Hydrobiologia 155: 65–74 (1987)
R. O. Brinkhurst and R. J. Diaz (eds) Aquatic Oligochaeta

# A comparison of phenetic and phylogenetic methods applied to the systematics of Oligochaeta

Ralph O. Brinkhurst & Amanda F. L. Nemec
*Ocean Ecology Laboratory, Institute of Ocean Sciences, P.O. Box 6000, 9860 West Saanich Road, Sidney, B.C., Canada V8L 4B2*

*Key words:* aquatic Oligochaeta, phylogeny, phenetics, cladistics, Naididae, Dorydrilidae

## Abstract

A comparative study of naidid subfamilies shows that a combination of ordination, Jaccard/Average Linkage cluster analysis and Wagner parsimony provides a useful basis for a rational phylogeny but that this does not differ markedly from the original proposed by Sperber nearly four decades ago. Hennig rules, modified by Wiley, permit a preliminary phylogeny and classification of the Annelida to be made by hand. An error in earlier versions suggested that the Dorydrilidae lacked prostate glands, and this is corrected.

## Introduction

At our previous meeting, the senior author discussed two competing views of the phylogeny and resulting classifications of the Oligochaeta (Brinkhurst, 1984a) following two other contributions on the same subject (Brinkhurst, 1982, 1984b). Since then, Kasprzak (1984) has discussed these ideas, and concluded that computer analyses based on Hennig's principles would be a new approach to classification of the oligochaetes. In this paper we address the various methods that can and have been used for this purpose, and we present data from computerized analyses at the generic level for Naididae, and hand calculated Hennig-style diagrams leading to a suggested higher classification for the Annelida.

## Materials and methods

A data matrix was compiled including 24 binary characters used in generic definitions of 24 naidid taxa. All naidid genera are included with the exception of *Wapsa*, now seen as a synonym of *Paranais* (Brinkhurst and Coates, 1985) and two recent genera described from immature specimens (*Neonais* Sokol'skaya, 1962 and *Rhopalonais* Grimm, 1974) both of which may be assumed to be related to *Nais* or *Dero*. The genus *Pristina* is treated as two genera (*Pristina* Ehrenberg, 1828 and *Pristinella* Brinkhurst, 1985) and *Slavina* is divided into A and B groups as the genus is probably not monophyletic. Also included is an hypothetical ancestor that has all the characters in the supposedly plesiomorphic condition, as determined by a manual comparison with the sister groups, the related microdrile families. The ancestor is used instead of an out-group, (Wiley, 1981). The characters for all genera were polarized (0 for plesiosmorphic state, 1 for apomorphic) as required by some, but not all, analyses. A reduced data set was obtained by deleting secondary characters, multistate characters, highly correlated characters and those that are unreliably coded. This process yielded a short matrix of 15 characters, used to see if a smaller data set would suffice.

The cluster analysis methods used were Simple Matching (SM) and Jaccard (J) coefficients with single linkage (SL), complete linkage (CL) and average linkage (AL) which, for both short and full data sets, yielded 12 possible combinations (Cormack, 1971). Jaccard compares only positive matches and may be held to be Hennigian in this regard if the characters are carefully coded so that "presence" is equal to "apomorphic" (i.e. the absence of a prostate might be coded "+" if it is thought to be the derived state of the character). A principal component analysis ordination (Gower, 1966) was obtained for both SM and J similarity matrices, for both short and full data sets, using IMSL routine OPRINC. The graphical extension of ordination described by Banfield & Gower (1980) and the eigen values associated with each coordinate were used to help reveal any distortion in the ordinations. Phylogenetic methods have been reviewed by Felsenstein (1982, 1983a). We used two parsimony methods, one according to Camin & Sokal (1965) where character states may only change from 0 to 1, and the other a Wagner analysis (Eck & Dayhoff, 1966; Kluge & Farris, 1969), which is the least restrictive method, allowing 0 to 1 or 1 to 0 changes which actually renders character polarization somewhat redundant. These two analyses were run with both short and long data sets using the Wagner command in PHYSYS (Farris & Mickevitch, 1982) with the global branch swapping option (.S) and retaining up to 10 equally parsimonious trees; Camin-Sokal was implemented using WIS in PHYSYS. The resultant Wagner Wiss trees were optimized using DIAGNOSE and IRIAGNOSE, respectively, with options, A, C, H and O. A character compatibility analysis was also performed using CLIQUE in PHYLIP 2.4 (Felsenstein, 1983b). Finally, the evolutionary tree presented by Sperber (1948) was re-drawn by calculating the distances between dichotomies and re-plotting the result as a cluster diagram. Tentative relationships noted by Sperber had to be firmed up in the process, and a confusion about antiquity of forms showing many apomorphies related to a predatory diet created an apparent discrepancy between the resulting figure and the opinions expressed in the body of Sperber's text. The programme DIAGNOSE was used to evaluate the resultant tree using the long data matrix with some changes (*Bratislavia* and the hypothetical ancestor omitted, a combination of *Pristina* and *Pristinella, Slavina A* used for *Slavina* to reduce the taxa to the 1948 conditions).

Tree length, the overall consistency index (Kluge & Farris, 1969) F-ratio and D. measure (Brooks *et al.* 1984 pers. comm.) were used to compare the relative merits of the Sperber tree and the trees produced by the various parsimony methods. All statistics were calculated using the LFIT command in PHYSYS or were hand calculated. The various statistics measure how well a tree fits the data, with respect to certain desirable charactericS, and are especially useful where they separate trees of equal length, since they then provide objective criteria for choosing between the multiple solutions that are typically produced by parsimony methods.

The parsimony methods allow for display of the locations of character state changes on the trees produced. From an examination of the naidid results it is obvious that most dichotomies involve 1 to 3 characters. The limitation of the number of characters used in the hand-drawn diagrams of the evolutionary schemes from Phylum to Family level, presented here as preliminary hypotheses, is not, therefore, a critical deficiency. A mixture of somatic and reproductive characters was employed, but no formal data matrix was prepared.

## The character states

The characters used in the analysis of naidid genera were as follows, the plesiomorphic state being indicated:

1. Hair chaetae present;
2. No obvious serrations on hair chaetae;
3. No elongate hair chaetae;
4. No strong difference between anterior and posterior ventral chaetae;
5. No slight difference in anterior ventral chaetae;
6. Needle chaetae curved proximally;
7. Penial chaeta present;
8. Spermathecal chaetae;
9. Nodulus on needle;
10. Atrium with prostate;
11. Vas deferens without prostate;
12. Vas deferens subapical on atria;
13. Male pore simple;
14. Clitellum absent between male pores;

15. Vascular system simple;
16. Proboscis absent;
17. Budding;
18. Dorsal chaetae from II;
19. Glands arond atrial duct;
20. Gills absent;
21. Dorsal chaetae less than 10;
22. Stomach present;
23. Dorsal chaetae present;
24. Diverticulum on pharynx.

Characters 2–4, 6, 8, 9, 14, 15 and 19 were omitted in the short data set, 17 on the revised Sperber tree.

## Results

In our version of the Sperber tree (Fig. 1) the ancestral form is forced to be apomorphic for characters 1, 10, 15 and 18, but this is largely due to the fact that Sperber drew in *Chaetogaster, Amphichaeta* and *Paranais* close to the ancestor because she supposed them to be ancient because they are highly derived forms. If she had expressed this in Hennigian terms they would be shown far from the ancestor, but we do not wish to start tampering with her tree and so

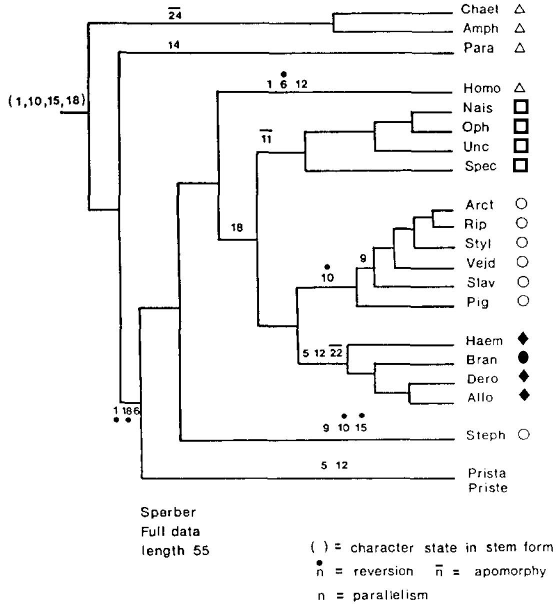


*Fig. 1.* Dendrogram of naidid genera derived from Sperber (1948) with positions of character changes inserted using DIAGNOSE. Symbols denote final groupings identified in this study.

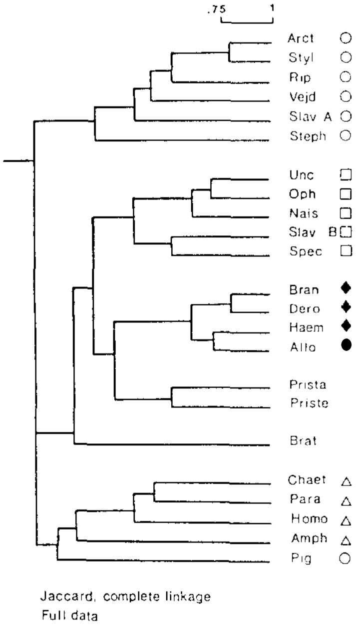


*Fig. 2.* An example of the result of a cluster analysis of naidid genera.

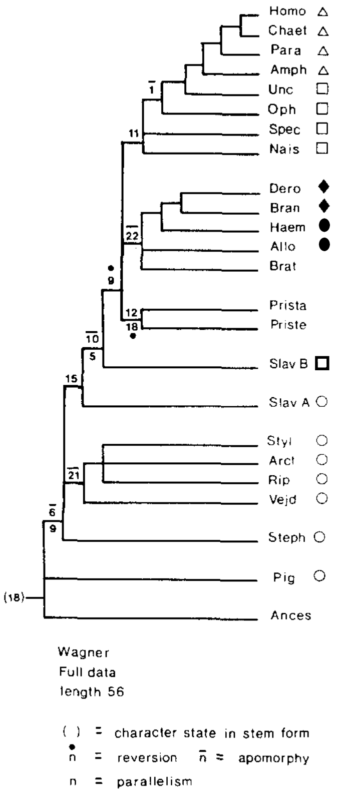


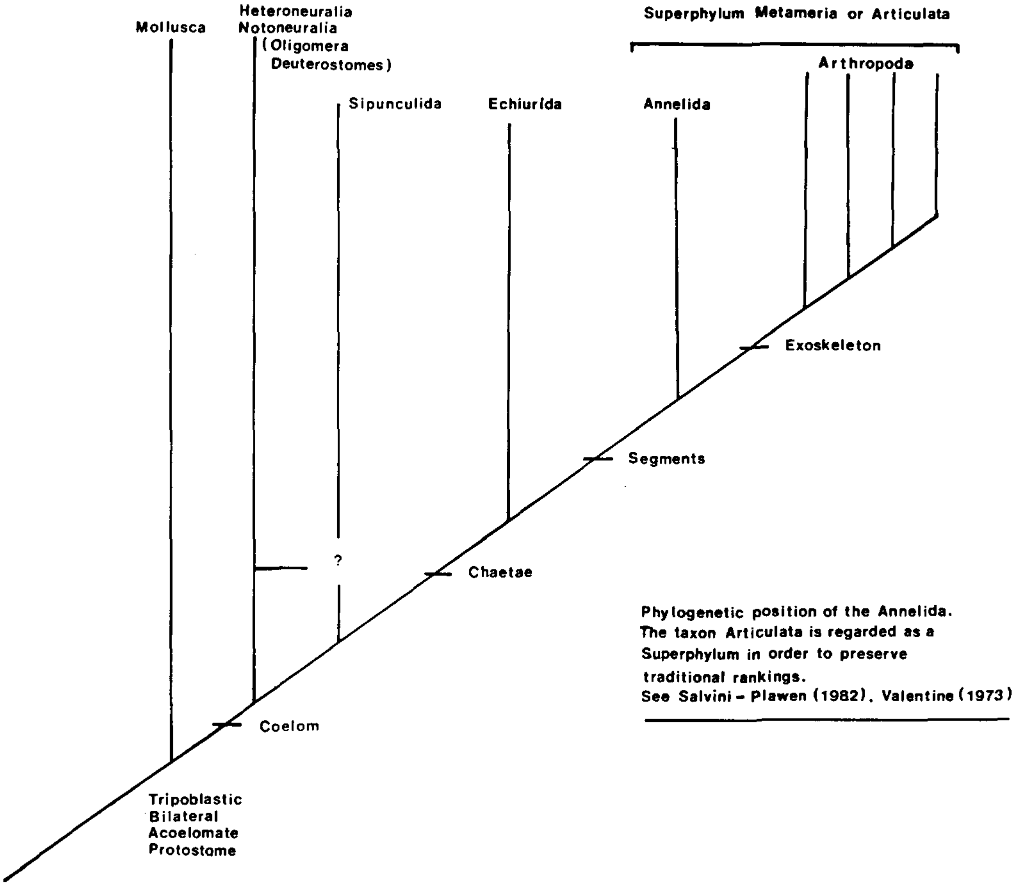
*Fig. 3.* An example of the result of a parsimony method analysis of naidid genera.

we simply recognize this distortion. These characters promptly revert to the plesiomorphic state quite early in the resultant tree. This alone indicates the value of recent methods in which such logical errors can be avoided. Four groupings of genera are marked by symbols and the fifth consists of *Pristina* and *Pristinella.*

These marks are introduced to enable the reader to compare the results of one analysis with another especially when all of the analyses are reported. It may still be useful here where only representative results are shown.

Essentially the same groupings as those seen in Sperber's tree were obtained from the 12 cluster analyses, but those using the single linkage approach (nearest neighbour) produced the chaining, rather than discrete clustering, that is an artefact of this method. All the other analyses tend to produce the same 5 groups, but some genera are less firmly associated with their groups than the rest (e.g. *Stephensoniella* and *Piguetiella* in the *Stylaria* group) and *Bratislavia* seems to have no fixed abode. The result for Jaccard and complete linkage with the full data set is shown here (Fig. 2) and the rest will be documented elsewhere. Note that in this diagram the *Stylaria* group is well separated from the rest. This fact is emphasized by the ordination analyses (not illustrated), in which the *Stylaria* group minus *Pristinella* is well separated, with 80% of the variance accounted for by the first pair of coordinates when the SM coefficient is used. Note also that the *Dero* and *Nais* groups are closely associated and that the *Chaetogaster* group forms a third axis. *Pristina* and *Pristinella* are placed close to *Dero* as they are in many of the results obtained.

In the parsimony methods (Camin-Sokal, Wagn-



Phylogenetic position of the Annelida.
The taxon Articulata is regarded as a Superphylum in order to preserve traditional rankings.
See Salvini - Plawen (1982), Valentine (1973)

*Fig. 4.* A tentative Hennigian phylogeny of the coelomate protostome Bilateria.

er) we can inspect both the groupings of genera and the placement of the character changes (which were also available in the analysis of our version of Sperber's tree, Fig. 1). There are some unresolved groups (indicated by three or more equal branches rather than dichotomies on the tree) in all the resulting trees. The placement of *Slavina B* is inconsistent in these analyses, and both *Stephensoniana* and *Piguetiella* are usually associated with the presumed ancestor, as in the Wagner – full data set example shown (Fig. 3) which is 1 of 5 equally parsimonious trees. The position of major conversions to apomorphy, reversals and parallelisms are shown. The tree shown here is shorter than the version provided by Camin-Sokal (56 versus 64), a product of the restrictive conditions of Camin-Sokal which allows no reversions, and which leads to many supposed parallelisms in character state change. It should be noted that the Sperber tree, even subject to our distortion, has a length of only 55. In terms of the other measures used to choose between trees, the Sperber and Wagner trees are very similar (consistency index, F-ratio, D measure as I, D and $D_2$). The analysis with the full data set produce better values for these indices than the short data set (though the length of the tree is obviously reduced with the short data) with the exception of the D measure.

In terms of characters, while there were 18 compatible cliques, none of them contained more than 9 of the 24 characters. Some characters are autapomorphies (occurring in only 1 genus i.e. 2, 23) and are not used at major dichotomies. Those characters with a low frequency of apomorphy (2–4) also fail to appear at major dichotomies (3, 4, 7, 8, 13, 14, 16, 17, 19, 20, 21, 23, 24) and all but one of the characters with high frequency of compatibility are found in this group, the exception being character 22 (5 apomorphies, compatible in 11 of the

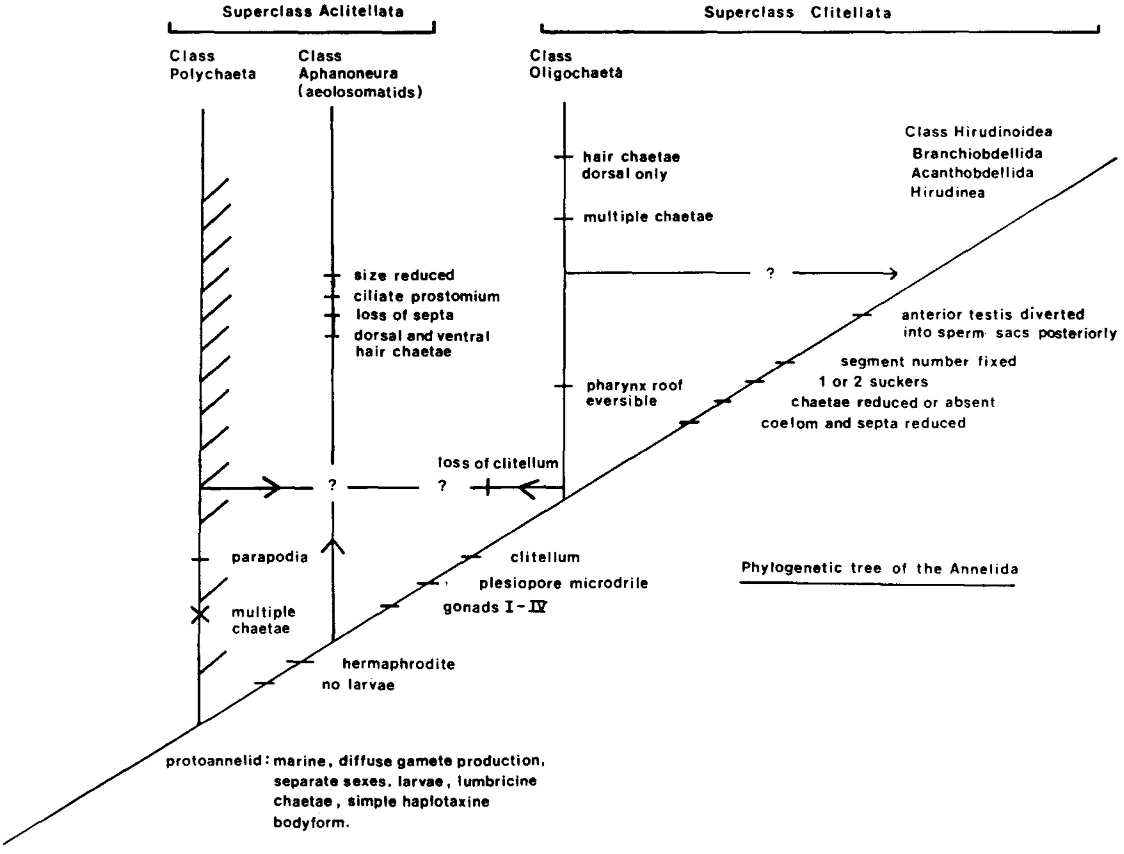


*Fig. 5.* A tentative Hennigian phylogeny of the Annelida.

18 largest compatible cliques). Characters with a frequency of apomorphy of 5–18 in the basic data matrix had a compatibility frequency of 0–4 (with the exception of character 22, noted above). The same suite of characters show up at the major dichotomies in all of the analyses, with only 1–3 found at each junction. All of this suggests that astute visual inspection of the original data matrix could lead to the production of a very similar tree, as one might expect. The relative placement of characters such as "loss of hair chaeta" will determine the relationship of *Uncinais* and *Ophidonais* to the *Chaetogaster* and *Nais* groups, for example, and the various methods produce slightly different results, or can be manipulated to determine the effect of different character selections and polarities.

The end result of this study will be discussed in detail elsewhere, but the suggestion is that the existing subfamilies are not adequate, and that the *Stylaria* group (circle symbols in the Figures) should be recognized as one subfamily with the rest constituting one large subfamily. The *Nais* and *Dero* groups are regarded as tribes within that subfamily. *Pristina* and *Pristinella* are as well separated from each other as most genera, more so than some (*Dero* – *Branchiodrilus*), and are regarded as a tribe, as is the *Chaetogaster* group. Some genera are only placed incertae sedis in this scheme (i.e. *Neonais, Rhopalonais, Bratislavia, Stephensoniana, Piguetiella*).

At the same time as the detailed investigation of the evolutionary relationships among the naidids was performed, an attempt was made to create higher level evolutionary trees of the Hennig style, by hand, beginning as one must with the placement of the Annelida in their proper context (or one possi-

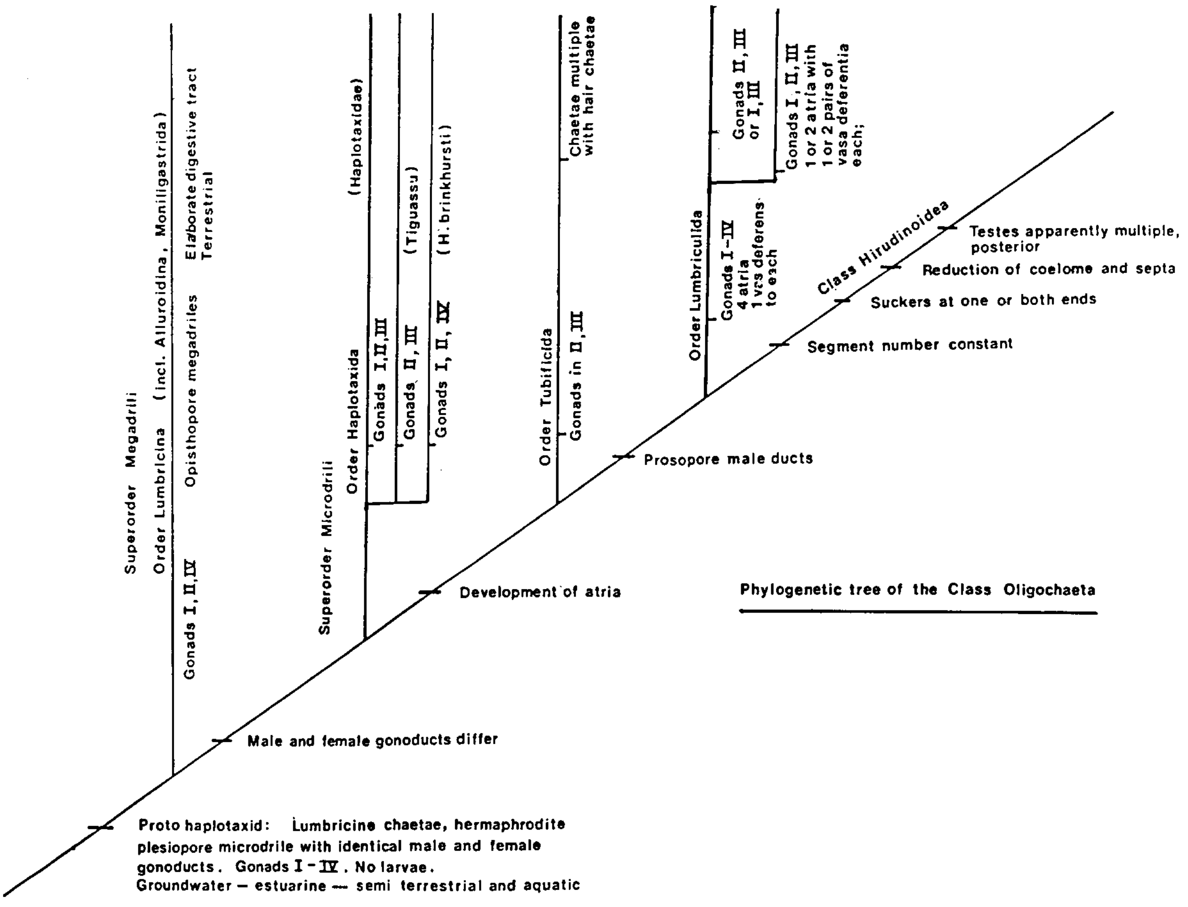


*Fig. 6.* A tentative Hennigian phylogeny of the Oligochaeta.

ble version of this – Fig. 4). This allows us to define the primitive annelid without reference to "chicken and egg" debates about hair chaetae, as the earliest worms clearly had simple chaetae. The resultant tree (Fig. 5) can be used to produce a classification which does not offend Hennig if we create Super Classes so that oligochaetes, leeches and polychaetes can be ranked as Classes. The aeolosomatids and their allies may be unique (a Class) or may be derived from polychaetes or oligochaetes by progressive miniaturization. However, the detailed tree of the oligochaetes (Fig. 6) does reveal a potential problem as leeches are now considered to be possible derivatives of the lumbriculids (unless the prosopore condition has arisen more than once by convergence) and as such should be part of the Order Lumbriculida, or some new taxon at some level perhaps called the Prosopora. Hence Fig. 5 could be drawn with a line to the leeches, originating well along the oligochaete axis, perhaps as indicated by the line broken by "?". Note that intriguing haplotaxids like *H. brinkhursti* and *Tiguassu* are not shown as megadrile and tubificine ancestors – in earlier accounts great care was taken to speak of "brinkhurstoid" and "tiguassoid" proto-haplotaxids as direct ancestors. The living forms are almost certainly parallelisms to the ancestors, or the Haplotaxidae are paraphyletic (Fig. 7) but they document the possibility of key changes in characters among this assemblage of primitive species.

The Families in the Tubificida are plotted on Fig. 8. *Propappus* belongs somewhere between the Enchytraeidae and either the Phreodrilidae or Haplotaxidae, but is now thought to be in a monotypic family (Coates, 1987). The Naididae and Tubificidae should probably not be Families at all (and hence Naidid sub-families discussed below would become demoted). Too little is known about the Opistocystidae to allow clear placement, and even less is known about the Lobatocerebrids, *Capiloventer* and other new forms such as the Randiellidae and Narapidae, nearly all of which are very

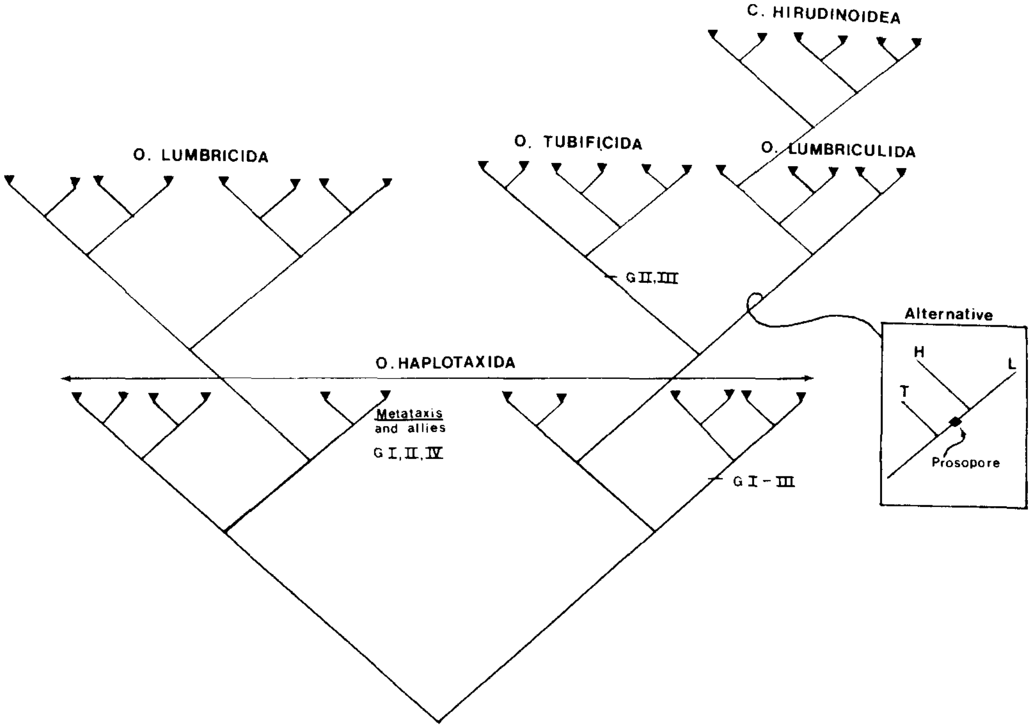


*Fig. 7.* A model showing the Haplotaxidae as a paraphyletic group, and two alternate possibilities for the origin of the leeches as members of the Lumbriculida. The number of dichotomies is not significant – the origins of major groups are shown as hypotheses for future investigation.

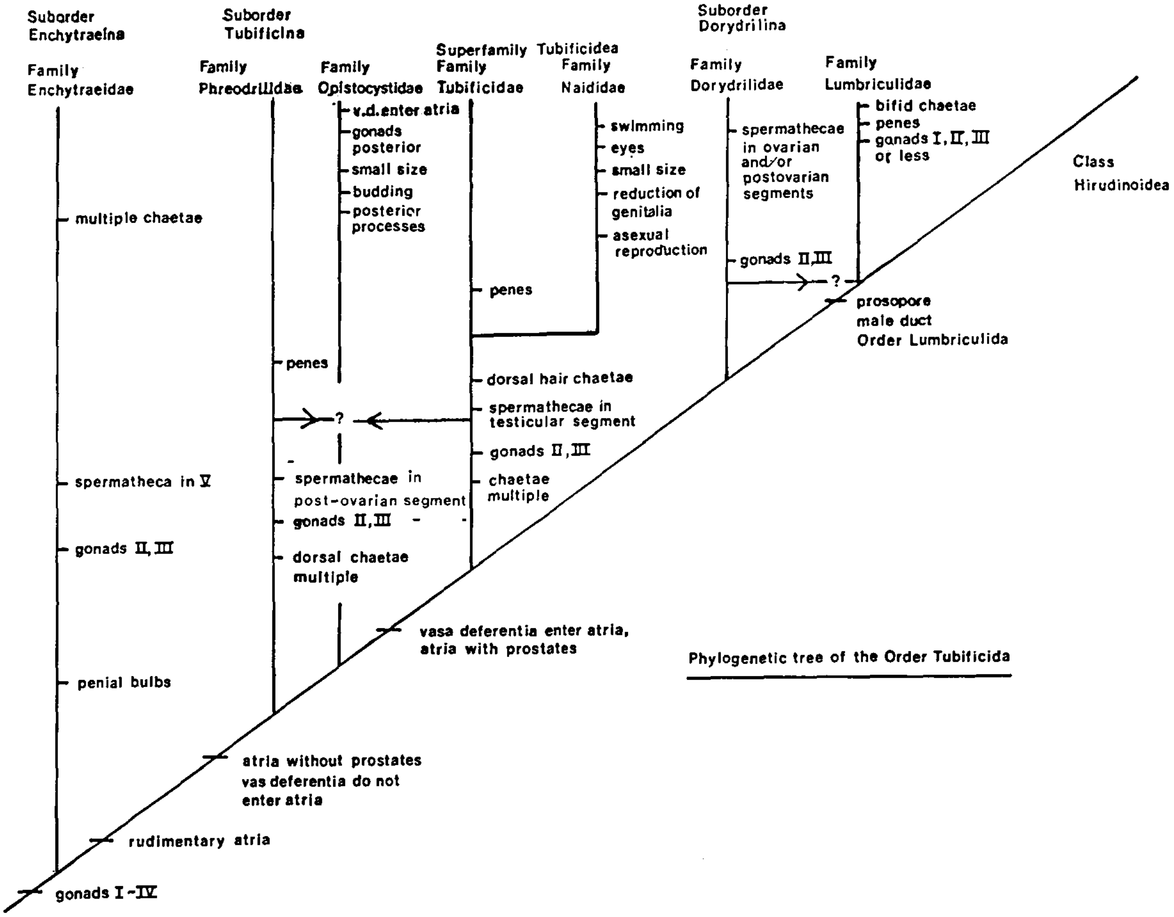


*Fig. 8.* A tentative Hennigian phylogeny of the Tubificida.

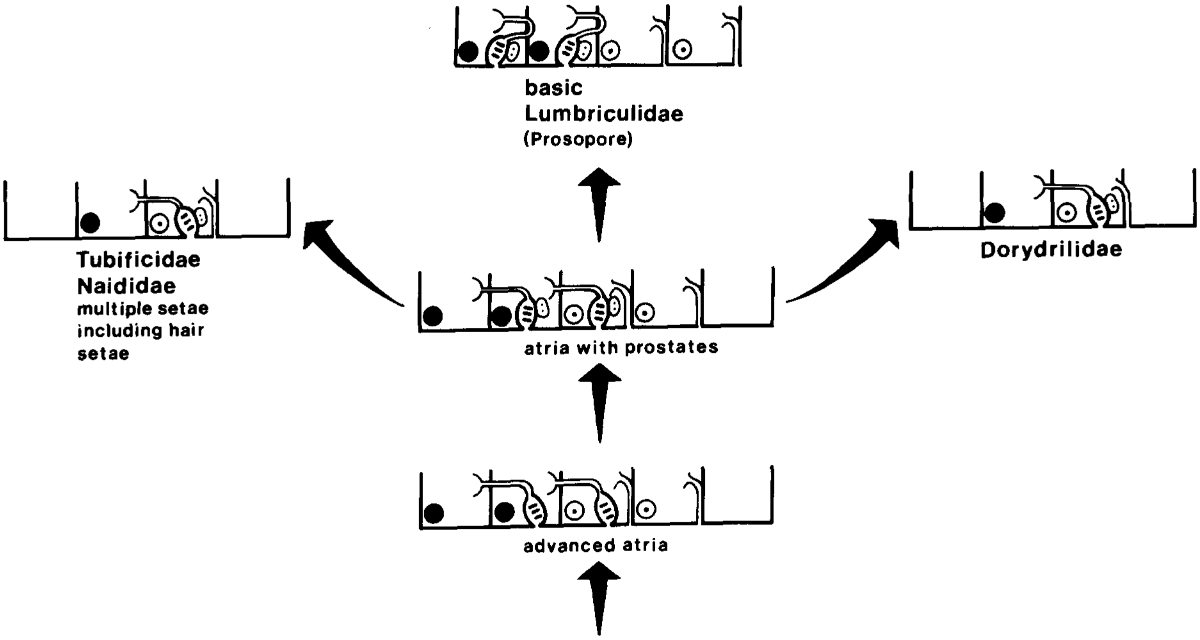


*Fig. 9.* A corrected version of the evolutionary trees showing the Dorydrilidae with prostates on the atria.

small and may as well be highly specialized as necessarily primitive.

In earlier accounts a misinterpretation of published illustrations of the atria in the Dorydrilidae led to their being shown as loacking prostates. This necessitates a correction to my earlier evolutionary-style models which is incorporated into Fig. 9.

## Discussion

Others have also attempted to use the recently developed methods of numerical taxonomy in oligochaete systematics. Sims in a series of papers (most recently Sims, 1980), used principal coordinates and cluster analyses (using single linkage and total resemblance i.e 0 – 0 matches included), to produce classifications that were a blend of relatively objective analyses and subjective input. Considerable ambivalence was expressed about the acceptance of analytical results. In the principal components analyses the first 6 axes accounted for only 4.8 to 17.7% of the observed variation (c.f. the values obtained in our analyses), which suggests a high degree of distortion that seriously detracts from the usefulness of the ordinations. Jamieson (1978, 1980) produced pioneering studies on opisthopore oligochaetes. Again single linkage cluster analysis was adopted because of its supposed Hennigian acceptability, but this has more to do with the similarity coefficient (in fact, Sims thought the chaining effect due to this method indicated continuity in the data series rather than recognizing it as an artefact). Jamieson too did not slavishily adopt the analytical results because of reservations about some of Hennig's assumptions. A minor criticism of this work is that too much is made of the relationship between the plesiopore microdrile haplotaxids and their close relatives (OTU's 1 – 9) that only attach to the Hennigram at a low level of similarity and are well separated from the rest in the stereophylogram. The similarity is so low it is worth noting that these groups had to be attached to the rest somewhere, but little information can be derived from their location. Jamieson's immediate conclusion (that the Tubificids are unlikely to intervene between the haplotaxids and the opisthopores) can be sustained, but the modern haplotaxids seem mostly to belong to the separate order Haplotaxida, with the Lumbricina, Haplotaxida and Tubificida all derived from a protohaplotaxid. The equivalence in rank is maintained by appealing to Conventions 2 and 3 of Wiley (1981), though the family could well be paraphyletic. As Jamieson rightly points out, numerical methods are of value for their relative objectivity, repeatability and heuristic qualities. They do not necessarily require slavish adherence to any one theory, especially where this leads to impractical or widely criticised classifications. Erseus (1984) used a Wagner analysis to assist with the process of classifying a large group of marine tubificids, and again quite openly showed where subjectivity entered into the process.

None of these computerized methods alone will automatically produce the "best" classification, but they do force one to declare all the assumptions made, to define the character states for all taxa and to avoid many errors in logic. (It is possible to manipulate a preconceived view of evolution to fit a Hennigian model, and so there is no unique solution). Computerised parsimonious methods do not produce a unique tree for each matrix, and will usually leave a choice between equally parsimonious trees. Total resemblance methods may not differ too much in their end results from Hennigian methods either, despite the controversy between supporters of the two schools, because we should anticipate true close resemblance and kinship to be the norm. Once homoplasy is recognized, earlier supposed resemblance usually appears superficial. While cluster analyse allow the taxa in the data to be placed in a coherent sequence, facilitating parsimony analysis (which can be subject to variation based on the sequence in which the taxa are entered into the matrix!), parsimony analyses may allow the identification of the changes in character state that create the tree. No one approach has an exclusive claim to pre-eminence. The very best classical scholars using traditional thought processes could and did provide models quite consistent with those derived from modern methods, as shown by the re-analysis of the work by Sperber. The quality of her work has not, in fact, been equalled by any of my generation, but better things may be in store as we use these methods to resolve issues such as the apparent paraphyletic status of the Haplotaxidae and the leeches, which seem to be the next major issues to be tackled.

## References

Banfield, C. F. & J. C. Gower, 1980. A note on the graphical representation of multivariate binary data. Appl. Statist. 29: 238–245.

Brinkhurst, R. O., 1982. Evolution in the Annelida. Can. J. Zool. 60: 1043–1059.

Brinkhurst, R. O., 1984a. The position of the Haplotaxidae in the evolution of the oligochaete annelids. Hydrobiol. 115: 25–36.

Brinkhurst, R. O., 1984b. Comments on the evolution of the Annelida. Hydrobiol. 109: 189–191.

Brinkhurst, R. O. & K. A. Coates, 1985. The genus *Paranais* (Oligochaeta, Naididae) in North America. Proc. Biol. Soc. Wash. 98: 303–313.

Camin, J. H. & R. R. Sokal, 1965. A method for deducing branching sequences in phylogeny. Evolution, 19: 311–326.

Coates, K. A., 1987. Redescription of the oligochaete genus *Propappus* and diagnosis of the new Family Propappidae (Annelida: Oligochaeta) (in press).

Cormack, R. M., 1971. A review of classification. J. R. Statist. Soc. A. 134: 321–353.

Eck, R. V. & M. O. Dayhoff, 1966. Atlas of protein sequence and structure 1966. National Biomedical Research Foundation, Silver Spring.

Erseus, C., 1984. Taxonomy and phylogeny of the gutless Phallodrilinae (Oligochaeta, Tubificidae), with descriptions of one new genus and twenty-two new species. Zool. Scr. 13: 239–272.

Farris, J. S. & M. F. Mickevitch, 1982. Phylogenetic analysis system (PHYSYS). Fortran V software system of cladistic and phenetic algorithms.

Felsenstein, J., 1982. Numerical methods for inferring evolutionary trees. Quart. Rev. Biol. 57: 379–404.

Felsenstein, J., 1983. Statistical inference of phylogenies. J. R. Statist. Soc. A. 146: 246–272.

Felsenstein, J., 1984. Phylogeny inference package (PHYLIP). Version 2.4. Documentation.

Gower, J. C., 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, 53: 325–338.

Jamieson, B. G. M., 1978. Phylogenetic and phenetic systematics of the opisthopore Oligochaeta (Annelida: Clitellata). Evol. Theory, 3: 195–233.

Jamieson, B. G. M., 1980. Preliminary discussion of an Hennigian analysis of the phylogeny and systematics of opisthoporous oligochaetes. Rev. Ecol. Biol. Sol. 17: 261–275.

Kasprzak, K., 1984. The previous and contemporary conceptions on phylogeny and systematic classifications of Oligochaeta (Annelida). Ann. Zool. 38: 205–223.

Kluge, A. G. & J. S. Farris, 1969. Quantitative phyletics and the evolution of anurans. Syst. Zool. 18: 1–32.

Sims, R. W., 1980. A preliminary numerical evaluation of the taxonomic characters of *Allolobophora* auct. and some allies (Lumbricida: Oligochaeta) occurring in France. Pedobiologia, 20: 212–226.

Sneath, P. H. A. & R. R. Sokal, 1973. Numerical taxonomy. W. H. Freeman, San Francisco, 573 pp.

Sperber, C., 1948. A taxonomical study of the Naididae. Zool. Bidr. Upps. 28: 1–296.

Wiley, E. O., 1981. Phylogenetics. The theory and practice of phylogenetic systematics. John Wiley, New York, 439 pp.