

# Paraphyletic species

Michael D. Crisp and Greg T. Chandler

## Abstract

Crisp, Michael D. and Chandler, Gregory T. (Division of Botany and Zoology, The Australian National University, Canberra ACT 0200, Australia) 1996. *Paraphyletic species* *Telopea* 6(4): 813–844. We present evidence, mainly from plants, that many recognised species and subspecies are paraphyletic. Whilst some cladists have argued that species are like other taxa, and should be monophyletic, it is clear that even cladists either implicitly or explicitly recognise non-monophyletic species. Moreover, species concepts such as the phylogenetic species concept and the composite species concept predict non-monophyly of many species. Whenever a monophyletic species is circumscribed, it is possible that a paraphyletic or metaphyletic 'residual' species is simultaneously recognised. Furthermore, attempts to place all organisms in a monophyletic taxon at every rank regress to the population level where monophyly is no longer applicable, leaving paraphyletic residuals. These groups of organisms can hardly be ignored, unless one wishes to define them out of existence (as in the monophyletic species concept). It has been argued that paraphyly is only an artifact of the Linnean system, which requires all organisms to be classified in certain ranks, e.g. species. However the phenomenon of regress shows that this is incorrect, because paraphyly is inherent in species. The solution to this conundrum is to recognise species as special taxa, which may be monophyletic or paraphyletic. (Higher taxa should always be monophyletic, and can be made so.) This requires the acceptance of a species concept that allows paraphyly, such as the phylogenetic species concept or the composite species concept. The monophyletic species concept, which does not allow paraphyly, is not acceptable. The special nature of species derives from their basal position in the phylogenetic system. Theoretically, the proportion of paraphyletic and metaphyletic species may be 50% or higher. Empirical estimates range from 20% to 50%. Use of non-monophyletic species in historical applications such as biogeography is widespread but may not be invalid, depending upon the assumptions made.

## Introduction

In recent years, systematists have sought a species concept that is compatible with a phylogenetic system. They have rejected the biological species concept because of its reliance on the single criterion of reproduction. Entities which are distinct in many evolutionary, biological, and ecological features are nevertheless capable of interbreeding (Endler 1989, pp. 629–30). The biological species concept has never dealt satisfactorily with the conundrum of potentially (but not actually) interbreeding allopatric populations. Above all, the biological species concept is based on contemporary micro-evolutionary processes and cannot be reconciled with a phylogenetic system, in which taxa are viewed as historical units, extended in time and the units of a nested hierarchy (Rosen 1979; Donoghue 1985; Cracraft 1989; Vrana & Wheeler 1992; Frost & Kluge 1995).

## Species as lineages

Systematists have debated whether species should be viewed as lineages or taxa (Nelson 1989b; Rieppel 1994; Frost & Kluge 1995). Recent views of species as lineages hark back to a model presented by Hennig (1966: fig. 6), showing a lineage of sexually reproducing organisms splitting into two daughter lineages. Each lineage is

made internally cohesive by reticulating ('tokogenetic') relationships among its component organisms, but no such connections exist between lineages — they are mutually exclusive. Species are the internodes of a phylogenetic tree and speciation is the point at which lineages split permanently. Hennig's model has been reproduced many times, with modifications to show details or complications such as temporary versus permanent splits, reticulation and extinction (Kornet 1993a; Kornet 1993b; O'Hara 1993; Frost & Kluge 1995; Graybeal 1995). Most importantly, a lineage species is a model of evolutionary process. It is viewed as a real entity that evolves in time and space, has a definite beginning and end, and may be the ancestor of lineages comprising one or many species. It has been called the 'evolutionary species concept' (Wiley 1981; Frost & Kluge 1995) and the 'internodal species concept' (Nixon & Wheeler 1990; Kornet 1993a).

Some authors have been preoccupied with 'exclusivity' of lineage species (Donoghue 1985; de Queiroz & Donoghue 1988; de Queiroz & Donoghue 1990a; de Queiroz & Donoghue 1990b; Baum 1992; Baum & Shaw 1995; Graybeal 1995). (This is often called 'monophyly' but strictly monophyly refers to a taxon diagnosed by an autapomorphy.) A lineage is exclusive only if all its members are more closely related to one another (by ancestry) than to any member of another lineage. For example, the lineage of descendants of Queen Victoria is not exclusive because some descendants are more closely related to members of other lineages (by marriage, e.g. cousins). This discussion can reduce to the absurd because any lineage may be shown to be non-exclusive if examined minutely enough, even body cells (Frost & Kluge 1995). In her formalisation of a lineage concept of species, Kornet (1993a) shows that internodal species are *mutually exclusive* partitions of the genealogical network. Whilst this is a different notion of exclusivity from that discussed above, Kornet shows the latter problem to be irrelevant by using descent rather than ancestry as the criterion of group membership.

### Species as taxa

A major problem with species conceived as lineages is that they have poor empirical content (Kornet 1993a). When we find two allopatric populations that are essentially similar, we have no way of judging their fate — whether they are the basis of new, historically separate lineages, whether either will become extinct, or whether they will reunite and become reproductively, tokogenetically cohesive again. Therefore, systematists have also proposed concepts of species that have an empirical basis. In this view, species are part of a pattern of similarity among organisms: the hierarchy of internested groups that are called taxa (Nelson & Platnick 1981; Nelson 1989b; Rieppel 1994). The internested groups or taxa are recognised by shared similarity in characters, known as synapomorphy or homology. This hierarchy is represented as a tree (cladogram or phylogeny), but it is an abstract representation of pattern. The branches of the tree represent taxonomic groups which are internested, static and do not evolve. Thus the stem at the base of the angiosperms represents not the ancestral species of all angiosperms, but the most inclusive set of all taxa that we call angiosperms, recognised by the set of characters that all angiosperms share, and marked on the stem. Rieppel (1994) suggests that species conceived as lineages and species conceived as taxa are 'complementary but incompatible'. (Frost & Kluge (1995) refer to this distinction as the 'scalar' hierarchy versus the 'specification' hierarchy.) If taxa are also considered to be ancestors and descendants, then we are confronted with a paradox (Nelson 1989b). For example, does the subordination of the angiosperms to the seed plants imply that the seed plants are the ancestors of the angiosperms? Surely not, because the angiosperms are also a part of seed plants,

and angiosperms are not ancestors of themselves, any more than I am part of my grandfather. However, it cannot be disputed that some member of the seed plants evolved into the first angiosperm. The resolution of this paradox is to recognise taxa as units of an hierarchical pattern, not as part of the evolutionary process. By logical extension, species belong to this hierarchy. This pattern, when reconstructed, may be used as a framework for hypotheses about the evolutionary process, e.g. that a seed plant with certain characteristics gave rise to the first angiosperm. Thus the role of ancestor is restricted to lineages and their subunits, such as individuals or populations (Rieppel 1994) or 'internodons' (Kornet 1993b). If species are treated as taxa, then they are not different in kind from higher taxa. They are simply the least inclusive units in the systematic hierarchy.

Recent concepts of species as phylogenetic taxa derive from Nelson & Platnick (1981: 12), who define species as 'the smallest diagnosable cluster of self-perpetuating organisms that have unique sets of characters'. This is known as the 'phylogenetic species concept', 'irreducible unit' or 'minimum diagnosable unit' (Cracraft 1989; Nixon & Wheeler 1990; Nixon & Wheeler 1992). However, a unique or diagnostic character may be either an apomorphy or a plesiomorphy, and a group diagnosed solely by the latter is not monophyletic, which is anathema to authors such as Nelson (1989a; 1989b). Such species are not simple internodal partitions of a phylogenetic tree. They 'survive' a speciation event in which an autapomorphic species branches off from the phylogenetic stem (Nixon & Wheeler 1992: fig. 4.7B). Contrast this with Hennig's (1966: fig. 6) methodological extinction of ancestral species at branch-points. Neither Cracraft nor Nixon and Wheeler confront the paraphyly issue, but instead imply that paraphyletic species are acceptable, provided that they manifest unique and fixed character combinations. Under the phylogenetic species concept, speciation is the point at which a lineage acquires an apomorphy, or more precisely when a new character is fixed (Nixon & Wheeler 1992: fig. 4.7). This is true even of species diagnosed by a plesiomorphy, because at some point earlier in history, every plesiomorphy was an apomorphy. A problem with the notion of an 'irreducible unit' is that it is prone to regress (cf. de Queiroz & Donoghue 1990b). Peripherally isolated populations in which trivial genetic characters can easily become fixed would be diagnosed as species, contrary to general practice. One solution to the paraphyly problem is the monophyletic species concept, under which species have at least one autapomorphy (Rosen 1979; Donoghue 1985; Nelson 1989a; Nelson 1989b). However, this concept is unsatisfactory because ultimately taxa are not necessarily divisible into monophyletic sister-taxa (Smith 1994b). The smallest autapomorphic unit may have as its sister-group an unresolved symplesiomorphic cluster of organisms. The autapomorphic species concept consigns these to limbo, outside any species, but they can scarcely be ignored.

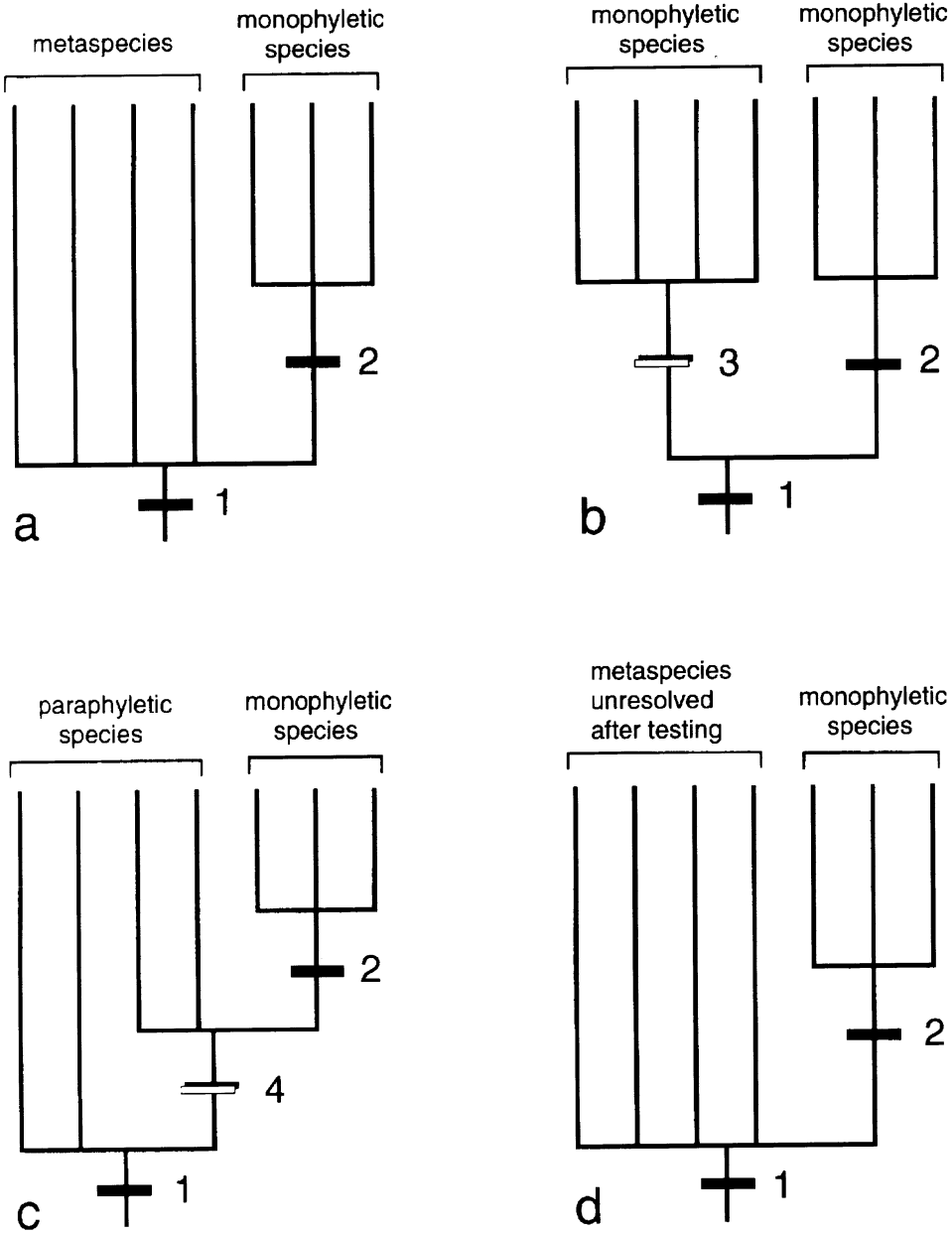
Some authors have taken the pragmatic view that phenetic clusters may be treated as species. This approach has been termed the 'phenetic species concept', although it is actually an empirical criterion, free of assumptions about evolutionary process. Such units have been termed 'phena' (Mayr 1969; Smith 1994b), to distinguish them from theoretically based 'species'. In fact, the phenetic species concept is the formal equivalent of the traditional 'taxonomic' or 'morphological' species concept, under which species are circumscribed by intuitively perceived similarity among their members (Sneath & Sokal 1973: 364–5). Sometimes this concept is conflated with the phylogenetic species concept; however, clustering by overall similarity is not the same as clustering by diagnostic (fixed) characters. Clusters in phenetic space may share no diagnostic character; usually they are circumscribed by a series of partially correlating (polythetic) characters. Nevertheless, some authors have argued that phenetic clusters may be equivalent in practice to phylogenetic species (Theriot 1992; Crisp & Weston 1993).

The composite species concept (Kornet 1993b), as its name implies, combines the lineage and taxon views of species. Kornet first formalises the internodal species concept (Kornet 1993a) then reveals a significant drawback with it: every isolated population is a potential new lineage and it can be made permanent by extinction. Thus, internodal species tend to be trivial units compared with those that are generally recognised as species, and are more akin to populations. Moreover, the internodal species concept is operationally intractable, because the fate of an isolated population cannot be determined. Instead, Kornet defines composite species as lineages of 'internodons' which begin with the fixation of a novel character in an ancestral internodon and end with another fixation in a descendant internodon (or extinction). Composite species are parts of lineages because internodons have ancestor–descendant relationships and are mutually exclusive. Because they are diagnosed by fixed novel characters (autapomorphies), they are also taxa and operationally equivalent to phylogenetic species. Thus the composite species concept seems to reconcile the tension between species-as-lineages and species-as-taxa (above). It should be noted that composite species are paraphyletic groups of internodons, unless they become extinct, in which case they become monophyletic (Kornet 1993b: 69).

### Paraphyly and metaphyly

Cladism has led to rejection of paraphyletic taxa on the grounds that they are not real phylogenetic units and lead to confusion about both the distribution of characters and the relationships of taxa (Donoghue & Cantino 1988; Humphries & Chappill 1988). Paraphyletic groups are considered 'artifactual' and qualitatively different from monophyletic groups, which are 'real' taxa (Nelson 1989b). For every monophyletic taxon recognised, any of a series of paraphyletic groups may be constructed by excluding the monophyletic taxon from more inclusive (higher-level) monophyletic groups. In this way, paraphyletic groups have been treated as taxa, diagnosed by symplesiomorphies or the absence of the autapomorphies which circumscribe the excluded monophyletic groups. When taxa are discovered to be paraphyletic, systematists are inclined to divide them into several more narrowly circumscribed, monophyletic taxa (monophyly can also be achieved by amalgamation). However, this process of division may regress to the species level, where a problem is encountered: species are not divisible into monophyletic subunits. Moreover, both the phylogenetic species concept and the related composite species concept predict that many, if not most, species are not monophyletic (above). Here is a conundrum: if species are to be considered taxa, logically the sanction against paraphyly should apply to them (Cracraft 1989; Nelson 1989a; Nelson 1989b).

Empirically, it has long been recognised that many accepted species are paraphyletic ('paraspecies': Ackery & Vane-Wright 1984). In a paraspecies, some (but not all) members are more closely related to members of another species than to other members of the paraspecies. Evidence for paraphyly would be a synapomorphy which some members of the paraspecies share with the other species (Fig. 1c). Some authors have pointed out that any species that lacks an evident autapomorphy is at least potentially paraphyletic; however, this is only an inference based on lack of evidence (it is also potentially monophyletic). The term 'metaspecies' has been coined (Donoghue 1985) to distinguish such species (whose phylogenetic status has not been resolved by cladistic analysis) from paraspecies (whose presumed monophyly has been tested and refuted). (Gauthier (1986) extends the metataxon concept to higher taxa but this is not relevant here.) Phylogenetic analysis of populations comprising a metaspecies may have one of three outcomes (cf. de Queiroz & Donoghue 1988: fig. 7): (i) a synapomorphy may be found for all populations, and



**Fig. 1.** Illustration of monophyletic species, paraspecies (paraphyletic) and metaspecies (unresolved). **a**, Initial phylogeny showing a metaspecies as sister-group to a monophyletic species. The metaspecies has no apomorphic characters except 1, which it shares with its sister species. The monophyletic species has an autapomorphy, character 2. **b-d**, Possible outcomes following cladistic analysis of populations in **a**. **b**, A new synapomorphy (character 3) is found for populations comprising the metaspecies, which is now recognised as a monophyletic species. **c**, A new synapomorphy (character 4) is found which is shared by two populations of the metaspecies and the monophyletic sister species. The original metaspecies is now recognised to be paraphyletic. **d**, No further apomorphies are found, and the metaspecies remains unresolved. These definitions apply equally to subspecies. Terminal branches represent populations. Solid bars represent original apomorphies; open bars represent additional apomorphies discovered following cladistic analysis.

the 'metaspecies' is shown to be monophyletic (Fig. 1b); (ii) a synapomorphy may be found (character 4, Fig. 1c) demonstrating that some populations of the metaspecies are more closely related to a recognised monophyletic species, in which case the 'metaspecies' is shown to be paraphyletic; (iii) no new apomorphy is found, and the species remains an unresolved metaspecies (Fig. 1d), diagnosed only by a symplesiomorphy (character 1, Fig. 1d). Both paraspecies and metaspecies are diagnosed by symplesiomorphy (character 1, respectively in Figs. 1c and 1d). However, they differ in that evidence exists to show that part of the paraspecies is more closely related to another species (character 4 in Fig. 1c), whereas no such evidence is found in a metaspecies (Fig. 1d). To summarise, depending upon the observed distribution of apomorphies among populations, the phylogenetic status of a species may be: unresolved (a metaspecies), non-monophyletic (a paraspecies) or monophyletic (an autapomorphic species). Note that irrespective of the phylogenetic relationship of their populations, all these species are diagnosable units consistent with the phylogenetic and composite species concepts. Therefore all are real, discoverable and corroborable entities. Moreover, the phylogenetic relationship of their parts (monophyletic, paraphyletic or metaphyletic) is also discoverable and corroborable (by the adducement of additional evidence).

## Objectives

In this paper, we present examples of paraspecies and metaspecies and empirically estimate their proportion of all species. We show that any attempt to purge the system of these is futile, because of the asymmetric distribution of apomorphic (relatively advanced) characters among basal lineages (species). Consequently we address the conundrum of paraspecies and metaspecies in a system to which these are anathema. Finally, we consider the implications for comparative methods such as cladistic biogeography of a false assumption of species monophyly.

For the purpose of this paper, we make no fundamental distinction between species and subspecies. This paper is concerned with lowest-level taxa, whether ranked as species or subspecies. The concepts monophyly, paraphyly and metaphyly apply equally to either, and to taxa of any rank. We do not consider the effects of reticulation, as this is a separate problem.

## Examples of paraspecies

The following five examples report cladistic analyses using as terminals either populations or geographic forms that do not have evident autapomorphies and may well be paraphyletic. Are these suitable units for cladistic analysis? Some authors suggest that using paraphyletic terminals invalidates phylogenetic analysis (Cracraft 1989; de Queiroz & Donoghue 1990a; de Queiroz & Donoghue 1990b; Nixon & Wheeler 1990; Wheeler & Nixon 1990; Vrana & Wheeler 1992). Moreover, because evolution among populations is likely to be reticulate, the strictly hierarchical model of cladistic relationships may be invalidated (Crisp & Weston 1993). However, these problems are not restricted to populations: the monophyly of most taxa (e.g. species and subspecies) is untested and thus uncertain (Nelson 1989b). Moreover, this paper shows that many such taxa are probably paraphyletic. An extensive literature attests to the frequency of reticulate evolution among recognised species (e.g., Funk 1985; Barton & Hewitt 1989; Harrison 1991; Arnold 1992; Grant & Grant 1992; Smith 1992). Thus, problems affecting cladistic analysis of populations apply at least in part to subspecies and species. Vrana & Wheeler (1992) advocate using as terminals individual organisms, whose monophyly can (perhaps) be safely presumed. However,

