

## Phylogenetics: An over-nourished tree?

K. N. Ganeshiah, K. Chandrashekhara, R. Uma Shaanker and A. R. V. Kumar

Promotion pays. Be it a soft drink or software its eventual survival in the market depends heavily on promotion. This is true for ideas in science as well. In science, as elsewhere, ideas are owned by those who propose them and rightly these owners are prompted to project and promote their ideas. Especially in situations where controversies and ambiguities persist despite repeated prescriptions, the ideas that are better nourished and promoted through journals have a longer life. The more an idea gets promoted the more its followers, who in turn add to its promotion (by citing it) and hence to the success (fame) of the promoter and of the idea. The feedback chain thus makes the promoter stronger to further promote his ideas over those that are less cared and undernourished. Phylogenetics appears to be an approach in systematics blossoming due to such promotion.

Of course this does not mean that even the worst ideas can get promoted in science. But this is certainly true in a number of situations where universally accepted philosophy does not, and perhaps cannot, exist and where several approaches are equally good or worse. One such branch is the systematics in biology.

### Systematics: A controversial field of biology

The turmoil in systematics arising out of several conflicting philosophies is being

repeatedly expressed by several workers over decades:

'Taxonomy is an ever-changing, controversial and exciting field of biology' (Quicke<sup>1</sup>).

'The field of systematics has been in considerable turmoil as various investigators developed different methods of classification and argued their merits. *I guarantee that no one method or view has all the good points*' (emphasis added) (Fitch<sup>2</sup>).

'Recent years have witnessed increasing dissatisfaction with the principles and practices of biological classification' (Sokal and Sneath<sup>3</sup>).

In fact from the beginning, systematics has had a very chequered history and there has been a perennial deficit of objectivity in the techniques of biological classification. The field has seen repeated attempts to inject objectivity (e.g. numerical or phenetic taxonomy<sup>3</sup>); however these attempts have either been continuously resisted by the conservative taxonomists and/or are multiply assassinated by others in an attempt to promote their own approach. Cladistic or phylogenetic taxonomy (see Boxes 1 and 2) is the most recent approach that acted as a ruthless mercenary of phenetic taxonomy (Box 2).

### Phylogenetics: An emerging field of systematics

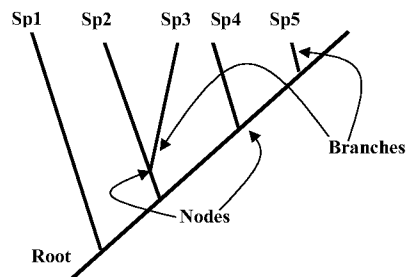
In the recent Avalanche meeting<sup>4</sup> (sixth in the series since 1989) held on 29th and 30th May 2000, a team of 35 participants met at Sangam along the banks of river Cauvery near Kanakapura, to discuss the present status of phylogenetics. The topic for this meeting was prompted by the recent spurt in the literature indicating a mad rush for constructing phylogenetic trees (PGTs) for almost all groups of organisms.

The most recent experience of this spurt of interest was seen at the International Botanical Congress, St. Louis, Missouri, USA during August 1999 (ref. 5). Among the 2500 or so posters presented, easily more than half were on construction of PGTs for a range of organisms, with several workers independently building them for the same set of organisms! Similarly oral sessions were also crowded with papers dealing with the PGTs and on the procedures to construct them. The purpose of the Avalanche meeting was to understand if this spurt is in any way justified.

Besides, most journals otherwise dealing with the classical (?) subjects of botany and zoology, of late, are also dedicating their pages to PGTs. Significantly, and surprisingly, this spurt does not seem to have caught up in India! There does not appear to be any strong

#### Box 1. Phylogenetic trees: Their structure and construction

Phylogenetic tree of a group of species is a graphical representation of their relation by evolutionary descent and hence has species as the tips of the branches. Branches represent the lineage and the node represents the point at which the lineage emerges. The distances between the nodes represent the evolutionary time separating the two lineages. These trees which are only the estimates of the actual phylogeny are constructed by delineating the derived from the ancient characters of the taxa under study (ingroup) and by following certain principles or rules. Among several protocols proposed for tree construction, a few such as parsimony and maximum likelihood are more frequently used. The parsimony method is based on the assumption that the tree that requires minimum steps or evolutionary changes represents the best estimate of phylogeny. The maximum likelihood approach attempts to estimate the likelihood of getting the observed data for a given tree and accordingly identifies the tree that has a high likelihood of arising from the data. For more details see B-Rao and Majumdar<sup>6</sup> and Eggleton and Vane-Wright<sup>11</sup>.



## Box 2. Phenetic and phylogenetic systems

Phenetic system is another term for numerical taxonomy<sup>19,20</sup> and was proposed by Sokal and Sneath during the 1960s. Their approach was aimed more at introducing a certain objectivity into the otherwise authoritarian and highly personalized field of traditional taxonomy. Perhaps the beginning of the numerical approach to taxonomy can be traced to Adanson who during the 1700s identified that the traits are continuous and hence categorical grouping requires an identification of ranges.

The phylogenetic approach to taxonomy (often referred to as cladistic taxonomy) began with Hennig during the 1950s but remained untapped almost for two to three decades. His suggestion was that the taxonomic organization of living systems can be more natural if we base it on their evolutionary history.

The phenetic approach relies on the distances and similarities among the taxa on a range of traits. The traits can be very exhaustive and often weighted depending upon their importance in discriminating the taxa. On the other hand, the phylogenetic approach lays emphasis on selecting the traits; derived traits are more important than the ancestral ones. Further, it also emphasizes the need to base the tree construction on homology (similarity due to common descent) but to avoid using the traits evolved by homoplasy (those evolved due to convergence among distant taxa or parallelly evolved among related taxa and those that have reverted).

Both phenetic and phylogenetic analyses end up in trees called dendrogram and phylogenetic trees, respectively. The former need not represent the phylogeny except by default or chance, while the latter is expected to represent the phylogeny.

Both the approaches seem to frequently result in the classificatory systems that are in direct conflict with the traditional Linnean system. Phenetic classification assorts the taxonomic groups based on the similarities among them in the characters studied and conflicts with the Linnean system because the latter uses a few key traits. The phylogenetic system has conflicts arising because the hierarchy of classification (genera, families, etc.) of the Linnean system becomes incompatible with that built by relation due to descent.

academic programme in this area (but see B-Rao and Majumdar<sup>6</sup>). The task of the meeting was also to evaluate the need for such a programme in the curriculum of our education as a major stream of taxonomy. This article attempts to reflect a summary 'feeling' of the participants on the present status and developments in phylogenetics.

### Hopes and promises of the phylogenetic approach

Promoters of phylogenetic analysis have been promising a lot. It is expected to reorganize the classification of the entire living system in an evolutionarily (and hence naturally) meaningful order; link the organisms on a tree of life that depicts their natural relation by descent. In other words, it is expected to offer us a family tree of life in which the evolutionary path and time distances between the organisms are depicted. The emerging cladistic trees (see Box 1) are also expected to offer an objective alternative for the traditional taxonomy (that does not assume any evolutionary relationship between the organisms). Finally, the phylogeneticists also claim to arrive at a solution for resolving the controversial concept of the 'species'<sup>1</sup>.

There were sufficient indications in the meeting that the emergence of phylogenetics as an alternative to traditional

taxonomic classification is not as much due to its own strength and objectivity, but rather because of the weaknesses of the traditional taxonomy. These weaknesses of traditional taxonomy are well known, purely because being the oldest, it is perhaps the most trampled field. On the other hand, phylogenetics being a relatively new area, is growing by capitalizing on the promised objectivity and hopes hidden in the wilderness of its unknown territory. While the spin-offs of the developments in phylogenetics are harvested in several other areas as well (e.g. predicting the emerging variants of influenza virus<sup>7</sup>), its growth as an alternative to traditional classificatory procedures, appears largely due to promotion.

### Breaks in the phylogenetic tree

In a way, the fault lines of phylogenetic monolith were visible from the very beginning. Phylogenetic trees are built on the premise that organisms with a common descent share common traits<sup>8</sup> – simple premise that has no problem in theory. But, in practice, it is perhaps the most complex premise to use. It implies that there are certain traits in a species that come from the ancestors and others that are derived. The challenge while constructing a tree is to define these ancestral and derived traits. Traits in organisms mutate continuously from one state to the other, revert, and diverge away

from each other. They can also arise by convergent evolution (as haemoglobin in legumes and mammals), or by parallel evolution (see Box 2). Hence the exercise of ascertaining a given trait as ancestral or derived suffers from a lot of ambiguity<sup>1,6</sup>. Identifying them from fossil records is not always possible, since fossils do not retain all the traits<sup>1</sup>.

But the issue of character selection is too crucial for the tree construction and the trees constructed are only as unambiguous as the selection of traits can be<sup>1,6,9</sup>. Phylogenetic trees constructed using molecular data differ widely from those built on morphological traits; those built on adaptive traits differ from those built on structural features<sup>10,11</sup>. To complicate matters further, it is difficult to decide if trees built by pooling all the types of data are better than those built using carefully selected sets of data<sup>12</sup>.

In fact there are more basic problems with respect to the selection of traits. It has been shown that the quality of information contained in the characters is likely to be influenced by the shape of the original tree and hence the information base we use for traits may be more confounded<sup>13</sup>. Thus 'these revelations should remind evolutionary investigators to question rigorously the output of the phylogenetic estimation software. The evolutionary histories we derive must always be considered provisional'<sup>13</sup>.

The ambiguity in PGTs arises also due to the protocols used in constructing them. There are a number of arguments, algorithms and associated programmes to construct trees but none are unequivocally accepted as each of them have their own limitations, problems or at least uncertainties<sup>6,11</sup>. For instance, the parsimony approach to build the PGTs is based on the assumption that best estimates of trees are those that require least number of steps or changes. Though more than 60 per cent of the trees built are based on this technique, there does not appear to be a strong justification for this assumption. Similarly, other techniques are not generally agreed upon as any better. In fact, even the assumptions inherent in different methods of phylogenetic tree construction are not clearly justified and it is often opined that the 'popularity of any particular algorithm has depended more on its easy availability than on scientific criteria'<sup>6</sup>.

Thus, there appears to be no way of knowing if we are capable of even constructing the nearest estimate of the PGTs. The structure and topology of PGTs change based on the traits used, organisms included for the study and the methodology followed<sup>14,15</sup>. Therefore, the resulting phylogenetic trees are not as robust as they appear<sup>11</sup> and are only as ambiguous as the phyletic dendrograms developed during the 1960s by the numerical taxonomists but ruthlessly rejected by the powerful sweep of the phylogeneticists. Strangely enough based on these ambiguous trees, phylogeneticists have begun to derive a supposedly more stable system of classification of the organisms and are even proposing a new 'Phylocode' for naming the taxa<sup>16</sup>. Besides adding more problems to the already confused state of taxonomic names, it is not clear how stable this system could be when the trees constructed themselves are unstable and suffer from unlimited range of ambiguities.

### The over-promoted approach

While it is being realized that it is difficult to arrive at the unambiguous trees and that, for most groups there are no readily available trees, the promoters of phylogenetic analysis have begun to impose the phylogenetic approach on to different spheres of biology. We had one such difficult experience. In a publication

from our group, we proposed an index of measuring biological diversity based on the traditional system of classification and using the existing techniques of measuring biodiversity<sup>17</sup>. One of the referees, perhaps with allegiance to phylogenetic analysis, had almost thrown away our suggestion as less sophisticated than his own which was based on PGTs. Only after our questioning of the availability and unambiguity of such PGTs did our manuscript pass through that critical acceptance gate of the publication fort of the journal. But not many would be as lucky as we were and the phylogenetic bandwagon can drum down those seeking alternate approaches and suppress their voices rendering them inaudible.

In summary, this Avalanche meeting created sufficient panic on the problems of blindly welcoming the phylogenetic trees as alternatives to traditional and phenetic taxonomy. Phylogenetics thus seems to be a tree over-nurtured by its guardians and suffers from all the syndromes of weakness its proponents identified with other approaches in systematics. There is a need for a thorough introspection before hitting the last nails on the coffins of numeric taxonomy that began almost two hundred years ago<sup>18</sup> and propagated during the 1960s and 1970s by Sokal and Sneath<sup>3</sup>. As viewed by Sneath<sup>20</sup>, phylogenetics should perhaps find a welcome place as another branch of numerical taxonomy and not as an overarching field beyond and above other schools of systematics.

1. Quicke, D. L. J., *Principles and Techniques of Contemporary Taxonomy*, Black Academic and Professional, London, 1993, p. 309.
2. Fitch, W. M., in *Cladistics: Perspectives on the Reconstruction of Evolutionary History* (eds Duncan, T. and Stuessy, T. F.), Columbia University Press, New York, 1984, pp. 221–252.
3. Sokal, R. R. and Sneath, P. H. A., *Numerical Taxonomy: The Principles and Practices of Numerical Classification*. W. H. Freeman & Co., 1963.
4. Avalanche meetings are a series of discussion workshops on the contemporary issues of evolutionary ecology organized by the 'Tuesday Group' of the University of Agricultural Sciences, Bangalore. The series began during 1989 at a hill resort called Avalanche (and hence the name), near Ooty and the meetings are held every one and a half to two years. The meetings are uniquely characterized by the choice of the places that are serene,

isolated, and poorly connected by the communicating channels so that the participants, by default, indulge in a committed, healthy and informal discussion on the topic chosen.

5. Ganeshiah, K. N., *Curr. Sci.*, 1999, **77**, 739–741.
6. B-Rao, C. and Majumdar, K. C., *J. Biosci.*, 1999, **24**, 121–137.
7. Bush, R. M., Bender, C. A., SubbaRao, K., Cox, N. J. and Fitch, W. M., *Science*, 1999, **286**, 1921–1925.
8. Hennig, W., *Grundzuge einer Theorie der Phylogenetischen Systematik*, Deutscher Zentralverlag, Berlin, 1950.
9. Harvey, P. H. and Pagel, M. D., *The Comparative Method in Evolutionary Biology*, Oxford University Press, 1991.
10. Patterson, C., Williams, D. M. and Humphrey, C. J., *Annu. Rev. Ecol. Syst.*, 1993, **24**, 153–188.
11. Eggleton, P. and Vane-Wright, in *Phylogenetics and Ecology* (eds Eggleton, P. and Vane-Wright), Academic Press, 1994.
12. Quicke, D. L. J. and Belshaw, R., *Syst. Biol.*, 1999, **48**, 436–459.
13. Salisbury, B. A., *Syst. Biol.*, 1999, **48**, 153–159.
14. Sanderson, M. J., Baldwin, B. G., Bharrath, G., Campbell, C. S., von Dohlen, C., Ferguson, D., Porter, M. J., Wojciechowski, M. F. and Donoghue, M. J., *Syst. Biol.*, 1993, **42**, 562–568.
15. Bremer, B., Jansen, R. K., Oxelman, B., Backlund, M., Lantz, H. and Kim, K.-J. *Syst. Biol.*, 1999, **48**, 413–435.
16. Cantino, P. D., Bryant, H. N., Queiroz, K. de., Donoghue, M. J., Eriksson, T., Hillis, D. M. and Lee, M. S. Y., *Syst. Biol.*, 1999, **48**, 790–807.
17. Ganeshiah, K. N. and Uma Shaanker, R., *Biodiver. Conserv.*, 2000, **9**, 953–963.
18. Adanson, M., *Histoire Naturelle du Senegal Coquillages. Avec la Relation Abreege d'un Voyage Fait en ce Pays, Pendant les Annees 1749, 50, 51, 52 et 53*. Coquillages, Preface, pp. xi, xx, xxix–lxxxviii. Bauche, Paris, 190 + xcvi + 175 pp.
19. Numerical taxonomy in a broad sense includes both phenetic and phylogenetic systematics<sup>20</sup>. However in general and in this article it is referred exclusively to phenetic taxonomy.
20. Sneath, P. H. A., *Syst. Biol.*, 1995, **44**, 281–289.

K. N. Ganeshiah\* is in the Department of Genetics and Plant Breeding, K. Chandrashekar and A. R. V. Kumar are in the Department of Entomology, and R. Uma Shaanker is in the Department of Crop Physiology, GKVK, Bangalore 560 065, India. \*For correspondence. (e-mail: kng@vsnl.com)