

Synbiologia Hungarica 6:1-42, 2010

Taxonomy in Evolutionary Perspective

An essay on the relationships between taxonomy and evolutionary theory

by

János Podani

Department of Plant Taxonomy and Ecology, Biological Institute, Eötvös University,
Pázmány P. s. 1/c, H-1117 Budapest, Hungary. E-mail: podani@ludens.elte.hu

© The author

ISSN 1219-9370

Electronically published June 7, 2010

Scientia, Budapest

Abstract. Evolution is the driving force behind biological diversity on Earth, therefore the taxonomic practice must adapt itself to the notion of change. In this context, the paper presents an overview of some unresolved problems in phylogenetic taxonomy and gives a road map to appreciation, interpretation and proper use of the main taxonomic systems. For organizing knowledge on biodiversity, we have a choice between synchronous and diachronous strategies. *Synchronous* classifications are concerned with a given time slice of the phylogeny, i.e., the result of evolution. Cladograms are synchronous tree representations of evolutionary history by showing sister-group relationships; they can be converted to dendrograms with some loss of information and therefore to a Linnaean hierarchy of extant taxa as well. In this hierarchy, morphological and genetic gaps increase along with increasing ranks, which is in agreement with requirements for a meaningful supraspecific classification. According to a new terminology, a taxon in a synchronous classification can be monoclادistic, paraclادistic or polyclادistic, with respect to the relative position of the members of this group in a reference cladogram. Monoclادy refers to the condition that the given taxon includes all and only extant descendants of a given hypothetical common ancestor. This criterion appears particularly relevant to contemporary molecular taxonomic surveys. The binominal nomenclature is still useful for naming extant organisms provided that the bacterial, botanical and zoological codes are harmonized. A *diachronous* evolutionary classification which includes all extinct and living organisms is based on segments delineated along the phylogenetic tree. Conceptually, the Linnaean hierarchy of all life is incompatible with the segmentation of this tree, because wide gaps necessary for separating supraspecific taxa are evolutionary absurdities in the spatiotemporal continuum of populations, given the Darwinian gradualist model of speciation. Higher taxa, if projected from our recent knowledge back into the past, can at best be fuzzy sets. Theoretically, the phylogenetic tree remains the only meaningful representation of ancestor/descendant relationships, i.e., the diachronous pattern of life, but its full reconstruction is absolutely impossible for obvious reasons. Monophyletic, paraphyletic and polyphyletic groups, if defined based on ancestors

and descendants, have relevance only in the context of phylogenetic trees. Paraphyly is a mathematical necessity in diachronous phylogenetic classifications. A practical significance is that the uninominal nomenclature in the Phylo-Code represents the only unambiguous and coherent way to name both extinct and living organisms and the branches of the phylogenetic tree. These issues are illustrated by artificial examples, whereas some consequences regarding the True Tree of Life and an assumed True Classification of Life are outlined. The main conclusion is that if phylogenetic taxonomy wishes to 1) maintain and improve the Linnaean hierarchy while 2) satisfying the Darwinian requirement that all descendants of an ancestor are kept together in every taxon, then 3) common ancestry is to be inferred by neo-Hennigian cladistic approaches, in order to 4) generate a synchronous (Crowsonian) classification of extant organisms.

Keywords. Cladogram, Diachronous classification, Monoclad, Monophyly, Paraclad, Phylogenetic tree, Synchronous classification, True Classification of Life, True Tree of Life.

Abbreviations. ADR–Ancestor Descendant Relationships; GSB–Great Scale of Being; GT–GradeTree; SGR–Sister Group Relationships; TCL–True Classification of Life; TTL–True Tree of Life.

1

Introduction

“The systematic structure of real phylogenies is cladistic (branching), irrespective of what principle of classifying taxa is adopted.”
(Griffiths 1974)

In a recent commentary, Brummitt (2008) summarizes the current debate on “evolutionary taxonomy” versus cladistics published in the past decades in the taxonomic-phylogenetic literature. The title of his short paper (“*Evolution in taxonomic perspective*”) apparently prefers taxonomy (i.e., classification of organisms) to evolution, reflecting literally the opinion that it is evolution that can be, and has to be viewed in the context of taxonomy. This gives me an opportunity to scrutinize the implications of this attitude because I think that a few issues regarding the relationships between biological taxonomy and evolutionary theory require clarification and deserve more attention than before¹.

Biologists agree that evolution is the one process that produces biological diversity and no question that it is a natural phenomenon that operated and continues to operate independently of the human observer. Accumulating morphological, cytological, genetic and molecular evidence demonstrates convergence of evolutionary information into an increasingly stable tree² (in part, a network), an approximation to the *True Tree of Life* (TTL) although we are still very-very far from a consensus³. Since evolution is associated to a unique pathway connecting populations in the spatio-temporal continuum, the TTL must be a unique summary of genealogical relationships (cf. “The one true tree of life”, Dawkins

1 In this paper, I will neither discuss nor criticize the transformed/pattern cladistic approach to taxonomy in which evolutionary considerations, phylogeny and ancestor-descendant relationships play much less role, if any. The Reader is referred to the literature on this subject matter, e.g., Dawkins (1986), Scott-Ram (1990), Panchen (1992), Ereshefsky (2001), Williams and Ebach (2007) and Podani (2010).

2 The Appendix summarizes some important terms used and proposed in this paper. This is given with the intention to reduce some terminological heterogeneity (and confusion) concerning different types of trees and related concepts in contemporary taxonomy.

3 Nevertheless, progress may be very well appreciated by comparing two books written on the same subject, *The hierarchy of life* (Fernholm et al. 1989) and *Assembling the tree of life* (Cracraft and Donoghue 2004).

1986) even if we shall never be able to reconstruct it even to the coarsest details. This certainly is not the case with biological taxonomy. Several hundred years of history, including the somewhat futile dispute over monophyly/paraphyly (see Podani 2010) illustrate lucidly that there is no universally valid classification of organisms upon which biologists should and are able to agree and to which accumulating evidence would converge. This is just the other way around: we are witnessing increasing disagreement among classifications and especially theories and views that lie behind them. There are never-ending controversies about the basic units of taxonomic classification, the species, for which at least 26 different definitions have been suggested thus far (Wilkins 2006), not to mention the nominalist views according to which species are just abstractions of our brain (“A species in my opinion is a name given to a group of organisms for convenience, and indeed of necessity”, Haldane 1956). Consequently, higher taxonomic categories are subject to even more serious criticism by theoreticians; they are a source of deeper conflicts of view among biologists and are victims of inconsistency over the different kingdoms. The disagreement upon any argument suggesting the existence of a single, universal classification is extremely strong – except perhaps the requirement of “naturalness”, which can be defined in many ways – thus complicating the issue even further. Thus, by paraphrasing Dobzhansky’s (1973) famous motto, *it is taxonomy that can, and has to be examined in the light of evolution, and not vice versa*. It is therefore imperative to overview our chances for achieving the final and ideal goal of many taxonomists, the supposed *True Classification of Life* (TCL), given our knowledge on biological evolution. The question is whether it is possible to continue and complete the work of Linnaeus, who was convinced that his classification system was a discovery of what God created and who considered himself as being the most appropriate person to make that discovery (Blunt 2002). I would like to express my contention that such a goal is illusory and almost impossible logically for several reasons. I am aware of the fact, of course, that thousands of papers and dozens of books have touched upon this problem already, from the perspective of various disciplines of biology, philosophy and mathematics. Admittedly, some of the points below may not be new to some readers although meant to be new, which only reflects my inability to cope with this enormous literature.

Each of the following four main sections focuses on an important aspect of phylogenetic taxonomy. First, I use artificial examples to demonstrate that the Linnaean hierarchy can only be meaningful for organisms living at a given point of time (*The Boundary Paradox*). The second part is concerned with the compatibility of the three basic types of tree-graphs used in evolutionary biology with

the Linnaean classification (*The Hierarchy Conflict*). Then, new light is shed onto the classical problem of monophyly/paraphyly, because I suggest that these conditions should be treated separately for phylogenetic trees and cladograms. This leads to a new terminology which hopefully clarifies some theoretical conflicts persisting in systematics, especially between “evolutionary taxonomy” and phylogenetic cladistics (*Monophyly vs Paraphyly – Or Something Else?*). Finally, some nomenclatural peculiarities and the way they confuse evolutionary biologists are discussed (*Nomenclatural Comments*). The paper is concluded with an Appendix which gives precise definitions of terms used in this essay.

2

The Boundary Paradox

“Among existing organisms there are real and evident discontinuities ... but if past as well as Recent organisms are included, these discontinuities will vanish.”
(Crowson 1970)

One of the fundamental sources of misunderstanding and disagreement between different schools of taxonomy is that, contrary to statements frequently encountered in taxonomic papers (like “family A evolved from family B”), supra-specific taxa are *not* evolvable entities. For example, Wiley (1976) wrote that

„...higher taxa do not evolve, they cannot be considered ancestral units. The reality of higher taxa is based solely upon whether they are accurate reflections of past speciation events and thus higher taxa are historical constructs and have no reality as active units of evolution”

Hörandl (2006) said that “...evolution is constituted and continued by individuals and populations”. Members of populations can interbreed, they produce offspring with or without genetic information recombined, multiplied and possibly mutated as compared to their parents, populations may become isolated in the topographic or ecological space, and they are subject to selection and extinction. The genealogy of populations may be summarized by phylogenetic trees, the most classical example of which was drawn by Darwin (1859) in order to illustrate his concepts on *speciation*, the gradual process that gives rise to variations, and then new species and genera. Clearly, Darwin’s tree-like diagrams show Ancestor-Descendant Relationships between populations (ADR trees, Dayrat 2005). New forms and varieties may appear during evolution in other ways as well (e.g., allopolyploidy, hybridization, horizontal gene transfer and – last but not least – endosymbiogenesis) but in any case, we have a complex spatiotemporal continuum of populations in which abrupt changes are the exception rather than the rule⁴. However, taxonomic classifications imply *hard (crisp)*

4 The theory of punctuated equilibrium is in fact a special variant of Darwinian gradualism and the slight differences between the two models do not concern the validity of the conclusions to be drawn in the present paper. Saltations can be excluded easily on the grounds of population genetics (Dawkins 1986),

partitions, that is, each population is assigned to only one species, each species into only one genus, and so on. Boundaries are therefore meant to be sharp *by definition*. Proposals on overlapping classifications (Jardine and Sibson 1971) are only of historical interest now, whereas the concept of fuzzy species which is being seriously considered among bacteriologists and virologists (see e.g., van Regenmortel 1997) apparently escaped the attention of other taxonomists (a noted exception is Hall, 1997) and was not even listed by Wilkins (2006). It is well-known that taxonomic distinction is not always unequivocal between extant plant populations; recall the case of *Rosa*, *Ribes*, *Sorbus* or *Hieracium*, just to mention a few genera which have many “problematic” (sibling) species. If we extend our interest to the past, we can get into more trouble with distinguishing species and especially higher taxa. The issue is more serious than previously thought, as illustrated by a simple mental experiment summarized in Figure 1. Let us start with the assumption that taxa are delimited using phenetic information – as usual – in three different points of time with intervals long enough to allow major evolutionary changes to happen. A taxonomist at Time 1 would recognize two species, A and B, differing in, say, two autapomorphic characters, and an imaginary cladist at the same time would recognize a simple dichotomy for two populations. Also, he would probably agree that they represent two species from the same genus (lowest box in the figure). Evolution continues and A and B become ancestors of several new species which develop many new character states by adaptation to changed environmental conditions. Speciation and anagenesis are so expressed that another imaginary taxonomist at Time 2 would recognize three species, A1 and A2 belonging to one genus, and B1 to another. A cladist would be able to reconstruct the phylogeny as shown in the figure up to Time 2. By examining extant populations at Time 3, an extant taxonomist describes three genera, two of them assigned to family X and the other to family Y. A cladist colleague reconstructs the phylogeny and characterizes both families as monophyletic, i.e., descendants of species A and B, respectively. Then, *according to the principles of taxonomic schools that wish to classify extant and extinct species together based on ancestry*, A and all of its descendants are assigned to the same family, just like B and its descendants to another. That is, based on information obtained at Time 3, *two extinct populations that differed at species level only at Time 1 are later separated at family level*. Obviously, our taxonomist at Time 1 did not put them into separate fami-

but see also my comments on the gymnosperm-angiosperm transition later. Other „models of evolution” (e.g., repeated extinctions due to catastrophic events followed by creations) should not even be mentioned here. Allopolyploids represent the only taxa that arise “instantaneously” (Cavalier-Smith 2010).

lies because he could not *predict* that divergence would be so strong: evolution is largely unpredictable. The taxonomic boundary (double line in Fig. 1) running between the two branches changed over time, first it was between species, then between genera and finally between families. This is a paradoxical situation arising clearly from the conflict between evolutionary and classificatory thinking, or more precisely, between notions of *change* and *steadiness*, respectively. This experiment is not entirely hypothetical; it can happen with a low probability (but see Foote 1996) that some fossils can be contrasted with recent populations from the same lineage to demonstrate this situation.

If the above train of thought is not false, we are faced with serious consequences as follows:

- Evolution overwrites classifications. Thus, the supposed TCL into *disjunct* supraspecific taxa can only be *synchronous* i.e., one referring to the living world at a given point of time. [In fact, many authors, including Darwin played with this idea by expressing theoretical and practical difficulties

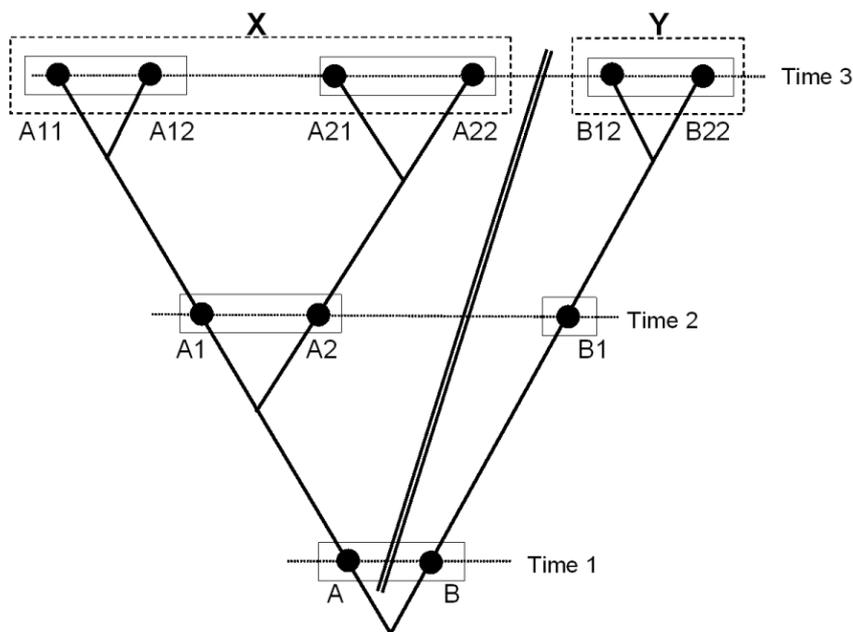


Figure 1. Hypothetical phylogenetic tree illustrating the boundary paradox. Black dots are populations distinguished by imaginary taxonomists as species at three different times (thin horizontal lines); other populations not shown. Small boxes correspond to synchronous genera, large dotted boxes to families. Double line separates *taxa* shown on the two major branches, which itself „evolves” from a species separator to a family separator.

with arranging extant and extinct taxa together into the same classification (see e.g., de Queiroz and Gauthier 1992 and references therein). Note also that Linnaeus placed fossils into the kingdom of rocks (Lapidaria) – although for reasons obviously other than those emphasized here.] It was a British entomologist, Crowson (1970) who suggested most explicitly that the taxonomy of fossils must be separated from the taxonomy of extant organisms.

- A single classification, as a static construct, cannot reflect faithfully the process of evolution, only the results of it. A long series of classifications could (see also Brundin 1966), a phylogenetic tree can. If a synchronous classification is to be contrasted with evolutionary history, then cladograms are the logical choice, because they are also synchronous representations of relationships among extant taxa (Sister Group Relationships, SGR trees, Dayrat 2005).⁵
- *Branches* of a phylogenetic or genealogical tree – as being representations of *diachronous* (= happening over time) phenomena – *cannot be conceived as taxa in a synchronous classification*. Ancestors and descendants can appear together only in a *diachronous classification*, which should be based on segmenting the phylogenetic tree. However, segmentation into higher taxa is difficult if not impossible because gaps between them are results of evolution, they are visible only at a given horizontal cross section of the tree and cannot be observed along the tree. According to Darwin, higher ranks reflect the absence of extinct, transitional forms descended from common ancestors (Padian 1999). As Crowson (1970) put it:

“Among existing organisms there are real and evident discontinuities which provide an objective basis for the divisions of our hierarchy – but if past as well as Recent organisms are included, these discontinuities will vanish.”

5 The distinction between phylogenetic (evolutionary) trees and cladograms is quite clear from Dayrat’s (2005) discussion, and I can only fully agree with that (see Appendix). Unfortunately, the taxonomic and phylogenetic literature is very eclectic and inconsistent in this regard: the majority of authors use these names interchangeably, whereas some theoreticians propose an unnecessarily restricted usage of terminology. For example, McManus (2009) suggests that cladograms are trees produced by maximum parsimony methods, while evolutionary trees are those derived from maximum likelihood analysis. As obvious from the Appendix, it is not the method that determines whether a tree is evolutionary or cladistic. For some (e.g., Scott-Ram, 1990), a “tree” implicitly means a phylogenetic tree, but obviously the latter is a special case of the former in graph theory terms (see Appendix). To find the background behind misunderstandings, we may go as far back as Hennig (1966) who did not distinguish clearly between phylogenetic trees and cladograms.

For him, discontinuities may be attributed to the scarcity of fossil record, “for which palaeontological systematists should be duly grateful”. According to Griffiths’ (1974) view:

“As a result of evolutionary processes of differentiation there are gaps in the array of objects at higher organization levels now existing on earth.”

- At best, species are the only entities with restricted spatiotemporal extension (for Hennig “species are chunks of the genealogical tree”, while some philosophers use the biologically misleading terminology that “species are individuals”, e.g., Ghiselin 2004), but fuzzy limits appear to have more general spatiotemporal validity in the living world than sharp boundaries, raising the possibility for a fuzzy classification (for maths, see Bezdek 1974) or fuzzy phylogeny reconstruction (Auyeung 2005).
- A Linnaean system of classification is still useful for arranging *extant* populations into taxa, for naming and cataloging them, while the PhyloCode (Cantino and DeQueiroz 2007, Dayrat et al. 2008) appears the only adequate – perhaps imperfect but promising – way for naming branches of both extant and extinct organisms on the phylogenetic tree. That is, if a species is considered from historical perspective, i.e., understood to comprise all of its component individuals that ever lived, we should apply only *uninomina*. It is because classification may change even at genus level within the lifetime of a given species.
- Evolutionarily minded biologists cannot say that “*Boraginaceae* occurred, say, 80 million years ago”, or that “*Angiospermatophyta* showed up at the end of the upper Jurassic”. These groups, all higher taxa, are generally defined on the basis of characters observed on extant organisms, and we do not know if the combination of defining characters was possessed at all by the last common ancestor population of any recent taxonomic group. Angiosperms are a good case in point: it is impossible that the carpel, double fertilization and the flower appeared simultaneously in a jump from one population to another. Indeed, the process probably lasted at least 100 millions years, as estimated by Stuessy (2004). Statements such as “the last common ancestor of populations comprising the extant family *Boraginaceae* diverged ca. 80 million years ago” or “the first fossil plant known to exhibit typical angiosperm characters is 130 million years old” are more concordant with evolutionary theory. I can refer to the beginning of this section: populations rather than taxa evolve, so that taxa

cannot “occur” anyway (no higher taxon can “bud” from another because these are not coherent entities, Ghiselin 2004).

One has to be careful with the time dimension, of course, otherwise the above conclusions may lead to absurdities. The time scale should be much coarser than what we can directly observe, but should be fine enough to reflect, in general, significant evolutionary change over time. I am afraid, however, that since the molecular clock, as well as the “morphological clock” differ with evolutionary lineages, no universal rules can be given for specifying the relevant time scale. Nevertheless, one can imagine a classification of life constructed separately for different epochs. For example, according to Crowson (1970, p. 67):

“Instead of trying to fit fossils into classifications devised for Recent plants and animals, we should have to construct a separate classification of each era of the geological past...The resulting overall system, however strange and cumbersome it might appear, would in effect have gained a new dimension and would partially escape the insoluble contradictions which confront the attempt to incorporate fossils and modern organisms in a single system.”⁶

This would readily solve the problem, for example, that early angiosperm-like plants (such as *Archaeofructus*, *Archaeanthus*, and *Caloda*) cannot be assigned to extant taxa of angiosperms, or that Devonian dichotomously branched plants have no place in classifications of extant pteridophytes. Epoch-level classifications would certainly involve more work by paleobiologists than the two systems suggested by Reif (2005; cit. in Hörandl 2007), one for extant and the other for both extant and extinct organisms.

There may still be readers who smell circularity in the above argumentation: I assumed that there were three taxonomies in different times and then showed that classifications are time dependent. However, I merely wanted to imitate the taxonomist’s practice, which would certainly produce different classifications when the objects to be classified are different (in number and in their properties as well), so the tautology is removed. For those still skeptical that *drawing boundaries is the key issue when a classification is contrasted with evolution*, two other mental experiments are recommended, in which the horizontal (spatial) and vertical (temporal) aspects are viewed separately.

6 Griffiths (1974), however, gives priority to Hennig’s historical definition by saying “Systematic presentations of organisms in a particular time horizon or section (geological period) are indeed valid, but of secondary importance.”

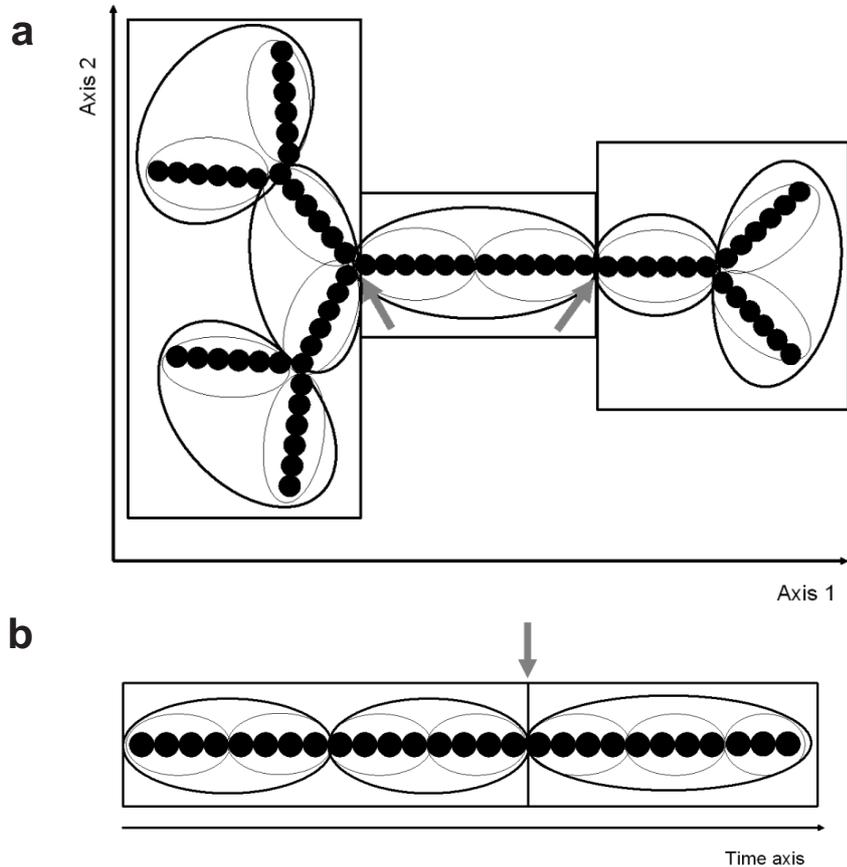


Figure 2. The boundary paradox for hypothetical populations (dots), species (small ellipses), arbitrary genera (large ellipses) and even more arbitrary families (boxes) in **a**: an hypothetical morphospace and **b**: anagenesis. Arrows point to boundaries where populations are “separated” at the family level.

Suppose first that all populations representing the lineages in Figure 1 are extant and are available for a taxonomist who wishes to set up an hierarchical classification. In this case, there is no temporal aspect of the study, nor phylogeny. This is purely hypothetical, contrary to the previous experiment, for obvious reasons. The populations can be arranged in a morphospace (ordination) as illustrated in Figure 2.a. Assume that the taxonomist is able to find minor characters that can be used to delineate species (small ellipses in the figure) in this continuum. At the same time, he concludes with good reason that some populations falling far apart from each other in the morphospace are sufficiently different to distinguish them at genus or even family level. However, in order to have a nested, Linnaean classification (an inclusive hierarchy, see below), genus lim-

its must coincide with certain species limits, family boundaries also with certain genus and in turn with species limits – and so on. The implications of nestedness are that there must be pairs of fairly similar populations *that are separated not only at species but also at higher taxonomic levels* and that *two populations from different families may be much closer to each other than two species from the same genus*. This is the result if an hierarchical classification is forced upon a spatial continuum of populations.

Similar conclusion can be drawn from the next hypothetical example in which only the temporal aspect is examined: a single population evolves from one species to the other and so on (anagenesis, Fig. 2.b). Let us accept the – otherwise disputed – condition that species are separable in time. However, if higher level taxa are superimposed over this single lineage of populations, then we shall encounter the same problem as in the morphospace. We end up with the absurd situation that the parent population is separated from the offspring at the family level. Such a paradox obtains if an hierarchical classification is superimposed forcefully upon a temporal continuum of populations. Foote (1996) has reached similar conclusions regarding identification of species, not higher taxa. As quoted by Dayrat (2005):

“...specimens that would be morphologically identical before and after a budding event would be named differently, whereas specimens that would be morphologically distinguishable in case of a phyletic transformation [anagenesis] would receive the same name”.

As said above, receiving different family assignment would be even more nonsensical.⁷

The three experiments discussed above are hopefully sufficient to illustrate one of the main points of this paper: a diachronous classification of the Linnaean type imposes an artefactual hierarchy onto populations in a spatio-temporal continuum⁸ and is therefore misleading. While separation between taxa in a hierarchical synchronous (i.e., Linnaean) classification increases with taxonomic ranks, this is just the other way in a diachronous classification. Separation is the best, albeit not too sharp, at the species level while the boundaries become fuzzier when the taxonomic ranks increase. For example, the morphological boundary between extant gymnosperms and angiosperms is wide (notwithstand-

7 After revising the manuscript several times, I came across a new book by Dawkins (2009) who illustrates brilliantly the problem with separating high taxa in time, using *Australopithecus* and *Homo* as examples.

8 As Hull (1979) notes, “the Linnaean hierarchy... lends itself more naturally to expressing discontinuous than continuous phenomena”.

ing gnetophytes), whereas the boundary between extinct gymnosperms and angiosperms was not sharp at all: if Stuessy's estimate is just 50% right, there must be a 50 million year "thick" boundary line between gymnosperms and angiosperms. What else would this be if not a fuzzy boundary, if boundary at all? Similar albeit shorter periods separate reptiles and birds, but many other examples could be listed by going as far back as the emergence of the cell from a stage when all life was a single gene pool (Woese 2000) with no group or tree structure at all⁹.

Now, I can try to answer a question asked by Brummitt (2006) who complained:

“...how it can be that when evolution is creating wider and wider variation, they [cladists] favour a classification which must give lower and lower ranks as evolution progresses. But I have had no answer.”

I think the response comes directly from the above arguments: Figure 1 shows that just the opposite is the case. When evolution is progressing, a meaningful classification is also “progressing”, and then the appearance of new species and increasing variation are well-reflected if the diverging descendants are delimited at higher and higher taxonomic levels. Brummitt's problem derives apparently from the proposition that descendants, some or all, must be nested within the old taxon to which their closest common ancestor population belongs – an ignorable issue if branches of the phylogenetic tree are not envisaged as taxa and classifications refer to a single time slice of the phylogeny.

9 See DeSalle et al. (2009) who pointed out that tree structure may have been prevailing during bacterial genomic evolution as well.

3

The Hierarchy Conflict

“...many biologists assume that species-poor lineages that appear »early branching« or »basal« are ancestral...”
(Omland et al. 2008)

The Linnaean classification corresponds to a so-called *inclusive* hierarchy: species are small boxes (with populations in them) which are included in larger boxes representing genera and so on, i.e., it is a completely nested system of partitions of organisms (Fig. 2ab, see e.g., Panchen 1992). In fact, such a classification system can be represented by a tree (a dendrogram), with species as terminals and higher taxa as internal nodes. All species have equal rank, all the genera are equal to one another, and so on. A higher taxon is an *equivalence class* (see “Relations” in the Appendix) for all included taxa that are one level lower. Due to nestedness, the equivalence relations are also nested in some sense (Jardine and Sibson 1969).

There is another kind of hierarchy, the *exclusive* hierarchy which is a partition of objects such that there is an ordering relation among the classes (Panchen 1992). One class is superior to all others, the next class is superior to the remaining ones and so forth. Good examples are the church, the caste system of India, the military ranks and most importantly the *scala naturae* (the Great Scale of Being, GSB) – a summary of relationships in the natural and even supernatural world, an arrangement quite popular in the XIXth century and earlier. The relationship is an equivalence relation for entities within a class *but subordinate* between classes, the latter having only property 3 of binary relations. (This is transitivity, i.e., if *a* is subordinate to *b* and *b* is to *c*, then it follows that *a* is also subordinate to *c*. Obviously, if *a* is subordinate to *b*, then *b* cannot be subordinate to *a*, and *a* is not subordinate to itself, so the “subordinate to” relation is neither symmetric nor reflexive. See Appendix, for the terminology of relations)

Why did I enter this long discourse on relations and hierarchy types? First of all, a hierarchy cannot be of both kinds, it is impossible in logic. *The relation between members of a higher taxon cannot be at the same time equivalence relation and subordinate relation.* In other words, two genera from the same family

cannot have equal rank and superior/subordinate relationship simultaneously in the *same* classification. Undoubtedly, most taxonomists strive for a TCL which is basically Linnaean, an inclusive hierarchy, and reject the idea to fit all organisms into a sequential system like the GSB. Nevertheless, fragmentary ideas inherited from the GSB are still present in our biological thinking (“lower plants” vs “higher plants”, “primitive” vs “advanced”, etc.) mostly because it was conceptually linked to evolution and phylogeny by Lamarck and especially Haeckel. Many of Haeckel’s “phylogenetic trees” were summaries of morphological progress, illustrating older/newer relationships among forms, body plans or *grades* (Dayrat 2003)¹⁰. Such a grade tree (GT), or at least its “trunk” demonstrated a linear sequence such as *Thallophyta* → *Thallobrya* → *Bryophyta* → *Pteridophyta* → *Gymnospermae* → *Angiospermae* for plants (see Fig. 3 in Dayrat 2003). Such trees imply subordinate relationships in a way reminiscent of GSB and should not be confused with Darwinian ADR trees, nor Hennigian cladograms (SGR trees), in which the subordinate relationships are of the ancestor-descendant type largely in the population genetic sense. It follows that GT (and GSB as well) should not be confused with anagenesis either, because the latter is a linear sequence of populations which may be illustrated by a special ADR tree (a lineage, as in Fig. 2b) and is *not* necessarily associated with any morphological *progress*. (Evolution may involve morphological reduction as a result of parasitism, for example, or a series of changes without any particular “directionality” in the morphological sense.) Then, the natural question arises whether one can find a logical construct which relies on GT as well as some TTL simultaneously to provide a consistent, dendrogram-like TCL. Apparently, proponents of evolutionary taxonomy feel that it is possible: “...all grades, including paraphyletic taxa are.....permissible and desirable but indeed inevitable” (Brummitt 2008) while cladograms still serve a profound basis for classification together with phenetics (Hörandl 2006, 2007). However, any attempt to summarize Darwinian speciation and Haeckelian progression in form of a Linnaean tree is not only eclectic but also logically very-very challenging, if not impossible. One immediately sees from what is said above that grades (subordination) and equivalent taxonomic ranks (co-ordination) are in conflict when used simultaneously within the same higher taxon. A Linnaean system does not imply that *Bryum* is “lower” than *Aster* while a GT does. Grades, being mostly linear structures also conflict greatly with

10 Most of Haeckel’s trees were not Darwinian, contrary to the general belief, but grade trees, while the famous “Monophyletischer Stammbaum der Organismen” from 1866 is in fact a cladogram (see Appendix). Other trees, presented as small inserts reflected morphological relationships (see e.g., Fig. 3 in Williams and Ebach 2009).

cladograms which handle the groups in a horizontal way. However, a cladogram can be easily transformed into a dendrogram, and then to a Linnaean hierarchy. A possible mode of operation is:

Step 1: cancel edge lengths in the rooted cladogram to obtain an ordered dendrogram;

Step 2: establish the number of ranks to be used in creating the taxonomic hierarchy;

Step 3: shift hierarchical levels in the dendrogram to the next level that corresponds to a rank, thus obtaining a multifurcating dendrogram which is nothing else but a Linnaean hierarchy.

That it is possible follows from the inclusion order of matrix representations of trees (Lapointe and Legendre 1992) which basically states that dendrograms, such as the Linnaean hierarchy are a subset of cladograms. The fact that a cladogram can be reduced to a dendrogram is a strong argument for neglecting *grade thinking* in the taxonomic practice – and for constructing classifications maximally congruent with cladograms. The obvious result is that, for example, green algae, bryophytes, pteridophytes, gymnosperms and dicots are no longer considered as taxa, although still useful for characterizing morphological categories.

To find an answer to Brummitt's question ("*Am I a bony fish?*", 2006), it is worth examining the case of vertebrates as well, although in a simple manner which is not meant to give an up-to-date report on the subject. Rather, it serves another simple and widely known example to illustrate the relationship between classification and different types of trees. The traditional Linnaean classification (Fig. 3.b) is purely phenetic, the evolutionary relationships among the five classes are not shown, of course. A possible grade tree (Fig. 3.a) shows "phylogenetic" advancement of body plans of the five major groups, but the names have no taxonomic meaning and only represent morphological/developmental levels. It should not mean for anyone that the five groups "evolved" from each other in that order.¹¹ The diagram of Figure 3.c is a simplified cladogram, a backbone tree of extant vertebrate taxa, showing hypothesized (mostly accepted, Meyer and Zardoya 2003) evolutionary relationships among major

11 Although a very naïve student might say something like this: "Fishes represent the lowest grade because they have only gills, amphibians are transitional between aquatic and terrestrial forms with their alternating lifecycle, reptiles are true terrestrial animals without being able to regulate their body temperature, birds can do that but they still lay eggs, egg-laying monotremes are a nice transition to mammals which are viviparous and therefore the most advanced." In the first tree diagram on phylogeny ever published (Lamarck 1809), this sequence appears except mammals ("Poissons → Reptiles → Oiseaux → Monotrèmes").

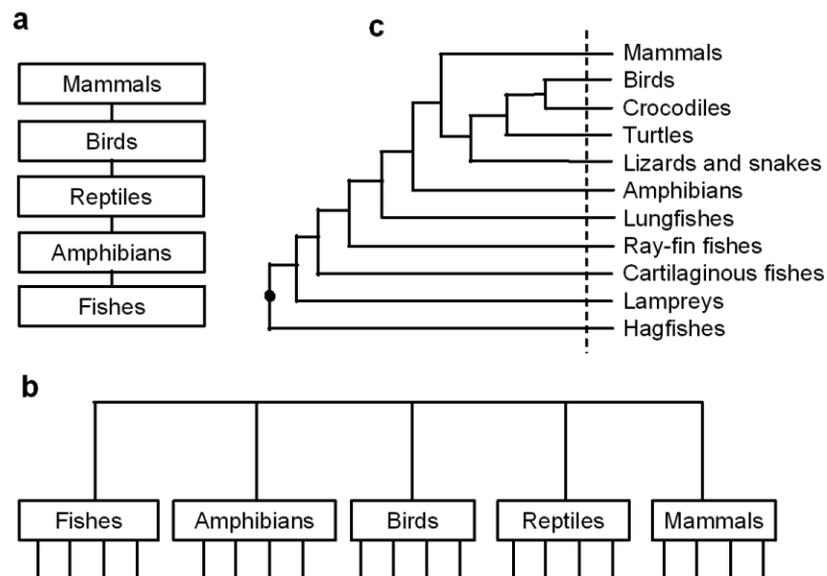


Figure 3. Tree diagrams for extant vertebrates illustrating different views on evolution and classification. **a:** Lamarckian-Haeckelian grade tree – or a chain – showing “phylogenetic progress”, **b:** Linnaean hierarchy for vertebrate classes (vertical bars merely indicate that there are orders and lower ranks below), **c:** rooted cladogram with dotted line indicating a possible class level cladistic classification of vertebrates. Raising some traditional orders to the class rank removes paraclady of fishes and reptiles.

groups which, as indicated by the dotted line, can be conceived as classes in a cladistic classification. The only price we have to pay for a classification most consistent with phylogeny is that we give up our GT-related ideas on fishes and reptiles as taxa, and delineate more than five classes for the vertebrates. Thus, humans need not be confused with bony fish, and will never be classified as such, if the classification is synchronous (branches are not meant to represent taxa) and it does not intend to show grades either.

Confusing grades with anagenesis and sister group relationships can have strange consequences – the bony fish “problem” was just one case in point. To see some botanical examples, we can also cite Farjon (2007) who commented on the closeness of *Gnetales* and *Pinaceae* as depicted by some molecular cladistic studies (e.g., Bowne et al. 2000) as follows: “...it is unlikely that [Gnetales evolved] from modern conifers. A sister group relationship to *Pinaceae* is an evolutionary absurdity...”. Obviously, no gnetophytes have evolved from *modern* conifers, nobody said that, because modern conifers live today, just as mod-

ern gnetophytes. The molecular hypothesis states no more than that gnetophytes and conifers may have a common ancestor population from which the two branches diverged, one perhaps more rapidly than the other. Misunderstanding a sister group relationship in a cladogram as if it were a subordinate relationship is a serious mistake; it is the absurdity. Omland et al. (2008) refer to the misconception that species poor sister lineages are ancestral as “primitive lineage fallacy”.

4

Monophyly vs Paraphyly – Or Something Else?

“...the co-descendants of the same form must be kept together in one group...”
(Darwin 1871)

Consider the following two statements: 1) “species *a-d* comprise a monophyletic group together with their most recent ancestor and all of its descendants”, and 2) “species *a-d* form a monophyletic group because their most recent hypothetical ancestor has no other extant descendants.” The difference between them is evident, yet it happens quite often that the first statement is meant while the second is the case, and vice versa. In both cases, monophyly is declared for a group of organisms, but in the first situation reference is made to their position in a phylogenetic tree. The second statement is based on a cladogram, because only extant populations are evaluated and the ancestor is only hypothetical. As said, these two conditions are obviously different and therefore referring to monophyly is ambiguous without explicitly stating what kind of a tree one is concerned with.¹² To resolve this ambiguity, I suggest to restrict the original definition of *monophyly* to *phylogenetic trees*, which means that monophyletic groups have temporal implications and can only be examined in a diachronous classification. For *cladograms*, a new term, *monoclady* is introduced, which means that a group of organisms comprises all terminals of a given clade.¹³ The condition of monoclady has to do with extant taxa and is meaningful for a synchronous classification. Paraphyly-paraclady, and polyphyly-polyclady are analogous pairs of terms, whose meaning is explained and illustrated in the Appendix.¹⁴

If we examine Darwin’s views on classification very closely (Padian 1999), we find a passage in his book on the descent of man (Darwin 1871; emphasis added):

- 12 Mishler (2009) has also pointed out that there are two basically different ways of defining monophyly, the synchronous one („all and only descendants of a common ancestor”) and a diachronous definition („an ancestor and all of its descendants”).
- 13 Bertrand (2008) suggested that the concepts of monophyly and paraphyly must be separated from that of clade, the first two referring to a *property of a group of populations*.
- 14 These terms are also introduced in two *Point of view* papers in *Taxon* (Podani 2009, 2010), while precise graph-theory definitions appear first in this paper.

"[The natural system,] it is now generally admitted, must be, as far as possible, genealogical in arrangement,— that is, the co-descendants of the same form must be kept together in one group, apart from the co-descendants of any other form; but if the parent-forms are related, so will be their descendants, and the two groups together will form a larger group."

I am tempted to say, even though it seems to be an ahistorical interpretation which is not always fortunate, that this may be conceived as the first formulation of monoclady, due to the clear synchronous implications in the prefix co- (e.g., coauthor, cooperation).

To see how these conditions relate to each other, let us examine three well-known groups (reptiles, gymnosperms and angiosperms) from the traditional classification which are well-established in everyday speech as well (I mean vernacular names which obviously ignore the continuity/discontinuity problem inherent in diachronous classifications). Assume that we know at least the main branches of the phylogenetic tree and the cladistic reconstruction of the sister group relationships is also correct. Then, reptiles turn out to be paraphyletic, because those ever lived include the ancestors of birds and mammal-like reptiles as well. The extant reptiles form a paracladistic group, as shown in Figure 3.c, since cladistic reconstruction suggests that crocodiles are sister to birds rather than to other reptiles. Gymnosperms, i.e., those ever lived are paraphyletic because angiosperms have gymnosperm origin, while the extant gymnosperms (conifers, cycads, ginkgo and gnetophytes together) are monocladistic (e.g., Bowe et al. 2000). Angiosperms are generally thought to be both monophyletic and monocladistic. The fourth combination, monophyly with paraclady is impossible if the phylogenetic tree and the cladogram are both correct. Nonetheless, paraclady may occur as a misleading result due to long branch attraction in a molecular cladogram (i.e., *a* and *b* are paracladistic on the cladogram, while they are in a monophyletic group in the true phylogenetic tree).

These definitions, which are hopefully precise enough, may help taxonomists to use their trees to cross-validate or explain an external classification or to make one. The example of gymnosperms shows that the judgment of a group depends on whether a phylogenetic tree or a cladogram is the reference. A Linnaean taxon – which is to be synchronous, as argued in this paper – may therefore be monocladistic, paracladistic or polycladistic. If a Linnaean taxon is evaluated with respect to a phylogenetic tree, the statement that it is mono-(para-, poly-) phyletic is thus imprecise. Nevertheless, we can examine if the members of that taxon are all and the only terminal nodes in a given branch. If it is so, this has to do with the monoclady condition again: the phylogenetic tree becomes implicitly a cladogram, since we restrict our interest to sister group re-

relationships in it. I understand, of course, that the proposed shift from a well-established, but somewhat misused terminology will not be easy because the concept of monophyly, understood irrespective of tree types, is rooted very deeply in our taxonomic thinking.

The distinction between phylogenetic and cladistic trees, and the associated changes in terminology will have some influence on the current monophyly/paraphyly debate, because statements will have to be expressed more carefully than before. Some of the consequences and a few related issues are examined below.

- The introduction of our new terminology clarifies an old issue, namely, when a species represents an “unchanged” lineage from which other lineages and branches arose by “budding” (Panchen 1992; Ereshefsky 2008). In the phylogenetic tree of Figure 4.a, the parent species (black) and its last derivative (gray) are *paraphyletic*, because the first population of the black species is the ancestor of the white branch as well. In a cladistic reconstruction (inset, same figure), however, the surviving parent species A and its living derivative B comprise a *monocladistic* group, which is acceptable intuitively because the recent populations of A and B do have a common ancestor population, whose “exact” species membership is irrelevant in a synchronous classification.
- If the Linnaean classification is restricted to extant organisms, as advocated in this paper, then the monophyly/paraphyly problem vanishes or, more precisely, shifts to the monoclady/paraclady contrast. Paraphyly was detected formerly because ancestors and descendants were classified together, while in cladistic analysis the extant descendants may nevertheless prove monocladistic (as in case of gymnosperms). The monophyly/paraphyly issue is thus left aside to handle by proponents of diachronous classifications. They will have to cope with an unavoidable difficulty: the higher the taxonomic rank in a diachronous classification, the more difficult to interpret the monophyly condition due to the increased fuzziness of taxonomic boundaries (or expanded transitional zones) – as explained in the Section on boundary paradox. Furthermore, in phylogenetic diachronous classifications paraphyletic groups (at least one) cannot be avoided (see Appendix). This is not so with synchronous classifications which can be constructed to include only monocladistic taxa.
- Essentially, paraphyly is a phenomenon often encountered when a traditional classification is not entirely validated by cladistic reconstruction –

that is, paraphyly may very well be an *indication of disagreement of a diachronous classification with a synchronous analysis*. This ambiguity is deepened by the contrast that the classification is phenetic in most of the cases, whereas the phylogeny reconstructions questioned are basically molecular or, at best, combined. Most, if not all examples mentioned by supporters of paraphyletic (more precisely, paracladistic) taxa appear to justify this statement (*Veronica/Hebe*, *Cupressaceae*, *Cactaceae/Portulacaceae*, *Scrophulariaceae*, *Primulaceae/Myrsine* etc). Contrasts between phenetic, character-based classifications and old-style, character-based cladograms have been less commonly mentioned in this context (e.g., Stuessy and König 2008). Therefore, the central tenet of contemporary taxonomy may *not be around monophyly and paraphyly* but rather

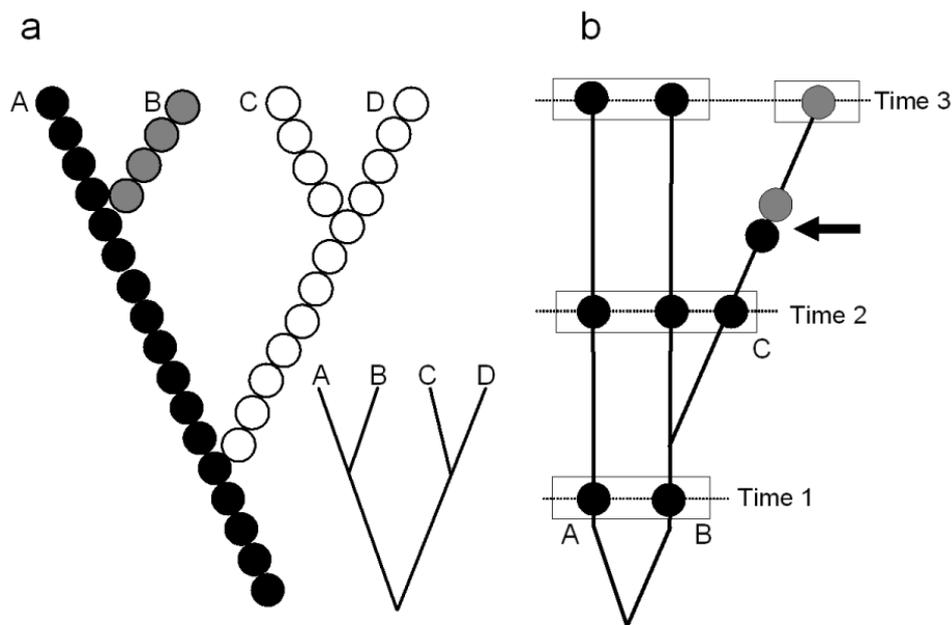


Figure 4. Some special cases of the paraphyly/paraclady problem. **a:** A single species (black) is older than the derived groups (white and gray). The group of terminals A and B is monoclادistic but the black-gray branch is paraphyletic since the populations on the white branch are also descendants of a black population. **b:** Species A and B must keep together in the same genus while C diverges from them into a different genus, which is the only possibility for paraclady at Time 3. Arrow points to a boundary between two populations at which shift to a new species must coincide with a shift to a new genus as well, thus immediately leaving a paraphyletic “ancestor genus” behind (cf. the boundary paradox).

around the contrast of synchronous versus diachronous thinking and around different data sets which offer more or less conflicting solutions. Needless to say, monophyly (or rather monoclady) of very many plant families established by traditional taxonomy (ranging from *Poaceae* through *Ranunculaceae* to *Asteraceae*) has been confirmed by molecular analysis (APG III, 2009), but *these* results were never questioned yet. Maybe I am wrong but one thing seems certain: this issue would not emerge at all if all higher taxa were discarded before reconstructing evolutionary pathways and were defined afterwards, using the procedure of cladogram→dendrogram reduction as described above.

- In a synchronous classification, then, the only problem that remains is whether *paracladistic taxa are allowed or not* (I do not think that polycladistic groups would ever be as crucial as paracladistic ones, but see Appendix for a precise distinction between them). Now let us examine how paraclady can occur. Monocladistic groups can be derived fairly easily – cutting the cladogram horizontally produces them and/or single species. Getting paracladistic groups from a tree is a bit more tricky: 1) select first a clade, cut it off and then remove a non-basal edge from it, or 2) remove any edge from the tree (except for those stemming from the root) in order to disentangle it into a monocladistic group (terminal taxa on the clade above) and a paracladistic group (all others). As a consequence, paraclady always has meaning only with reference to a clade, or to what it automatically implies: a monocladistic group. Nevertheless, obtaining a classification from the tree will always imply subjective decisions. For monocladistic taxa the question is where to cut the tree or at what levels to reduce the cladogram into a dendrogram to obtain ranks. But the definition of paracladistic taxa requires an extra decision: why is that edge or clade removed from the cladogram? To me it suggests that monoclady has logical¹⁵ priority over paraclady and that a classification based on monoclady requires fewer arbitrary decisions.
- In order to speak of paraclady at all, we need to have at least two species that are not in sister relationship on the clade. Hörandl's first example (her Fig. 1, 2006) illustrates the most often disputed case, when these

15 For Hörandl (2007), inclusiveness of a monophyletic group is „*just* a logical component” (italics added) in its definition, implying that „evolutionary considerations” occasionally supersede what logic dictates. However, I do not think in science there is stronger argumentation than what results from a clear logic, especially in classification which is basically a logical construct. At the same time, Hörandl apparently has no problem with the fact that a Linnaean hierarchy is also inclusive.

two species are basal, one after the other. My argumentation will be tricky, indeed, but I hope I can capture the reader's attention by the following discussion. Let us examine by another mental exercise (see also the „Jacob and Esau” example in Dawkins 1986) how a series of practical taxonomists would handle this problem over time, and what the evolutionary implications would be. In the phylogenetic tree of Figure 4.b, species A and B are the only direct descendants of the ancestor population, and an observer at Time 1 considers them as members of the same genus. Shortly after, a new lineage buds from the right lineage and then an imaginary taxonomist finds three species at Time 2. Hörandl (2006) suggests that it is often possible that the „progenitor species (B) does not change” while the derivative population runs into a new species (C) through anagenesis. Allowing this possibility for the moment, we have to settle two things: 1) the new species remains fairly close phenetically to its progenitor for a while so that at Time 2 they still belong to the same genus as the other two (box in the figure), i.e., the branching event does not imply an immediate switch into a different genus (which, I think is reasonable from what I said above) – so there is no paraphyly nor paraclady *yet*, and 2) up to Time 2 it is not only species B that does not change, but species A is also fairly constant to remain within the generic limits with B. In other words, A and B must keep together up to Time 2 while C continues evolving. If it does not happen because A also suffers anagenesis or speciates into a new genus some other way, paraclady *can never occur* anyway. Consequently, in order to detect paraclady later, at Time 3, i.e., separation of C at genus level *two conditions must be met*: species C evolves rapidly further on and species A and B still keep together. Only in such cases can we speak of paraclady as Figure 4.b shows. Now, this is the point where an evolutionary biologist can ask the question: how can happen that anagenesis on the rightmost lineage reaches genus level separation from another group of populations which cease evolving in unison and does not split any further? I do not suggest that it is absolutely impossible (see e.g., the *Ammobium/Nablonium* example in Stuessy and König 2008) but it is definitely not common enough to justify in general that paraclady is “inevitable” in classification. In fact, the more species the progenitor group includes, the less likely is that evolution stops for all lineages and consequently the less likely to make trouble for an extant taxonomist. Of course, two crucial questions remain which concern the *extra decision* always associated with paraclady: how to define

unequivocally that the difference between two species necessitates their separation into different genera and is it possible at all to provide standards that are valid on every part of the TTL? Probably there is no positive answer (multiply the number of species concepts by the number of genus delimitation concepts), therefore the definition of taxa only on the basis of common descent offers the only logical alternative – even if the result is sometimes contradictory with our expectation or the tradition, both deeply anchored in vernacular names and folk classifications.

5

Nomenclatural Comments

“...the genus rank does not exhibit any unique features distinguishing it from other Linnaean higher taxon ranks and is thus biologically meaningless...”
(Lee 2002)

Nomenclature is inevitable for classification and, without naming at least the terminal nodes, phylogenetic trees and cladograms are also meaningless. An efficient nomenclature is essential for approaching the TTL and therefore an ideal TCL. Since, as argued in Section 2, uninomina are the only adequate names if we examine evolutionary history; the binominal nomenclature is the biggest obstacle to name organisms from the past. (I do not want to discuss historical details, because the philosophical, logical and botanical origins of binominal nomenclature are well-known.) I would modify Lee’s statement (motto above) to say more explicitly that genera have no meaning in the context of evolution. If species are the only entities which have *spatiotemporal* individuality (but see Mishler 2009) and therefore more or less sharp boundaries (forget now about the prokaryotes), while genera and those above do not, why do we attach the name of a species to the name of a higher taxon if the contents of that taxon varies in time? Why is the belief of taxonomists so strong that the name of a species must also bear the name of a taxonomic unit one (or more, cf. subgenus) level higher? If species are the UNITS of evolution and, fortunately, basic UNITS of classification as well, then UNInomina would be absolutely logical and sufficient to identify populations¹⁶. I mentioned already a few genus names like

16 For comparison: languages bear uninomina (English, German, Slovakian, etc.), they have evolutionary history approximated by cladograms and the extant ones are arranged in a hierarchical classification as well. We may see now quite clearly that the worst decision made by Linnaeus was to make nomenclature and classification *dependent on each other* (Ereshefsky 2001; Podani 2009). A new species cannot be named without finding its place in the system, while the Linnaean classification includes a rank (genus) which is part of the name of every species. In fact, the UNITS of nomenclature are the genera: species names (e.g., *sylvestris*, *purpurea*) are meaningless by themselves. However, I do not think that uninomina will ever replace binomina for historical reasons and insurmountable practical difficulties; therefore the only solution is perhaps to *discard the genus level from the hierarchical classification of life*. That is, the old double names can be retained as if they were uninomina (compare this with Mich-

uninomina in Section 2 and I am sure all paleobotanists knew what I was talking about.

As I argued elsewhere (Podani 2009), using binomials is not the only tradition in nomenclature which goes back to pre-evolutionary thinking and is therefore in serious conflict with evolutionary theory. The existence of three codes (for plants, animals and bacteria separately, not to mention the code for viruses) reflects pre-Darwinian concepts; classification of life into these three groups is completely inconsistent with our present-day knowledge on the tree of life (see Fig. 3 in Podani 2009 to appreciate the artificiality of this categorization). There have been different naming conventions over the branches of TTL: 1) the codes differ whether tautonyms (*Vulpes vulpes*) are allowed or not, 2) grammatical rules are inconsistent (e.g., ...*idae* is the ending for animal families as well as for plant subclasses) 3) many groups of organisms can have the same name on different parts of the TTL, e.g., *Aotus* refers to both monkeys and leguminous plants (for hundreds of “interkingdom” homonyms, visit http://species.wikimedia.org/wiki/List_of_valid_homonyms, others not listed are given in Podani 2007), 4) there are 13 extreme cases in which the same binomen is used validly for two remote species (e.g., *Trigonidium grande* refers to both a cricket and an orchid). Even within the same code, supraspecific taxa can have identical names (e.g., *Acanthocephala* is a genus and a phylum, *Articulata* is a subordo and a subclass among metazoans). The nomenclature is thus chaotic for evolutionary biologists dealing with TTL (and the supposed TCL) *in their entirety*, although I understand that for a specialist of a given group it is the contrast between traditional arrangement and cladogram-based classifications that presents some sort of chaos. Without a common basis in naming conventions, i.e., harmonization of the three codes of nomenclature, however, no universal tree and classification are attainable at all.

ener’s old suggestion from 1964), species need not be renamed just because a new genus level classification would be more concordant with hypothesized evolutionary history, and only suprageneric taxa need to be continuously revised by cladistic analysis. See the rank-free “praenomen concept” (Griffiths 1974) and the suggestion that the second part of a species name be used as uninomen (Cantino and De-Queiroz 1997) to appreciate the difficulties with finding a compromise between traditional rank-based and the uninominal nomenclature (Dayrat et al. 2008).

6

Closing Remarks

I listed several arguments supporting that some prevailing ideas about biological taxonomy need revision if it is agreed that evolution, a unique spatiotemporal process, is the driving force behind biodiversity and therefore classification. Furthermore, many problems and debates in taxonomy derive from the multiple use of terms that are not defined unambiguously anyway, so terminological clarification is also in order. Table 1 provides a brief summary of my recommendations.

In my view, if an hierarchical classification is meant to reflect the “true pattern of life”, it is to be restricted to a temporal slice of the phylogenetic tree. That is an hierarchical TCL, if exists, can only be synchronous. The diachronous pattern of life could only be depicted by the phylogenetic tree or, in part, a network, which otherwise must enjoy priority to classification (Panchen 1992). The major argument is that the Linnaean hierarchy assumes increasing gaps along with increases in taxonomic rank, which is not satisfied in the spatiotemporal continuum of populations. Identifying branches of the phylogenetic tree as taxa is therefore burdened with much arbitrariness. Another argument comes from hierarchy theory: while a cladogram (SGR tree) can be mapped into a dendrogram and thus into a Linnaean hierarchy to delimit synchronous taxa with equal rank, trees assuming subordinate relationships among groups of populations (morphological chains) are in logical conflict with inclusive hierarchies.

The temporal aspect is one major reason behind current discrepancies regarding monophyletic versus paraphyletic or rather monoclادistic and paracladistic taxa. I propose here that the conditions of “-phyly” must be distinguished on the basis of the type of the reference tree. I consider another factor just as important: paraphyly in many cases reflects no more than disagreement of a classification (which is usually phenetic) with a cladogram (most commonly molecular) so that the monophyly/paraphyly contrast reduces to the conflict between the sources of taxonomic information. Truly paracladistic groups are not encountered as commonly as generally thought; this condition arises only if parental lineages remain in quasi stasis until the derived lineage “sufficiently” diverges from them to qualify for separation at a higher rank. Otherwise, with

Table 1. Comparison of two strategies for organizing knowledge of biodiversity based on evolutionary theory. The first four rows relate to trees, the others to taxonomy in general.

	Synchronous	Diachronous
Tree type	Cladogram – Steiner tree	Phylogenetic (genealogical, evolutionary) tree – Directed spanning tree
Relationships shown	Sister group	Ancestor/descendant
Subtree name	Clade	Branch and lineage
Conditions for groups based on ancestry	Mono-Para-clady Poly-	Mono-Para-phyly Poly-
Increasing ranks associated with...	Expanding gaps	Enhanced fuzziness
Linnaean hierarchy	Logical and compatible	Much less logical and incompatible
Relevance to the True Tree of Life	Across (horizontal)	Across and along (both horizontal and vertical)
Nomenclature	Any, but uninominal preferred	Uninominal

divergence by Darwinian speciation, no such problems occur because the resulting groups need not be nested into any parental group in a synchronous classification.

The “True Tree of Life” would be a phylogenetic tree (Dayrat 2005), an obvious simplification because of reticulate events in evolution. However, it is absolutely unattainable in its entirety due to the lack of full access to extinct organisms. Its backbone tree, the “True Cladogram of Life”, may show the sister group relationships of extant taxa correctly and is therefore a more realistic goal of contemporary evolutionary biology¹⁷. A highly respected “True Synchronous Classification of Life” into monoclادistic groups could be a logical derivative of this cladogram, yet it is illusory because there will always be subjective deci-

17 In the book edited by Cracraft and Donoghue (2004), there are exactly 200 tree diagrams. The majority (190) are cladograms, no matter how called in the figure captions (e.g., “phylogeny”, “tree”, “interrelationship”). Of these, 26 cladograms include both extant and extinct taxa, mostly vertebrates, while the others (164) refer to extant organisms exclusively. There are 5 trees of historical interest (e.g., Darwin’s and Haeckel’s) and only 5 trees tend to depict some partial hypotheses on ancestor/descendant relationships. The most typical synchronous classification is the one proposed by the Angiosperm Phylogeny Group (APG III 2009 and earlier versions).

sions to be made in delineating supraspecific taxa. The idea of a “True Diachronous Classification of Life” is a lot more deceptive: even if the complete phylogenetic tree were known, one would still have insurmountable difficulties with delimiting taxa at all hierarchical levels – raising the necessity of a fuzzy diachronous classification. To summarize major steps and transitions in evolution, a “True Grade Tree of Life” may be much easier to construct (see Cavalier-Smith 2010, for examples).

In my view, suggestions by four leading figures in the history of biology are implicit in the above outlined phylogenetic/cladistic classification of all extant life. Synchrony allows us to retain and improve the Linnaean inclusive hierarchy and their ranks which are so deeply rooted in our biological thinking and cannot be changed easily. Defining taxa as monoclادistic groups satisfies the criteria set out by Darwin, namely that the most natural classification is based on genealogy and that all co-descendants of a given ancestor are to be kept together in a given taxon. Hennigian cladistics in a revised form, i.e., no matter if cladograms are derived by traditional synpomorphy-based optimization, parsimony, maximum likelihood, Bayesian methods or distance-based algorithms (e.g., neighbor joining or the recently suggested Boolean procedure, Jakó et al. 2009), provides a sound methodological basis to detect monoclادistic groups using data on extant organisms. Finally, the end result we can achieve this way is a Crowsonian classification, i.e., a complete system of all recent life on Earth.

Acknowledgements

I thank A. Engloner, J. Garay, T. Kalapos, G. Kovács, L. Orlóci and I. Zachár for constructive comments and discussions. Comments by referees of earlier versions of the manuscript, most remarkably E. K. Lienau and M. Haber, are gratefully acknowledged. It does not necessarily mean that we always agreed upon all points, of course. The author’s research was supported by Hungarian Research Grant OTKA No. NI68218.

References

- APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linnean Soc.* 161:105-121.
- Auyeung A. 2005. A new phylogenetic tree model for fuzzy characters. In: *International Conference on Information Technology: Coding and Computing (ITCC'05) Volume I*. IEEE Computer Society, Washington. pp. 2-7.
- Bertrand Y. 2008. Contrasting the general with the particular in phylogenetics—a proposal to keep meanings of mono/paraphyletic and clade/grade separated. *Taxon* 57:705-708.
- Bezdek J.C. 1974. Numerical taxonomy with fuzzy sets. *J. Math. Biol.* 1:57-71.
- Blunt W. 2002. *Linnaeus. The compleat naturalist*. Princeton Univ. Press, New York.
- Bowe L.M., Coat G., dePamphilis C.W. 2000. Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proc Natl. Acad. Sci. USA* 97:4092-4097.
- Brummitt W. 2006. Am I a bony fish? *Taxon* 55:268-269.
- Brummitt W. 2008. Evolution in taxonomic perspective. *Taxon* 57:1049-1050.
- Brundin L. 1966. Transantarctic relationships and their significance, as evidenced by chironomid midges with a monograph of the subfamilies of Podonominae and Aphroteniinae and the austral Heptagylae. *Kungliga Svenska Vetenskapsakademiens Handlingar* 11:1-472.
- Cantino P.D., De Queiroz K. 2007. International code of phylogenetic nomenclature. Version 4b. <http://www.ohio.edu/phylocode/PhyloCode4b.pdf>.
- Cavalier-Smith T. 2010. Deep phylogeny, ancestral groups and the four ages of life. *Phil. Trans. Roy. Soc. B* 365:111-132.
- Cracraft J., Donoghue M. eds. 2004. *Assembling the tree of life*. Oxford Univ Press, Oxford.
- Crowson R.A. 1970. *Classification and biology*. Heinemann Educational, London.
- Darwin C. 1859. *On the origin of species*. Murray, London.
- Darwin C. 1871. *The descent of man, and selection in relation to sex*. Vol. 1. Murray, London.
- Dayrat B. 2003. The roots of phylogeny: How did Haeckel build his trees? *Syst. Biol.* 52:515-527.
- Dawkins R. 1986. *The blind watchmaker*. W. W. Norton, New York.
- Dawkins R. 2009. *The greatest show on Earth: The evidence for evolution*. Transworld, London.
- Dayrat B. 2005. Ancestor-descendant relationships and the reconstruction of the Tree of Life. *Paleobiology* 31:347-353.
- Dayrat B., Cantino P.D., Clarke J.A., de Queiroz K. 2008. Species names in the PhyloCode: The approach adopted by the International Society for Phylogenetic Nomenclature. *Syst. Biol.* 57:507-514.
- de Queiroz K., Gauthier J. 1992. Phylogenetic taxonomy. *Annu. Rev. Ecol. Syst.* 23:449-480.

- DeSalle R., Lienau E.K., Rosenfeld J.R., Planet P.J. 2009. The twin phylogenomic challenges. In: *NIBR Symposium 2009*. National Institute of Biological Resources, Korea. pp. 23-31.
- Dobzhansky T. 1973. Nothing in biology makes sense except in the light of evolution. In: *The American Biology Teacher* 35:125-129.
- Ereshefsky M. 2001. *The poverty of Linnaean hierarchy: A philosophical study of biological taxonomy*. Cambridge Univ. Press, Cambridge.
- Ereshefsky M. 2008. Systematics and taxonomy. In: Sarkar S., Plutynski A., editors. *A companion to the philosophy of biology. Chapter III.2*. Wiley-Blackwell, New York. pp. 99-118.
- Fajon A. 2007. In defence of a conifer taxonomy which recognizes evolution. *Taxon* 56:639-641.
- Farris J.S. 1974. Formal definitions of paraphyly and polyphyly. *Syst. Zool.* 26: 258-554.
- Farris J.S. 1976. Phylogenetic classification of fossils with recent species. *Syst. Zool.* 25:271-282.
- Fernholm B., Bremer K., Jörnwall H. 1989. *The hierarchy of life*. Excerpta Medica, Amsterdam.
- Foote M. 1996. On the probability of ancestors in the fossil record. *Paleobiology* 22:141-151.
- Ghiselin M.T. 2004. Mayr and Bock versus Darwin on genealogical classification. *J. Zool. Syst. Evol. Res.* 42:165-169.
- Griffiths G.C.D. 1974. On the foundations of biological systematics. *Acta Biotheor.* 23: 85-131.
- Haeckel E. 1866. *Generelle Morphologie der Organismen*. Reimer, Berlin.
- Haldane J.B.S. 1956. Can a species concept be justified? In: Sylvester-Bradley P.C., editor. *The species concept in palaeontology: A symposium*. The Systematics Association, London. pp. 95-96.
- Hall A.V. 1997. A generalized taxon concept. *Bot. J. Linnean Soc.* 125:169-180.
- Hennig W. 1966. *Phylogenetic systematics*. Univ. Illinois Press, Urbana.
- Hennig W. 1975. 'Cladistic analysis or cladistic classification?': a reply to Ernst Mayr. *Syst. Zool.* 24:244-256.
- Hörandl E. 2006. Paraphyletic versus monophyletic taxa—evolutionary versus cladistic classifications. *Taxon* 55:564-570.
- Hörandl E. 2007. Neglecting evolution is bad taxonomy. *Taxon* 56:1-5.
- Hull D.L. 1979. The limits of cladism. *Syst. Zool.* 28:416-440.
- Jakó É., Ari E., Ittész P., Horváth A., Podani J. 2009. BOOL-AN: A method for comparative sequence analysis and phylogenetic reconstruction. *Mol. Phylogenet. Evol.* 52:887-897.
- Jardine N., Sibson R. 1969. A logical basis for biological classification. *Syst. Zool.* 18:37-52.
- Jardine N., Sibson R. 1971. *Mathematical taxonomy*. Wiley, London.
- Lamarck J.-B. 1809. *Philosophie zoologique, ou Exposition des considérations relatives à l'histoire naturelle des animaux*. Paris.
- Lapointe F.-J., Legendre P. 1992. A statistical framework to test the consensus among additive trees (cladograms). *Syst. Biol.* 41:158-171.
- Lee M.S.Y. 2002. Species and phylogenetic nomenclature. *Taxon* 51:507-510.
- McManus F.G. 2009. Rational disagreements in phylogenetics. *Acta Biotheor.* 57:99-127.
- Meyer A., Zardoya R. 2003. Recent advances in the (molecular) phylogeny of vertebrates. *Annu. Rev. Ecol. Syst.* 34:311-338.
- Michener C.D. 1964. The possible use of uninominal nomenclature to increase the stability of names in biology. *Syst. Zool.* 13:182-190.
- Mishler B.D. 2009. Species are not uniquely real biological entities. In: Ayala F.J., Arp R., editors. *Contemporary debates in philosophy of biology*. Wiley-Blackwell, New York. pp. 110-122.

- Nelson G.J. 1971. Paraphyly and polyphyly: redefinitions. *Syst. Zool.* 20:471-472.
- Omland K.E., Cook L.G., Crisp M.D. 2008. Tree thinking for all biology: the problem with reading phylogenies as ladders of progress. *BioEssays* 30:854-867.
- Padian K. 1999. Charles Darwin's views of classification in theory and practice. *Syst. Biol.* 48:352-364.
- Panchen A.L. 1992. *Classification, evolution and the nature of biology*. Cambridge Univ. Press, Cambridge.
- Platnick N.I. 1977. Paraphyletic and polyphyletic groups. *Syst. Zool.* 26: 195-200.
- Podani J. 2007. *Evolutionary systematics of land plants*. 2nd ed. Eötvös Kiadó, Budapest (in Hungarian).
- Podani J. 2009. Taxonomy versus evolution. *Taxon* 58:1049-1053.
- Podani J. 2010. Monophyly and paraphyly: A discourse without end? *Taxon* 59: in press.
- Posada D., Crandall K.A. 2001. Intraspecific gene genealogies: trees grafting into networks. *TREE* 16:37-45.
- Reif W-E. 2005. Problematic issues of cladistics: 16. Taxonomic groups and the definition of paraphyly. *Neues Jahrb. Geol. Paläontol.* 238:191-229.
- Scott-Ram N.R. 1990. *Transformed cladistics, taxonomy and evolution*. Cambridge Univ. Press, Cambridge.
- Sneath P.H.A., Sokal R.R. 1973. *Numerical taxonomy*. 2nd ed. Freeman, San Francisco.
- Stuessy T.F. 2004. A transitional-combinational theory for the origin of angiosperms. *Taxon* 53:3-16.
- Stuessy T.F., König C. 2008. Patrocladistic classification. *Taxon* 57:594-601.
- van Regenmortel M.H.V. 1997. Viral species. In: Claridge M.F., Dawah H.A., Wilson M.R., editors. *Species. The units of diversity*. Chapman and Hall, London. pp. 17-24.
- Wiley E.O. 1976. The phylogeny and biogeography of fossil and recent gars (Actinopterygii: Lepisosteidae). *Mus. Nat. Hist. Univ. Kansas Misc. Publ.* 64:1-111.
- Wilkins J.S. 2006. A list of 26 species „concepts”. http://scienceblogs.com/evolvingthoughts/2006/10/a_list_of_26_species_concepts.php.
- Williams D.M., Ebach M.C. 2007. *Foundations of systematics and biogeography*. Springer, New York.
- Williams D.M., Ebach M.C. 2009. What, exactly, is Cladistics? Re-writing the history of systematics and biogeography. *Acta Biotheor.* 57: 249-268.
- Woese C.R. 2000. Interpreting the universal phylogenetic tree. *Proc. Natl. Acad. Sci. USA* 97:8392-8396.

Appendix

This is a summary of important mathematical, mostly graph theoretical terms as used, and some new definitions as proposed in this paper. The order follows a logical sequence of definitions and is therefore not alphabetical. Some of them require graphical illustration; these are collected in Figures 5-6.

Set: a collection of distinct objects or elements which can be anything (e.g., individual plants, species, numbers, letters or even sets). The elements may be defined by *extension*, that is, by listing each member (for example, $S = \{1, 3, 5\}$). The *intensional* definition is based on a rule (for example, S is the set of all individuals of a species living at the same time). Whereas in these conventional sets membership is of the yes/no type, in *fuzzy sets* membership is expressed by the degree of belonging measured on a continuous scale, from 0 to 1.

Relation: a property possessed by an *ordered pair* from a set of objects (i.e., I mean only binary relation in this paper). For objects a and b , it is denoted by aRb . Binary relations can be characterized in many ways; for the present discussion three properties deserve particular attention. *Reflexivity* means that an object is related to itself (aRa , e.g., $a = a$). *Symmetry* holds if aRb implies also that bRa (e.g., a is sister to b). A relation is *transitive* when aRb and bRc imply that aRc (e.g., if a is ancestor to b and b is ancestor to c then it follows that a is also ancestor to c). *Equivalence relations* are reflexive, symmetric and transitive (e.g., *taxon membership*; that is, a is in the same taxon as itself, if a is in the same taxon as b , then b is in the same taxon as a , and if a is in the same taxon as b and b is in the same taxon as c , then a is in the same taxon as c). Such relations define *equivalence classes* within a set, i.e., a *classification* (hard *partition*) into disjunct subsets.

Graph: an abstract construct visualized by a diagram composed of *vertices* (noda, dots), representing a set of objects, and connecting *edges* (lines) representing relations between pairs of objects. The *degree* of a vertex is the number of edges incident to it. A *path* is an alternating sequence of vertices and edges such that no vertices and edges are repeated. A graph G is *connected* if there is a path from every vertex to any other vertex. If the starting

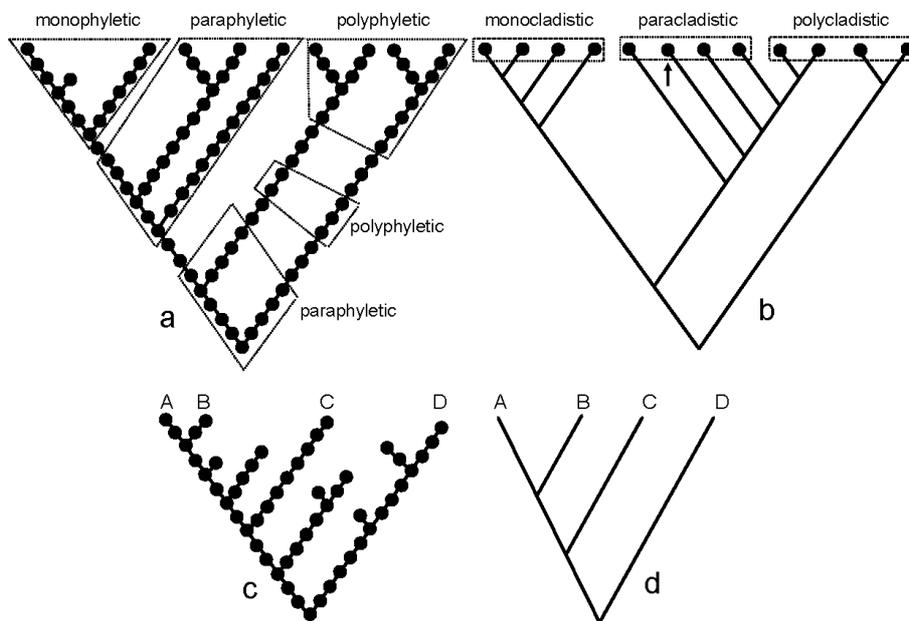


Figure 5. Conditions for groups on **a**: phylogenetic trees and **b**: cladograms, and **c**: a phylogenetic tree with **d**: its backbone cladogram. In **b**, arrow points to a vertex whose removal makes the strictly paracladistic group weakly paracladistic.

and ending vertex in the path is identical, we have a *cycle*. A connected graph with cycles is also called a *network*. A *subgraph* of G includes vertices that represent a subset of the vertices in G , plus the edges associated only to this subset. For more, see http://en.wikipedia.org/wiki/Glossary_of_graph_theory

Tree: a connected graph without cycles (“loops, reticulations”). In this, vertices of degree 1 are called *leaves* (terminal nodes); all other vertices are *internal*. If there is a particular internal vertex (the root) distinguished from the others, then we say that the tree is *rooted*. As a result, the edges point away from the root, i.e., the relation is *not* symmetric (such as the ancestor-descendant relation). A *subtree* is a connected subgraph of a tree, and is obtained by removing edge(s) from that tree.

Phylogenetic tree: rooted tree in which each vertex corresponds to a set of individuals belonging to the *same* species, extant or extinct, and edges represent Ancestor-Descendant Relations (ADR tree, Dayrat 2005) between the

sets. Accordingly, higher taxa cannot be vertices in such trees. When the associated time scale is refined, the vertices decompose into single individuals and the relations change to parent-offspring relations (which is the ultimate case, the *tokogenetic tree* for asexual populations, or a *tokogenetic network* for sexually reproducing populations; Posada and Crandall 2001). Phylogenetic and tokogenetic trees are *diachronous representations* of evolutionary-genealogical processes and may also be called *evolutionary* or *genealogical* trees. Darwin has popularized this kind of trees first, in accordance with his gradualist views on speciation. A *branch* is understood here as a *subtree* obtained by removing the internal edge incident to it within the tree (I was unable to find a better term than this, although the word “branch” may also refer to a single edge in the graph in the mathematical literature. “Phylum” would have been more unequivocal, but this term has been taken as a rank in Linnaean classifications). A *lineage* is a subtree of the phylogenetic tree in which all vertices have a degree of 2 except for the starting and ending vertices which have 1. Mathematically, the phylogenetic tree is a *directed spanning tree* (or an *arborescence*).

Cladogram: unrooted or – most commonly – rooted tree in which leaves (i.e., terminal vertices) represent individuals, populations, species or higher taxa. Internal vertices correspond to hypothetical evolutionary stages (e.g., hypothetical ancestor sequences in maximum likelihood cladograms) and are not to be interpreted in the same way as terminal nodes – and are not even shown. In a rooted cladogram, edges represent hypothetical ancestor/descendant relations. In a fully resolved (dichotomous) cladogram, all internal vertices have a degree of 3, but the root has 2. Calculated from data describing contemporaneous organisms, a rooted cladogram is a summary of hypothesized Sister Group Relationships (SGR tree, Dayrat 2005). In this, an internal vertex corresponds to a pair of subsets of leaves (sister groups). If all sister group relationships are correct, then a rooted cladogram – as a *backbone tree* – is a faithful summary of a phylogenetic tree (without all extinct lineages and branches). Thus, a cladogram may be a correct *synchronous representation* of evolutionary history¹⁸. In a *weighted* cladogram, or phylogram (Omland et al. 2008), there is a number associated to each edge (e.g., the number of character changes). The sum of weights is the *length* of the tree. If the condition that the sum of weights along the path between two leaves equals their evolutionary distance satisfies for all pairs of leaves, then we have an *additive tree*. In general, the leaves are *not equidistant* from the

18 Cladograms may also be derived for extant and extinct (fossil) organisms taken together, as suggested by several authors (e.g., Farris 1976). However, we cannot exclude the possibility that fossils are on the same lineages as extant organisms (Foote 1996). Such cladograms are no longer SGR-trees, because potential ancestors cannot be sister to descendants. For pattern cladists, cladograms have an even more general meaning (Williams and Ebach, 2007), detached from evolutionary theory, with use not restricted to biological entities at all (cf. Dawkins 1986). For a comparison of pattern cladistic trees with other tree types, see Podani (2010).

root (i.e., the sum of weights from a leaf to the root is not a constant). A *clade* refers to a subtree from the cladogram obtained by cutting the internal edge incident to it. Haeckel's Monophyletischer Stammbaum (Haeckel 1866, reproduced as Fig. 1 in Dayrat 2003) is cladogram-like in most parts, while it was Hennig who first proposed cladograms for routine use in the taxonomic study of extant organisms. Mathematically, a cladogram is a special *Steiner tree*.

Dendrogram: rooted tree in which leaves (i.e., terminal vertices) represent elements of a set S , i.e., practically anything, while internal nodes represent clustering steps or ranks. In a weighted dendrogram, there is a value associated to every edge, and all leaves are equidistant from the root (implying the ultrametric property). Such dendrograms result from hierarchical clustering (e.g., UPGMA popularized first by the numerical taxonomic approach led by Sneath and Sokal 1973) and may also serve as a special cladogram if computed from appropriate evolutionary distances. An ordered (or ranked) dendrogram has no weights, only the sequence of internal vertices matters. Unweighted rooted cladograms are often illustrated in ranked dendrogram form. A Linnaean hierarchy (in fact, any inclusive hierarchy) can be represented by a ranked dendrogram in which several internal vertices have identical ranks, corresponding to taxonomic categories.

Monophyly: the condition for a given subset of vertices in a phylogenetic tree such that the group comprises all and only descendants of a single ancestral population of individuals, and the ancestor population itself. [This definition has been the most commonly accepted one in biology, and is used in the present paper as well. The term "monophyly", however, "has had an extremely varied history" (Hull 1979), a note emphasizing the importance of precise definitions.] The group satisfying this condition is then a *monophyletic* group. By definition, the set of all nodes in a *branch* provides a monophyletic group. *Paraphyly* is detected in a classification when a given subset of connected vertices does not include all descendants of their latest common ancestor: such a group is called *paraphyletic*. Finally, *polyphyly* refers to a group of 2 or more populations whose nearest common ancestor is not included (*polyphyletic* group, i.e., without systematic continuity of populations, Griffiths 1974). That is, these terms reflect diachronous evolutionary relationships, they are meaningful with reference to diachronous classifications and do not refer only to extant populations.

For those who prefer exact mathematical statements, we can formulate the following

Lemma: Paraphyly is a mathematical necessity in *any* partition of n vertices into $k \geq 2$ groups, given a directed spanning tree for these n vertices.

Proof. Clearly, partitioning implies removal of edge(s) from the spanning tree, thus leaving at least one subtree which contains ancestor(s) cut off from its (their) descendant(s), and that subtree corresponding to a paraphyletic group.

Monocladly: this and its derivatives are new terms. It is the condition for a given subset of terminal vertices in a rooted cladogram such that all have the same hypothetical ancestor vertex and that ancestor has no other descendants. [Nelson's (1971) definition of "monophyly" is in fact the same ("A group into which have been placed all species or groups of species that are assumed to be descendants of a single hypothetical ancestral species, that is, a complete *sister-group system*", emphasis added). This illustrates lucidly why there is so much confusion about monophyly in current taxonomic literature, cf. Podani (2010).] This group is then *monocladistic*. By definition, the set of all terminal nodes on a *clade* represents a monocladistic group. *Paraclady* occurs when a given group is incomplete with reference to its corresponding clade, and the group is called *paracladistic*. It may be useful to distinguish between two cases of paraclady:

a) *strict* paraclady when the group is obtained by removing one smaller clade [which corresponds with Nelson's (1971) definition of "paraphyly" which refers to an "...incomplete sister-group system lacking one species or one monophyletic species group"], and

b) *weak* paraclady when the removed group does not form a single clade (Fig. 5.b). If the group of vertices comprises two or more subgroups with different closest ancestors none of them being ancestor to the other and none of them sister to the other, then we can speak of *polyclady* (*polycladistic* groups). These terms have relevance with respect to synchronous classifications only. [The corresponding definition by Nelson (1971) is that a

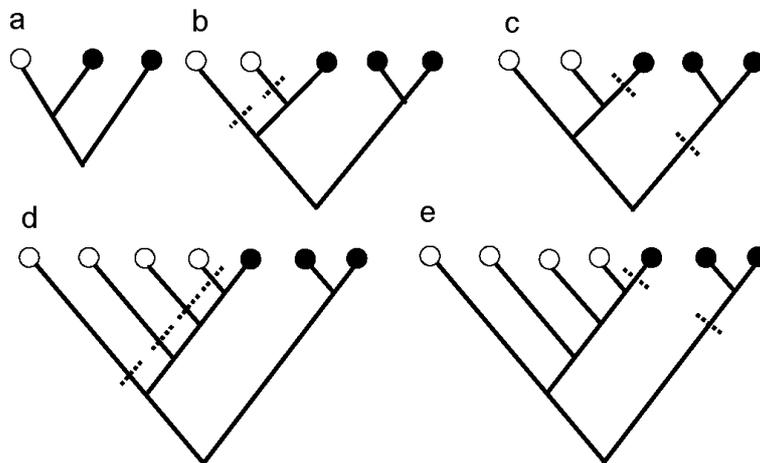


Figure 6. Paraclady and polyclady for a group (black symbols) depend on tree shape, number of vertices and different evolutionary scenarios: reversals (**b** and **d**) or convergence (**c** and **e**), except for very small trees (**a**). Dotted lines show edges where the group membership character (*sensu* Farris 1974) changed.

“polyphyletic” group is an “incomplete sister group system lacking two or more species (or monophyletic species groups) that together do not form a monophyletic group”. Nelson’s “polyphyletic” groups include those termed weakly paracladistic here but, as I argue here, true polyclady and weak paraclady must be distinguished because they may imply different evolutionary scenarios (convergence and multiple reversals, respectively).]

From the above description and the example of Figure 5.b, the meaning of polyclady¹⁹ may be understood only intuitively. In fact, clear distinction between weak paraclady and polyclady depends on the number of vertices and tree shape, and any decision requires careful scrutiny of the tree and the group in question. As Figure 6.a illustrates, for three vertices only strict paraclady is possible. For five nodes, it can happen that weak paraclady and polyclady cannot be distinguished on cladistic grounds: if weak paraclady is assumed it means that the group membership variable (sensu Farris 1974, Platnick 1977) changed twice for non-members of the group (“reversal” to the outgroup state). However, polyclady is also possible by changing twice the group membership character for the members of the group, reflecting “convergence” (Fig. 6.b-c). Then, for larger numbers of nodes, weak paraclady can be much less parsimonious than polyclady (4 versus 2 steps, Fig. 6.d-e) so that we can assume with good reason that the group membership character of the group was not uniquely derived (Farris 1974), and the group is polycladistic. The example from Hennig (1975) repeatedly discussed in the literature (e.g., Fig. 4 in Platnick 1977, Fig. 2.a in Cavalier-Smith 2010) is another case when weak paraclady and polyclady cannot be distinguished based merely on tree topology.

Grade tree: a simple chain- or ladder-like graph with no or only a few furcations. The vertices represent morphological types or grades, and even if they are higher taxa the edges show direction of progression of form or some other one-dimensional ordering relation. Any exclusive hierarchy may be represented by a grade tree. The *scala naturae*, and most diagrams of Lamarck, Haeckel and Bessey (his “cactus diagram”) are variants of grade trees. Diachronous classifications of life often appear in form of grade trees. They should be interpreted with care in taxonomy because they seem to suggest that one *group* as a whole “evolved” from the other.

19 Polyclady has been rarely used as a morphological term, referring to the “multibranching condition” of thalloid plants, for example, but this will not cause confusion, hopefully.