Beyond cladistics: Extending evolutionary classifications into deeper time levels

Elvira Hörandl

Department of Systematic and Evolutionary Botany, University of Vienna, Rennweg 14, 1030 Vienna, Austria; elvira.hoerandl@univie.ac.at

INTRODUCTION

The controversy over principles of classification in Taxon, started by Brummitt (e.g., 2002, 2006) and Nordal & Stedje (2005), continues. The need for the recognition of paraphyletic taxa has been argued from theoretical perspectives (Hörandl, 2006, 2007; Van Wyk, 2007; Farjon, 2007; Brummitt, 2008) and practical considerations (e.g., Brickell & al., 2008). Counter-arguments from a cladistic perspective have, for example been proposed by, e.g., Ebach & al. (2006) and Albach (2008).

Novel methodical approaches of patrocladistic analyses have been proposed for recognition of evolutionary change at the tips of the phylogeny (Stuessy & König, 2008) and have stimulated renewed discussions (Stuessy, 2009a; Wiley, 2009). In previous papers I outlined that strict application of cladistic principles for grouping of taxa fails to recognize the diversity of evolutionary processes. The first paper focused on species-level classifications (Hörandl, 2006); later, I tried to summarize theoretical principles of grouping for an evolutionary classification based on phylogenetic reconstructions, but including a broader spectrum of methodical approaches and a pluralistic view of theoretical concepts (Hörandl, 2007). In this present paper I (1) demonstrate more explicitly theoretical and methodical problems of cladistic classifications; (2) discuss paraphyly at deeper time levels; (3) show the effects of asynchronous evolution of characters at deeper time levels; and (4) discuss the implications of these aspects for classification.

THEORETICAL AND METHODICAL PROBLEMS OF CLADISTIC CLASSIFICATIONS

Nowadays there is a broad consensus that grouping of organisms should be conceptualized to reflect common ancestry, but opinions differ regarding how this should be achieved. Most cladistic classifications restrict the concept of grouping to a single process of evolution, that of dichotomous splitting of a lineage into two new lineages, resulting in extinction of the ancestor (Hörandl, 2006). This is also the only hypothesis inferred from a cladogram, and only in this case does a cladogram actually reflect the true phylogeny. A cladogram presents one hypothesis of a phylogeny, nothing more. Whether other evolutionary processes (e.g., budding and merging) are involved is neither tested, nor verified, nor falsified, but rather is neglected a priori. Thus, cladograms do not test evolutionary hypotheses, but force the analysis and the conclusion into a single hypothesis. Consequently, cladograms alone do not allow for a conclusion other than accepting a clade appearing in a cladogram as constituting a formal taxon. In this respect, cladistic classifications follow the same pattern-oriented principle as previous hierarchical ordering systems that have chosen selected characters or phenetic similarity as grouping criteria. Taxonomy has not actually made much progress in paradigms since the time of pre-Linnean classification (Stuessy, 2009a).

Evolutionary classifications are based on phylogenetic principles and the criterion of common ancestry, but consider various kinds of evolutionary processes: splitting, budding (or split-offs), and merging of lineages. Evolutionary studies conducted at the population and species level have well demonstrated that budding (split-offs), and merging do exist as regular evolutionary processes (reviewed by Hörandl, 2006). These processes often result in paraphyly, which is a normal phase in evolution. Evolutionary classifications do not make any a priori decision on the kind of process that has shaped a lineage. For evolutionary classifications it is necessary to formulate a working hypothesis on the kind of evolutionary process before the analysis, and the analytical tools should be selected specifically to test this hypothesis. The conclusion is then drawn from the results of the study. This is actually the normal working procedure of evolutionary biologists (see, e.g., Futuyma, 2005: XV). Consequently, in an evolutionary classification, the taxonomic decision should be drawn only after an understanding of the evolutionary process involved.

Much of the controversy is due to the preferred use of cladograms for reconstruction of phylogenies. Cladistic classifications regard a cladogram as a direct representation of a phylogeny, but do not recognize that a cladogram is a tool for recognition of lineage splits, and nothing else. If evolutionary history was shaped by budding or reticulate evolution, then a cladogram is an inappropriate tool to elucidate this process, because (1) the data structure is no longer hierarchical, and (2) taxa that occur on the nodes of phylogenies trees are forced onto the tips of cladograms. The latter problem applies to all types of phylogenetic tree reconstructions. Podani (2009) recently suggested a new terminology for discriminating results of cladograms (mono-, para- and polycladistic) and the inferred phylogenies (mono-, para- and polyphyletic). Although this approach helps to separate results of a cladogram from a phylogeny, it does not help to reconstruct phylogenies that are not strictly dichotomous. If rapid evolutionary change occurs...
toward the tips of the phylogeny, then a cladogram is not informative. Testing statistical support of tree topologies is no test of the validity of the hypothesis; expanding of sampling gives more information and a better resolution of trees, but does not broaden the concept. Total evidence approaches bear the danger that the strongest phylogenetic signal of one dataset or marker overrides the others.

Studies of evolutionary processes can be of course based on previous cladistic or phylogenetic tree-building analyses to develop a working hypothesis. But then a specific analysis is needed to gain insights into the evolutionary processes involved. If reticulate evolution is suspected, then, e.g., a network analysis could be tried to test whether or not the data structure is hierarchical. This information can provide insights whether processes other than dichotomous splits were involved in evolutionary history. Meanwhile a broad range of analytical tools is available for studying reticulate evolution (e.g., Huson & Bryant, 2006). Statistical approaches may help to better understand the background of a non-tree like structure of data: e.g., Joly & al. (2009) presented statistical tools for the discrimination of hybridization from ancient lineage sorting using coalescent times. Pisani & al. (2007) used an interesting approach of “signal-stripping” in a supertree approach for testing the hypothesis that eukaryotes had an endosymbiotic origin. Their results overall confirmed the chimeric nature of eukaryotes by revealing different sister relationships to different groups of prokaryotes by disentangling contrasting phylogenetic signals. If a taxon is suspected to have originated through rapid anagenetic change, then a patrocladistic analysis could be attempted (Stuessy & König, 2008).

This approach also starts with a cladogram reconstruction, but proceeds with an algorithm to specifically test for divergence between taxa. Critics of this approach (e.g., Wiley, 2009) overlook that patrocladograms are methodical tools for testing the amount of divergence, but not a tool for reconstructing character evolution. All of these analyses should serve as examples for the many hypothesis-based approaches.

Paraphyly is not restricted to species or genus level. Current evidence suggests that, e.g., polyploidization, as a frequent speciation process (e.g., Wood & al., 2009) potentially creates paraphyly because of synchrony and even sympaty of progenitors and derivatives (Hörandl, 2006). It is therefore useful to separate polyploidization events from other speciation processes in species-level systematics (e.g., Hörandl & al., 2009). Polyploidization has occurred frequently in almost all groups of eukaryotes, and several times in the history of angiosperms (e.g., Van de Peer & al., 2009). Whole genome duplications may enhance species diversification and evolutionary novelty, but are not necessarily connected to mass extinctions. The budding of new groups out of older groups without extinction of the latter is probably ubiquitous in evolutionary histories. Reticulate evolution can cause paraphyly on species level, if the parental species exist synchronously with their hybrid derivatives. On deeper time levels, reticulate evolution can occur via symbiogenesis or smaller-scale horizontal gene transfer. Also in this case, the parental lineages do not necessarily disappear. Therefore, paraphyly is present at all levels of the hierarchy. At deeper time levels, the problems with cladograms increase in many ways: First, the necessity arises to include taxa of different ages and even fossils in the analysis; therefore, a cladogram squeezes different time levels artificially together, and many different evolutionary processes that have occurred in the course of time are forced into an artificial dichotomy. Second, the markers used for phylogenetic reconstruction evolve at different speeds. Whenever evolutionary history was not strictly dichotomous, then bifurcating cladograms are prone to producing methodological artifacts. Third, incomplete sampling becomes a crucial issue. Molecular analyses based on extant species only suffer from an incomplete sampling of taxa, because fossils have to be excluded; morphological analyses including fossils have an incomplete sampling of well-resolving markers. Such sampling effects are, e.g., obvious in phylogenetic reconstructions of seed plants, where morphological data reveal gymnosperms s.l. (including fossils) as paraphyletic (Bateman & al., 2006; Hilton & Bateman, 2006; Rudall & Bateman, 2007), whereas molecular data imply an (perhaps artificial) monophyly of extant gymnosperms (cycads, extant conifers, Ginkgo and Gnetales) and a sister relationship to extant angiosperms (Burleigh & Matthews, 2004; Hajibabaei & al., 2006).

Figure 1 shows model phylogenies involving paraphyly, i.e., evolution of a group of derivatives (A–D) without complete extinction of the progenitor. This might happen because of speciation before or after the budding of the group A–D (Fig. 1a–b), or after reticulate evolution without extinction of parental taxa (not shown). In these scenarios, the “grade” inferred from the cladogram (Fig. 1d) is a tree-building artifact, because X and Y are in Fig. 1a an unrecognized monophyly, in Fig. 1b an older monophyletic group that has become paraphyletic. For completeness, it should be added that many “paraphyletic” groups may in fact be unrecognized monophyla only because their synapomorphies (whether morphological or molecular) have not yet been sampled in the analysis (Fig. 1c). In deep time levels, originally shared characters of X and Y may have already evolved further on each branch and cannot be scored any more as synapomorphies in extant samples. A cladogram alone (Fig. 1d) fails to reconstruct their sister relationship because XY apparently lack synapomorphies.

Thus, a series of branches in a cladogram may have three different meanings: (1) that of a grade, i.e., a series of bifurcating splits that has separated first X, then Y, then ABCD (not shown); (2) that of an older monophyletic group, which has given rise to a younger descendant group without going completely extinct and has become paraphyletic (Fig. 1a–b); (3) that of a monophyletic group, that has not been recognized because an incomplete marker set was used (Fig. 1c). A cladogram alone does not allow for disentangling these different cases, and fails to reconstruct the sister relationship of X and Y. Additional information is needed before a taxonomic conclusion can be drawn. Fossils of transitional forms from X to Y to ABCD would suggest a gradual change, whereas a lack of transitions would rather infer paraphyly. The alternative monophyletic
groupings inferred from the cladogram alone (Fig. 1d) imply either a loss of information by lumping (ABCDXY) or an incorrect sister relationships (YABCD), thus inferring a serious misconception of the true phylogeny.

## ASYNCHRONOUS EVOLUTION OF MARKERS IN DEEP TIME

Characters (whether molecular or morphological) evolve in their own speed and may be shaped by many transition series over time. Characters do not evolve out of nothing—they evolve out of previous characters of the ancestor. The “phylogenetic signal” of markers is often the result of lineage-specific substitution patterns and changing structural and functional constraints. Long-term histories potentially confer multiple transitions; rapidly evolving characters such as morphological structures are likely to hardly ever remain stable over the whole evolutionary history of an old clade. At the origin of the clade a single character is a synapomorphy, but it turns to a symlesiomorphy when a new clade originates; at the tips of the phylogeny the former synapomorphy may change into various different autapomorphies. Moreover, characters also show paralogy in the sense that the ancestral state does not necessarily disappear in bifurcating phylogenies: a character may remain unchanged in one descendant and change only in the other descendant of a dichotomous split.

Figure 2 demonstrates how long-term evolution of characters used as markers affects phylogenetic reconstructions and interpretations of cladograms. In Fig. 2a, the character “grey box” changes into “white box” in the common ancestor of extant taxa ABCD, and finally into “dashed box” in extant taxon A, whereas BCD retain the “white box”. A cladogram based on the character set “boxes” will reveal an incorrect reconstruction of the phylogeny. The cladogram reconstruction infers not only a wrong common ancestor of the clade BCD but also a misconception of synapomorphies, because the “white box” is not in fact a synapomorphy of BCD. If the phylogeny is reconstructed after the character set “ellipses” (cladogram not shown), the two datasets will reveal incongruent results. If, e.g., “boxes” represent a morphological character, and “ellipses” a molecular marker, then the grouping AB and CD will have no shared morphological character—a common result in molecular phylogenetic analyses. A cladogram based on a molecular dataset will always reveal a morphological character as homoplasious if it has not diversified synchronously with the marker that is used to reconstruct the tree. This asynchronous marker evolution may contribute significantly to the frequently observed lack of congruence between morphological and molecular datasets.

This case assumes a dichotomous phylogeny. The situation becomes more complicated if paraphyly is involved as illustrated in Fig. 1a–b. Figure 2 illustrates an example of paraphyly combined with asynchronous character transitions: “ellipses” change during the split of X from the ancestor of the other taxa (YABCD). Later, the transition of the character “boxes” from black to white occurs during the budding of ABCD; this monophyletic group therefore has the synapomorphy “white box”. If the lineages X and Y persist until Present, and no change occurs in the characters used, then the cladogram reconstruction using “ellipses” groups Y with ABCD, whereas the cladogram using “boxes” groups X and Y together. One might misinterpret this as a case of reticulate evolution in the origin of Y. If “boxes” represent a molecular dataset and “ellipses” a morphological dataset, than one could misinterpret this as a

---

**Fig. 1.** Possible evolutionary pathways (a–c) and corresponding cladogram (d) of paraphyletic groups, based on five characters (outgroups and characters within the clade ABCD not shown); F1† and F2† are fossils. a, Split-off (budding) of the clade (A–D) without extinction of the progenitor lineage, which later speciates into X and Y; b, speciation of the sister group (X, Y) before the split-off of the clade A–D; c, unrecognized monophyly X, Y because of undersampling of characters. In all three cases, the cladogram fails to reconstruct the monophyly of X and Y (= paraphyly in 1a and b), but suggests a grade (a series of dichotomous splits).
parallel evolution of character “ellipses” in Y and in ABCD. In fact, the pattern could be simply an effect of asynchronous marker evolution plus emergence of new lineages without extinction of progenitor lineages. Taken together, the frequently observed “homoplasy” of morphological characters may to some extent be due to methodological artifacts.

These schemes should demonstrate that a better understanding of evolutionary processes is needed before a taxonomic conclusion can be confidently drawn, because paraphyly, gradual transitions and asynchronous character evolution impact negatively on the results of cladistic reconstructions. If paraphyly and sequential evolution over long time periods is suspected, then it could be useful to separate time zones (e.g., analyze fossils of certain time periods separately) to recognize older monophyletic groups before the extant modern groups appeared and rendered them paraphyletic. A vertical study of character evolution might reveal transitions through time, and the resulting trees of different time levels could be grafted onto each other according to transitions observed in key characters. Basically, if one group has arisen out of the other, then ideally an autapomorphy of a taxon of the older group should be homologous with a synapomorphy of a younger group. Such a stepwise process could disentangle the sequence of monophyletic groups that have become retrospectively paraphyletic. Fossils may thus help to understand the temporal diversification of morphological characters.

### CLASSIFICATION INSTEAD OF CLADIFICATION

Taxonomies should fulfill four main criteria, in the following order of priority: theoretical foundation on a natural process, predictivity, information content and practicability (e.g., Mishler, 2009). The theoretical foundation based on descent with modifications is a shared feature of cladistic and evolutionary classifications, just that the former accept only a priori a single kind of process and take cladograms as a representation of the true phylogeny. Whenever the evolutionary history was not only shaped by equally bifurcating splits, then a classification based on a cladogram alone indeed becomes a “cladification” (sensu Mayr & Bock, 2002). Evolutionary classifications consider various processes and are therefore broader in their theoretical background, but they do require an explicit methodical approach for testing specific hypotheses. The conclusion depends on the outcome of this analysis and not on the restricted dogma that evolution should be always dichotomous.

Evolutionary classifications do allow for predictivity, because groupings include all descendants of a common ancestor that have originated from the same evolutionary process, as it acted at the origin of the group. With an explicit methodical approach for testing an evolutionary hypothesis, the grouping criteria become testable and reproducible. Cladistic classifications do have only a limited predictivity because of the restriction to a single evolutionary process.

The information content of cladistic classifications is low because of the restriction to synapomorphies. In contrast, evolutionary classifications can be based on a broader range of characters by regarding combinations of characters (synapomorphies, plesiomorphies + lack of synapomorphies, autapomorphies; Hörandl, 2007). Molecular and morphological data can complete each other for this purpose. Molecular data are efficient for reconstruction of descent; with appropriate analytical tools, they are very useful for the recognition of evolutionary processes. However, the commonly used molecular markers have only limited information content for grouping concepts in an evolutionary context. Morphological characters are useful for recognition of modification (e.g., Alexander, 2006) and do
have a high information content for grouping of evolutionary entities. Morphological characters relate to structure and function, and therefore they contribute significantly to the phenotype that responds to evolutionary processes like selection, adaptation, co-evolution, etc. But, morphological characters do not follow cladistic principles to act in isolation as synapomorphies; in contrast, they usually function in combination in an evolutionary context. Actually the success of big groups is often more referable to character combinations than to a particular synapomorphy (e.g., Stuessy, 2004, on angiosperms). Hilton & Bateman (2006) noted that in fact most synapomorphies in land plants themselves have a composited character. The morphological phenotype as a whole is composed of plesiomorphies and apomorphies, and the latter are based on the former. If the success of angiosperms is referred to flowers, carpels and double fertilization, then it should be held in mind that these structures are based on features that have evolved earlier in their progenitor groups: e.g., double fertilization for endosperm development could not have evolved without an earlier evolution of seeds. The use of character combinations in classification has therefore a higher predictivity and information content for evolutionary success than the use of synapomorphies alone. Therefore, evolutionary classifications should recognize groups that have a common ancestry, that have originated from the same evolutionary process (at the time of the origin and therefore at the base of the group), and that share evolutionary significant phenotypic characters. If rapid morphological divergence occurs in single lineages, than a classification of monotypic taxa is more informative about this divergence than classifying a heterogeneous clade as a taxon (Stuessy & König, 2008; see also Emadzade & al., in press, on genera of Ranunculaceae).

Morphological synapomorphies congruent with molecular tree topologies are often scarce, because morphological characters evolve asynchronously to molecular data (see discussion above). This problem is obvious in cladistic classifications (e.g., Cantino & al., 2007). If grouping cannot be apomorphy-based, then it must be bound to specific tree topologies that are dependent on sampling of markers and species, tree-building algorithms, and preference of one topology among several equally parsimonious reconstructions. Species as specifiers of clades (as suggested by classifications following the PhyloCode; De Queiroz, 2006) give neither information about the actual content of the clade nor of its inclusiveness. Härlin (2003) pointed out that a there is a difference between ancestry (the whole evolutionary history) and an ancestor of a clade: a monophyletic group can have the same content of descendants, but different internal tree topologies. Thus, specifiers do neither define nor circumscribe a clade if it contains more than two species. Without morphological synapomorphies, clades do have low information content outside a circle of specialists who are familiar with the species composition of clades and the special literature where continuously new tree topologies are being published. Classifications following the PhyloCode therefore lack information content and practicability.

The loss of morphological synapomorphies applies to molecular-based cladistic classifications within the hierarchical (Linnaean) system as well; here, predefined categories, partly with a standardized naming (“flagged hierarchies”), are used, which is practical to get an overview of diversity, and to effectively locate a single entity in a complex world. Flagged hierarchies serve effectively the needs of identification, documentation, and communication (Stevens, 2006). For Linnaean classifications the problem arises that classifying monophyletic groups only neglects paraphyletic groups as the former monophyla that exist at all levels of the hierarchy. Therefore, the hierarchical framework becomes less informative about evolutionary change because internal information levels get lost, and character combinations of evolutionary significance are neglected. For instance, if gymnosperms s.l. are abandoned in cladistic classifications, because they are probably paraphyletic, then the information level between angiosperms and seed plants is lost, with all its relation to morphology, reproductive biology, physiology, ecology, etc. If Crustacea are abandoned because they are paraphyletic with regard to hexapods (Richter & al., 2009), then an information level is lost that fundamentally relates to structure, function and different ecological niches, etc. This is actually the justified criticism of users (e.g., Brickell & al., 2008), who face the impracticability to work with morphologically heterogeneous giant taxa (e.g., Potentilla including Fragaria and Alchemilla; Portulacaceae including Cactaceae, among others) with low information content. This protest is not only a matter of a loss of tradition and change of names, but a loss of information related to structure, function, biological features, ecology and evolutionary history within the hierarchical system.

Evolutionary classifications can serve better practicability and communication by using not only the hierarchical Linnaean system, but also a broader spectrum of characters for classification. Beside scarce synapomorphies, also apomorphies + lack of autapomorphies, and in cases of anagenetic change, autapomorphies can be regarded. Character combinations of evolutionary relevance can be used. Taxa of an evolutionary classification correspond to classes as organizational levels and evolutionary entities (Mayr & Bock, 2002). They necessarily include monophyletic and paraphyletic taxa, and provide therefore more internal evolutionary relevant information levels. Classes based on a broader spectrum of characters are also more suitable for perception, and broad communication, including teaching. Some authors discuss classes as a valuable parallel reference system for higher taxa additionally to strictly clade-based classifications (e.g., Knoop & Müller, 2006). In fact, many proponents of cladistic classifications propose the use of informal names for paraphyletic groups (e.g., Richter & al., 2009). However, what is the use of formal cladistic classifications if they require a parallel reference system?

Finally, it should be noted by taxonomists that nobody is forced to re-classify a group only because of the result of a cladogram or a molecular phylogenetic tree; in fact, for evolutionary classifications, a cladogram is actually just the start of the work, not the end (see also the recommendations for workflow by Stuessy, 2009b). We need to achieve a better understanding of evolutionary processes before formal taxonomic conclusions can be drawn. To users it should be noted that nobody is forced to follow cladistic classifications if they
appear to be uninformative or premature “cladifications”: there is no must, no law and no code that one has to follow cladistic principles of classifications. Nomenclatural codes set rules for naming of taxa, but not for grouping concepts (e.g., De Queiroz, 2006).

■ ACKNOWLEDGEMENTS

Financial support from the Austrian Academy of Sciences (ÖAW), Commission for Interdisciplinary and Ecological Studies (KIÖS; Ranunculus phylogeny project), and the Austrian Science Foundation FWF (P-19006-B03) is gratefully acknowledged. I thank Tod F. Stuessy (Vienna) and Richard M. Bateman (Kew) for valuable comments on the manuscript, and many colleagues for stimulating discussions.

■ LITERATURE CITED


