Chapter 9

Unleashing the Force of Cladistics?
Metazoan Phylogenetics and Hypothesis Testing

Ronald A. Jenner

Submitted to American Zoologist
UNLEASHER THE FORCE OF CLADISTICS? METAZOAN PHYLOGENETICS AND HYPOTHESIS TESTING

"The matrix is the world that has been pulled over your eyes to blind you from the truth." - Morpheus in *The Matrix* (1999) by A. Wachowski & L. Wachowski.

SYNOPSIS. The accumulation of multiple phylogenetic hypotheses for the Metazoa invites an evaluation of current progress in the field. I discuss three case studies from the recent literature to assess how cladistic analyses of metazoan morphology have contributed to our understanding of animal evolution. The first case study on cleavage cross patterns examines whether a decade of unanimous character scoring across different cladistic studies can be considered a reliable indicator of accumulated wisdom. The two remaining case studies illustrate how the unique strength of cladistic analyses to arbitrate between competing hypotheses can be crippled when insufficient attention is directed towards the construction of the data matrix. The second case study discusses a recent morphological cladistic analysis aimed at providing insight into the evolution of larval ciliary bands (prototrochs) in the Spiralia, and the third case study evaluates how four subsequent morphological cladistic analyses have contributed to our understanding of the phylogenetic placement of a problematicum, the Myzostomida. I conclude that current phylogenetic analyses of the Metazoa have not fully exploited the power of cladistics to test available alternative hypotheses. If our goal is to generate genuine progress in understanding rather than stochastic variation of opinions through time, we have to shift our attention from using cladistics as an easy tool to generate “novel” hypotheses of metazoan relationships, towards employing cladistics more critically as an effective instrument to test the relative merit of available multiple alternative hypotheses.

MORPHOLOGICAL CLADISTICS OF THE METAZOA: ASSESSING CURRENT PROGRESS

“Study of the higher level phylogeny of multicellular animals has been something of a backwater for decades, largely because all the morphological clues had been pushed beyond their limits, and mutually contradictory speculations led only to dead ends” (Patterson, 1990: 199). Patterson’s perspective painted a rather discouraging picture for expected future progress in morphological metazoan phylogenetics a decade ago. Luckily, Patterson’s prophecy did not discourage the next generation of phylogeneticists. Higher-level animal phylogenetics is currently bustling with activity. Seminal work in molecular systematics jump started renewed interest in animal relationships in the late 1980s (Field *et al.*, 1988; Raff *et al.*, 1989), and this initial impulse was reinforced at the beginning of the 1990s when a new era of phylogenetic research on the Metazoa was ushered in by the publication of the first comprehensive computer-assisted cladistic analyses of animal
morphology (Brusca and Brusca, 1990; Meglitsch and Schram, 1991; Schram, 1991). These analyses set a new standard for metazoan phylogenetics, especially through the use of explicit taxon by character data matrices (Meglitsch and Schram, 1991; Schram, 1991), the use of which greatly facilitated the transparency of cladistic analyses.

Since then, a substantial number of comprehensive morphological and total evidence (combined molecular and morphological data) cladistic analyses of the Metazoa have been published (Jenner and Schram, 1999), including no less than five analyses published in the first two years of the new millennium (Zrzavy et al., 2001; Giribet et al., 2000; Sørensen et al., 2000; Nielsen, 2001; Peterson and Eernisse, 2001). Phylogenetic analyses of 18S rRNA/DNA sequences quickly monopolized the field of molecular metazoan cladistics, and the discovery of apparently conflicting phylogenetic signals inherent in morphological and molecular data sets has since generated an exciting dialogue between the molecular and morphological subdisciplines of metazoan phylogenetics. Discussions about the phylogenetic placements of acoels and nemertodermatids (basal-most extant bilaterians or derived lophotrochozoans?), brachiopods and phoronids (lophotrochozoan protostomes or deuterostomes?), and arthropods and annelids (monophyly of Ecdysozoa or Articulata?) are among the more conspicuous of the current debates (e.g., Lüter and Bartolomaeus, 1997; Ruiz-Trillo et al., 1999; Baguñà et al., 2001; Wägele et al., 1999; Wägele and Misof, 2001; Zrzavy, 2001; De Rosa, 2001). The resolution of these debates will eventually be dependent upon the reconciliation of molecular and morphological phylogenetic evidence. However, before we can hope for such an overarching consensus, we first have to secure a more modest goal: to establish the contribution of a decade of morphological cladistic research towards our current understanding of metazoan relationships.

One way to assess progress in the field is to discuss topological congruence between different studies to identify taxa with uncertain phylogenetics positions. This strategy is most commonly employed in recent studies. Another possible strategy, however, instead focuses on character assessment across different studies to examine whether characters shared between analyses are interpreted in the same way. For example, one may investigate whether unanimous character scoring across different cladistic analyses is a reliable indication of accumulated understanding. This issue will be explored in the first case study by a comparison of character scoring for cleavage cross patterns across different cladistic analyses. Another tactic is to explore how cladistic analyses are used to test available hypotheses of body plan evolution or phylogenetic relationships. In the second case study, I explore in how far a cladistic analysis specifically designed to assess hypotheses for the evolution of larval ciliary bands (prototrochs) in the Spiralia (Rouse, 1999) has aided our understanding of body plan evolution. In a last case study, I explore how subsequent cladistic analyses have contributed to our understanding of the phylogenetic position of a phylogenetic
problematicum, the Myzostomida. These three examples are specifically selected for the clear insights they offer into the strengths, and especially several seemingly unrecognized limitations of morphological cladistic analyses of the Metazoa.

THREE CASE STUDIES: STRENGTHS AND LIMITATIONS OF MORPHOLOGICAL METAZOA CLADISTICS

Dissolving unanimity: the phylogenetic insignificance of cleavage cross patterns

When one compares the scoring of characters shared between two or more morphological cladistic analyses of the Metazoa, it becomes apparent that scoring conflicts are not difficult to find. For example, the study of synapomorphies supporting alternative sister group hypotheses for the ‘acoelomate’ worms (Platyhelminthes, Nemertea, Gnathostomulida), which have been proposed by cladistic analyses published over the last decade, reveals that roughly 35% of these characters shared between studies exhibit scoring conflicts for one or more taxa (unpublished data, R. A. Jenner). In contrast, other characters are scored identically in all recent phylogenetic studies that included them. Characters coding for spiralian cleavage cross patterns belong to this category.

The cleavage cross pattern refers to a stereotypical pattern of blastomeres organized around the apical pole in the embryos of various taxa with spiral quartet cleavage (see Brusca et al., 1997 for terminology of spiral cleavage). Two geometries are distinguished: the molluscan or radianect cross with cross arms composed of the micromeres $1a^{12-1d^{12}}$ and $2a^{11-2d^{11}}$ and their derivatives; and the annelid or interradiate cross with cross arms composed of the micromeres $1a^{112-1d^{112}}$ and $1a^{2-1d^{2}}$ and their derivatives (Raven, 1966; Siewing, 1969; Verdonk and Biggelaar, 1983). Recent studies universally score the molluscan cross present in molluscs and sipunculans, whereas the annelid cross is scored as a potentially unique synapomorphy of annelids and echinoderms. In fact, cleavage cross patterns have featured as textbook examples (Ruppert and Barnes, 1994; Pilger, 1997) of ontogenetic characters that allowed an escape from the difficulty of resolving phylogenetic relationships between phyla with strikingly disparate adult body plans.

Study of the literature reveals that at least 12 recently published cladistic analyses of variable scope coded one or two characters for cleavage cross pattern, all with identical character scorings (Eernisse et al., 1992; Rouse and Fauchald, 1995; Haszprunar, 1996; Salvini-Plawen and Steiner, 1996; Zrzavy et al., 1998; Giribet, 1999; Rouse, 1999; Almeida and Christoffersen, 2000; Edgecombe et al., 2000; Giribet et al., 2000; Haszprunar, 2000; Peterson and Eernisse, 2001), while the phylogenetic significance of these characters is also attested in other works (Rice, 1985; Scheltema, 1993; McHugh, 1997). The finding of such identical character scorings across different studies naturally inspires
confidence in their status as faithful representations of observational data. However, appearances may be deceptive.

First, it is surprising that apparently none of the above mentioned cladistic studies has adopted the proper definition of cross patterns rooted in cell lineage identity as a basis for their character scoring. Instead, character scoring is based on the differential dimensions and positions of the cells making up the arms of the crosses and the interradial cells, despite the fact that several of these analyses properly discuss character coding with respect to cell lineage identity. This can clearly be seen for the molluscan cross in both molluscs (Fig. 1a) and sipunculans (Fig. 1c), where the size differences of the blastomeres can be quite conspicuous. However, by properly focussing on cell lineages one can identify both the molluscan and annelid crosses in the embryos of the various taxa with spiral quartet cleavage, including molluscs (Fig. 1a), annelids (Fig., 1b), and sipunculans (Fig. 1c), even when differences in cell size may create the false impression that only a single cross pattern is present. This is a striking observation, because it contradicts the character scoring observed in all cladistic data matrices published so far. Moreover, using cell lineage as a diagnosis, we can also identify cross patterns (or the precursors cells) in taxa that are never scored for either of the cross patterns. For example, as illustrated in Fig. 1d, both a molluscan and an annelid cross can be mapped onto an entoproct cleavage-stage embryo. Clearly, this invites a reappraisal of all taxa with spiral quartet cleavage. For example, Riedl (1969) reported the presence of a cross pattern in Gnathostomulida comparable to that found in annelids, however, without any subsequent mentioning in recent cladistic studies that included both the character and taxon (Zrzavy et al., 1998; Peterson and Eernisse, 2001). Interestingly, almost 20 years ago Hennig (1983) came very close to the revised interpretation presented here when he emphasized that a double cleavage cross pattern could be observed in both mollusc and annelid embryos (Hennig, 1983: fig. 162). However, he maintained an additional emphasis on cell dimensions (for example, reporting the presence of only an annelid cross in sipunculans), and he did not explore whether cross patterns could be observed in spiralian other than the neotrochozoans (molluscs, annelids, sipunculans, echiurans).

Second, although the standard schematized illustration of cross patterns shows a clearly outlined cross with outstretched arms (Fig. 1a), the degree of character variation within phyla can be quite considerable (Nielsen, 2001: 135-136), so that a cross pattern diagnosed by differential blastomere dimensions may not be discernable in certain taxa at all.
FIG. 1. Cross patterns indicated on cleavage-stage embryos of four spiralian phyla. A: Mollusca (*Lymnaea stagnalis*, Gastropoda); B: Annelida (*Amphitrite* sp., Polychaeta); C: Sipuncula (*Golfingia vulgaris*); D: Entoprocta (*Pedicellina cernua*). Black circles mark the molluscan cross, while gray squares mark the annelid cross. (A modified from Verdonk and Biggelaar, 1983; B modified from Anderson, 1973; C and D modified from Nielsen, 2001).

Third, apart from the identified problems in character scoring, difficulties with character coding are also evident. Annelid and molluscan crosses are composed of different cells with distinct cell lineage identities and cell fates. This lack of homology between the cells of the molluscan and annelid crosses removes the basis for considering them “the same but different,” and consequently for coding them as alternative states of a single cleavage cross pattern character as in Eernisse *et al.* (1992) and Rouse and Fauchald (1995), and as suggested by Peterson and Eernisse (2001). Furthermore, the coding of a character for only the molluscan, but not the annelid cross in the analyses of

I conclude that the molluscan and annelid crosses cannot be upheld as independent potential synapomorphies of molluscs/sipunculans and echiurans/annelids, respectively. Using cell lineage as a proper diagnostic criterion, both patterns can be identified within individual embryos of taxa with spiral quartet cleavage, including taxa traditionally thought to lack the patterns completely, such as entoprocts. An undue emphasis on the differential dimensions of blastomeres combined with a failure to properly consider cell lineage identity, and a neglect of intra-taxon character variation mistakes a fortuitous superficial resemblance for a phylogenetically significant character. A more appropriate perspective would consider cross patterns to be direct consequences of having spiral quartet cleavage, without implying that all taxa with spiral quartet cleavage necessarily feature such patterns. Clearly, the phylogenetic significance universally attributed to cleavage cross patterns in recent cladistic analyses is misleading, and there appears to be no basis for defending the currently adopted cladistic character scorings. This is a striking conclusion because the repeated inclusion of cross pattern characters in different data matrices over the last decade offered maximal opportunity to properly evaluate and reevaluate their phylogenetic significance.

Some bare bones fundamentals: the epistemology of cladistics and hypothesis testing

Cladistics embraces a set of powerful methods to study phylogenetic relationships and character evolution. To borrow Dennett’s (1995) apt metaphor for natural selection, cladistics can be thought of as the “universal acid” of systematic biology that has permeated every nook and cranny of the discipline. But if such a potent tool is used uncritically, we may reap confusion rather than a deeper understanding of nature. To effectively employ cladistics to further our understanding of the macroevolution of animal body plans, it is critical to master the epistemology of cladistic knowledge-claims.

Cladistics is a multifunctional tool. A cladistic analysis can be used to generate novel phylogenetic hypotheses, but also to test available hypotheses. Although this distinction may appear trivial, it is in fact critical to understand because these different uses of cladistics require quite different procedures.

When cladistics is employed to produce a first estimate of the phylogeny of a particular group of organisms, we encounter no great procedural difficulties. Any “find ‘em and grind ‘em” approach to data collection and processing will suffice to provide an initial phylogenetic estimate. However, any subsequent effort to further evaluate or test the phylogeny of these taxa is more complicated because it needs to take previous efforts explicitly into account if genuine progress in understanding, rather than mere change of opinion, is our goal. It is
important to recognize that the epistemology of cladistics is predicated on Popperian testability (Kluge, 1997a, b; Siddall and Kluge, 1997; see also Systematic Biology 50[3] for recent discussions of the significance of Popper's ideas for the philosophy of phylogenetic inference). In fact, the notion that Popperian testability is at the heart of cladistic knowledge-claims is widely endorsed by cladists, and it is embodied within the term "sophisticated falsification," which has been used to label this philosophy of phylogenetic inference (Kluge, 1997b; Siddall and Kluge, 1997).

Central to an understanding of the strength of sophisticated falsification is the notion of competing hypotheses, available for testing (Siddall and Kluge, 1997). In cladistics, testing is done through a character congruence or parsimony test. The first step of a morphological cladistic test is always an analysis of anatomical similarity (Patterson, 1982), upon the basis of which a data matrix is constructed filled with propositions of primary homology. For an effective test between competing hypotheses, the data matrix should be so structured that all available potentially corroborating and falsifying (disconfirming) evidence for all competing hypotheses is included, so that all alternatives could in principle be vindicated or refuted. The most parsimonious cladogram resulting from the cladistic analysis then embodies the most highly corroborated, most severely tested hypothesis, with the highest explanatory power (in need of the least number of ad hoc explanations). Characters congruent with the topology of the most parsimonious tree(s) are provisionally accepted as corroborated (secondary) homologies with a single evolutionary origin, whereas the homology of incongruent characters with multiple independent evolutionary origins, i.e. homoplasies, is provisionally refuted. It should be noted, however, that not all results of a cladistic analysis are therefore necessarily the corroborated results of a congruence test. Only unbiased and careful selection of input data allows an efficient cladistic test. In practice this means that one has to pay special attention to three key ingredients that determine testing efficacy: taxon selection, character selection, and primary homology determination.

Although the above is generally endorsed in theory, dissection of recent morphological cladistic analyses of the Metazoa clearly reveals indications that the unique strength of cladistics to arbitrate between competing phylogenies or hypotheses of character evolution has not yet been fully exploited (see also Jenner, 2001a; unpublished data). The next example provides an illuminating illustration.

**Evolving prototrochs: interpreting ciliary bands in the Spiralia**

Rouse has published a series of valuable papers in which he attempts both to reconstruct the phylogenetic relationships within the Spiralia (a taxon of varying composition depending upon the source consulted, but often defined to include Mollusca, Annelida, Echiura, Sipuncula,
Gnathostomulida, Nemertea, Entoprocta, Platyhelminthes, and Panarthropoda), and to test hypotheses for the evolution of larvae and ciliary bands (Rouse and Fauchald, 1995, 1997; Rouse, 1999, 2000a, b, c). Rouse's cladistic analyses are remarkable in the breadth of their scope (including most of the more than 80 accepted polychaete families), and the attention paid to constructing a robust morphological data matrix. His studies have provided a first cladistic estimate of polychaete phylogeny, and his results have shed important light on the evolutionary dynamics of larvae and ciliary bands in spiralians.

I will discuss Rouse's first comprehensive cladistic study that was aimed at assessing the overall homology and evolution of ciliary bands and larvae in spiralians (Rouse, 1999). Because this study is specifically aimed at resolving a well-defined problem with a comprehensive cladistic analysis, it is particularly suited for yielding insights into the strengths and weaknesses of a cladistic analysis as a test of competing hypotheses.

Among the main hypotheses that Rouse set out to test are the status of the trophophore larva as a plesiomorphy for Spiralia, and the attendant hypothesis that all protostome taxa that lack such a trophophore larva must have lost it. Evolution of a trophophore larva at the base of the Spiralia, or even Protostomia, has been defended by several zoologists at different times during the history of evolutionary zoology, and today it features most prominently in the work of C. Nielsen (e.g., Nielsen, 1998; Nielsen, 2001). Trophophore larvae are chiefly diagnosed by possession of a prototroch, structurally defined as a pre-oral horseshoe or ring of (usually) compound cilia on multiciliate cells. However, additional features such as an apical organ, prototroch, metatroch, ciliated food groove (and potentially several other ciliary bands), and a pair of proctonephridia, may diagnose a more strictly defined trophophore larva (Nielsen, 2001). Rouse (1999) drew several important conclusions from his cladistic analysis, including that the overall homology of strictly defined trophophore larvae cannot be upheld, but that a less strictly defined trophophore larva (diagnosed by the sole possession of a prototroch) supports a clade Trochozoa (Annelida, Echiura, Sipuncula, Mollusca, Entoprocta). He also concluded that nemerteans, platyhelminths, and rotifers never had a prototroch. The cladogram apparently supporting these conclusions is depicted in Figure 2. I will assess the robustness of Rouse's conclusions by evaluating the treatment of the main ingredients that determine efficacy of a cladistic test: taxon selection, character selection, and primary homology determination.

Character selection—Rouse (1999) based his analysis on 140 morphological characters. However, 124 of these were specifically selected to resolve phylogenetic relationships within polychaetes (Rouse and Fauchald, 1997), nine characters coded for different larval ciliary bands, one character coded for a strictly defined trophophore larva, one character coded for downstream larval feeding, one character coded for larval proctonephridia, and two characters coded for the
problematic cleavage cross characters discussed above. This leaves just two characters not specifically selected to resolve polychaete relationships or based on larval morphology to resolve spiralian phylogeny: one coding for coelom and one coding for ventral nerve cord with circumoesophageal nerve ring. This scanty selection of morphological characters scarcely does justice to the complex problem of reconstructing spiralian phylogeny, and a failure to discuss or incorporate morphological evidence supporting different spiralian relationships in previous cladistic analyses, such as Haszprunar (1996) and Nielsen et al. (1996), clearly shows that Rouse's cladistic analysis fails as a proper test of competing hypotheses. This conclusion is in stark contrast with the exemplary treatment of character data in a previous analysis, Rouse and Fauchald (1995), where all pertinent data from earlier studies were carefully discussed and re-analyzed.

![Cladogram](image-url)

**FIG. 2.** Cladogram depicting the strict consensus based on the successive weighting analysis of Rouse (1999). Note that all polychaete subtaxa treated separately in Rouse's analysis are collapsed into a monophyletic Polychaeta.

Taxon selection—Rouse (1999) chose as terminal taxa for his analysis most of the accepted polychaete families, and in addition
Clitellata, Rotifera, Platyhelminthes, Nemertea, Mollusca, Sipuncula, Echiura, and Entoprocta. The deliberate exclusion from the analysis of a variety of other taxa, specifically those that lack larvae and prototrochs, even when they have previously been hypothesized to be closely related to the selected ingroup taxa, is defended by the following statement (p. 414): “I maintain that the exclusion of certain taxa does not invalidate the assessment of the hypotheses under examination. Even if they are subsequently placed among the taxa selected here for study, they will have either lost the features in question (e.g. a prototroch) or be primitively lacking them.” This statement immediately reveals that Rouse does not fully grasp how to exploit the strength of cladistics to test competing hypotheses of phylogenetic relationships and character evolution. A cladistic analysis can only test whether a character has a unique evolutionary origin, or has evolved more than once independently, if, and only if, all pertinent taxa are included in the analysis, also those hypothesized to primitively or secondarily lack the feature of interest. The true evolutionary dynamics of prototrochs and trochophore larvae will never become apparent when all taxa that lack these characteristics are excluded from the analysis. In such a case, character losses will always remain unsupported conjecture, convergent evolution an unexplored myth, and consequently, the phylogenetic significance of the feature in question will always remain essentially unstudied.

Previously performed morphological and molecular phylogenetic analyses can profitably be used as guides for deciding which additional taxa to explore. For example, Rouse (p. 412) mentioned Gastrotricha as a member of the currently formulated clade Eutrochozoa, but nevertheless decides to exclude them from the analysis. Strikingly, Rouse did not consider Gnathostomulida, even though they have often been united with other taxa included in Rouse’s analysis, in particular platyhelminths (together comprising Plathelminthomorpha: Zrzavy et al., 1998; Peterson and Eernisse, 2001). Similar considerations would necessitate the inclusion of other taxa, including Panarthropoda, Cyclophora, and Ectoprocta. Lophophorates are excluded because their inclusion “would bias the study away from any global homology of the ‘trochophore’” (p. 415). This curious statement can only be understood by assuming that Rouse wants to impress upon the reader that the overall homology of a strictly defined trochophore larva is so unlikely that it will be refuted even when the data matrix is structured in such a way as to bias the results towards accepting the overall homology of this larval type. Indeed, Rouse concludes that his analysis does not support the homology of a strictly defined trochophore larva with an apical organ, prototroch, metatroch, ciliated food groove, and a pair of protonephridia. Instead his defines a clade Trochozoa on the basis of possessing a less strictly defined trochophore larva: a larva with a prototroch. However, given that various potentially closely related taxa that lack larvae and ciliary bands have been excluded from the analysis, this means that the overall homology of the prototroch in trochozoans
has not really been tested, but merely assumed. In this respect it should be noted that although Rouse (p. 417) explicitly states that “an aim of this paper is to examine this hypothesis [that all protostome taxa lacking a trophophore larva have secondarily lost it]”, the exclusion of virtually all these taxa from his analysis simple does not allow this test! Instead, his puzzling conclusion (p. 440) that “groups such as the Arthropoda, Brachiopoda, Chaetognatha, Ectoprocta, Gastrotricha, Phoronida may have lost a prototroch” is totally gratuitous, and remains unaddressed by his own data.

Primary homology assessment—A prototroch (or a larva possessing one) is coded as a character in several cladistic analyses in addition to that of Rouse (Eernisse et al., 1992; Nielsen, 2001; Peterson and Eernisse, 2001). When the character scoring between these studies is compared, one finds no debate about character scoring for phyla such as Mollusca, Sipuncula, Echiura, and Entoprocta, all scored present for a prototroch. However, opinions do differ for the scoring of three phyla in particular: Platyhelminthes, Nemertea, and Rotifera. Since Rouse’s (1999) analysis is specifically aimed at resolving these issues, it is worth asking whether he has effectively used cladistics to help resolve these interpretational disagreements.

Rouse’s (1999: 440) conclusion is clear: “taxa such as Nemertea, Platyhelminthes and Rotifera appear to have never had a prototroch…” (italics added). Is this conclusion the corroborated result of a cladistic congruence test? Within Platyhelminthes, it is the pre-orval band of long cilia (not compound) found in polyclad Götte’s and Müller’s larvae that has been interpreted as a prototroch (Nielsen, 2001). In nemerteans the band of long compound cilia that surrounds the mouth field of heteronemertean pilidium larvae has been interpreted as a prototroch (Nielsen, 2001; Peterson and Eernisse, 2001). Finally, the adult rotiferan trochus, which is composed of compound cilia, has been interpreted as a prototroch (Nielsen, 2001; Peterson and Eernisse, 2001). In order to use a cladistic analysis to shed light on the potential homology of these bands with accepted prototrochs in taxa such as molluscs and entoprocts, the first step that needs to be undertaken is to accept primary homology of all these ciliary bands. This homology proposal can subsequently be corroborated if the cladistic analysis indicates that the distribution of ciliary bands in all taxa possessing them, including platyhelminths, nemerteans, and rotifers, is consistent with a single evolutionary origin. In contrast, initial homology is refuted when the distribution of ciliary bands indicates independent evolutionary origins.

In contrast, Rouse simply denies the presence of a prototroch in platyhelminths, nemerteans, and rotifers (p. 421). This decision not to accept primary homology of pre-orval ciliary bands across the Spiralia entirely removes any power his cladistic analysis may have in testing competing hypotheses for the evolution of ciliary bands in these taxa. A congruence test can only be performed on data actually included in the analysis. Consequently, his conclusion that platyhelminths, nemerteans, and rotifers “appear to have never had” a prototroch is not supported
by his analysis. In fact, given the structure of Rouse’s data matrix, it would be impossible to disconfirm the potential homology of the ciliary bands in these three problematic taxa with the prototrochs of other trochozoans! The cladogram (Fig. 2) suggests that Rouse’s parsimony analysis is singularly unhelpful in informing the interpretation of ciliary bands in platyhelminths, nemertean, and rotifers because the three problematic phyla are located as the nearest outgroups of a clade defined by possession of a prototroch. Consequently, tree length would not change even if these taxa would have been scored as possessing prototrochs. A single character change would explain the distribution of prototrochs on the cladogram.

It is of course possible that Rouse’s decision to score platyhelminths, nemertean, and rotifers as lacking a prototroch is defensible, and perhaps even preferable over scoring them as possessing a prototroch. Indeed, ciliary bands that might be interpreted as prototrochs are restricted to polyclads among platyhelminths, and heteronemertean among nemertean, creating uncertainty about the primitive states for the phyla. However, Rouse did consider the morphology of these larvae to justify his scores for a character on larval protonephridia, so the restricted distribution of these larval forms cannot be used as an effective argument for not scoring prototrochs for these two phyla. Another reason why platyhelminths, nemertean, and rotifers are scored as not having a prototroch may be based upon an argument (Rouse, 1999: 421) that the potential prototrochs in these taxa are not developed from special blastomeres, called trochosblasts, from which the prototrochs in other taxa such as molluscs and annelids develop. However, this argument is unconvincing because Rouse accepts a broad structural definition for the scoring of a prototroch in most polychaete taxa where cell lineage data are lacking. Thus Rouse did not supply any convincing reasons for not scoring a prototroch in platyhelminths and nemertean. The only reason that could explain why rotifers are scored as lacking a prototroch in Rouse’s analysis is that the trochos is an adult rather than a larval structure.

Rouse’s various explicit statements about evaluating hypotheses with a cladistic analysis clearly reveal that he does not fully grasp the fundamentals of constructing an effective cladistic test, even though he presents his analysis as such. As a consequence, readers who just focus on Rouse’s results rather than the design of his analyses may falsely ascribe special significance to his conclusions as being based on the corroborated results of a cladistic test. Restrictive taxon selection did not allow a proper assessment of the evolutionary dynamics of larval ciliary bands across different phyla, because all taxa lacking larvae and ciliary bands were left out of the analysis. Consequently, the possibility of convergence of prototrochs between phyla could not be assessed with the data set compiled by Rouse. The professed design of Rouse’s analysis as a cladistic test aimed to examine the hypothesis that all taxa lacking a trophophore larva must have lost it, is in reality little more than a misleading mask overlying conclusions that are merely recycled
assumptions. It should be noted that his conclusions concerning the absence of prototrochs in platyhelminths, nemerteans, and rotifers may be entirely justified by careful morphological study, but it is critical to realize that such conclusions are not novel insights emerging from the cladistic analysis.

It is essential to realize that "not all tests of cladograms are what they seem to be" (Kluge, 1997a: 90), and it is necessary to carefully consider the design of a cladistic analysis when studying its results in order to be able to understand the merit of the conclusions. I trust that the reader will not regard these critical remarks as being in any way depreciative of Rouse's efforts to shed light on the evolution of ciliary bands in the Spiralia. However, these remarks are intended as being helpful for the design of future cladistic analysis as effective tests. Rouse's seemingly integrated analysis seems an uneasy joining of two parts. His analysis is admirable in its unprecedented coverage of polychaete taxa, and the attention directed towards researching their character coding and scoring (see also Rouse and Fauchald, 1997). This part has yielded valuable insights into the evolutionary dynamics of larval types and ciliary bands within the annelids. However, the part dealing with the remaining spiralian phyla can be improved substantially, primarily through consideration of additional characters and taxa, and careful attention to the relationship of cladistic testing efficacy and primary homology scoring of the characters of interest.

Placing a problematicum: Myzostomida

Myzostomids join rank with taxa such as chaetognaths, pogonophores, and gastrotrichs in their ability to resist reliable phylogenetic placement on the basis of morphological data. Myzostomids are commensals or parasites on, or in echinoderm hosts, and it is reasonable to assume that their unique anatomy is reflective of a long evolutionary history spent in association with their hosts. Furthermore, as is typical of many phylogenetic problematica, myzostomids feature a mixture of characteristics that suggest affinities with disparate taxa (Rouse and Fauchald, 1995; Eeckhaut et al., 2000; Grygier, 2000; Zrzavy, 2001). However, the current majority opinion suggests that myzostomids are highly modified polychaetes (Fauchald and Rouse, 1997; Rouse and Fauchald, 1995, 1997; Westheide, 1996; Grygier, 2000; Nielsen, 2001; Zrzavy, 2001; Rouse and Pleijel, 2001), but uncertainty remains. Since 1996, there have been four attempts to use morphological cladistic analyses to position Myzostomida within the Metazoa (Haszprunar, 1996; Rouse and Fauchald, 1997; Zrzavy et al., 1998, 2001), and during the last half decade the first molecular sequence data have been extracted from myzostomids to be used in phylogenetic analyses. Here, I will not discuss the molecular results. The interested reader is referred to Eeckhaut et al. (2000), Littlewood et al. (2001), and Zrzavy et al. (2001).

Phylogenetic problematica such as Myzostomida should force workers to exert special care in the construction of their morphological
data matrices in order to prevent the results from being determined by character selection bias. The sequential publication of four cladistic analyses based upon four different morphological data matrices within a short time span, therefore creates an ideal circumstance for investigating whether cladistic analyses have generated any progress in our understanding of the phylogenetic position of Myzostomida.

To provide a context for the following discussion, I summarized in Table 1 those characters that have figured most prominently in recent debates about the position of myzostomids within the Metazoa: spiral cleavage, trochosoma larva, nectochaete larva (postmetamorphic juvenile with chaetae), chaetae, aciculae (internalised chaetae functioning as support rods inside the parapodia), marginal cirri (assumed sensory organs), muscular eversible axial pharynx, parapodia, anteriorly attached sperm flagellum, coelom, and segmentation. The table summarizes for each of the four cladistic analyses which of the characters have been included and excluded in order to assess the logic of character selection. It is not the issue here whether all imaginable data that might bear on the phylogenetic position of myzostomids are included or not. What is important, however, is to determine whether all available data that have been discussed explicitly in previous studies, are included or at least discussed in these analyses. The following will primarily focus on character selection rather than character interpretation and scoring. Figure 3 illustrates the possible affinities of myzostomids.

### Table 1. Comparison of character selection between recent cladistic analyses for those characters that have been considered most important for systematizing Myzostomida.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Spiral cleavage</td>
<td>Spiralians</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Trochosoma larva</td>
<td>Trochozoans</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Nectochaete larva</td>
<td>Most polychaetes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Chaetae</td>
<td>Various taxa including annelids, echiu-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>rans, brachiopods</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aciculae</td>
<td>Aciculate polychaetes</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Cirri</td>
<td>Aciculate polychaetes</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Muscular axial</td>
<td>Various aciculate</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>
I—Motivated chiefly by doubts about the presence of a coelomic body cavity and segmentation in myzostomids, Haszprunar included them as a taxon in his cladistic analysis. He found myzostomids to be a sister group of a clade (Sipuncula (Echiura (Polychaeta Clitellata))). What is most striking with respect to character selection in Haszprunar’s study, is that none of the characters that could indicate affinities with or within Polychaeta are coded in the data set (nor is Polychaeta divided into subtaxa). It should be noted, however, that the potential homology of the marginal cirri of myzostomids with the ventral and dorsal cirri of polychaetes was uncertain at the time of Haszprunar’s analysis, and only with the publication of the detailed study of the nervous system of the myzostomid *Myzostoma cirriferum* by Müller and Westheide (2000) do we now have a more robust basis for proposing this homology. Müller and Westheide (2000) found the innervation pattern of myzostomid parapodia and marginal cirri to be identical to that of polychaete parapodia and cirri. Haszprunar also left out the character on sperm ultrastructure which had previously been used to suggest a close relationship of myzostomids and acanthocephalans, and he in fact concluded (p. 23) that future inclusion of this character might revise his results.

Rouse and Fauchald (1997): myzostomids as polychaetes—Rouse and Fauchald included myzostomids in some of the analyses that were part of their comprehensive morphological cladistic study that included most polychaete families. Although the lack of many ‘typical’ polychaete features in myzostomids may have caused some instability in their placement among polychaetes in Rouse and Fauchald’s analyses, their results nevertheless indicate that they could be members of a clade of polychaetes with hypertrophied axial pharynges within a larger clade Aciculata (defined by possession of aciculae). This result is not surprising given the use of most characters that could indicate polychaetan affinities (with the exception of a nectochaete larva). However, the restricted focus of the study did not allow Rouse and Fauchald to test whether other, excluded, characters could suggest a position outside the polychaetes.
Zrzavy et al. (1998): myzostomids as neotrochozoans II—Zrzavy et al. (1998) used the largest morphological data matrix compiled to date, for assessing the placement of Myzostomida among the Metazoa. Their analysis suggested myzostomids to be the sister group of a clade (Echiura Pogonophora Polychaeta Aphanoneura Clitellata), a result broadly similar to that of Haszprunar (1996). However, with the exception of parapodia, none of the characters that could indicate a relationship with all or some polychaetes were included in the analysis.

Zrzavy et al. (2001): myzostomids as prosomastigozoans—The analysis of Zrzavy et al. (2001) was specifically aimed at resolving the long lasting controversy about myzostomid affinities, and they found the surprising result that myzostomids formed part of a clade they christened Prosomastigozoa, named on the basis of possessing sperm with the flagellum attached anteriorly to the cell body and curving posteriad. The remaining members of this novel metazoan clade were Rotifera, Acanthocephala (together Syndermata), and Cycliophora. The
result of this analysis is especially noteworthy because the clade is also supported by Zrzavy et al.'s molecular data set. Such congruence of independent data sources is especially gratifying in our current age of conflicting placements of various phyla on the basis of morphological and molecular data. However, in accord with the previously discussed studies, Zrzavy et al. (2001) did not include all available morphological data. Similar to Zrzavy et al. (1998), none of the characters that could indicate a relationship of myzostomids with Polychaeta or some of its members were included in the analysis, nor was Annelida split up at least in Polychaeta and Clitellata. For a proper representation of all phylogenetically informative aspects of myzostomid morphology, it would be necessary to revise the matrix substantially. However, Zrzavy et al. (2001) do use the morphological data set of Rouse and Fauchald (1997) in combination with their 18S sequence data, but even this combination never yielded a monophyletic clade comprising myzostomids and polychaetes, even when the molecular and morphological data were weighted 1:10. However, the efficacy of this experiment should be regarded with caution. Only 16 annelids for which 18S sequences were available were included in this analysis. Only four of these species represent members of the clade Aciculata within which myzostomids were placed in Rouse and Fauchald's (1997) analysis.

Zrzavy et al. (2001: 186) conclude that the grouping of myzostomids with Syndermatata and Cycliophora is never contradicted by any of their analyses “regardless of character combinations, character weights, species sampling and combinations, and tree-building methods.” Naturally, this only indicates robustness of the results for their particular data set. However, are the results equally robust in the face of changes to information content of the morphological matrix? I performed one experiment. I added one character to the matrix and scored it present only for Annelida and Myzostomida. This character can, for example, be interpreted as a nectochaete larva, or as aciculae or cirri. In the first case the taxon Annelida can be interpreted as a taxon Polychaeta (assuming nectochaete larvae as primitive for polychaetes), while in the latter two cases the taxon Annelida can be interpreted as a polychaete subclade (Aciculata). Adding this one character results in the separation of myzostomids from the other prosomastigozoans, as they end up in an large polytomy with other protostome taxa, including the neotrochozoans which do not form a clade any more (strict consensus). A control analysis yielded a monophyletic Prosomastigozoa including myzostomids as in Zrzavy et al. (2001). Both analyses are heuristic searches of 100 random addition replicates, TBR branch swapping, with the “outgroup” taxon of Zrzavy et al. (2001) excluded. Character 3 (segmentation) was scored present for myzostomids as suggested by Zrzavy et al. (2001). Although Zrzavy et al. (2001) set out to specifically test the position of myzostomids within the Metazoa with a cladistic analysis, their results are determined by bias in character selection.
Addition of a single character incompatible with the favored results leaves the position of myzostomids unresolved.

The above discussion clearly indicates that none of the morphological cladistic analyses published so far has exhaustively used all readily available morphological evidence for placing myzostomids. This conclusion is also relevant for interpreting the results of total evidence analyses (Zrzavy et al., 1998, 2001) that were based on the same morphological data matrices. Although all necessary morphological evidence was available for all four studies (with the exception of new data supporting primary homology of myzostomid marginal cirri with polychaete cirri which became available only in 2000), none of the examined data matrices is expressly designed for testing the phylogenetic placement of the myzostomids. Instead, the different studies merely provided differing perspectives of uncertain merit on a vexing phylogenetic problem. Such practices are not diagnostic of an efficiently operating sophisticated falsification research program aimed in producing progress in understanding. Table 1 may be helpful in guiding the construction of future data matrices that give alternative hypotheses equal chances of corroboration or refutation.

STANDING ON THE SHOULDERS OF GIANTS? STEPS TOWARDS A COORDINATED CLADISTIC RESEARCH PROGRAM FOR THE METAZOON

"If I have seen further, it is by standing on the shoulders of giants." This statement, most famously attributed to Sir Isaac Newton, became a popular motto during the Scientific Revolution (Gould, 1995), as it aptly underlines a maxim that remains at the heart of our modern conception of scientific progress: we can only see further if we explicitly use the results of past efforts as a foundation for further explorations. I discussed three case studies to explore whether current practices in metazoan cladistics bear any witness to the worth of this insight.

The least ambiguous signal that recent morphological cladistic analyses of the Metazoa use each other's data is not necessarily a reason for exaltation. Some of the most recent and most comprehensive studies have used data matrices compiled by different workers rather uncritically, with the result that shortcomings of these data sets continue to have detrimental effects on the results of new analyses (Jenner, 2001a; unpublished data). The first case study of spiralian cleavage cross patterns illustrates this general problem by showing that even unanimous character scoring across different studies may not be a reliable indicator of the phylogenetic significance of a character.

The two other case studies point to a more difficult problem. When one wants to effectively use cladistics to arbitrate between competing hypotheses, special care must be taken in the construction of the data matrix. A matrix is a mosaic of characters each with their own specific phylogenetic significance, and not all characters will be equally important for testing a given hypothesis. It is therefore critical to know from the beginning which data can potentially support or refute a given
hypothesis, and to ensure that none of these data are excluded without explicit justification. A large matrix that contains no scoring errors may nevertheless cripple the testing powers of a cladistic analysis when it is uncritically compiled. The second and third case studies dealing with the evolution of spiralian ciliary bands and the placement of myzostomids among the Metazoa clearly illustrate that cladistic analyses need not be effective tests, even though they are presented as such.

If we can regard the case studies discussed in this paper as a reasonable representative sample of current practices in metazoan cladistics, a conclusion strengthened by the results of a more comprehensive evaluation of metazoan phylogenetics that will appear elsewhere (unpublished data, R. A. Jenner), then a reconsideration of our current research strategies is opportune (see also Jenner, 2001b). In general, recent cladistic analyses have not yet fully exploited the power of cladistics to test available competing hypotheses. Instead, different workers have used largely independent data matrices to plant different trees of uncertain merit in the already densely grown forest of metazoan phylogenies. Consequently, it is very difficult to distinguish true progress in our understanding of metazoan macroevolution, and the mere change of opinions with the passage of time. This is strikingly illustrated if we just consider character choice in two subsequent publications by the same group, for example, for placing the 'acelomate' worms, Platyhelminthes, Nemertea, and Gnathostomulida. Of the 18 morphological synapomorphies supporting sister group relationships of these three taxa included in the data matrix of Zrzavy et al. (1998), only two were included in the data matrix of Zrzavy et al. (2001). Two additional characters in the latter matrix are reminiscent but not identical to characters included in the matrix of the first study. Amazingly, Zrzavy et al. (2001) supply no justification whatsoever for excluding more than 75% of the characters that their previous analysis revealed as being those most important for placing platyhelminths, nemerteans, and gnathostomulids.

The key to progress in cladistic research lies in repetition of research cycles that make maximal use of previously performed studies (Kluge, 1997b). A first cladistic analysis of the phylogenetic relationships of a given group of organisms only produces a foundation upon which to built new cladistic analyses. For many taxa not even a first estimate of their phylogenetic relationships is yet available, and in other instances only a single attempt has been made to reconstruct their phylogeny. However, higher-level metazoan phylogenetics finds itself in a fortunate situation. The past decade has witnessed the generation of multiple alternative phylogenies on the basis of both molecular and morphological evidence. All these hypotheses are available for testing. For morphological cladistics that means one has to carefully consider all proposed sister group hypotheses for each terminal taxon, in order to insure that all known characters that can potentially corroborate or refute any of these sister group hypotheses for all taxa are included in a
single data matrix. In other words, data matrix construction has to become much more explicit, and data selection has to be clearly justified in terms of the hypotheses to be tested. Only then will a cladistic analysis be an effective test of alternative hypotheses, and only then will the results of cladistic analyses be more readily interpretable in terms of their strengths and weaknesses through a more transparent link between data matrix and cladogram.

REFERENCES


