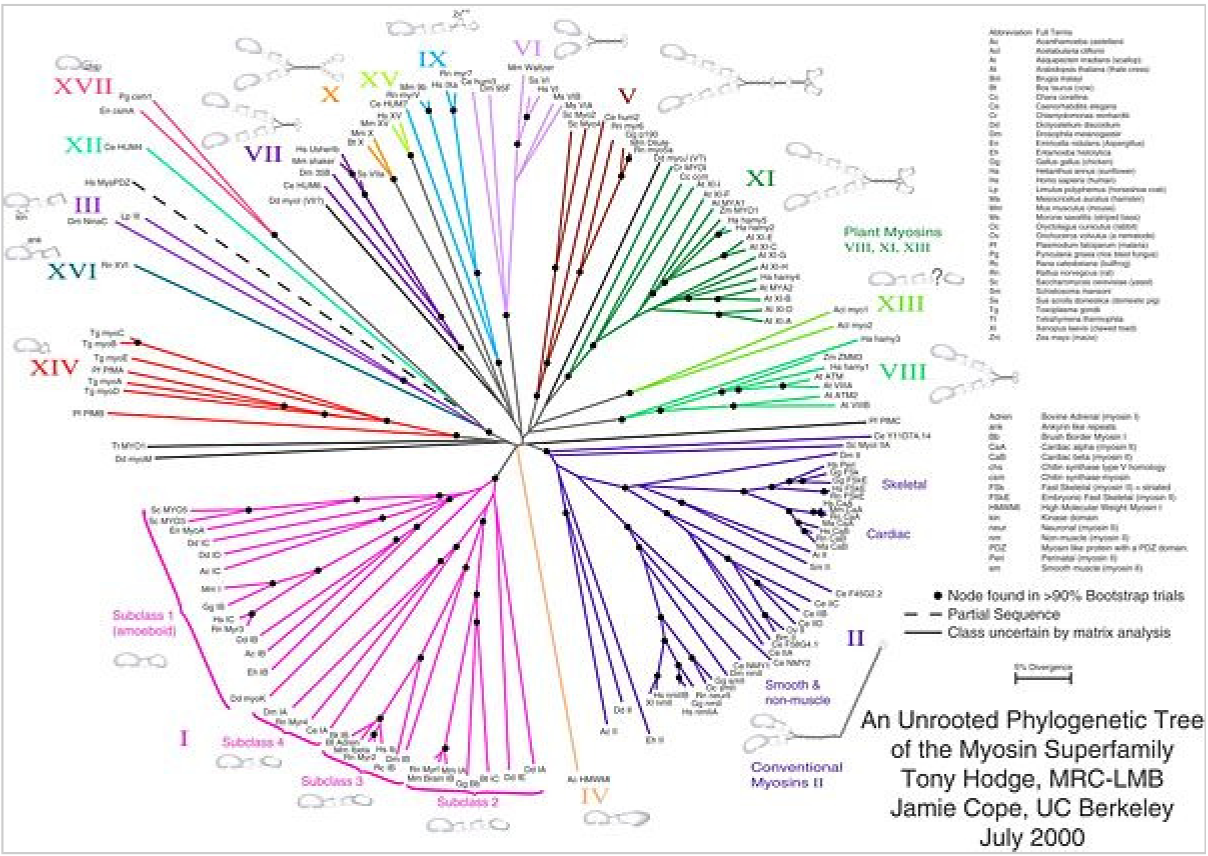
Cladistics

From Wikipedia, the free encyclopedia

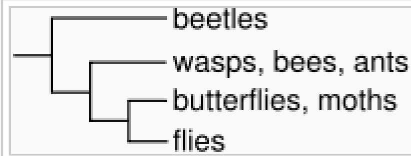
Cladistics the systematic classification of entities on the basis of the order of their divergence from ancestors. Cladistics is distinguished from other classification systems because it requires the notion of ancestry, and it places heavy emphasis on quantitative analysis. Cladistics originated in the field of biology but in recent years has found application in other disciplines. In the realm of biology, cladistics is used to generate tree-like diagrams called *cladograms* that represent the evolutionary tree-of-life.

A major contributor to this school of thought was the German entomologist Willi Hennig, who referred to it as **phylogenetic systematics** (Hennig, 1966). The word *cladistics* is derived from the ancient Greek *κλάδος*, *klados*, or "branch."



Unrooted cladogram of the myosin supergene family[1]

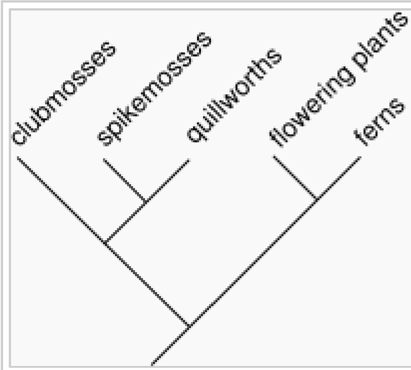
Cladograms



This cladogram shows the evolutionary relationship among various insect groups inferred from a dataset.

As the end result of a **cladistic analysis**, tree-like relationship-diagrams called "cladograms" are drawn up to show hypothesized relationships. [2] A cladistic analysis can be based on as much or as little information as the researcher selects. Modern systematic research is likely to be based on a wide variety of information, including DNA-sequences (so-called "molecular data"), biochemical data and morphological data.

In a **cladogram**, all organisms lie at the leaves[3], and each inner node is ideally binary (two-way). The two taxa on either side of a split are called *sister taxa* or *sister groups*. Each subtree, whether it contains one item or a hundred thousand items, is called a clade.



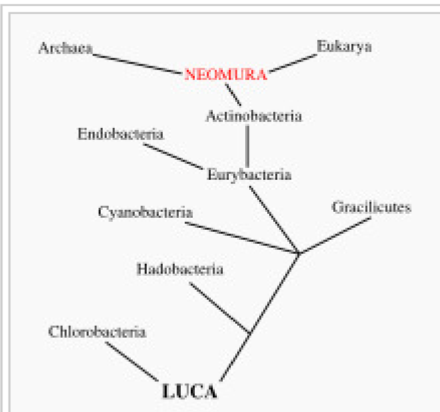
Cladograms are trees in the graph theoretic sense.

2-Way vs 3-Way Forks

Cladistics suggests that all forks in a cladogram be 2-way forks. Trees can be drawn with 3-way forks or 4-way forks, but many cladists discourage forks larger than 2-way. See Phylogenetic tree for more information about forking choices in trees.

Depth of a Cladogram

If a cladogram represents N species, the number of levels (the "depth") in the cladogram is on the order of $\log_2(N)$[4] . For example, if there are 32 species of deer, a cladogram representing deer will be around 5 levels deep (because $2^5=32$). A cladogram representing the complete tree of life, with about 10 million species, would be about 23 levels deep. This formula gives a lower limit: in most cases the actual depth will be a larger value because the various branches of the caldogram will not be uniformly deep. Conversely, the depth may be shallower if forks larger than 2-way forks are permitted.



A cladogram showing how Eukaryota and Archaea are more closely related to each other than to Bacteria. Note

Number of Distinct Cladograms

For a given set of species, the number of distinct rooted cladograms that can be drawn (ignoring which cladogram best matches the species characteristics) is[5]:

the 3-way fork in the middle of the cladogram.

Number of Species	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**N**
Number of Cladograms	1	3	15	105	945	10,395	135,135	2,027,025	34,459,425	1*3*5*7*...*(2N-3)

This exponential growth of the number of possible cladograms explains why manual creation of cladograms becomes very difficult when the number of species is large.

Represenation of Extinct Species in Cladograms

Cladistics makes no distinction between extinct and non-extinct species[6], and it is appropriate to include extinct species in the group of organisms being analyzed. Cladograms that are based on DNA/RNA generally do not include extinct species because DNA/RNA samples from extinct species are rare. Cladograms based on morphology, especially morphological characteristics that are preserved in fossils, are more likely to include extinct species.

Incorporation of Time in Cladogram

The cladogram tree has an implicit time axis[7], with time running forward from the base of the tree to the leaves of the tree. If the approximate date (for example, expressed as millions of years ago) of all the evolutionary forks were known, those dates could be captured in the cladogram. Thus, the time axis of the cladogram could be assigned a time scale (e.g. 1 cm = 1 million years), and the forks of the tree could be graphically located along the time axis. Such cladograms are called *scaled cladograms*.

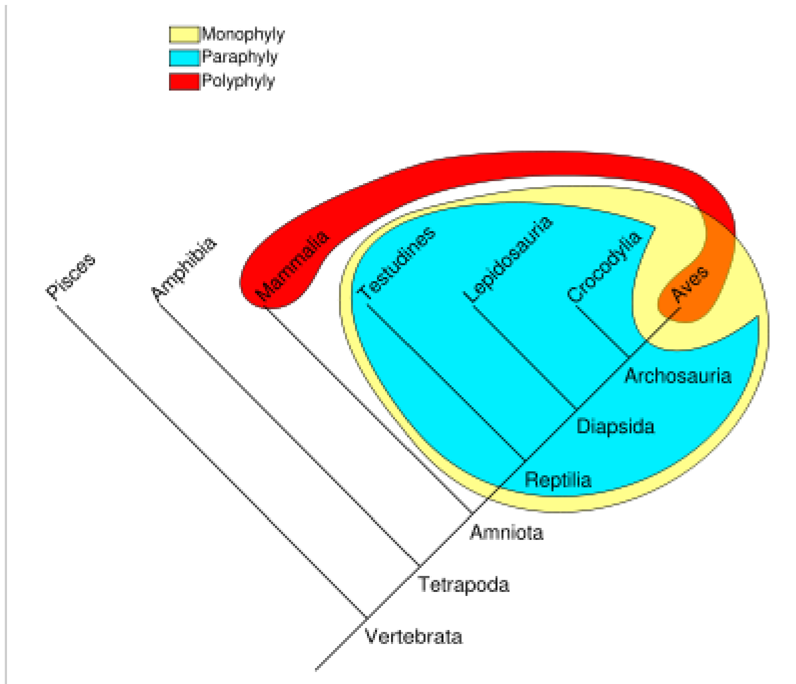
Many cladograms are not scaled along the time axis, for a variety of reasons. First, most cladistic methods that generate cladograms do not produce dating information (although DNA/RNA data often does provide approximate dates for the forks). Second even if the dating information were available, positioning the cladogram's forks along the time axis in proportion to their dates may cause the cladogram to become difficult to understand, or hard to fit within a human-readable format.

Terminology

- A *clade* is an ancestor species and all of its decencents
- A *monophyletic* group is a clade
- A *paraphyletic* group is a monophyletic group that excludes some of the descendants (e.g. reptiles)
- A *polyphyletic* group is a group consisting of members from two non-overlapping monophyletic groups (e.g. flying animals)
- An *outgroup* is an organism that is considered not to be part of the group in question, but is closely related to the group.
- A characteristic that is present in both the outgroups and in the ancestors is called a *plesiomorphy* (meaning "close form", also called an ancestral state).
- A characteristic that occurs only in later descendants is called an *apomorphy* (meaning "separate form", also called a "derived" state) for that group. Note: The adjectives *plesiomorphic* and *apomorphic* are used instead of "primitive" and "advanced" to avoid placing value-judgments on the evolution of the character states, since both may be advantageous in different circumstances. It is not uncommon to refer informally to a collective set of plesiomorphies as a *ground plan* for the clade or clades they refer to.
- A species or clade is *basal* to another clade if it holds more plesiomorphic characters than that other

clade. Usually a basal group is very species-poor as compared to a more derived group. It is not a requirement that a basal group be extant. For example, palaeodicots are basal to flowering plants.

- A clade or species located within another clade is said to be *nested* within that clade.

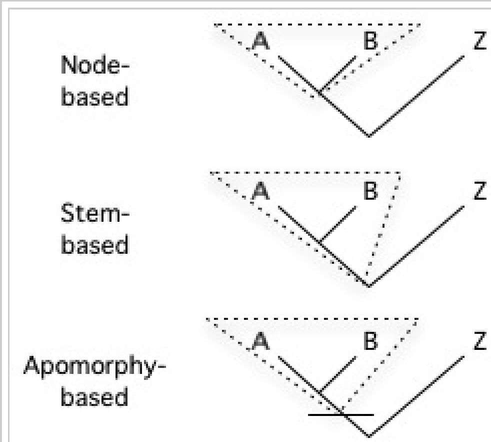


Phylogenetic groups, or *taxa*, can be monophyletic, paraphyletic, or polyphyletic.

Origin of the term "Cladistics"

Hennig's major book, even the 1979 version, does not contain the term 'cladistics' in the index. He referred to his own approach as phylogenetic systematics, implied by the book's title (Hennig, 1979). A review paper by Dupuis[8] observes that the term 'clade' was introduced in 1958 by Julian Huxley, 'cladistic' by Cain and Harrison in 1960 and 'cladist' (for an adherent of Hennig's school) by Mayr in 1965.

Three Definitions of Clade



There are three ways to define a clade for use in a cladistic taxonomy[9].

- *Node-based*: the most recent common ancestor of A and B and all its descendants.

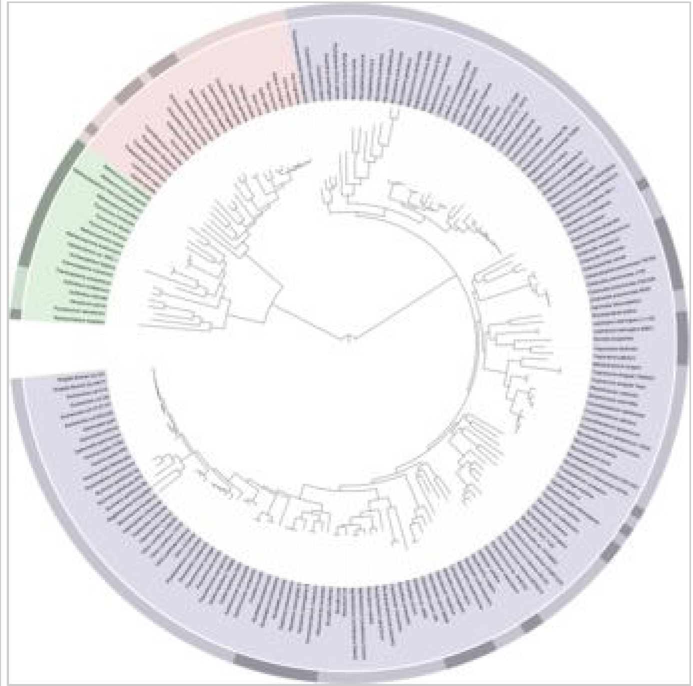
- *Stem-based*: all descendants of the oldest common ancestor of A and B that is not also an ancestor of Z.

- *Apomorphy-based*: the most recent common ancestor of A and B possessing a certain apomorphy (derived character), and all its descendants. This definition is generally discouraged by most cladists.

Cladistics Compared to Linnaean Taxonomy

Prior to the advent of Cladistics, most taxonomists used Linnaean taxonomy (a Kingdom-based approach) to organizing lifeforms. That traditional approach used several fixed levels of a hierarchy, such as Kingdom, Phylum, Class, Order, and Family. Cladistics does not use those terms, because one of the fundamental premises of Cladistics is that the evolutionary tree is very deep and very complex, and it is not meaningful to use a fixed number of levels.

Linnaean taxonomy insists that groups reflect phylogenies, but (in contrast to cladistics) allows both monophyletic and paraphyletic groups as taxa. Although, since the early 20th century, Linnaeian taxonomists generally attempted to make genus- and lower-level taxa monophyletic (even though the word might not have been used).



A highly resolved, automatically generated Tree Of Life, based on completely sequenced genomes [10]

The school of thought now known as cladistics took inspiration from the work of Willi Hennig, and since that time, there has been a spirited debate about the relative merits of cladistics versus Linneaen classification.[11]. Some of the debates that the cladists engaged in had been running since the 19th century, but they entered these debates with a new fervor[12], as can be learned from the *Foreword* to Hennig (1979) in which Rosen, Nelson and Patterson wrote the following:

> Encumbered with vague and slippery ideas about adaptation, fitness, biological species and natural selection, neo-Darwinism (summed up in the "evolutionary" systematics of Mayr and Simpson) not only lacked a definable investigatory method, but came to depend, both for evolutionary interpretation and classification, on consensus or authority. (Foreword, page ix).

Proponents of cladistics enumerate key distinctions between cladistics and Linnaean taxonomy as follows[13]:

Cladistics	**Linnean Taxonomy**
Handles arbitrarily-deep trees.	Prefers a tree about 10 levels deep.
Treats all levels of the tree as equivalent.	Treats each tree level uniquely. Uses special names (Family, Class, Order, etc) for each level. Often must invent new level-names (such as superorder, suborder, infraorder, parvorder, magnorder) to accomodate new discoveries.
Discourages naming or use of groups that are not monophyletic	Acceptable to name and use paraphyletic and polyphyletic groups
Primary goal is to reflect actual process of evolution	Primary goal is to group species based on morphological similarities
Assumes that the shape of the tree will change frequently, with new discoveries	New discoveries often require re-naming or re-levelling of Classes, Orders, and Kingdoms
Aims to be an objective process, free from personal interpretation	Some aspects are subjective (e.g., see the conflict between 3-Kingom, 5-Kingdom, 6-Kingdom schemes in Kingdom)
Taxa, once defined, are permanent (e.g. "Taxa X consists of the most recent common ancestor of species A and B, and its descendents)	Taxa can be renamed and eliminated (e.g. Insectivora is one of many taxa in the Linnaean system that have been eliminated).

Proponets of Linnaean taxonomy contend that it has some advantages over cladistics, such as[14]:

Cladistics	**Linnean Taxonomy**
Does not include a process for naming species	Includes a proces for giving unique names to species
Difficult to understand the essence of a clade, because clade definitions emphasize ancestory at the expense of meaningful characteristics	Taxa definitions based on tangible characteristics
Ignores sensible, clearly-defined paraphyletic groups such as reptiles	Permits clearly-defined groups such as reptiles
Difficult to determine if a given species is in a clade or not (e.g. if clade X is defined as "most recent common ancestor of A and B, and its decendents", then the only way to determine if Y is in the clade is to perform a complex evolutionary analysis)	Straightforward process to determine if a given species is in a taxon or not
Is limited to groupings that are rooted in evolution	Process does not depend on evolution
Limited to organisms that evolved by inherited traits; not applicable to organisms that evolved via complex gene-sharing or lateral transfer (e.g. viruses)	Applicable to all organisms, regardless of evolutionary mechanism

Cladistics Compared to Phenetics

For some decades in the mid-late 20th century, the a commonly used methodology was phenetics ("numerical taxonomy"). This can be seen as a precedessor[15] to some methods of today's cladistics (namely distance matrix methods like neighbor-joining), but made no attempt to resolve phylogeny, only similarities. Considered cutting-edge at its time as they were among the first bioinformatics applications, phenetic methods are today superseded by cladistic analyses due to their inability of phenetics to provide an evolutionary hypothesis, except by chance.

Step by Step Procedure

A typical procedure for generating a cladogram is[16]:

- 1) Gather data
- 2) Organize data
- 3) Create possible cladograms
- 4) Select cladogram most consistent with data

1) Gather data

A cladistic analysis begins with a certain set of information, usually a group of species, and a set of characteristics for each species.

To organize this information a distinction is made between *characters,* and *character states*. Consider the color of feathers, this may be blue in one species but red in another. Thus, "red feathers" and "blue feathers" are two character states of the character "feather-color."

Molecular vs. Morphological Data

The characteristics used to create a cladogram can be roughly categorized as either morphological (synapsid skull, warm-blooded, notochord, unicellular, etc) or molecular (DNA, RNA, or other genetic information). Prior to the advent of DNA sequencing, all cladistic analysis used morphological data.

As DNA sequencing has become cheaper and easier, molecular systematics has become a more and more popular way to reconstruct phylogenies[17]. Using a parsimony criterion is only one of several methods to infer a phylogeny from molecular data; maximum likelihood and Bayesian inference, which incorporate explicit models of sequence evolution, are non-Hennigian ways to evaluate sequence data. Another powerful method of reconstructing phylogenies is the use of genomic retrotransposon markers, which are thought to be less prone to the problem of reversion that plagues sequence data. They are also generally assumed to have a low incidence of homoplasies because it was once thought that their integration into the genome was entirely random; this seems at least sometimes not to be the case however.

Ideally, morphological, molecular and possibly other (behavioral etc.) phylogenies should be combined into an analysis of **total evidence**: all have different intrinsic sources of error. For example, character convergence (homoplasy) is much more common in morphological data than in molecular sequence data, but character reversions that cannot be noticed as such are more common in the latter (see long branch attraction). Morphological homoplasies can usually be recognized as such if character states are defined with enough attention to detail.

2) Organize Data

The researcher decides which character states were present *before* the last common ancestor of the species group (*plesiomorphies*) and which were present *in* the last common ancestor (*synapomorphies*) by considering one or more *outgroups.* This makes the choice of an outgroup an important task, since this choice can profoundly change the topology of a tree. Note that only synapomorphies are of use in characterising clades.

Avoid Homoplasies

A homoplasy is a character that is shared by multiple species due to some cause *other* than common ancestry [18]. Typically, homoplasies occur due to convergent evolution. Use of homoplasies when building a cladogram is sometimes unavoidable, but is to be avoided when possible.

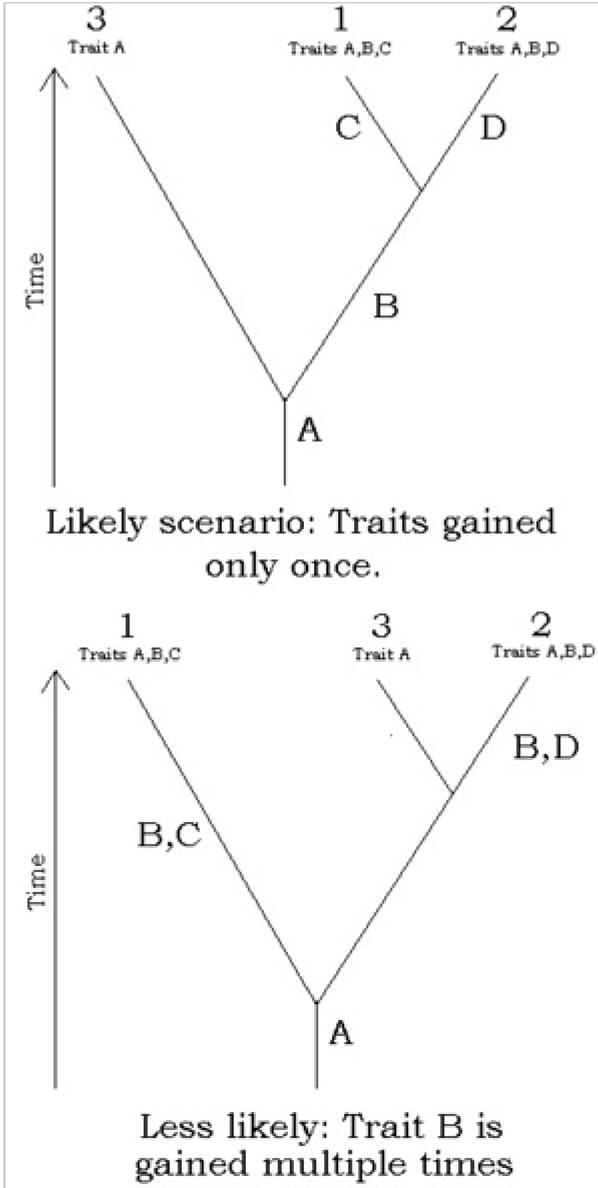
A well-known example of homoplasy due to convergent evolution would be a character "presence of wings". Though the wings of birds, bats, and insects serve the same function, each evolved independently, as can be seen by their anatomy. If a bird, bat, and a winged insect were scored for the character "presence of wings", a homoplasy would be introduced into the dataset, and this confounds the analysis, possibly resulting in a false evolutionary scenario.

Homoplasies can often be avoided outright in morphological datasets by defining characters more precisely and increasing their number: in the example above, e.g. utilizing "wings supported by bony endoskeleton" and "wings supported by chitinous exoskeleton" as characters would avoid the homoplasy. When analyzing "supertrees" (datasets incorporating as many taxa of a suspected clade as possible), it may become unavoidable to introduce character definitions that are unprecise, as otherwise the characters might not apply at all to a large number of taxa. The "wings" example would be hardly useful if attempting a phylogeny of all Metazoa as most of these don't have wings at all. Cautious choice and definition of characters thus is another important element in cladistic analyses. With a faulty outgroup and/or character set, no method of evaluation is likely to produce a phylogeny representing the evolutionary reality.

3) Consider All Possible Cladograms

Next, different possible cladograms are drawn up and evaluated. When there are just a few species being organized, it is possible to do this step manually, but most cases require a computer program, such as PAUP*.

Computer programs that generate Cladograms use algorithms that are very computationally intensive[19], because the cladogram algorithm is NP-hard. The programs that generate cladograms are trying to minimize a metric. Because of the astronomical number of possible cladograms, minimization programs cannot guarantee that the solution is the overall best solution. A non-optimal cladogram will be selected if the program settles on a "local minimum" rather than the desired "global minimum". To help solve this problem, many cladogram programs use a simulated annealing approach to increase the likelihood that the selected cladogram is the optimal one. See Phylogenetic tree for more information about tree-generating computer programs.



Letters indicate traits, species are numbered. Branchings show divergence from a common ancestor. The position of a letter indicates on which branch it evolved.

Distance-matrix methods such as neighbor-joining or UPGMA, which calculate genetic distance from multiple sequence alignments, are simplest to implement, but do not invoke an evolutionary model. Many sequence alignment methods such as ClustalW produce both sequence alignments and phylogenetic trees. Methods including maximum parsimony, maximum likelihood and Bayesian inference apply an explicit model of evolution to phylogenetics. Identifying the optimal tree using many of these techniques is NP-hard, so heuristic search and optimization methods are used in combination with tree-scoring functions to identify a reasonably good tree that fits the data.

Tree-building methods can be assessed on the basis of several criteria:[20]

- efficiency (how long does it take to compute the answer, how much memory does it need?)
- power (does it make good use of the data, or is information being wasted?)
- consistency (will it converge on the same answer repeatedly, if each time given different data for the same model problem?)
- robustness (does it cope well with violations of the assumptions of the underlying model?)
- falsifiability (does it alert us when it is not good to use, i.e. when assumptions are violated?)

4) Select Best Cladogram

Many cladograms are possible for any given set of taxa, but one is chosen based on the principle of parsimony: the most compact arrangement, that is, with the fewest character state changes (synapomorphies), is the hypothesis of relationship accept here (see Occam's razor for a discussion of the principle of parsimony and possible complications).

The chosen cladogram is selected because it is optimal as defined by some metric. There are several metrics available to determine which cladogram - of all possible cladograms - best fits the raw data. Least Squares is one commonly used metric: the selected cladogram is the cladogram with the smallest error metric, where the error metric is the sum of the squares of all the differences between the cladogram and the data. Using different metrics can - in some circumstances - yield different optimal cladograms for the same data set.

An issue faced when using a computer program is that some programs can be sensitive to the order the input data (list of species and their characteristics) is presented. Inputting the data in various ways can cause such programs to produce multiple "best" cladograms. In these cases, the users should input the data multiple times and compare the results.

Preference for Monophyletic Groups

A clade is a complete subtree of a cladogram. Some groups of organisms are not clades, for example paraphyletic groups and polyphyletic groups. A clade is also called a monophyletic group.

Following Hennig, cladists argue that paraphyly is as harmful as polyphyly. The idea is that monophyletic groups can be defined objectively, in terms of common ancestors or the presence of synapomorphies. In contrast, paraphyletic and polyphyletic groups are both defined based on key characters, and the decision of which characters are of taxonomic import is inherently subjective. Many argue that they lead to "gradistic" thinking, where groups advance from "lowly" grades to "advanced" grades, which can in turn lead to teleology. In evolutionary studies, teleology is usually avoided because it implies a plan that cannot be empirically demonstrated.

Going further, some cladists argue that ranks for groups above species are too subjective to present any meaningful information, and so argue that they should be abandoned. Thus they have moved away from Linnaean taxonomy towards a simple hierarchy of clades. The validity of this argument hinges crucially on how often in evolution gradualist near-equilibria are punctuated. A quasi-stable state will result in phylogenies, which may be all but unmappable onto the Linnaean hierarchy, whereas a punctuation event that balances a taxon out of its ecological equilibrium is likely to lead to a split between clades that occurs in comparatively short time and thus lends itself readily for classification according to the Linnaean system.

Other evolutionary systematists argue that all taxa are inherently subjective, even when they reflect evolutionary relationships, since living things form an essentially continuous tree. Any dividing line is artificial, and creates both a monophyletic section above and a paraphyletic section below. Paraphyletic taxa are necessary for classifying earlier sections of the tree – for instance, the early vertebrates that would someday evolve into the family Hominidae cannot be placed in any other monophyletic family. They also argue that paraphyletic taxa provide information about significant changes in organisms' morphology, ecology, or life history – in short, that both taxa and clades are valuable but distinct notions, with separate purposes. Many use the term *monophyly* in its older sense, where it includes paraphyly, and use the alternate term *holophyly* to describe clades (*monophyly* in Hennig's sense). As an unscientific rule of thumb, if a distinct lineage that renders the containing clade paraphyletic has undergone marked adaptive radiation and collected many synapomorphies - especially ones that are radical and/or unprecedented -, the paraphyly is usually not considered a sufficient argument to prevent recognition of the lineage as distinct under the Linnaean system (but it *is* by definition sufficient in phylogenetic nomenclature). For example, as touched upon briefly above, the Sauropsida ("reptiles") and the Aves (birds) are both ranked as a Linnaean class, although the latter are a highly derived offshoot of some forms of the former which themselves were already quite advanced.

Phylocode Approach to Naming Species

A formal code of phylogenetic nomenclature, the PhyloCode[21], is currently under development for cladistic taxonomy. It is intended for use by both those who would like to abandon Linnaean taxonomy and those who would like to use taxa and clades side by side. In several instances (see for example Hesperornithes) it has been employed to clarify uncertainties in Linnaean systematics so that in combination they yield a taxonomy that is unambiguously placing the group in the evolutionary tree in a way that is consistent with current knowledge.

Applying Cladisitics to Other Disciplines

The processes used to generate cladograms are not limited to the field of biology[22]. The generic nature of cladistics means that cladistics can be used to organize groups of items in many different realms. The only requirement is that the items have chararacteristics that can be identified and measured.

For example, one could take a group of 200 spoken languages, measure various characteristics of each language (vocabulary, phonemes, rhythms, accents, dynamics, etc) and then apply a cladogram algorithm to the data. The result will be a tree that may shed light on how, and in what order, the languages came into existence.

Thus, cladistic methods have recently been usefully applied to non-biological systems, including determining language families in historical linguistics, culture, and filiation of manuscripts in textual criticism.

Footnotes

1. ^ Hodge T, Cope M (2000). "A myosin family tree (http://jcs.biologists.org/cgi/content/full/113/19/3353)". *J Cell Sci* **113 Pt 19**: 3353-4. PMID 10984423 (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=10984423).
2. ^ See, for example, pp. 45, 78 and 555 of Joel Cracraft and Michael J. Donaghue, eds. (2004). *Assembling the Tree of Life*. Oxford, England: Oxford University Press.
3. ^ Albert, Victor (2006). *Parsimony, Phylogeny, and Genomics*. Oxford University Press, 3-55. ISBN 0199297304.
4. ^ Aldous, David (1996), "Probability Distributions on Cladograms", *Random Discrete Structures*, Springer, p. 13
5. ^ Lowe, Andrew (2004). *Ecological Genetics: Design, Analysis, and Application*. Blackwell Publishing, 164. ISBN 1405100338.
6. ^ Scott-Ram, N. R. (1990). *Transformed Cladistics, Taxonomy and Evolution*. Cambridge University Press, 83. ISBN 0521340861.
7. ^ Freeman, Scott (1998). *Evolutionary Analysis*. Prentice Hall, 380. ISBN 0135680239.
8. ^ Dupuis, Claude (1984). "Willi Hennig's impact on taxonomic thought". *Annual Review of Ecology and Systematics* **15**: 1-24. ISSN 0066-4162 (http://worldcat.org/issn/0066-4162).
9. ^ de Queiroz, K. and J. Gauthier (1994). "Toward a phylogenetic system of biological nomenclature". *Trends in Research in Ecology and Evolution* **9** (1): 27-31.
10. ^ Letunic, I (2007). "Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=17050570)" (Pubmed). *Bioinformatics* **23(1)**: 127-8.
11. ^ Benton, M. (2000). "Stems, nodes, crown clades, and rank-free lists: is Linnaeus dead?". *Biological Reviews* **75** (4): 633-648.
12. ^ Hull, David (1988). *Science as a Process*. University of Chicago Press, 232-276. ISBN 0226360512.
13. ^ Hennig, Willi (1975). "'Cladistic analysis or cladistic classification': a reply to Ernst Mayr". *Systematic Zoology* **24**: 244-256.
14. ^ Mayr, Ernst (1976). *Evolution and the diversity of life (Selected essays)*. Cambridge, MA: Harvard Univ. Press. ISBN 0-674-27105-X.
15. ^ Mayr, Ernst (1982). *The growth of biological thought: diversity, evolution and inheritance*. Cambridge, MA: Harvard Univ. Press, 221. ISBN 0-674-36446-5.
16. ^ DeSalle, Rob (2002). *Techniques in Molecular Systematics and Evolution*. Birkhauser. ISBN 376436257X.
17. ^ Hillis, David (1996). *Molecular Systematics*. Sinaur. ISBN 0878932828.

18. ^ West-Eberhard, Mary (2003). *Developmental Plasticity and Evolution*. Oxford Univ. Press, 353-376. ISBN 0195122356.
19. ^ Hodkinson, Trevor (2006). *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa*. CRC Press, 61-128. ISBN 0849395798.
20. ^ Penny, D., Hendy, M. D. & M. A. Steel. 1992. Progress with methods for constructing evolutionary trees. *Trends in Ecology and Evolution* **7**: 73-79.
21. ^ Pennisi, E. (2001). "Evolutionary Biology: Preparing the Ground for a Modern 'Tree of Life'". *Science* **293**: 1979-1980.
22. ^ Mace, Ruth (2005). *The Evolution of Cultural Diversity: A Phylogenetic Approach*. Routledge Cavendish. ISBN 1844720993.

See also

- Bauplan
- Bioinformatics
- Biomathematics
- Clade
- Coalescent theory
- Dendrogram
- EDGE of Existence Programme
- Evolution of Mollusca for a cladistic illustration
- Evolutionary tree
- Last common ancestor
- Important publications in phylogenetics
- Language family
- Maximum parsimony
- Molecular phylogeny
- PhyloCode
- Phylogenetics
- Phylogenetic tree
- Phylogenetic network
- Phylogenetics software
- Phylogeography
- Phylogenetic comparative methods
- Scientific Classification
- Systematics

References

- Ashlock, Peter D. (1974). "The uses of cladistics". *Annual Review of Ecology and Systematics* **5**: 81-99. ISSN 0066-4162 (http://worldcat.org/issn/0066-4162).
- Cuénot, Lucien (1940). "Remarques sur un essai d'arbre généalogique du règne animal". *Comptes Rendus de l'Académie des Sciences de Paris* **210**: 23-27. Available free online at http://gallica.bnf.fr (No direct URL). This is the paper credited by Hennig (1979) for the first use of the term 'clade'.
- Cavalli-Sforza, L.L. and A.W.F. Edwards (Sep., 1967). "Phylogenetic analysis: Models and estimation procedures (http://links.jstor.org/sici?sici=0014-3820%28196709%2921%3A3%3C550%3APAMAEP%3E2.0.CO%3B2-I)". *Evol.* **21** (3): 550-570.
- de Queiroz, Kevin (1992). "Phylogenetic taxonomy". *Annual Review of Ecology and Systematics* **23**: 449–480. ISSN 0066-4162 (http://worldcat.org/issn/0066-4162).
- Dupuis, Claude (1984). "Willi Hennig's impact on taxonomic thought". *Annual Review of Ecology and Systematics* **15**: 1-24. ISSN 0066-4162 (http://worldcat.org/issn/0066-4162).
- Felsenstein, Joseph (2004). *Inferring phylogenies*. Sunderland, MA: Sinauer Associates. ISBN 0-87893-177-5.
- Hamdi, Hamdi; Hitomi Nishio, Rita Zielinski and Achilles Dugaiczyk (1999). "Origin and phylogenetic distribution of *Alu* DNA repeats: irreversible events in the evolution of primates". *Journal of Molecular Biology* **289**: 861–871. PMID 10369767 (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=10369767).
- Hennig, Willi (1950). *Grundzüge einer Theorie der Phylogenetischen Systematik*. Berlin: Deutscher Zentralverlag. .
- Hennig, Willi (1982). *Phylogenetische Systematik (ed. Wolfgang Hennig)*. Berlin: Blackwell Wissenschaft. ISBN 3-8263-2841-8.
- Hennig, Willi (1975). "'Cladistic analysis or cladistic classification': a reply to Ernst Mayr". *Systematic Zoology* **24**: 244-256. The paper he was responding to is reprinted in Mayr (1976).
- Hennig, Willi (1966). *Phylogenetic systematics (tr. D. Dwight Davis and Rainer Zangerl)*. Urbana, IL: Univ. of Illinois Press (reprinted 1979 and 1999). ISBN 0-252-06814-9. Translated from manuscript and so never published in German.
- Hull, David L. (1979). "The limits of cladism". *Systematic Zoology* **28**: 416-440.
- Kitching, Ian J.; Peter L. Forey, Christopher J. Humphries and David M. Williams (1998). *Cladistics: Theory and practice of parsimony analysis*, 2nd ed., Oxford University Press. ISBN 0-19-850138-2.
- Luria, Salvador; Stephen Jay Gould and Sam Singer (1981). *A view of life*. Menlo Park, CA: Benjamin/Cummings.

ISBN 0-8053-6648-2.
- Mayr, Ernst (1982). *The growth of biological thought: diversity, evolution and inheritance*. Cambridge, MA: Harvard Harvard Univ. Press. ISBN 0-674-36446-5.
- Mayr, Ernst (1976). *Evolution and the diversity of life (Selected essays)*. Cambridge, MA: Harvard Univ. Press. ISBN 0-674-27105-X. Reissued 1997 in paperback. Includes a reprint of Mayr's 1974 anti-cladistics paper at pp. 433-476, "Cladistic analysis or cladistic classification." This is the paper to which Hennig (1975) is a response.
- Patterson, Colin (1982). "Morphological characters and homology". Joysey, Kenneth A; A. E. Friday (editors) *Problems in Phylogenetic Reconstruction*, London: Academic Press.
- Rosen, Donn; Gareth Nelson and Colin Patterson (1979), Foreword *provided for Hennig (1979)*
- Shedlock, Andrew M; Norihiro Okada (2000). "SINE insertions: Powerful tools for molecular systematics". *Bioessays* **22**: 148–160. ISSN 0039-7989 (http://worldcat.org/issn/0039-7989). PMID 10655034 (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=10655034).
- Sokal, Robert R. (1975). "Mayr on cladism -- and his critics". *Systematic Zoology* **24**: 257-262.
- Swofford, David L.; G. J. Olsen, P. J. Waddell and David M. Hillis (1996). "Phylogenetic inference", in Hillis, David David M; C. Moritz and B. K. Mable (editors): *Molecular Systematics*, 2. ed., Sunderland, MA: Sinauer Associates. ISBN 0-87893-282-8.
- Wiley, Edward O. (1981). *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*. New York: Wiley Interscience. ISBN 0-471-05975-7.
- Zwickl DJ, Hillis DM (2002). "Increased taxon sampling greatly reduces phylogenetic error". *Systematic Biology* **51**: 588-598.

External links

- The Compleat Cladist (pdf) (http://www.amnh.org/learn/pd/fish_2/pdf/compleat_cladist.pdf)
- Tree of Life illustration (http://tellapallet.com/tree_of_life.htm) - A high-level cladogram showing the complete tree of life.
- Example of cladistics used in textual criticism (http://rjohara.net/darwin/files/bmcr)
- Journey into Phylogenetic Systematics (http://www.ucmp.berkeley.edu/clad/clad4.html)
- Phylogenetics Primer (http://www.talkorigins.org/faqs/comdesc/phylo.html) from Talk.Origins
- Willi Hennig Society (http://www.cladistics.org/)
- Cladistics: The International Journal of the Willi Hennig Society (http://www.blackwellpublishing.com/journal.asp?ref=0748-3007) (ISSN 0748-3007 (http://dispatch.opac.ddb.de/DB=1.1/LNG=EN/CMD?ACT=SRCHA&IKT=8&TRM=0748-3007))
- Phylogenetic inferring on the T-REX server (http://www.trex.uqam.ca/)
- A list of cladogram-generating programs (http://evolution.genetics.washington.edu/phylip/software.pars.html)