogroups were distinguished consistently by at least 0.6% sequence divergence in mtDNA, whereas mean within-clade sequence divergences normally were less than ≈0.2%. However, in our interpretations, of greater concern were visually evident discontinuities in the mtDNA phylogenies (or distance matrices), even when a few mtDNA haplotypes within a phylogroup exceeded 0.6% (see Fig. 1 for several such examples). Even if the stated numbers of salient mtDNA phylogroups recognized were to be increased 2-fold or 3-fold (under alternative interpretations that might be possible for some of the phylogenies, such as those in Fig. 1), the primary conclusions of this review would remain essentially unchanged.

DISCUSSION

Biotic Discontinuities. The striking empirical finding noted here is that most taxonomically recognized vertebrate species surveyed to date are subdivided into at most only a small number of highly distinctive intraspecific mtDNA phylogroups. Furthermore, by this same mtDNA yardstick, nearly all taxonomic congeners are distinguished by conspicuous phy-

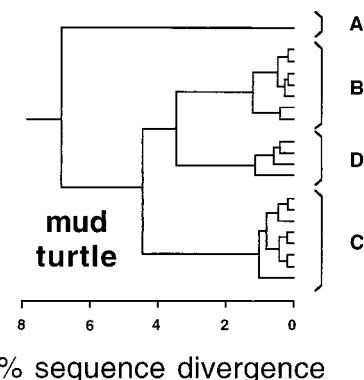


FIG. 3. Example of a situation in which four distinctive, geographically oriented mtDNA phylogroups were reported, in this case within the mud turtle (*Kinosternon subrubrum*; patterned after data presented in refs. 40 and 43).

genetic gaps (notably, about 90% of putative sister species show mtDNA sequence divergences greater than 2%; ref. 11). Taken together, these observations suggest that, for the vertebrates, phylogenetic units as demarcated by evident matrilineal disjunctions tend to agree reasonably well in number and composition to the traditionally perceived biotic discontinuities registered as taxonomic species. If these results can be generalized, then the number of vertebrate species recognized in current taxonomies is fairly close (certainly within an order-of-magnitude) to the number of phylogenetically highly distinctive matrilineal units in the biotic world.

In conducting such tests and interpreting current evidence, opposing biases arise. In some of the original molecular surveys, limited geographic and genomic sampling may have resulted in substantial underestimates of the number of principal intraspecific phylogroups actually present within a taxonomic species. Thus, further empirical study of these taxa may identify additional mtDNA phylogroups not yet recognized in the available material. On the other hand, most of the species included in this review were sampled across major portions of their respective ranges. Furthermore, they often were chosen by the authors for phyleogeographic analysis, precisely because they have exceptionally large geographic ranges whose histories of occupancy were of special interest. If broadly distributed as opposed to locally endemic species are especially prone to consist of multiple historical units (as seems likely), then the estimated numbers of phylogroups per taxonomic species may be biased upward in the literature reviewed here. Also, some major mtDNA phylogroups (particularly in species with male-biased dispersal) may well not register genome-wide population subdivisions. For this reason also, the current number of salient intraspecific units as estimated by mtDNA data could be biased upward. Although we do not know the relative magnitudes of these potential biases, the fact that they are in opposition (and hence should partially cancel one another) suggests that the net outcome might not depart dramatically from the general picture painted in Fig. 2.

In any event, the main conclusion from this survey of the available literature is that a general qualitative agreement seems to exist between the number of conspicuous biotic partitions as indicated in mtDNA lineages and those registered in existing species taxonomies. This compatibility of outcomes probably reflects an underlying historical reality to many of the biotic units traditionally recognized as taxonomic species. This overview is not intended to supersede detailed appraisals of possible agreement between taxonomic and molecular boundaries in particular taxa. However, it does suggest from a broad perspective that molecular and traditional organismal assessments of biotic diversity at the species level yield roughly similar outcomes.

Under the current analysis, this conclusion is most relevant to well studied vertebrates in temperate regions, where molecular phyleogeographic efforts thus far have been concentrated. Outcomes might differ under other circumstances, such as among various invertebrates or in tropical biotas. Wake (17) suggested that patterns of phyleogeographic differentiation in several of the temperate faunas surveyed might be caused by Pleistocene-mediated range restrictions and population extinctions that "sharpened borders between groups of populations and heightened the genetic cohesion of units." To test more broadly for any general agreement between mtDNA phyleogeographic discontinuities and current taxonomic boundaries between species, similar molecular phyleogeographic summaries should be extended to invertebrates and to organisms inhabiting other geographic regions such as the tropics. Perhaps, for example, previously unrecognized sibling species will prove to be far more common within some invertebrate groups (18).

Phylogenetic vs. Reproductive Criteria in Species Recognition. The current summary suggests that molecular phyloge-

neticists can join New Guinean and Amazonian natives as well as traditional academic systematists with respect to generally shared perceptions of biological discontinuities in nature. What the mtDNA perspectives add is an important emphasis on the historical (i.e., phylegetic) aspect of biotic pattern. In other words, phylegetic bridges and chasms registered in extant mtDNA genealogies presumably describe recent connections and relatively ancient gaps, respectively, among surviving matrilines through extended pedigrees. Thus, for sexually reproducing organisms, patterns of historical, vertical phylegetic connection (coalescence; refs. 19 and 20) may be at least as important as contemporary, horizontal reproductive relationships *per se* in accounting for the unities and separations of the living world.

Coalescent patterns in gene trees are related intimately to historical patterns in population demography (7, 21, 22). In particular, tight connections among nonanastomose genotypes suggest recent lineage coalescence to a shared ancestor, likely because of relatively small evolutionary effective population sizes that cause extant lineages to have shallow temporal depth. Conversely, large genetic gaps between gene-tree branches suggest long-standing historical population separations. In support of this likelihood, nearly all of the deep phylegetic disjunctions registered in the intraspecific mtDNA gene trees in this review involved regionally separate populations.

Recently, many authors have called for a complete abandonment of the BSC and its replacement with a phylegetic species concept (PSC) that emphasizes historical and genealogical aspects of biodiversity (refs. 23–32; reviewed in ref. 33). Although some of the original PSC formulations have serious difficulties in implementing species diagnosis (34), the idea of infusing species concepts with more explicit genealogical perspectives has considerable merit.

If (for sake of discussion) the major intraspecific mtDNA phylogroups discussed in the current report were to be elevated to full species status (as might be recommended under some versions of a PSC), then the resulting number of vertebrate species would be somewhat less than double the current number of recognized taxonomic species. Current species assignments, based primarily on morphological and behavioral appraisals, presumably have been erected under the general philosophical orientation of the BSC (which has dominated evolutionary biology throughout most of this century). Why might the BSC and at least some plausible versions of a PSC yield roughly similar estimates of the number of vertebrate species?

In our opinion, the reason is that reproductive and phylegetic underpinnings of biodiversity are related concepts; they are far from mutually exclusive as explanations for biotic continuities and discontinuities in sexually reproducing organisms (7, 34). Long-standing reproductive barriers (extrinsic or intrinsic) enable deep genealogical differentiation via their demarcation of historical population units whose temporally extended demographies have direct impact on the divergent branching structure of a mtDNA gene tree. Conversely, substantial reproductive links within and among geographic populations help to define the temporally extended demographic units within which historical genealogical ties are likely to be relatively recent (compared with those between such units), particularly if evolutionary effective population sizes are fairly small (as usually seems to be the case; refs. 35 and 36). Thus, notwithstanding the current adversarial relationship between the BSC and PSC, reproductive and genealogical (i.e., phylegetic) underpinnings of biotic diversity in sexually reproducing organisms are intertwined intimately.

We thank Andrew DeWoody, Anthony Fiumera, Beth McCoy, Bill Nelson, Devon Pearse, and Christine Spencer for helpful comments on the manuscript. This work was supported by the Pew Fellowship Program and by funds from the University of Georgia.

1. Dobzhansky, T. (1937) *Genetics and the Origin of Species* (Columbia Univ. Press, New York).
2. Mayr, E. (1963) *Animal Species and Evolution* (Harvard Univ. Press, Cambridge, MA).
3. Diamond, J. D. (1966) *Science* **151**, 1102–1104.
4. Pires, J. M., Dobzhansky, T. & Black, G. A. (1953) *Bot. Gaz.* **114**, 467–477.
5. Coyne, J. A. (1994) *Evolution* **48**, 19–30.
6. Avise, J. C., Arnold, J., Ball, R. M., Jr., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. & Saunders, N. C. (1987) *Annu. Rev. Ecol. Syst.* **18**, 489–522.
7. Avise, J. C. (1999) *Phylogeography: The History and Formation of Species* (Harvard Univ. Press, Cambridge, MA), in press.
8. Mayr, E. (1940) *Am. Nat.* **74**, 249–278.
9. Avise, J. C. & Walker, D. (1998) *Proc. R. Soc. London Ser. B* **265**, 457–463.
10. Avise, J. C., Walker, D. & Johns, G. (1998) *Proc. R. Soc. London Ser. B* **265**, 1707–1712.
11. Johns, G. & Avise, J. C. (1998) *Mol. Biol. Evol.* **15**, 1481–1490.
12. Taberlet, P. (1996) in *Molecular Genetic Approaches in Conservation*, eds. Smith, T. B. & Wayne, R. K. (Oxford Univ. Press, Oxford), pp. 125–142.
13. Baker, A. J. & Marshall, H. D. (1997) in *Avian Molecular Evolution and Systematics*, ed. Mindell, D. P. (Academic, New York), pp. 51–82.
14. McMillan, W. O. & Palumbi, S. R. (1997) *J. Mol. Evol.* **45**, 473–484.
15. Kimura, M. (1980) *J. Mol. Evol.* **16**, 111–120.
16. Avise, J. C. (1994) *Molecular Markers, Natural History and Evolution* (Chapman & Hall, New York).
17. Wake, D. B. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 7761–7767.
18. Knowlton, N. (1993) *Annu. Rev. Ecol. Syst.* **24**, 189–216.
19. Hudson, R. R. (1990) *Oxford Surv. Evol. Biol.* **7**, 1–44.
20. Harding, R. M. (1996) in *New Uses for New Phylogenies*, eds. Harvey, P. H., Leigh Brown, A. J., Maynard Smith, J. & Nee, S. (Oxford Univ. Press, Oxford), pp. 15–22.
21. Avise, J. C., Neigel, J. E. & Arnold, J. (1984) *J. Mol. Evol.* **20**, 99–105.
22. Harvey, P. H., Leigh Brown, A. J., Maynard Smith, J. & Nee, S., eds. (1996) *New Uses for New Phylogenies* (Oxford Univ. Press, Oxford).
23. Rosen, D. E. (1979) *Bull. Am. Mus. Nat. Hist.* **162**, 267–376.
24. Nelson, G. J. & Platnick, N. I. (1981) *Systematics and Biogeography: Cladistics and Vicariance* (Columbia Univ. Press, New York).
25. Cracraft, J. (1983) *Curr. Ornithol.* **1**, 159–187.
26. Donoghue, M. J. (1985) *Bryologist* **88**, 172–181.
27. Mishler, B. D. & Brandon, R. N. (1987) *Biol. Philos.* **2**, 397–414.
28. de Queiroz, K. & Donoghue, M. J. (1988) *Cladistics* **4**, 317–338.
29. McKittrick, M. C. & Zink, R. M. (1988) *Condor* **90**, 1–14.
30. Frost, D. R. & Hillis, D. M. (1990) *Herpetologica* **46**, 87–104.
31. Wheeler, Q. D. & Nixon, K. C. (1990) *Cladistics* **6**, 77–81.
32. Mallet, J. (1995) *Trends Ecol. Evol.* **10**, 294–299.
33. Claridge, M. F., Dawah, H. A. & Wilson, M. R. (1997) *Species: The Units of Biodiversity* (Chapman & Hall, New York).
34. Avise, J. C. & Wollenberg, K. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 7748–7755.
35. Avise, J. C., Ball, R. M., Jr., & Arnold, J. (1988) *Mol. Biol. Evol.* **5**, 331–344.
36. Bucklin, A. & Wiebe, P. H. (1998) *J. Heredity* **89**, 383–392.
37. Dowling, T. E. & Brown, W. B. (1993) *Mar. Mamm. Sci.* **9**, 138–155.
38. Van Wagner, C. E. & Baker, A. J. (1990) *J. Mol. Evol.* **31**, 373–382.
39. Walker, D., Burke, V. J., Barák, I. & Avise, J. C. (1995) *Mol. Ecol.* **4**, 365–373.
40. Walker, D. & Avise, J. C. (1998) *Annu. Rev. Ecol. Syst.* **29**, 23–58.
41. González-Villaseñor, L. I. & Powers, D. A. (1990) *Evolution* **44**, 27–37.
42. Billington, N. & Strange, R. M. (1995) *Trans. Am. Fish. Soc.* **124**, 770–776.
43. Walker, D., Moler, P. E., Buhlmann, K. A. & Avise, J. C. (1998) *Herpetologica* **54**, 174–184.
44. Stepien, C. A. & Faber, J. E. (1998) *Mol. Ecol.* **7**, 1757–1769.