

## Phylogenetic systematics perspective and problems with ancestral species – theoretical considerations

*Katarzyna A. Kaszycka*

**ABSTRACT** The paper summarizes the main assumptions and procedures of phylogenetic systematics (cladistics), and presents problems with identifying common ancestral species from this perspective.

*Katarzyna A. Kaszycka*, 1999; *Przegląd Antropologiczny – Anthropological Review*, vol. 62, Poznań 1999, pp. 67–74, figs 2. ISBN 83-86969-44-X, ISSN 0033-2003

### Phylogenetic approach

The use of phylogenetic systematics in comparative biology has become more widespread only over the last twenty or so years. Phylogenetic systematics, also known as cladistics, is a method of analysis that was first developed by the German entomologist Willi Hennig [HENNIG 1965, 1966]. Hennig invented it as an empirical method for discovering and justifying phylogenetic relationships.

There are two basic assumptions underlying phylogenetic systematics: (1) all species are connected through ancestor-descendant relationships, and (2) species change their features over time, and these changes are passed on to descendant species. All species, therefore, are mosaics of primitive and derived traits, and the phylogenetic relationships in a group of species can be deduced from the pattern of their shared derived traits. The phylogenetic method is thus based on: (1) de-

termining which traits are relatively primitive and which are relatively derived with respect to the group being studied, and (2) grouping various species according to their shared derived traits.

Two other methods of classification and reconstruction of relationships are known: traditional systematics and phenetics. Traditional systematics, as pointed out by WILEY *et al.* [1991], is based on character weighting and intuition. The systematist selects the characters believed to be important (i.e., conservative) and classifies species according to these characters, which themselves do not have to have any evolutionary significance. As such, these classifications might be artificial. The phenetic method of classification [SOKAL & SNEATH 1963; SNEATH & SOKAL 1973] was invented as an attempt to apply empirical methods to establishing relationships between taxa. The method is based on the comparison of the number of characters in common, i.e., the greater the number of such characters, the closer the relationship. However, according to

WILEY *et al.* [1991], phenetics has no advantage over traditional systematics in determining taxonomic relationships, because it reflects the total similarity of the organisms in question, and thus possibly creates artificial groups too.

Phylogenetic systematics employs strictly empirical methods in reconstructing genealogical (common ancestry) relationships. As mentioned before, the evidence of common ancestry comes from shared derived characters. Phylogenetic relationships are expressed in the simple statement that two taxa are more closely related to each other than either is to a third.

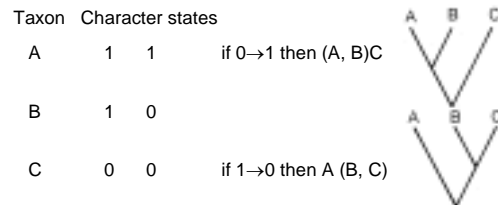
Phylogenetic hypotheses are expressed at different levels of complexity: (1) in the form of a cladogram (a graphic, atemporal representation of a sequence of branching events), and (2) in the form of a phylogenetic tree (adding a specification as to which taxa gave rise to which, i.e., depicting the sequence of ancestry and descent). There are a number of different trees derivable from each cladogram. This is because there are two types of relationships: ancestral-descendant and sister group.

The ancestral (primitive) character is called the plesiomorphic character. The descendant (derived or evolved) character is called the apomorphic character. In sum, three different kinds of characters are distinguishable:

- (1) symplesiomorphies – shared primitive,
- (2) synapomorphies – shared derived,
- (3) autapomorphies – unique derived.

Relationships between taxa cannot be established on the basis of unique characters. Nor are shared primitive characters useful. This is because characters can remain unchanged during a number of speciations, and common possession of a

primitive feature cannot be evidence of a close relationship between the taxa. Such a character, therefore, has no classificatory value. For this reason, the only traits potentially useful in phylogenetic analysis are shared derived ones. However, to be able to discriminate between derived and primitive homologies, one needs to determine the polarity of the characters. The polarity concept and polarity criteria constitute a fundamental concern in phylogenetic inference. This is because polarity has a significant effect on sister group relationships, which can be depicted by a simple example (Fig. 1):



There are four basic criteria of character polarization: the outgroup comparison, the ontogeny criterion, common equals primitive, and ancient equals primitive. The most widely accepted method, and one considered to be the most reliable [e.g., WATROUS & WHEELER 1981; MADDISON *et al.* 1984], is the outgroup comparison. In the outgroup comparison the rule is that for a given character with two or more states within a group under study (ingroup), the state occurring in a related group (outgroup) is assumed to be the plesiomorphic state [WATROUS & WHEELER 1981]. The most critical outgroup comparisons involve the sister group (the closest relative) of the taxa studied. The outgroup criterion ensures a more globally parsimonious solution [MADDISON *et al.* 1984], i.e., it requires the fewest hy-

potheses of parallelisms and reversals both within the ingroup and among the outgroups.

There are two ways in which the ontogeny criterion has been used: (1) the state that appears earlier in the development is plesiomorphic (refers to an ontogenetic sequence, e.g., Haeckel's biogenetic law); and (2) the character state that is more generally observed among two or more ontogenies is plesiomorphic (refers to generality, e.g., von Baer's law). The common equals primitive criterion is an ingroup comparison. The idea is that a character state distributed commonly among members of an extant taxon is primitive (which is, however, not necessarily true). The basis for ancient equals primitive comes from fossil evidence. The assumption made here is that all the character states found on the oldest fossil are the most primitive. (For a further review of these methods see, e.g., NELSON & PLATNICK [1981]).

SIMPSON [1953] called attention to different rates of character evolution, both within and between lineages. Characters, which evolve sporadically, will be then the best species defining characters. On the other hand, those characters, which change frequently and are variable within a species, will be poor species defining characters.

There are different approaches to dealing with character data in cladistic analysis [FINK 1986]: parsimony, probability, and compatibility. Parsimony methods seek trees that require the fewest character changes, thus minimizing instances of homoplasies (parallelisms, convergences, or reversals) as the basis for choosing among alternative hypotheses of sister species relationships. According to Farris [FARRIS 1983, FARRIS

& KLUGE 1985], one justification for using parsimony in phylogenetic inference is that the simplest cladogram (with the minimal number of evolutionary transformations) has the greatest explanatory power. Probability algorithms are based on some initial assumptions concerning changes in the evolutionary rate (these are rather used for molecular data). Compatibility methods build trees based only on those sets of characters that appear not to have homoplasy.

Phylogenetic analysis requires that character states be defined as discrete rather than continuous values. Therefore, morphometric (continuous) traits must either be coded into discrete states, or excluded from the analysis. Some authors, e.g., PIMENTEL & RIGGINS [1987], raised a number of objections to the use of continuous traits in cladistic analyses. Morphological traits may provide useful information, and various methods of coding quantitative characters into discrete states have been developed, e.g., MICKEVICH & JOHNSON [1976], ALMEIDA & BISBY [1984], ARCHIE [1985], BAUM [1988], CHAPPILL [1989], but these may result in some distortion of the relationships between taxa. Therefore, it may be risky to draw conclusions about the phylogenetic relationships between taxa based on quantitative characters only. The results of any cladistic analysis will be influenced not only by the data set used, but also by polarity criteria, methods of coding the characters, the ordering of multistate characters, and the composition of taxonomic units.

Some authors, e.g., HARRISON [1993], TRINKAUS [1995], WOLPOFF & CRUMMETT [1995] raise the issue that cladistic analysis cannot be justifiably applied to taxa below the species level. Other

authors, however, disagree. LIEBERMAN [1995: p. 189], for example, says: "Theoretically, cladistics can resolve relationships in an array of taxonomic units as long as the members of each taxon are more closely related to each other than they are to members of other taxa." The phylogenetic systematic method, as presented by WILEY [1981], allows reconstructing genealogical relationships at population, species and higher taxa levels. TATTERSALL & ELDRIDGE [1977: p. 207] assert that: "At the level of the cladogram there is no necessity to define formal taxa." In fact, so-called "operational taxonomic units" are quite frequently used, e.g., in CHAMBERLAIN & WOOD [1987], WOOD [1991].

### Ancestor – descendant relationships

That ancestral species existed follows from the assumption that all new species have originated from other species. The process of species formation (cladogenesis) is the splitting of one lineage into two.

In a cladistic model (e.g., HENNIG [1966]) it is expected that the ancestral

species becomes extinct at the time of speciation, irrespective of how similar it might be to one of its daughter species. This is because HENNIG [1966] views the process of speciation as a subdivision of the entire species into two large populations, during which the integration of its gene pool becomes lost. As pointed out by HULL [1979], HENNIG'S [1966] intention was also to avoid the terminological confusion, which he thought would arise when a species before and after speciation would be called by the same name. However, HULL [1979] argues and demonstrates in his example at the level of an organism (this point is illustrated in Fig. 2.) that there is no reason for this sort of confusion to arise also at the species level. The author says [HULL 1979: p. 433]: "When a single *Paramecium* splits down the middle to form two new organisms, each is considered a distinct organism. If we were prone to name such entities, we would give each a separate name. However, *Hydra* can continue to exist while budding off other *Hydra*. Once again, if we were inclined to, we could give each of these organisms its own name. The parent *Hydra* would retain its name even though it budded off other descendant *Hydra*."

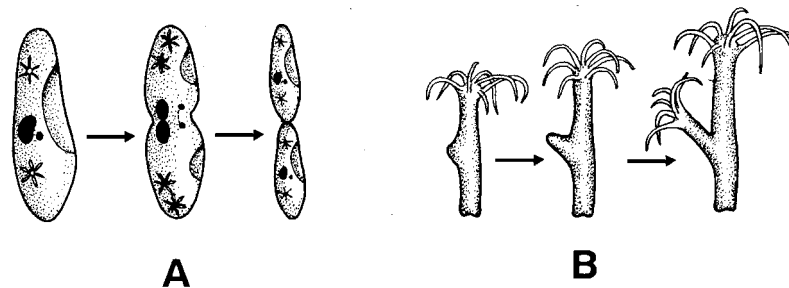


Fig. 2. Asexual ways of reproduction in animals: A. binary fission (*Paramecium*), B. budding (*Hydra*)

In the evolutionary species concept, as described by WILEY [1978, 1981] (see also HULL [1978, 1979]), the parent species may survive the split if it does not lose its identity (distinctness) and evolutionary tendencies. Following the work of MAYR [1963] and ELDREDGE & GOULD [1972], Wiley views the process of speciation also in terms of isolation and subsequent divergence of a small peripheral population which develops into a new species without affecting much the integration of the gene pool of the ancestral species. However, as WILEY [1978: p. 22] asserts: "the survival of the same species through more than a few of these 'buds'... would appear unlikely because it is unlikely that any one species could stand the loss of geographically or ecologically unique gene combination without its role and tendencies being changed." BELL [1979] also suggested, from the examples of fishes, that the survival of the ancestral species in some cases might have been quite common.

Several authors, e.g., NELSON [1972, 1973], CRACRAFT [1974], FARRIS [1976], ENGELMANN & WILEY [1977], and PLATNIK [1977] have argued that ancestral species cannot be empirically identified. Furthermore, while sister group relationships are testable on morphological evidence, ancestral-descendant relationships are not. Theoretically, from a purely cladistic perspective and that of the phylogenetic species concept (e.g., KLUGE [1992]), if a fossil taxon 'C' were to be considered as a candidate for the most recent common ancestral species of taxa 'A' and 'B', it would be expected to fulfil the following conditions:

- It would exhibit all the synapomorphies that 'A' and 'B' have as a group;

- It would not share apomorphies with one of the above mentioned species alone; therefore, for every character in which 'A' and 'B' differ, 'C' would be plesiomorphic;
- It would not exhibit apomorphies of its own (autapomorphies);
- It must be older than any part of the group 'A' and 'B' of which it is supposed to be the common ancestor.

However, according to cladists, even if taxon 'C' meets all these criteria, it does not mean that it is a real common ancestor, but only a hypothetical common ancestor (e.g., NELSON [1973]). This is because common ancestral species do not have any peculiarities that mark them as such. In fact, as the cladists argue, one can say for sure what is not a common ancestor, but not what is (this would be arguing from the negative evidence). In reality, therefore, the process of discovering common ancestors seems to be quite problematic.

ENGELMANN & WILEY [1977], following POPPER'S [1959] philosophy of science, claimed that the statements about ancestor-descendant relationships are not objective (and therefore unscientific) because they are not falsifiable. In fact, as ENGELMANN & WILEY [1977] argue, these statements can neither be corroborated nor rejected. The supposed corroborating evidence would be to find the supposed ancestor to be plesiomorphic. The supposed falsifiable evidence would be to find an autapomorphy in the supposed ancestor. However, the problem is that phylogenetic relationships between taxa can neither be based on plesiomorphic nor autapomorphic characters.

HULL [1979], in discussing the limits of cladism, pointed out that the difficul-

ties in establishing ancestor-descendant relationships are not empirical but methodological. In HULL'S [1979: p. 430] opinion: "cladists have exaggerated the difficulties in making reasonable inferences about probable ancestors." The author reminds us that statements about sister group relationships also carry with them the possibility of error, since they depend on polarity of characters and species recognition.

Other authors, e.g., BOCK [1973], BRETSKY [1975], GINGERICH & SCHOENINGER [1977], DELSON [1977], maintain that identification of ancestors is both possible and testable by incorporating the character distribution and stratigraphic and biogeographic evidence. Even though the fossil record is scarce and fragmentary, there is no reason to believe that no common ancestral species have yet been discovered. The important point is which type of relationship, ancestor-descendant or sister group, seems more probable given all the evidence. Following BRETSKY'S [1975] comments on the stratigraphic evidence, one can say that it seems more parsimonious to postulate ancestor-descendant relationships between time-successive taxa than to assume otherwise. In this way an unknown common ancestral species need not be postulated to explain the relationships. DELSON [1977] argued that in terms of morphological criteria it is more parsimonious to hypothesize ancestry when a taxon is intermediate, i.e., primitive with respect to its supposed descendants and derived relative to an ancestral condition. As HULL [1979: p. 437] stresses: "perhaps ancestors cannot be discerned with the same degree of confidence as sister-group relations, but the contrast is not between fact and fantasy."

## References

- ALMEIDA M. T., F. A. BISBY, 1984, *A simple method for establishing taxonomic characters from measurement data*, *Taxon*, **33**, 405-409
- ARCHIE J. W., 1985, *Methods for coding variable morphological features for numerical taxonomic analysis*, *Syst. Zool.*, **34**, 326-345
- BAUM B. R., 1988, *A simple procedure for establishing discrete characters from measurement data applicable to cladistics*, *Taxon*, **37**, 63-70
- BELL M. A., 1979, *Persistence of ancestral sister-species*, *Syst. Zool.*, **28**, 85-88
- BOCK W. J., 1973, *Philosophical foundations of classical evolutionary classification*, *Syst. Zool.*, **22**, 375-392
- BRETSKY S. S., 1975, *Allopatry and ancestors: A response to Cracraft*, *Syst. Zool.*, **24**, 113-119
- CHAMBERLAIN A. T., B. A. WOOD, 1987, *Early hominid phylogeny*, *J. Hum. Evol.*, **16**, 119-133
- CHAPPILL J. A., 1989, *Quantitative characters in phylogenetic analysis*, *Cladistics*, **5**, 217-234
- CRACRAFT J., 1974, *Phylogenetic models and classification*, *Syst. Zool.*, **23**, 71-90
- DELSON E., 1977, *Catarrhine phylogeny and classification: Principles, methods and comments*, *J. Hum. Evol.*, **6**, 433-459
- ELDREDGE N., S. J. GOULD, 1972, *Punctuated equilibria: An alternative to phyletic gradualism*, [in:] T. J. M. Schopf ed., "Models in Paleontology", Freeman, Cooper & Co., San Francisco, 82-115
- ENGELMANN G. F., E. O. WILEY, 1977, *The place of ancestor-descendant relationships in phylogeny reconstruction*, *Syst. Zool.*, **26**, 1-11
- FARRIS J. S., 1976, *Phylogenetic classification of fossils with recent species*, *Syst. Zool.*, **25**, 271-282
- FARRIS J. S., 1983, *The logical basis of phylogenetic analysis*, [in:] N. I. Platnik & V. A. Funk eds., "Advances in Cladistics", Vol. 2, Columbia Univ. Press, New York, 7-36
- FARRIS J. S., A. G. KLUGE, 1985, *Parsimony, synapomorphy, and explanatory power: A reply to Duncan*, *Taxon*, **34**, 130-135
- FINK W. L., 1986, *Microcomputers and phylogenetic analysis*, *Science*, **234**, 1135-1139
- GINGERICH P. D., M. J. SCHOENINGER, 1977, *The fossil record and primate phylogeny*, *J. Hum. Evol.*, **6**, 483-505

- HARRISON T., 1993, *Cladistic concepts and the species problem in hominoid evolution*, [in:] W.H. Kimbel & L.B. Martin eds., "Species, Species Concepts, and Primate Evolution", Plenum Press, New York, 345-371
- HENNIG W., 1965, *Phylogenetic systematics*, Ann. Rev. Entomol., **10**, 97-116
- HENNIG W., 1966, *Phylogenetic Systematics*, Univ. of Illinois Press, Urbana
- HULL D. L., 1978, *A matter of individuality*, Philosophy of Science, **45**, 335-360
- HULL D. L., 1979, *The limits of cladism*, Syst. Zool., **28**, 416-440
- KLUGE A. G., 1992, *Principles of Systematics*. Biology 491 course-pack, Univ. of Michigan, Ann Arbor, Michigan
- LIEBERMAN D. E., 1995, *Testing hypotheses about recent human evolution from skulls*, Curr. Anthropol., **36**, 159-197
- MADDISON W. P., M. J. DONOGHUE, D. R. MADDISON, 1984, *Outgroup analysis and parsimony*, Syst. Zool., **33**, 83-103
- MAYR E., 1963, *Animal Species and Evolution*, Harvard Univ. Press, Cambridge, Massachusetts
- MICKEVICH M. F., M. F. JOHNSON, 1976, *Congruence between morphological and allozyme data in evolutionary inference and character evolution*, Syst. Zool., **25**, 260-270
- NELSON G., 1972, *Comments on Hennig's "phylogenetic systematics" and its influence on ichthyology*, Syst. Zool., **21**, 364-374
- NELSON G., 1973, *Classification as an expression of phylogenetic relationships*, Syst. Zool., **22**, 344-359
- NELSON G., N. I. PLATNIK, 1981, *Systematics and Biogeography: Cladistics and Vicariance*, Columbia Univ. Press, New York
- PIMENTEL R. A., R. RIGGINS, 1987, *The nature of cladistic data*, Cladistics, **3**, 201-209
- PLATNICK N. I., 1977, *Cladograms, phylogenetic trees, and hypothesis testing*, Syst. Zool., **26**, 438-442
- POPPER K. R., 1959, *The Logic of Scientific Discovery*, Basic Books, New York
- SIMPSON G. G., 1953, *The Major Features of Evolution*, Columbia Univ. Press, New York
- SNEATH P. H. A., R. R. SOKAL, 1973, *Numerical Taxonomy*, W. H. Freeman, San Francisco
- SOKAL R. R., P. H. A. SNEATH, 1963, *The Principles of Numerical Taxonomy*, W. H. Freeman, San Francisco
- TATTERSALL I., N. ELDRIDGE, 1977, *Fact, theory, and fantasy in human paleontology*, Am. Sci., **65**, 204-211
- TRINKAUS E., 1995, *Comment*, Curr. Anthropol., **36**, 185-186
- WATROUS L. E., Q. D. WHEELER, 1981, *The outgroup comparison method of character analysis*, Syst. Zool., **30**, 1-11
- WILEY E. O., 1978, *The evolutionary species reconsidered*, Syst. Zool., **27**, 17-26
- WILEY E. O., 1981, *Phylogenetics. The Theory and Practice of Phylogenetic Systematics*, J. Wiley & Sons, New York
- WILEY E. O., D. SIEGEL-CAUSEY, D. R. BROOKS, V. A. FUNK, 1991, *The Compleat Cladist: A Primer of Phylogenetic Procedures*, Univ. of Kansas, Lawrence
- WOLPOFF M. H., T. L. CRUMMETT, 1995, *Comment*, Curr. Anthropol., **36**, 186-188
- WOOD B. A., 1991, *Koobi Fora Research Project, Volume 4. Hominid cranial remains*, Clarendon Press, Oxford

## Streszczenie

Na tle głównych założeń i procedur taksonomii filogenetycznej (kladystyki) – klasyfikacji odzwierciedlającej genealogię (poprzez cechy wspólne zaawansowane), przedstawiono teoretyczne rozważania dotyczące przeżywania bądź wymierania gatunków w procesie specjacji oraz problemów z identyfikacją gatunków stanowiących wspólnych przodków innych gatunków.

Hipotezy filogenetyczne można wyrażać na dwu poziomach złożoności: w formie kladogramu lub w formie drzewa filogenetycznego. Kladogram pokazuje najbliższe pokrewieństwa między taksonami i porządek, w którym taksony ulegały specjacji. Drzewo filogenetyczne natomiast zawiera informacje z kladogramu, uzupełnione o sekwencje przodków i potomków. W kladogramie wyrażony jest tylko jeden typ związków pokrewieństwa zwany „siostrzanym”, tymczasem związki między taksonami siostrzanymi mogą być konsekwencją dwu sytuacji: (1) jeden z nich może być przodkiem drugiego, (2) oba mają wspólnego przodka. Z każdego kladogramu można wyprowadzić kilka różnych drzew, ponieważ drzewo może przedstawiać oba typy wspomnianych związków pokrewieństwa: przodek-potomek, i siostrzane.

Według kladystów, nie można empirycznie zidentyfikować gatunków będących wspólnymi przodkami innych gatunków. Na dodatek, podczas gdy związki siostrzane są możliwe do testowania na podstawie dowodów morfologicznych, nie można testować związków przodek-potomek. Z punktu widzenia kladystyki, jeśli jakiś gatunek kopalny miałby być uważany za kandydata na ostatniego wspólnego przodka dwu innych gatunków, musi spełnić szereg warunków. Ale nawet wówczas, jest on tylko hipotetycznym (a nie rzeczywistym) wspólnym przodkiem. Jest tak, ponieważ wspólni przodkowie nie posiadają żadnych właściwości, które wyróżniają je jako takie. W związku z tym, proces odkrywania gatunków-przodków wydaje się dość problematyczny.

ENGELMANN & WILEY [1977] w oparciu o filozofię POPPERA [1959] twierdzą, że stwierdzenia o związkach przodek-potomek nie są obiektywne (i w związku z tym nienaukowe) ponieważ nie można ich sfalsyfikować (w rzeczywistości, jak uzasadniają, nie można ich ani potwierdzić, ani im zaprzeczyć). Ewentualnym dowodem potwierdzającym byłoby stwierdzenie plezjomorfii takiego przodka, a dowodem zaprzeczającym - znalezienie u niego jakiejś autapomorfii. Jednakże problemem jest to, że filogenetyczne relacje pomiędzy taksonami nie mogą bazować ani na plezjomorficznych, ani na autapomorficznych cechach.

HULL [1979] dyskutując ograniczenia kladyzmu wskazuje, że trudności w określaniu związków przodek-potomek nie są empiryczne, ale metodologiczne. Autor ten przypomina, że stwierdzenia o związkach siostrzanych również obarczone są ryzykiem błędu, ponieważ zależą one od polaryzacji cech i rozpoznawania gatunków. Wielu innych autorów również utrzymuje, że identyfikacja przodków jest zarówno możliwa, jak i testowalna poprzez uwzględnienie rozkładów cech (prymitywnych i zaawansowanych), danych stratygraficznych i biogeograficznych. Chociaż skamieniałości są rzadkie i fragmentaryczne, nie ma powodu przypuszczać, że nie znaleziono dotychczas gatunków będących wspólnymi przodkami innych gatunków. Ważną sprawą jest tu pytanie, który typ związków pokrewieństwa: przodek-potomek, czy siostrzany wydaje się bardziej prawdopodobny na podstawie wszystkich dowodów. Jak HULL [1979] podkreśla „być może przodkowie nie mogą być rozpoznani z tym samym stopniem pewności co związki siostrzane, ale różnica ta nie ma charakteru kontrastu między faktem a fantazją”.