Cladistics (1993) 9:1–39

ARTHROPOD PHYLOGENY: A COMBINED APPROACH

Ward C. Wheeler[1,3], Paulyn Cartwright[1,2] and Cheryl Y. Hayashi[1,2]

[1] *Department of Invertebrates, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024-5192, U.S.A. and*
[2] *Department of Biology, Yale University, U.S.A.*

Received for publication 5 August 1992; accepted 24 November 1992

Abstract—Ribosomal and ubiquitin protein coding sequence data are generated from 20 arthropods and five close relatives. These molecular data are combined with morphological characters derived from the literature to approach arthropod phylogenetics from the perspective of total evidence. Trilobita were included in the analysis through morphological comparison alone. The overall data strongly support arthropod monophyly. Recent molecular analyses which have yielded different results are shown to have been based on few characters, few taxa or both. The most parsimonious explanation of the data is (Annelida + (Onychophora + ((Trilobita + Chelicerata) + (Crustacea + (Myriapoda + Hexapoda))))). The data are largely concordant both internally among data sets and externally with previous cladistic anatomical analyses.

Introduction

The question of the origin of the arthropods has lately become increasingly vexing. Early discussion of the problem was based solely on hard features of external morphology. More recent studies have extended analysis from external attributes to internal, developmental, locomotory and molecular features in search of unequivocal evidence. These efforts have not, however, been successful in generating consensus for several reasons. Foremost among these is the use of single-character hypotheses. Many workers have trumpeted a single character (or character system) and relegated all other evidence to convergence or misidentification. A second problem has arisen most recently with molecular analyses—depauperate sampling. Arthropods may be the most diverse creatures on earth, yet many analyses have examined a pittance of taxa. A sample of four or five taxa with major lineages entirely absent is unlikely to yield a stable result, especially when the arguments hinge on the placement of taxa not sampled (such as the velvet worms, Onychophora). Although it is hopeless to survey these groups exhaustively, it is possible through judicious choice to create a robust taxonomic sample which would represent the major lineages in arthropod phylogeny well enough to estimate groundplan features of higher taxa.

Monophyly Versus Polyphyly

Long thought to be monophyletic (Fig. 1; Siebold and Stannius, 1848; Snodgrass, 1938), Tiegs (1947) proposed that arthropods had originated twice (Fig. 2). A Myriapoda + Hexapoda + Onychophora (="Uniramia") lineage was proposed and apposed by a group comprised of trilobites, crustaceans and chelicerates (later called "TCC" by Cisne, 1974). This diphyletic theory supposed that the attributes commonly used to argue for common ancestry (such as a hardened exoskeleton,

[3] To whom correspondence should be addressed.

0748-3007/93/010001+39$08.00/0

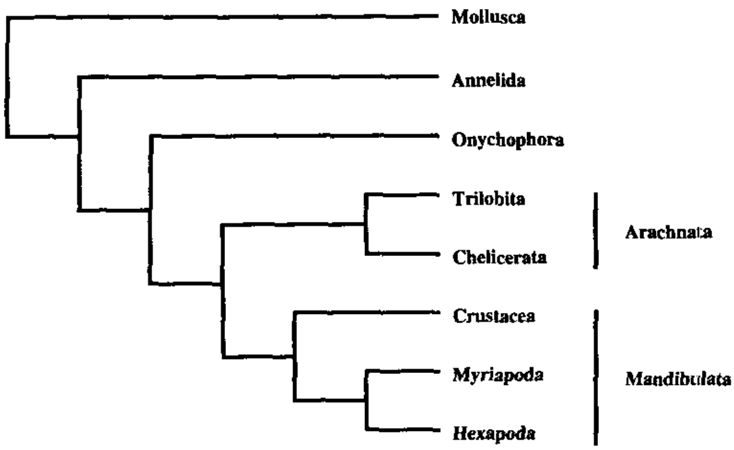

Fig. 1. Monophyletic hypothesis of arthropod origins showing subgroups (Snodgrass, 1938; with a monophyletic Myriapoda).

compound eyes and paired jointed appendages) were convergent due to similarities in lifestyle. The uniramians and TCC taxa were said to have been derived from unique annelid-like ancestors and independently wandered into the arthropod-like grade.

Manton (1964, 1966, 1972, 1973, 1979) with Anderson (1973, 1979) and others (Tiegs and Manton, 1958) extended the diphyletic theory into one of polyphyly (Fig. 3). While they upheld the uniramian lineage envisioned by Tiegs (1947), the TCC group was dismissed. This necessitated at least four independent originations. Manton argued that the mandibles of crustaceans were not homologous to those in myriapods and hexapods. The basis for the argument was that crustacean mandibles are multisegmented and use the gnathobasic endites to masticate food, whereas the uniramians posses a single-piece structure which bites with the tip of the whole limb ("jaws"). Although Manton did not propose any specific kinship for crustaceans or chelicerates with other non-articulates, she nonetheless discounted the character evidence for monophyly as convergence.

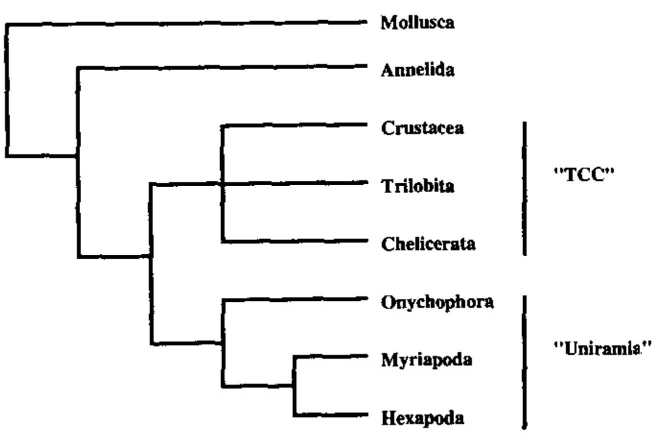

Fig. 2. Diphyletic theory of arthropod origins showing the subgroups Trilobita + Crustacea + Chelicerata ("TCC") and "Uniramia" (Tiegs and Manton, 1958).

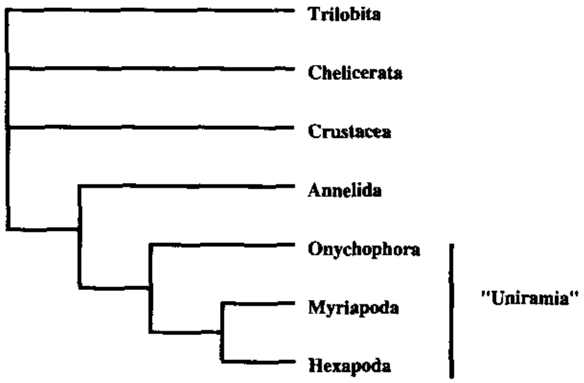

Fig. 3. Hypothesis of Lake (1990) showing arthropods as paraphyletic with respect to annelids and molluscs. Onychophora were not considered in this analysis.

Anderson added support to this position by introducing developmental fate maps. These features show specific similarity among annelid worms, onychophorans and tracheates (Myriapoda + Hexapoda). The "Uniramia" and annelids share a similar ordering of three fate map tissues with the stomodaeum anterior to the midgut and mesoderm as opposed to posterior in crustaceans and chelicerates. On this basis, Anderson united Annelida with the Uniramia but, like Manton, did not propose specific relationships for the remaining arthropods. Schram (1978) extended these observations to the pycnogonids (sea spiders), showing that they fit the general chelicerate pattern.

Weygoldt (1979, 1986) and others (Kristensen, 1975, 1981) have criticized both the observations and argumentation of Manton and Anderson. Manton's arguments on the non-homology of arthropod mandibles are predicted on their being entirely *too* different to be homologous. Of course, degree of apomorphy is no argument against homology (Hennig, 1966), it may simply be that a great deal has happened to these structures since the taxa diverged. To counter the actual character information that remains (mainly the embryological fate map information), there have been two complementary approaches. Weygoldt (1979) has directly confronted the interpretation of the fate maps themselves; questioning whether the proposed arrangements are characteristic of the groups sampled.

The second approach comes in the enumeration of large numbers of derived characters which unite chelicerates, crustaceans, myriapods and hexapods. Weygoldt (1986) listed eight features uniquely shared by his Arthropoda (Onychophora + Euarthropoda) and seven by his Eurathropoda (the usual arthropods—Trilobita + Chelicerata + Crustacea + Myriapoda + Hexapoda). To unite the Onychophora with arthropods he cited the following derived features: (1) cuticle containing chitin and protein (annelid chitin contains collagen), presence of resilin and regular molts involving ecdysone; (2) mixocoel; (3) dorsal blood vessel with paired ostia; (4) pericardial sinus and septum; (5) brain from the fusion of the acron with two to three postoral neuromeres; (6) nephridia with sacculi; (7) appendages with both extrinsic and intrinsic muscles that terminate in sclerotized claws; and (8) centrolecithal egg with superficial cleavage. His characters for the arthropods proper include: (1) hardened exoskeleton from sclerotized cuticle; (2) cephalon from one preoral and four postoral appendages; (3) biramous appendages with coxal endites; (4) a single pair of compound eyes; (5) four pairs of simple median eyes;

Table 1

Character support for competing hypotheses of arthropod origins. The numbers represent steps (character state originations) required by the most parsimonious resolution of the ideas of relationship put forth by various authors. The "Best" column represents the most parsimonious solution of the data. The "Arthropoda" column contains the number of steps required for the most parsimonious explanation of the data in light of arthropod monophyly (only in the case of Ubiquitin was this distinct from the overall most parsimonious solution). "Uniramia" is the hypothesis of Manton—that Onychophora + Tracheata be monophyletic (1973), "Uniramia + Annelida" is this proposal of Anderson (1979) incorporating Manton's ideas and "TCC" that of Cisne (1974). The columns do not appear to "add up" because the most parsimonious explanation of the various data sets are not necessarily identical.

	Steps				
	Best	Arthropoda	Uniramia	Uniramia + Annelida	TCC
Morphology	126	= 126	139	149	127
18S rDNA	478	= 478	489	494	481
Ubiquitin	387	≠ 391	414	414	410
Molecular	908	= 908	919	922	911
Total	1037	= 1037	1058	1071	1038

(6) nephridia present in only the last four cephalic and first two post-cephalic segments; and (7) cilia present only in spermatozoa. Variations in these features within arthropods are thought to be secondary modifications. Other workers, such as Boudreaux (1979), have lists that exceed 20 characters uniting arthropods (see morphological data matrix in Table 1).

Although Manton (1973) and Anderson (1979) and Schram (1986) have attempted to gainsay these attributes as "prone to convergence", no one has as yet produced morphological characters which would tilt the balance in favor of multiple arthropod origins.

Molecular Examinations

In the study of animal relationships published by Field et al. (1988), arthropods and their relatives were a minor section. Even so, their data contained 18S ribosomal RNA sequences of several arthropod taxa (*Artemia, Drosophila, Spirobolus* and *Limulus*) and several arthropod relatives such as annelids and molluscs. These were sprinkled among other protostome taxa. Although the authors stated that a polyphyletic theory of arthropod origins cannot be supported by their data, the results are difficult to interpret. These data have been variously reanalyzed and augmented with inconsistent results.

Lake (1990) reanalyzed a subset of these sequences and came to conclusions which differed from those of the original authors. His scheme had arthropods as a paraphyletic slurry stretched between non-arthropod protostomes, worms and deuterostomes (Fig. 4). This scheme is wildly incongruent with previous morphological work but, like the Field et al. scheme, had arthropods at least adjacent to one another.

In their examination of the placement of tongue worms (Pentastomida) within crustaceans, Abele et al. (1989) briefly examined other arthropods. Abele et al. also examined 18S rRNA sequences, but used a small subset (four sequences) of the taxa previously analyzed. This study found that the arthropod sequences studied were most parsimoniously monophyletic. Turbeville et al. (1991) also found a

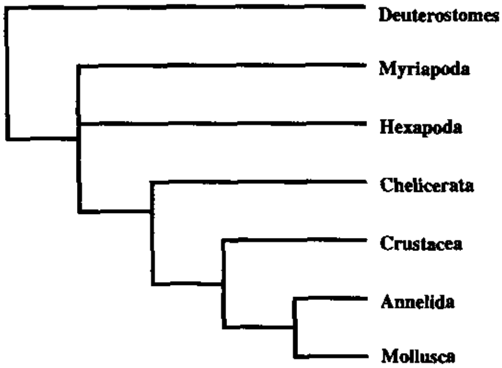

Fig. 4. Polyphyletic theory of arthropod origins (Manton, 1973; Anderson, 1973, 1979).

monophyletic Arthropoda. They did this by adding a few chelicerate sequences to their previous data set (Field et al., 1988) and removing some of the less well behaved taxa (like *Drosophila*).

These analyses of the same molecular data disagree. This is most likely due to several factors. The first of which is the unique sampling regimes followed in these studies; after Field et al., specific taxa have been included or excluded for reasons which are hard to fathom. In some cases, "fast clock" taxa have been removed and in others the reasoning is unclear. Since the data are the same, subsampling of exemplars from exemplars must be partially to blame. The second potential cause of disagreement is in the analysis of the sequences. The original analysis was a distance-based method (Fitch and Margoliash, 1967). These methods have been criticized at length for the unnecessary problems and inconsistencies they introduce (Swofford, 1981; Farris, 1981, 1985, 1986). There certainly is no reason not to perform character-based analysis, especially when it allows direct comparison with other (perhaps non-molecular) data. Lake used his own method to analyze the taxa and introduced a number of additional untested assumptions. Additionally, in none of these analyses has more than a cursory attention been made to sequence alignment. As the primary homology test used in molecular analysis, it is an important factor. This is especially true in light of the fact that the sequences previously used have been derived from RNA sequencing. RNA can be sequenced in only a single direction and is prone to many sequencing artifacts, hence the data have a high error rate compared with DNA analysis.

Adding to these problems is the poor sampling employed in these studies. Certain potentially important taxa (such as Pycnogonida and Onychophora) were unsampled and those groups which were examined are represented by a single member.

The use of a single molecule and lone exemplars seems likely to lead to an unstable situation. Such small data sets are unlikely to withstand the smallest analytical or empirical perturbation. In this study we have attempted to address these concerns through the use of multiple (if minimal) exemplars for large groups, DNA sequencing (with more objective alignment) and adding another molecule to the basket of data bearing on the arrangement of these taxa.

In sum, previous analyses present three cohorts of data. First is the large group of derived homologies which are used to unite the Arthropoda. Second is the fate map and "jaw" characters which unite the "Uniramia". Third is the 18S rRNA data which are used to support both of these and other morphologically unconsidered options. These data have not been analyzed together, except in an explanatory context, and no general attempt to distill a pattern from this seeming incongruence has been made. This is a general shortcoming of these analyses. Molecular studies, especially, have not incorporated other evidence into their schemes or assayed the relative levels of support for alternate hypotheses. Here, we attempt to redress some of these problems through robust sampling of both taxa and characters, permitting the discussion of both overall levels of support and of systematic relationships among groups.

Methods

TAXONOMIC SAMPLING

It is impossible to avoid exemplars in studies of such vast groups. It is possible, however, to choose these exemplars wisely, such that maximum cladistic diversity is sampled at minimum cost. This is done two ways. At a minimum, it is necessary to sample all the major groups. Here this means examining representatives not just of the standard groups (worms, chelicerates, crustaceans, myriapods and insects), but also important basal groups such as Onychophora and Pycnogonida. We were unable to sample the potentially important tardigrades. Though they may be close relatives of arthropods, fortunately, none of the hypotheses examined here (monophyly, "TCC" and polyphyly) center on the placement of this group.

Although fossil taxa (for the most part) are not amenable to molecular analysis, they may be scored for a large fraction of hard external anatomical characters. This is especially true for trilobites. One of the best-known fossil lineages, these arthropods occupy an important position in overall arthropod relationships (Weygoldt and Paulus, 1979). For this reason we have included this group (*sensu* Ramsköld and Edgecombe, 1991) in the morphological and combined analyses.

A second sampling requirement is that higher groups (like Crustacea, Hexapoda or Chelicerata) must be minimally represented by taxa which span the maximum cladistic diversity in the groups which can be sampled. An example of this would be that insects would be better represented by a combination of *Drosophila* and some basal apterous taxon than by 10 holometabolous lineages. By mapping taxa as closely to the basal-most split as possible, distinctions between autapomorphous and groundplan characteristics can be better made. Without this minimal sample, it is impossible to determine whether an attribute is unique to that lineage (or some higher taxon within the group) or is representative of the common ancestor of the entire group. Character states which are shared between (or among) these representatives form a parsimonious estimate of groundplan characters. Obviously, the greater the sample (including progressively more derived divergences) the better this estimate will be.

Molluscs were used as an outgroup in assessing articulate (annelids, onychophorans and arthropods) relationships. Since annelids were determined to be the most basal articulate lineage they then served as another outgroup for arthropodan relationships. Between ingroup and outgroup taxa, a total of 25 individual taxa were examined for both molecular and morphological characters, while the trilobites were scored for the morphological data alone (see Appendices 2 and 3).

Molecular Sampling

In order to achieve a more robust molecular sample, sequences were determined for two loci, the 18S nuclear ribosomal DNA and Polyubiquitin. Each of the loci was amplified using PCR and directly sequenced. All DNA fragments were sequenced in both directions to ensure accurate results.

The 18S rDNA sequences yielded 669 bases sequenced in three pieces. The rDNA is present in several hundred copies per genome and codes for the small subunit ribosomal RNA.

Polyubiquitin is a nuclear locus which occurs in low copy number (4–10, among most known repeats) (Sharp and Li, 1987; Lee et al., 1988). The repeat length is 228 bases, of which 188 were determined (the remainder was primer). Since ubiquitin has several repeats, a mixed PCR product was amplified and sequenced (variation was scored by IUPAC codes). Differences among the sequences are then observed as sequence ambiguities which were observed in both directions of sequencing (an artifact would most likely be present in only a single direction).

The rationale behind choosing these two molecules is in their diversity in molecular life-style. The 18S rDNA is present in hundreds of copies and is not translated. The ubiquitin is present in few copies and codes for protein. Presumably, any pattern common to these two disparate systems is representative of history and not some immediate adaptation to an internal or external environment.

There are then two types of sampling in this study, molecular and taxic. In both cases, care has been taken to ensure a robust sample.

Morphological Characters

Morphological characters were collected from the literature (see references in Appendix 2). This survey yielded 100 morphological characters summarized in Appendices 2 and 3.

Molecular Characters

The procedure for garnering sequence data consisted of isolating DNA from specimens collected in liquid nitrogen or alcohol. The DNA segments of interest were then amplified via the polymerase chain reaction and sequenced.

DNA Isolation

A phenol-based DNA purification procedure was employed. The samples, if dry, were rehydrated >12 hours in a solution of 0.25 M EDTA (pH 8.0), 0.05 M Tris (pH 7.0), 0.25% sodium dodecyl sulfate and 0.1 M NaCl. If the samples were preserved in alcohol, the sample was washed three times in >1000 volumes absolute ethanol over 24–48 hours. This was done to ensure the removal of impurities in the preserving ethanol. The sample was then vacuum-dried for >12 hours and rehydrated as with the dry specimens and both alcohol and dried material protocols proceed identically.

After rehydration, the tissue was homogenized and incubated with 0·1 mg/ml proteinase K at 55°C for 3 h. The solution was then extracted with equal volumes of phenol, phenol : chloroform : isoamyl alcohol (25 : 24 : 1) and chloroform :

isoamyl alcohol (24 : 1). The nucleic acids were then ethanol/ammonium acetate precipitated, washed in 70% EtOH and resuspended in TE. This procedure yielded fairly clean (OD 260/280 > 1·5) DNA.

Amplification

The procedures followed are based on combinations and modifications of the methods specified by Perkin–Elmer–Cetus and Allard et al. (1991). The 18S rDNA sequences were amplified and sequenced as three fragments "A", "B" and "C". Fragments B and C are contiguous. Ubiquitin was amplified as one 228 base pair fragment.

Six primers were used to amplify and sequence the 18S rDNA fragments. The "A" fragment was spanned by 5'-CCTGAGAAACGGCTACCACATC-3' and 5'-TAACCG-CAACAACTTTAAT-3', the "B" fragment by 5'-GGTGAAATTCTTGGACCGTC-3' and 5'-GTTTCAGCTTTGCAACCAT-3' and the "C" by 5'-ATGGTTGCAAAGCT-GAAAC-3' and 5'-GAGTCTCGTTCGTTATCGGA-3'. Ubiquitin was amplified and sequenced by 5'-TTGACCGGAAAGACCATCAC-3' and 5'-GGTCTTCACGAA-GATCTGCA-3'.

Double-stranded Amplification

Fifty picomoles of each of the 5' and 3' primers were used to insure geometric, double-stranded amplification. The buffers used were those specified and included by Perkin–Elmer. Amplification first employed 25 cycles with 1 min denaturation at 94°C, 1 min annealing at 50–55°C (depending on the primer) and 1 min 15 s elongation at 72°C. This was performed in a Perkin–Elmer–Cetus automated thermal-cycler. The only variation from the manufacturers' procedures came in the use of 1.25 units of *Taq* DNA polymerase rather than the recommended 2.5. After the final amplification cycle, samples were cooled to 4°C and five of the 100 microliter reaction were visualized via agarose electrophoresis.

After visualization the product was concentrated and purified with either Centricon 30 microconcentrator tubes (Amicon) or Gene-Clean (Bio 101) following the manufacturers protocols. This yielded approximately 40 microliters product volume. The purified product was either reamplified for single-stranded product or sequenced directly.

Single-stranded Amplification

Two microliters of the purified double-stranded product was reamplified in 100 microliters with 50 picomoles of each of the two primers in turn (to generate both strands individually). All other aspects of the amplification were unchanged except that only 20 cycles were used.

DNA Sequencing

Sequencing proceeded via a modified, Sanger et al. (1977) dideoxy (^{35}S labelled) protocol using Sequenase (United States Biochemical). Each template was annealed with 50 picomoles of the appropriate sequencing primer. The samples were then treated as specified in the Sequenase version 2.0 protocols. Gels were prerun for 1 h prior to loading and run for 3 h at 2000 V. Gels were then dried at 80°C for 45 min under vacuum. The gels were then exposed directly against Kodak X-AR5 X-ray film for 16–36 hours. All sequences were determined and verified by sequencing in both directions.

The corresponding data of Field et al. (1988) and Kim and Abele (1990) were not used for two reasons. Although there is overlap in the sequence locations used here, they are not completely coincident. Second, the RNA sequences of these published data possess a great deal of missing and ambiguous nucleotide assignments. In several cases, when the rDNA was sequenced, rRNA "autapomorphies" disappeared (although this did not happen with the Kim and Abele data). The DNA sequences also contain fewer gaps. Since many of these sites were phylogenetically informative, we were forced to sequence these taxa ourselves.

The *Drosophila* sequences were taken from Tautz et al. (1988) for 18S rDNA and Lee et al. (1988) for Ubiquitin.

Alignment

Alignment of DNA sequences was accomplished via a modified Needlemen and Wunsch (1970) method implemented in the program MALIGN (Wheeler and Gladstein, 1992). The aim of this program is to produce alignments which yield parsimonious cladograms. The alignment options searched for multiple equally costly alignments using cladogram length as an optimality criterion. The ratio of insertion (gap) to change (base substitution) costs was four to one.

Phylogenetic Analysis

Hennig86 (Farris, 1988) was used to construct parsimonious schemes of relationship of these taxa. Ambiguous characters were treated as missing. The morphological characters were examined with employing exact solution option (ie*) with characters 50, 52 and 89 non-additive (see Appendix 2). The options mh* bb* were used with DNA-based characters, all non-additive. When morphological characters were included together with the molecular data the characters were pooled and the overall most parsimonious solution(s) determined.

Congruence

Several phylogenetic analyses were performed. First the morphological data were analyzed independently. This was done to establish the supported hypothesis of arthropod relationships based on morphology alone. Second, the 18S rDNA was examined and phylogenetic hypotheses erected on this basis. The ubiquitin data were then used in a third analysis. The molecular data were pooled in order to compare these data to the morphological results. Finally, all the data (morphological, 18S rDNA and Ubiquitin) were combined and the overall best-supported hypothesis determined. The congruence statistic of Mickevich and Farris (1981) was calculated for each of these comparisons and these values are shown in Table 2.

Table 2

Levels of incongruence among data sets measured by the statistic of Mickevich and Farris (1981). The integer values in the lower triangle are the lengths of the single and combined data derived cladograms, while the upper fractions are the Mickevich and Farris values.

	Incongruence (Mickevich and Farris, 1981)			
	Morphology	18S rDNA	Ubiquitin	Molecular
Morphology	126	0.004 9	0.067	0.002 9
18S rDNA	607	478	0.047	—
Ubiquitin	550	908	387	—
Molecular (18S rDNA + Ubiquitin)	1 037	—	—	908

Incongruence overall three data sets = 0.044.

Results

Morphological

There were 100 characters which were garnered from the literature (Appendices 2 and 3). These characters concerned many aspects of the morphology of the organism from the specifics of eye structure to the organization of fate maps. Where homologies were unclear (characters 50, 54 and 55), states were coded in such a way as to be biased against arthropod monophyly. Of these assignments, 7% were either missing or inappropriate to the taxon. Almost half of these missing entries were due to trilobites alone (missing data 3.8% without trilobites which had

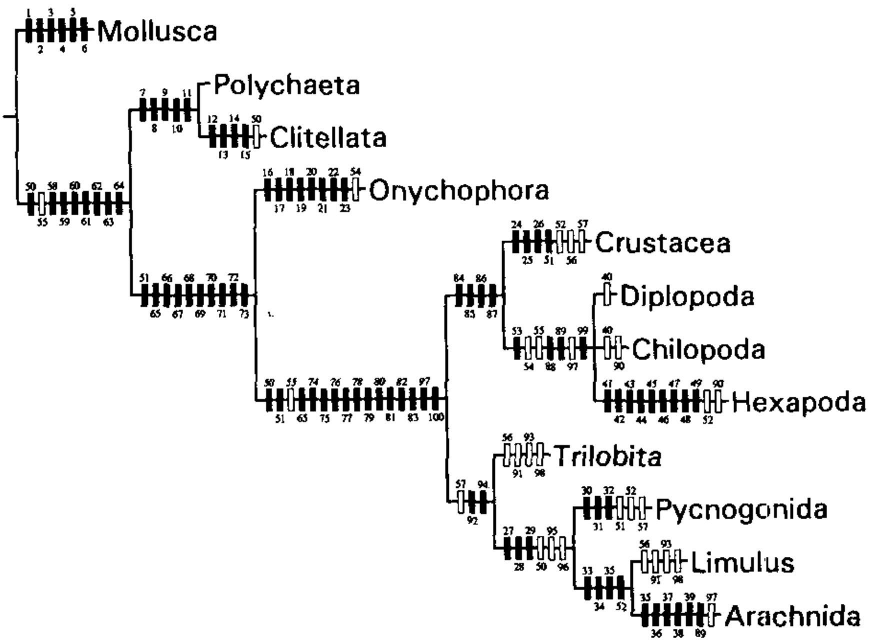

Fig. 5. Cladogram derived from the morphological characters of Appendices 2 and 3. This is the strict consensus of the two most parsimonious cladograms at 126 steps (CI = 0.84, RI = 0.85). The changes are numbered by character. The solid boxes represent non-homoplastic states and the open homoplastic. Characters were optimized using default (in essence delayed transformation of characters) with the "Squeeze" option of CLADOS (Nixon, 1992). This biases the character optimizations against single origins.

35% missing entries for the morphological characters), and these were mainly internal characters. Cladograms constructed with and without Trilobita were identical with respect to the remaining taxa. Arthropod monophyly is upheld in this scheme as is that of Articulata, Onychophora + Arthropoda, Arachnata (Trilobita + Chelicerata), Chelicerata (Pycnogonida + Euchelicerata), Euchelicerata (Xiphosura + Arachnida), Arachnida, Mandibulata (Crustacea + Tracheata) and Tracheata (Hexapoda + Myriapoda). There were two most parsimonious arrangements which differed in support for myriapod monphyly (strict consensus; Fig. 5), each of which required 126 steps, had a retention index of 0.85 and a consistency index of 0.84. The most parsimonious arrangement of taxa consistent with the "Uniramia" hypothesis of arthropod polyphyly required 13 additional steps for these data (139 steps). The inclusion of the annelids with "Uniramia" required yet another 10 steps (149 steps).

Molecular

The 18S rDNA yielded a total of 669 aligned positions (629 of which were analyzed phylogenetically—the other 40 were primer sequences). Strict consensus of the 17 most parsimonious arrangements based on these data (length 478, CI = 0.60, RI = 0.57) yielded the scheme shown in Fig. 6. These data are fairly concordant

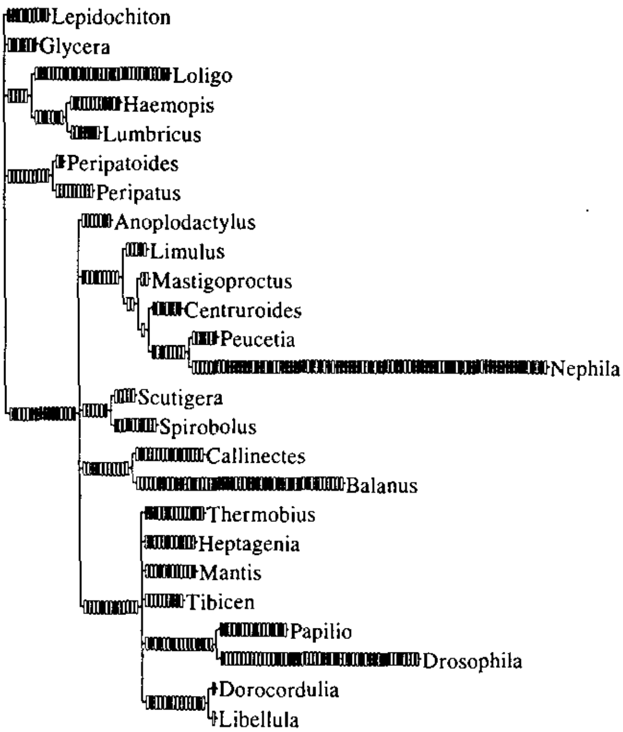

Fig. 6. Consensus cladogram derived from 18S rDNA sequence data. This is the strict consensus of 17 equally parsimonious cladograms at 478 steps (CI = 0.60, RI = 0.57). The solid boxes represent non-homoplastic states and the open homoplastic. Characters were optimized as in Fig. 5. The optimization of characters after polytomies should not be given much credence, but are included for illustrative purposes.

with the morphological information. The insects, arthropods, onychophorans and clitellate worms are supported as monophyletic groups. Furthermore, the Onychophora are the sister group to a monophyletic Arthropoda. Although there is a great degree of irresolution, the only component which is discordant with the morphological results is the placement of the squid (*Loligo*) with the clitellates (*Lumbricus* and *Haemopis*). Successive approximation weighting (Farris, 1969; Carpenter, 1988) chooses three of the 18 equally supported cladograms with considerably more resolution. This scheme supports several additional components concordant with morphological analysis: Crustacea, Araneae, Arachnida, Euchelicerata and Tracheata. The relationships among the Pycnogonida (as represented by *Anoplodactylus*), Chelicerata, Crustacea and Tracheata are in complete disagreement with morphological results. As mentioned, this most parsimonious hypothesis of arthropod monophyly required 478 steps. The minimum cost explanation of these data with "Uniramia" as a group requires 11 extra steps and 16 more to include Annelida as in the scheme of Anderson (1973).

The Ubiquitin data behaved very differently from the ribosomal. There were three equally parsimonious cladograms of length 387 with CI of 0.31 and RI of 0.42. These data, on their own support few groups unequivocally (Fig. 7), and almost all of those which are supported are non-traditional (at best). One of these arrangements was chosen by successive approximation weighting which increased resolution but decreased the meager levels of concordance even further. These data did, however, offer differential support for the various proposed hypotheses of

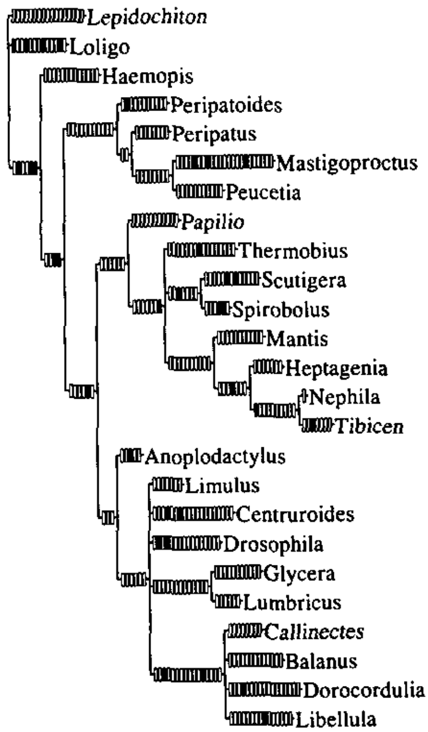

Fig. 7. Consensus cladogram derived from ubiquitin sequence data. This is the strict consensus of 14 equally parsimonious cladograms of 387 steps (CI = 0.31, RI = 0.42). The solid boxes represent non-homoplastic states and the open homoplastic. Characters were optimized as in Fig. 6.

relationship among these taxa (Table 1). The most parsimonious arrangement of these taxa subject to arthropod monophyly required four additional steps at 391. The most parsimonious arrangement of taxa consistent with the Uniramia hypothesis requires 23 additional steps (adding Annelida to Uniramia also requires 23 more steps).

The molecular evidence taken *in toto* increased concordance to the point where only the placement of one group is in disagreement (Fig. 8). There were seven equally parsimonious arrangements of these taxa at 908 steps with a CI of 0.45 and a RI of 0.44. The only group which stands out is the Crustacea. The placement of this taxon within the Hexapoda is curious to say the least. This seems to be the result of polarity difficulties within the arthropods more than actual character support for this placement. There are few character states which are shared between the Onychophora and some but not all of the arthropods. Hence, there are few opportunities for polarity assessment. An increase in length of only four steps will move the Crustacea out of the hexapods to their more traditional position with the Tracheates constituting the Mandibulata. Successive weighting did not substantively alter the conclusions except by pulling the pycnogonid out of the chelicerates and placing it at the base of the arthropods (not found in any of the unweighted cladograms) and supporting the grouping of the three worms as monophyletic. The molecular characters in sum favor arthropod monophyly over "Uniramia" by 11 steps and over Annelida + "Uniramia" by 14.

Total Analysis

When all characters are taken together, support for arthropod monophyly is strong. Each of the major groups is upheld and previous schemes of arthropod

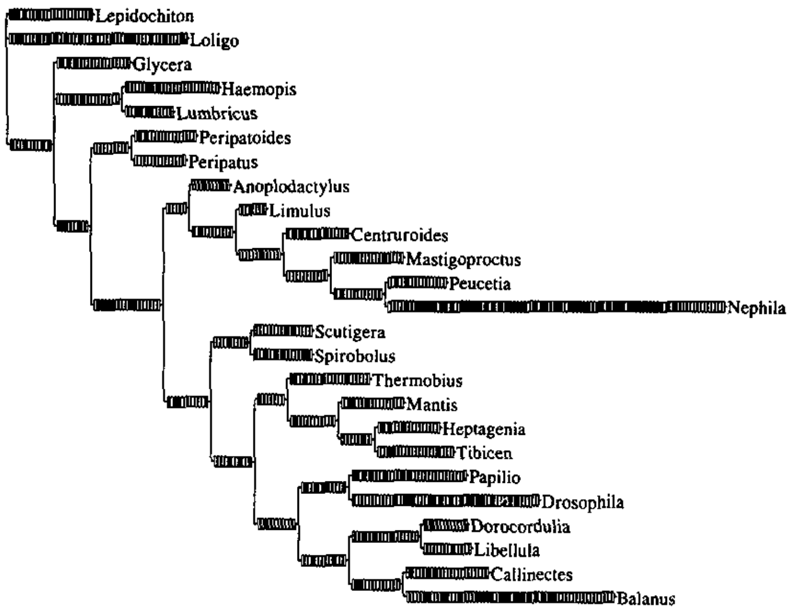

Fig. 8. Consensus cladogram derived from molecular sequence data. This is the strict consensus of two equally parsimonious cladograms of 908 steps (CI = 0.45, RI = 0.44). The solid boxes represent non-homoplastic states and the open homoplastic. Characters were optimized as in Fig. 6.

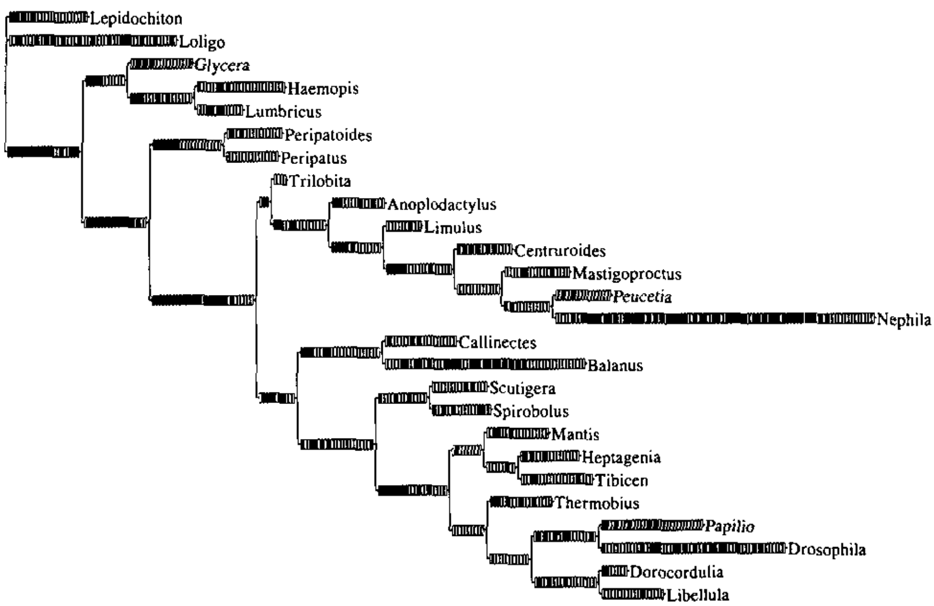

Fig. 9. Consensus cladogram derived from molecular and morphological data. This is the single most parsimonious cladogram (all the disagreement was with respect to the relative placement of insect taxa) of 1037 steps (CI = 0.50, RI = 0.61). The solid boxes represent non-homoplastic states and the open homoplastic. Characters were optimized as in Fig. 6.

relationships are supported—including Arachnata (Fig. 9). The total data yield seven equally parsimonious arrangements requiring 1037 steps (CI = 0.50, RI = 0.61). This overall scheme is robust to successive approximations weighting (although the interrelations among the insect taxa vary). Overall, 21 extra steps are required to support the "Uniramia" and 34 for Annelida + "Uniramia". The relative levels of support for each of the major hypotheses is shown in Table 1. In most cases (with the exception of Ubiquitin alone), the most parsimonious arrangement of taxa supports arthropod monophyly. All data sets—individual and combined—support arthropod monophyly as the most parsimonious hypothesis of relationships (see Appendix 6 for a summary of character support).

Most prominent among the unsupported groupings is the "Uniramia". The union of Onychophora with Tracheates (Tiegs and Manton, 1958) is overwhelmingly refuted. The only support there is for this group comes in fate map features (really a single character) and "jaw" structure. Both of these are most parsimoniously optimized as independent gains or losses. Even the supposed homology of the onychophoran "antennae" really only excludes a potential arthropod feature as opposed to supporting "Uniramia". Cisne's (1974) Trilobita + Chelicerata + Crustacea is also not supported (but the margin is very small—a single step). This group was always based on admitted character plesiomorphy. Even though the Trilobita are unavailable for molecular analysis, the implied relationships among the living groups alone (i.e. Crustacea + Chelicerata) is contradicted.

The hypothesis of Lake (1990), though not directly tested (i.e. no deuterostomes), is at least indirectly falsified. Although a definitive test would require sampling other protostome phyla (such as echiurans and nemerteans), it seems that Lake has the deuterostomes misplaced and the rest of the relationships upside down (Fig. 4). His scheme would require the reversal of nearly all the polarities of the characters in Appendix 2. Such a scheme would require that each of the 13 uncontroverted features which support Arthropoda as monophyletic would have to be both originated and then lost, as well as adding steps to the more equivocal states (at a cost of 33 steps). Rotating a single branch (articulates + Molluscs + Sipunculata) would yield a scheme in near complete concordance with morphologically-based ideas of invertebrate relationships. Much greater sampling will be required and more generally accepted methods of analysis applied to the data to substantiate this unique view.

Discussion

These data support arthropod monophyly unequivocally. Each of the data sets individually also support monophyly over the "Uniramia" hypothesis. The preferred relationships of the taxa largely agree with previous morphological analyses even with the addition of almost 900 nucleotide bases of comparative DNA sequence data. The groups Articulata, Arthropoda, Arachnata, Chelicerata and Tracheata are each upheld by both molecular and morphological analysis. Even though the two molecular sets differ in their specific conclusions (irresolution on the part of the 18S rDNA and non-standard groupings with the Ubiquitin), their sum conclusions were completely resolved (at least for the higher taxa) and largely concordant with morphological studies (Table 2). Only Mandibulata and Annelida (unresolved—not contradicted) were not supported by the molecular data, but this is most likely the result of polarization problems with the non-additive multistate characters (the inclusion of tardigrades might have been informative here).

This resolution is especially striking in light of the fact that the two molecular data sets are individually largely unresolved. The reason for this efficient dovetailing comes from the fact that the irresolution is not due to character conflict as much as to lack of support for groups. This is supported by the low levels of incongruence shown in Table 2. The data do not clash. Irresolution in Ubiquitin data is further exacerbated by heteroplasmies coded as missing.

Although the overall data strongly support this single hypothesis, there is some discordance among the individual sources of information (Table 2). The disagreement is most likely due to the historical effects which have molded the changes of these characters in these lineages. Notably, the greatest degree of discordance is not between the morphological and molecular information, but between the molecular sources themselves. The ribosomal sequence data seem to agree more with morphological analysis than do the ubiquitin. This is most likely the result of three factors.

Ubiquitin is relatively short (228 bases, of which 188 are used). The ribosomal DNA present over four times as many nucleotide positions. Second, internal sequence variation can mask phylogenetic signal. Characters which vary within an individual might be more properly called traits (Nixon and Wheeler, 1990), hence are inappropriate for historical analysis (this was almost entirely unobserved in the 18S sequences). Unfortunately, not all taxa are heteroplasmic at the same positions.

Additionally, these cases of multiple bases at a sequence site were scored as missing data, a treatment which is, though currently unavoidable, undesirable. Ubiquitin sequences may prove more informative when the entire repeats are cloned and the copies mapped and sequenced individually. Finally, all the Ubiquitin sequence variants we observed are silent (no amino acid replacements—there can be none with rDNA). While constraining change greatly, this attribute also allows for nucleotides to change relatively freely within their codon sets.

Ribosomal DNA have been implicated in having their own selectively-driven peccadillos (Wheeler and Honeycutt, 1988), and these data do have some unique aspects (placing the squid *Loligo* with clitellate worms). The gross discordance of the Ubiquitin data, however, is largely absent.

Conclusions

The strength of this analysis is not that the data sets chosen here are perfect for the problem, but that the foibles of the individual sources of information are not generally shared. Since Ubiquitin follows one set of rules and 18S another, the only source of shared information is history. The disagreements between these data reflect the unique selective regimes of the molecules themselves and not some multiplicitive and confusing origin. The addition of morphological evidence into the fray only strengthens this effect. It is highly unlikely that some factor would be canalizing protein coding, ribosomal and morphological features in such a way that they are all fooled identically. This is, of course, the rationale and justification for a phylogenetic approach based on total information (Kluge, 1989). In this way we can erect the hypothesis which best explains all the data.

Successive weighting was used to test the degree of character conflict. This method of character weighting was not used to establish relationships but test the sensitivity of the results to perturbation. As more data were added, weighting had diminished influence. Successive weights were able to change Ubiquitin-based schemes greatly, affected the summed molecular evidence less and the total data not at all. While not formally so, this aproach tells us about the stability of our conclusions (Farris et al., in prep). The inability to dislodge hypotheses by this procedure is a measure of their robustness. Arthropod monophyly is not easily perturbed.

The disagreement of this analysis with some molecular analyses most likely comes from the very features which give it stability—multiple sources of data sampled in multiple taxa. Previous analyses have used very few taxa and some of the methods employed are limited to almost absurdly small numbers of taxa. This has led to inconsistent and inexplicable results. As predicted by simulation (Wheeler, 1992), the increased taxonomic representation here had a positive effect on the analysis as measured by overall congruence among all data. The great success of taxonomic representation also points to a potential shortcoming of this discussion of arthropod relationships—extinct taxa.

Although it is possible to examine many of the morphological features of extinct taxa, molecular samples are almost entirely unavailable. The inclusion of trilobites in this analysis (with 95% missing data) turned out to be crucial to understanding character transformations, even though it did not affect the topology of living taxa. Trilobites had antennae, yet living Arachnata (chelicerates) do not. Without the inclusion of this Paleozoic group, antennae would be thought of as a property of mandibulates and not arthropods. The absence of antennae in chelicerates would then be thought of as primitive. Trilobites show the antennal absence in chelicerates to be a derived condition. Hypotheses of the homology of onychophoran "antennae" [though postoral—but Boudreaux (1979) suggests that they are partly acronal] would also have been affected [the inclusion of tardigrades—which have no antennae—and "armored" lobopods such as *Hallucigenia*, *Microdictyon* and *Luolishania* (Ramsköld and Xianguang, 1991) in this analysis might also have helped to throw light on this problem].

The inclusion of this single fossil taxon revealed an effect which can only grow with the inclusion of more such taxa. These results are tantalizing; inclusion of a greater sample of other arthropod and "proto-arthropod" fossils (Briggs and Fortey, 1989; Emerson and Schram, 1990; Briggs et al., 1992) will no doubt revise additional hypotheses of character support.

Interestingly (and not surprisingly), the relationships proposed here support almost every group elaborated by Snodgrass (1938), including Arachnata (though not named as such), Mandibulata and Tracheata (Fig. 10). Although he did not place Pycnogonida (other than within Chelicerata), the only real disagreement between this study and that of Snodgrass is the support here for a monophyletic Myriapoda.

Based on all the available evidence collected and presented here, the Arthropoda are monophyletic. Hypotheses proposing multiple origins of this group are shown to be entirely unsupported—based on single characters, non-empirical notions of character transformation or plesiomorphy. The support for this conclusion comes from samples, though not complete by any means, which are at least not single taxa. The molecular sources of information have distinct origins in function, organization and fate. The combination of these data with pre-existing morphological evidence hopefully will make this a more stable analysis of arthropod relationships.

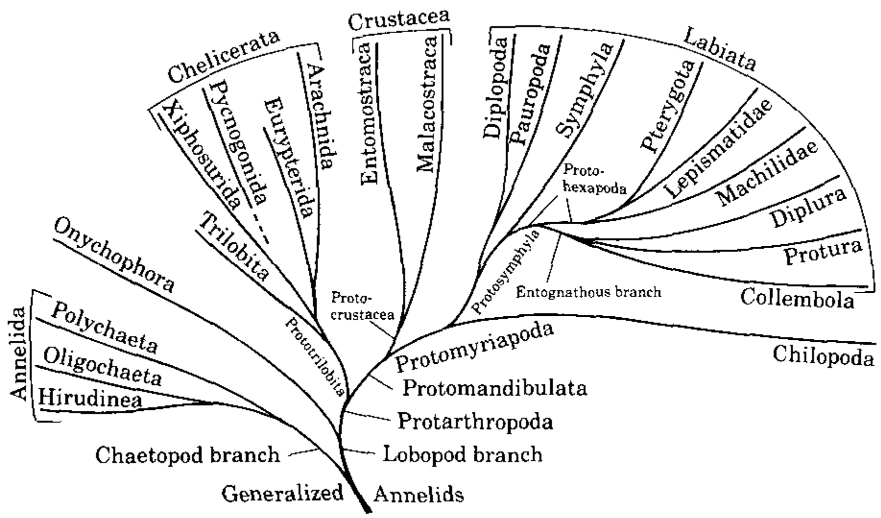

Fig. 10. Phylogeny of Snodgrass (1938). This scheme is completely concordant with the arrangement proposed here with the exception of Snodgrass's support of a paraphyletic Myriapoda (note the dichotomous nature of Snodgrass's arrangement).

Acknowledgments

We would like to acknowledge the help of E. Broach, M. Broach, J. Carpenter, R. DeSalle, G. Edgecombe, N. Eldredge, K. Fitzhugh, D. Jacobs, J. Gatesy, N. Landman, C. Michelski, M. Nitabach, K. Nixon, N. Platnick, M. Stowe, P. Vrana, L. Washer and two anonymous reviewers for aid in research and discussion in the preparation of this manuscript.

REFERENCES

ABELE, L. G., W. KIM AND B. E. FELGENHAUER. 1989. Molecular evidence for the inclusion of the phylum Pentastomida in the Crustacea. Mol. Biol. Evol. 6: 685–691.

ALLARD, M. W., D. L. ELLSWORTH AND R. L. HONEYCUTT. 1991. The production of single-stranded DNA suitable for sequencing using the polymerase chain reaction. BioTechniques 56: 779–827.

ANDERSON, D. T. 1973. Embryology and Phylogeny in Annelids and Arthropods. Pergamon, Oxford, New York, Toronto, Sydney, Braunschweig.

ANDERSON, D. T. 1979. Embryos, fate maps, and the phylogeny of arthropods. *In:* A. P. Gupta (ed.). Arthropod Phylogeny. Van Nostrand, New York, pp. 59–106.

BERGSTRÖM, J. 1979. Morphology of fossil arthropods as a guide to phylogenetic relationships. *In:* A. P. Gupta (ed.). Arthropod Phylogeny. Van Nostrand, New York, pp. 3–58.

BOUDREAUX, H. B. 1979. Arthropod Phylogeny with Special Reference to Insects. J. Wiley & Sons, New York.

BRIGGS, D. E. G. AND R. A. FORTEY. 1989. The early radiation and relationships of the major arthropod groups. Science 246: 241–243.

BRIGGS, D. E. G., R. A. FORTEY AND M. A. WILLS. 1992. Morphological disparity in the Cambrian. Science 256: 1670–1673.

BRUSCA, R. C. AND G. L. BRUSCA. 1990. Invertebrates. Sinauer, Sunderland, Massachusetts.

CARPENTER, J. M. 1988. Choosing among equally parsimonious cladograms. Cladistics 4: 291–296.

CISNE, J. L. 1974. Trilobites and the origin of arthropods. Science 186: 13–18.

CUTLER, B. 1980. Arthropod cuticle features and arthropod monophyly. Experientia 36: 953.

ELDREDGE, N. 1974. Revision of the suborder Synziphosurina (Chelicerata, Merostomata), with remarks on merostome phylogeny. Amer. Mus. Nov. 2543: 1–41.

EMERSON, M. J. AND F. R. SCHRAM. 1990. The origin of crustacean biramous appendages and the evolution of Arthropoda. Science 250: 667–669.

FARRIS, J. S. 1969. A successive approximations approach to character weighting. Syst. Zool. 18: 374–385.

FARRIS, J. S. 1981. Distance data in phylogenetic analysis. *In:* V. A. Funk and D. R. Brooks (eds). Advances in Cladistics: Proceedings of the First Meeting of the Willi Hennig Society. New York Botanical Garden, New York, pp. 3–22.

FARRIS, J. S. 1985. Distance data revisited. Cladistics 1: 67–85.

FARRIS, J. S. 1986. Distances and statistics. Cladistics 2: 144–147.

FARRIS, J. S. 1988. Hennig86, Version 1.5, Program and Documentation. Port Jefferson, New York.

FARRIS, J. S., A. G. KLUGE AND J. M. CARPENTER. Successive and spectral weight, in prep.

FIELD, K. G., G. J. OLSEN, D. J. LANE, S. J. GIOVANNONI, M. T. GHISELEN, E. C. RAFF, N. R. PACE AND R. A. RAFF. 1988. Molecular phylogeny of the animal kingdom. Science 239: 748–753.

FITCH, W. M. AND E. MARGOLIASH. 1967. The construction of phylogenetic trees. Science 155: 279–284.

HENNIG, W. 1966. Phylogenetic Systematics. Univ. of Illinois Press, Urbana.

HENNIG, W. 1981. Insect Phylogeny. J. Wiley & Sons, New York.

HYMAN, L. H. 1967. The Invertebrates. Vol. 6. Mollusca I. Aplacophora, Polyplacophora, Monoplacophora, Gastropoda. The Coelomate Bilateria. McGraw-Hill, New York.

JAMIESON, B. G. M. 1988. On the phylogeny and higher classification of the Oligocheata. Cladistics 4: 367–410.

KIM, W. AND L. G. ABELE. 1990. Molecular phylogeny of selected decapod crustaceans based in 18S rRNA nucleotide sequences. J. Crust. Biol. 10: 1–13.

KING, P. E. 1973. Pycnogonids. Hutchinson University Library.

KLUGE, A. 1989. A concern for evidence and a phylogenetic hypothesis for relationships among *Epicrates* (Boidae, Serpentes). Syst. Zool. 38: 1–25.

KRISTENSEN, N. P. 1975. The phylogeny of hexapod "orders". A critical review of recent accounts. Z. Zool. Evol. Forch. 13: 1–44.

KRISTENSEN, N. P. 1981. Phylogeny of insect orders. Ann. Rev. Entomol. 26: 133–157.

LAKE, J. A. 1990. Origin of the Metazoa. Proc. Nat. Acad. Sci. USA 87: 763–766.

LEE, H. J., A. SIMON AND J. T. LIS. 1988. Mol. Cell. Biol. 8: 4727–4735.

MANTON, S. M. 1964. Mandibular mechanisms and the evolution of arthropods. Phil. Trans. R. Soc. (B) 247: 1–183.

MANTON, S. M. 1966. The evolution of arthropodan locomotory mechanisms. Part 9. Functional requirements and body design in Symphyla and Pauropoda and the relationships between Myriapoda and Pterygota. J. Linn. Soc. (Zool.) 46: 103–141.

MANTON, S. M. 1972. The evolution of arthropodan locomotory mechanisms. Part 10. Locomotory habits, morphology and evolution of the hexapod classes. Zool. J. Linn. Soc. 51: 203–400.

MANTON, S. M. 1973. Arthropod phylogeny—a modern synthesis. J. Zool. 171: 111–130.

MANTON, S. M. 1979. Functional morphology and the evolution of the hexapod classes. *In:* A. P. Gupta (ed.). Arthropod Phylogeny. Van Nostrand, New York, pp. 387–466.

MICKEVICH, M. F. AND S. J. FARRIS. 1981. The implications of congruence in *Menidia*. Syst. Zool. 30: 351–370.

NEEDLEMAN, S. B. AND C. D. WUNSCH. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48: 443–453.

NIXON, K. 1992. CLADOS, Version 1.2. Program and Documentation. Trumansburg, New York.

NIXON, K. AND Q. D. WHEELER. 1990. An amplification of the phylogenetic species concept. Cladistics 6: 211–223.

PAULUS, H. F. 1979. Eye structure and the monphyly of Arthropoda. In: A. P. Gupta (ed.). Arthropod phylogeny. Van Nostrand, New York, pp. 299–383.

POCOCK, R. J. 1983. On the classification of the tracheate Arthropoda. Zool. Anz. 16: 271–275.

RAMSKÖLD, L. AND G. D. EDGECOMBE. 1991. Trilobite monophyly revisited. Hist. Biol. 4: 267–283.

RAMSKÖLD, L. AND H. XIANGUANG. 1991. New early Cambrian animal and onychophoran affinities of enigmatic metazoa. Nature 351: 225–228.

SANGER, F., S. NICKLEN AND A. R. COULSON. 1977. Sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. 74: 5463–5467.

SCHRAM, F. R. 1978. Arthropods: a convergent phenomenon. Fieldiana Geology 39: 61–108.

SCHRAM, F. R. 1986. Crustacea. Oxford University Press.

SHARP, P. M. AND W.-H. LI. 1987. Molecular evolution of ubiquitin genes. Trends. Ecol. Evol. 2: 328–332.

SHULTZ, J. W. 1990. Evolutionary morphology and phylogeny of Arachnida. Cladistics 6: 1–38.

SIEBOLD, C. T. W. VON AND H. STANNIUS. 1848. Lehrbuch der vergliechenden Anatomie der Wirbellosen Tiere. Veit, Berlin.

SNODGRASS, R. E. 1938. Evolution of the Annelida, Onychophora and Arthropoda. Smiths. Misc. Coll. 97: 1–159.

SNODGRASS, R. E. 1952. A Textbook of Arthropod Anatomy. Cornell Univ. Press, Ithaca, New York.

SWOFFORD, D. L. 1981. On the utility of the distance Wagner procedure. *In:* V. A. Funk and D. R. Brooks (eds). Advances in Cladistics: Proceedings of the First Meeting of the Willi Hennig Society. New York Botanical Garden, New York, pp. 25–43.

TAUTZ, D., J. M. HANCOCK, D. A. WEBB, C. TAUTZ AND G. A. DOVER. 1988. Complete sequence of the rRNA genes in *Drosophila melanogaster*. Mol. Biol. Evol. 5: 366–376.

TIEGS, O. W. 1947. The development and affinities of the Pauropoda, based on a study of *Pauropus silvaticus*. Q. J. Micr. Sci. 82: 165–336.

TIEGS, O. W. AND S. M. MANTON. 1958. The evolution of the Arthropoda. Biol. Rev. 33: 255–337.

TURBEVILLE, J. M., D. M. PFEIFER, K. G. FIELD AND R. A. RAFF. 1991. The phylogenetic status of the arthropods, as inferred from 18S rRNA sequences. Mol. Biol. Evol. 8: 669–702.

WEYGOLDT, P. 1979. Significance of later embryonic stages and head development in arthropod phylogeny. In A. P. Gupta (ed.). Arthropod Phylogeny. Van Nostrand New York, pp. 107–136.

WEYGOLDT, P. 1986. Arthropod interrelationships—the phylogenetic—systematic approach. Z. Zool. Syst. Evol. 24: 19–35.

WEYGOLDT, P. AND H. F. PAULUS. 1979. Untersuchungen zur Morphologie, Taxonomie und Phylogenie der Chelicerata. II. Cladogramme und die Enfaltung der Chelicerata. Z. Zool. Syst. Evol. 17: 117–200.

WHEELER, W. C. 1992. Extinction, sampling, and molecular data. *In:* M. Novacek and Q. D. Wheeler (eds). Extinction and Cladistic Analysis. Columbia Univ. Press, New York, pp. 205–215.

WHEELER, W. C. AND R. L. HONEYCUTT. 1988. Paired sequence difference in ribosomal RNAs: evolutionary and phyogenetic implications. Mol. Biol. and Evol. 5: 90–96.

WHEELER, W. C. AND D. L. GLADSTEIN. 1992. MALIGN, Version 1.01. Program and Documentation. New York, New York.

Appendix 1

Table A1
Taxa used in the study.

Mollusca	
Cephalopoda	*Loligo pealei*
Polyplacophora	*Lepidochiton cavernae*
Annelida	
Polycheata	*Glycera sp.*
Oligocheata	*Lumbricus terrestris*
Hirudinea	*Haemopis marmorata*
Onychophora	
Peripatoidae	*Peripatus trinitatis*
Peripatopsidae	*Peripatoides novozealandia*
Trilobita	groundplan of Ramsköld and Edgecombe (1991) (morphological analysis only)
Chelicerata	
Pycnogondia	*Anoplodactylus portus*
Xiphosura	*Limulus polyphemus*
Scorpiones	*Centruroides hentzii*
Uropygi	*Mastogoproctus giganteus*
Araneae	*Nephila clavipes*
Araneae	*Peucetia viridans*
Crustacea	
Cirrepedia	*Balanus sp.*
Malacostraca	*Callinectes sp.*
Myriapoda	
Chilopoda	*Scutigera coleoptera*
Diplopoda	*Spirobolus sp.*
Hexapoda	
Zygentoma	*Thermobius sp.*
Ephemerida	*Heptagenia sp.*
Odonata	*Libellula pulchella*
Odonata	*Dorocordulia lepida*
Dictyoptera	*Mantis religiosa*
Auchenorrhyncha	*Tibicen sp.*
Lepidoptera	*Papilio sp.*
Diptera	*Drosophila melanogaster*

Appendix 2. Morphological Characters

Synapomorphies of Mollusca:

1. Reduction of coelom and development of an open haemocoelic circulatory system.
2. Mantle from closed body wall.
3. Mantle shell gland produces spicules (and shell).
4. Ventral body wall muscles developed into a muscular foot.
5. Radula.
6. Chambered heart with separate atria and ventricles (Brusca and Brusca, 1990; Hyman, 1967).

Synapomorphies of Annelida:

7. Annelid head.
8. Epidermal paired setae (or bundles).
9. Longitudinal muscles broken into bands instead of sheets (Brusca and Brusca, 1990; Fitzhugh, pers. comm.).
10. Annelid nephridial system.
11. Cuticle with collagen but no chitin except in setae and stomodaeum (Bourdreaux, 1979).

Synapomorphies of Clitellata:

12. Clitellum.
13. Hermaphroditism.
14. Direct development without intervening larval stages.
15. Cerebral ganglion moved into anteriormost trunk segment (Brusca and Brusca, 1990; Jamieson, 1988).

Synapomorphies of Onychophora:

16. Suppression of external segmentation.
17. Oblique muscle layer in body wall.
18. Subcutaneous haemal channels.
19. Oral papillae.
20. Body papillae and scales.
21. Slime glands.
22. Non-migratory gastrulation.
23. Lobopods with pads and claws (Brusca and Brusca, 1990; Boudreaux, 1979).

Synapomorphies of Crustacea:

24. Two pairs of antennae.
25. Antennae biramous.
26. Nauplius or egg-nauplius stage in ontogeny (Schram, 1986).

Synapomorphies of Chelicerata:

27. Tagmosis into prosoma and opisthosoma without distinct head.
28. First appendages chelicerae (or cheliphores) of three articles.
29. "Typically" four pairs of walking legs (Weygoldt, 1986; Weygoldt and Paulus, 1979).

Synapomorphies of Pycnogonida:

30. Opisthosoma reduced.
31. Proboscis (Weygoldt and Paulus, 1979; King 1973).
32. Inverse retina in four median eyes (Paulus, 1979).

Synapomorphies of Euchelicerata:

33. Prosoma a carapace-like shield.
34. First or second opisthosomal segment modified into a genital somite.
35. Opisthosomal respiratorial lamellae absent (0), as book gills (1) or enclosed to form book lungs (2).

Synapomorphies of Arachnida:

36. Extraintestinal digestion.
37. Five simple lateral eye.
38. Slit sensillae (Weygoldt and Paulus, 1979).
39. Vitreous body present in median eyes (Paulus, 1979). 36 and 37 are questioned by Schultz (1990) but he also adds several features uniquely derived in arachnids.

Synapomorphies of Myriapoda:

40. Palps on first and second maxillae absent. Other features mentioned as myriapod synapomorphies may be more broadly distributed (such as the organ of Tömösvary) or have unclear homology relationships ("stink" glands) (Brusca and Brusca, 1990).

Synapomorphies of Hexapoda:

41. Thorax divided into three segments each with a pair of limbs.
42. Locomotory limbs six-segmented.
43. Abdomen with 12 segments (except for Collembola).
44. Distinct thorax and abdomen.
45. "Knee" as joint versus segment.
46. Labium.
47. Hexapod-type cephalization.
48. Abdominal cerci.
49. Two primary pigment cells in ommatidia (Hennig, 1981; Kristensen, 1975; Boudreaux, 1979; Paulus, 1979).

Other characters:

50. Antennae and palps: 0 = nothing, 1 = prostomial/preoral palps, 2 = postoral antennae.
51. Lateral eyes 0 = no, 1 = simple, 2 = compound, 3 = stalked compound ordered because 1, 2 and 3 all formed from a homogeneous secretion from a subcornegenous cell layer (Paulus, 1979).
52. Median eyes 0 = none, 1 = four, 2 = two—unordered (Weygoldt and Paulus, 1979).
53. Tracheae. *Onychophora are scored as unknown "?" due to the structural* disimilarities between the tracheae in myriapods and hexapods. This feature also occurs in some araneomorph spiders, but again, the structures, though superficially similar, appear quite different in detail.

54. Whole limb feeding structure. Although this homology statement is almost surely erroneous (Boudreaux, 1979; and many others), this character is included to force a rigorous test of arthropod monophyly (Manton, 1979).

55. Ordering of fate map tissues (anterior—stomodeum–midgut–mesoderm—posterior vs. anterior—midgut–mesoderm–stomodeum—posterior) (Pycnogonid from Schram, 1978; Anderson, 1979).

56. Fundamentally biramous post-antennal appendages (Tiegs and Manton, 1958).

57. Digestive diverticula.

58. Segment origination in caudal elongation or proliferation zone (Weygoldt, 1986).

59. Schizocoelous metamerism between preoral acron (prostomium) and the non-metamerism telson (periproct).

60. Acronal protocerebrum serving the eyes and containing an association center connected with pedunculate bodies.

61. Double ventral somatic nerve cord.

62. Dorsal and ventral longitudinal muscles.

63. Coelomoducts, their vestiges and derivatives.

64. Dorsal blood vessel with forward-going peristalsis (Boudreaux, 1979).

65. Loss of ectodermal cilia 0 = many tissues ciliate, 1 = no cilia except in photoreceptor and sperm cells, 2 = cilia except in sperm (although some sensillae may be cilia derived) (Paulus, 1979).

66. Elongated dorsal gonads.

67. Development of ventro-lateral appendages.

68. Reduction of coelom—haemocoel and circulating system with dorsal blood vessel with paired ostia and pericardial sinus.

69. Ecdysis.

70. Cuticle of alpha-chitin (as opposed to collagen in annelids) and protein (Cutler, 1980).

71. Resilin protein present.

72. Superficial blastoderm formation (Brusca and Brusca, 1990; Weygoldt, 1986; Boudreaux, 1979).

73. Rhabdomeric retinular structure in eye facets (Paulus, 1979).

74. Nephridia in at most, first four cephâlic and first two post-cephalic segments (Weygoldt, 1986).

75. Hard exoskeleton.

76. Articulating jointed appendage with arthroidal membrane.

77. Fully segmented sclerites.

78. Cephalic (at least anterior) ecdysis glands.

79. Cephalon with one pair of preoral and four pairs postoral appendages (Brusca and Brusca, 1990; Weygoldt 1986; and Paulus 1979).

80. All muscles striated.

81. Suppression of all circular body wall muscle.

82. Similar inter-segmented tendon system (Boudreaux, 1979).

83. Nephridia with sacculi (Weygoldt, 1986).

84. Specialized ommatidial structures—two corneagene cells, four semper cells and a cone with four parts, retinula with eight cells (versus variable and higher number of subunits) (Paulus, 1979).

85. Tripartite brain.

86. Mandibles (main feeding appendage) with strong coxal endites on third post-acronal head segment (Weygoldt, 1986).

87. Two pairs of maxillae on segments four and five (Brusca and Brusca, 1990).

88. Loss of mandibular palpus.

89. Malpighian tubules 0 = none, 1 = endodermally derived, 2 = ectodermally derived (Weygoldt, 1986).

90. Posterior gonopore (terminal in Chilopoda and only the Ellipura among the hexapods) (Pocock, 1893).

91. Exoskeleton with hard and strong dorsal side and soft ventral.

92. Trilobation.

93. Widened and broadened front end (Eldredge, 1974; Weygoldt, 1986; Weygoldt and Paulus, 1979).

94. Embryos with four gangliar post-oral segments.

95. Separation of cephalic and locomotor functions onto different tagma.

96. Coxal endites on appendage of second somite (primitively on all) (Boudreaux, 1979).

97. Anterior (0) vs. posterior (1) directed mouth (posterior in TCC) (Cisne, 1974).

98. Lamellar spines on appendages (Bergström, 1979).

99. Pretarsal segment of leg (dactylopodite) has only a single muscle (Snodgrass, 1952).

100. Single anatomically preoral limb bearing segment in larva (if not adult).

? = unknown, unclear or inapplicable:

50, 52 and 89 non-additive; others additive.

Appendix 3

Table A2
Morphological data matrix

	10	20	30	40	50	60	70	80	90	100
Mollusca	1111110000	0000000000	0000?00000	0?00?000?0	0000000000	0000000000	0000000000	0000000000	000000?00	0000?????0
Polychaeta	0000001111	1000000000	0000?00000	0?00?000?0	0000000001	0000100111	1111000000	0000000000	000000?00	0000?????0
Clitellata	0000001111	1111100000	0000?00000	0?00?000?0	0000000000	0000100111	1111000000	0000000000	000000?00	0000?????0
Onychophora	0000000000	0000011111	1110000000	0?00?000?0	0000000001	10?1100111	1111111111	1110000000	000000?00	0000??00?0
Trilobita	?0000?000?	00?000?000	0?00000000	0?00??00?0	0000000?2	20?0?111??	?1????1?1?	???111???	11???00??0	1111111?1
Pycnogonida	0000000000	0000000000	0000?01111	1100000000	0000000000	0100000111	1111211111	1111111111	1110000?00	0?01001001
Limulus	0000000000	0000000000	0000?01110	0011100000	0000000000	2200011111	1111211111	1111111111	1110000?00	1111001101
Arachnida	0000000000	0000000000	0000?01110	0011211110	0000000000	2200001111	1111211111	1111111111	1110000?10	0101000001
Crustacea	0000000000	0000000000	0001110000	0000000000	0000000002	3100011111	1111211111	1111111111	1111111000	0000111001
Diplopoda	0000000000	0000000000	0000000000	0?000000?1	0000000002	2011100111	1111211111	1111111111	1111111120	0000110011
Chilopoda	0000000000	0000000000	0000000000	0?000000?1	0000000002	2011100111	1111211111	1111111111	1111111121	0000110011
Hexapoda	0000000000	0000000000	0000000000	0000000000	1111111112	2111100111	1111211111	1111111111	1111111121	0000110011

Appendix 4

Table A3
Aligned 18S rDNA sequences

"A" Fragment	
Lepidochiton	CCTGAGAAACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGCAAATTACCCACTCCTGGCACGGGGAGGTAGTGACGAA
Loligo	....C......................
Glycera	....
Haemopis	....AT..A................
Lumbricus	....A...C.A....................
Peripatoides	....A..
Peripatus	....A...................
Anoplodactylus	....A.................C......................
Limulus	....CA.A...................
Centruroides	....C......................
Mastigoproctus	....-.C......................
Peucetia	....CA.A...................
Nephila	............G.........T................A.............A...C......................
Callinectes	......................TC...............A.................C......................
Balanus	....A.....A.T.A.............A.G
Scutigera	....A.................C......................
Spirobolus	....A................TCA.A...A...............
Thermobius	....C......................
Heptagenia	....C......................
Dorocordulia	....C......................
Libellula	....C......................
Mantis	....C......................
Tibicen	....A.................C......................
Papilio	......................T..................................C......................
Drosophila	......................T...................T..............CA..T..................

```
Lepidochiton     AAATAACAATACGGGATCT--C-TTCGAGGCCCCGTAATTGGAATGAGTACACTTTAAATCCTTTAACGAGGATCTATTG
Loligo           .......G.CR.....CTC..T.G...........CG..C...............C......G..........C.A.CC.
Glycera          ................CTC..A...........G..............................................
Haemopis         ................CTC..T..A...........G......................................A....
Lumbricus        ................CTC..T...T.................................................A....
Peripatoides     ................CTC..T..A.......T...............................................
Peripatus        ................CTC..T..AT......T..........................................C....
Anoplodactylus   ................CTC..T..........................................................
Limulus          .......G........CTC..T.A..............................C.........................
Centruroides     ................CTC..T.A.T....................G.....A.C....C...........T...C....
Mastigoproctus   ................CTC..TCA.T......T.....................C.........................
Peucetia         ................CTC..T.C.T...A........................C....................C....
Nephila          ..........G.....CTC..T.A.....A...A..G..C.........G.G..C...........G........C....
Callinectes      .......G..G..A..CTC..A..C.......T..C...C........................................
Balanus          ............A...CTC.....C....T.TTT..................ACCC...................C....
Scutigera        .............A..CTC..T..C.......T...............................................
Spirobolus       ..........G.....CTC..T..A........................C.......................C......
Thermobius       .......G........CTC..A..C........M.....C.......A...................T...A......C.
Heptagenia       .......G........CTC..A..C........M.....C........A............T......A..T.C...A..
Dorocordulia     .......G........CTC..A..C..............C........A...............................
Libellula        .......G........CTC..A..C..............C........A...............................
Mantis           .......G........CTC..A..C..............C...............................T...C....
Tibicen          .......G........CTC..A..C..............C...............................T.C.A....
Papilio          .......G........CTCATA..C..............C.........................G.......A.A....
Drosophila       ............A...CTCATA..C........T..................................A...C.A....
```

continued

Appendix 4—*contd.*

```
Lepidochiton      GAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCAGTTA
Loligo            .......................................................................
Glycera           .......................................................................
Haemopis          .......................................................................
Lumbricus         .....................................................C.................
Peripatoides      ..................................................................G....
Peripatus         ..................................................................G....
Anoplodactylus    ..................................................................G....
Limulus           ..................................................................G....
Centruroides      ................................................................TG....
Mastigoproctus    ..................................................................G....
Peucetia          ..................................................................G....
Nephila           .........A..AA.......C............................................G....
Callinectes       ...................CGG............................................G....
Balanus           ...................CG.............................................G....
Scutigera         ..................................................................G....
Spirobolus        ..................................................................G....
Thermobius        .....................C....................G.......................G....
Heptagenia        ..................................................................G....
Dorocordulia      ..................................................................G....
Libellula         ..................................................................G....
Mantis            ..................................................................G....
Tibicen           ..................................................................G....
Papilio           ................................C.................................G....
Drosophila        ..................................................................G....
```

"B" Fragment

```
Lepidochiton     GGTGAAATTCTTGGATCGCCGTAAGACGAACTACTGCGAAAGCATTTGCCAAGAATATTTTCATTAATCAAGAACGAAAG
Loligo           ...................................C....................G.......CG.......G......
Glycera          ........................................................G.........G.............
Haemopis         ........................................................G.......................
Lumbricus        ........................................................G.......................
Peripatoides     ...............C..T.....................................G.........G.............
Peripatus        ...............C..T.....................................G.........G.............
Anoplodactylus   ...............C..T..C..................................G.......................
Limulus          ...............C..T..C..................................G.......................
Centruroides     ...............C..T..C..................................G.......................
Mastigoproctus   ...............C..T..C..................................G.......................
Peucetia         ...............C..T..C..................................G.....R.................
Nephila          ...............C..T..G......C..C........................G...C.G..G.C............
Callinectes      ...............C..T..C...............................G..G.......................
Balanus          ................TTA..A..................................G.......................
Scutigera        ...............C..T..C.........C.AA.....................G.......................
Spirobolus       ...............C..T..C.........C.AA.....................G.......................
Thermobius       ..................T..C.....T...CGA......................G.......................
Heptagenia       ..................T..C......G..CGA.........G............G.....GC................
Dorocordulia     ...............C..T..C.........CGA.....................CG.....G..G..............
Libellula        ..................T..C.........CGA.....................CG.....G..G..............
Mantis           ...............C..T..C......G.TCGAA.....................G.....TC.G..............
Tibicen          ...............C..T..C.........CTAA..................T..G..C..G.................
Papilio          ..................T..C.........AT.A.................AGG.G.......C...............
Drosophila       ...............C..T........T....TAA.................AG..G.......................
```

continued

Appendix 4—*contd.*

Lepidochiton	TCAGAGGTTCGAAGACGATCAGATACCGTCGTAGT-TCTGACCATAAACGATGCCAACTAGCGATCCGCCGGAGTTGTTT
Loligo	..G...................................C....G...........GC....................C..
Glycera	............................A.C........................G........................
Haemopis	....T.G...G.........G..........
Lumbricus	....T.G...G...............C.C..
Peripatoides	..............................C...C
Peripatus	..G...................................C.................................C......C
Anoplodactylus	.T............G.............C.C........A..................C.................CC..
Limulus	.T............G.............C.C........A..................C...........T.....CC.C
Centruroides	.T............G.............C.C........A..................C...A.......T.....CC.C
Mastigoproctus	.T............G.............C.C........A..................C...........T.....CC.C
Peucetia	.T............G.............C.C........A................G.C.........G.T.T...CC.C
Nephila	.GT.G..C..A....A...T.......T.T.....C.TGT...C........A..CT.C.......GCG.T.C..A-CGA
Callinectes	.T............G.............C.C........A..............TG..C.............T...A..C
Balanus	.T.....A......................C........A...............G.....G....A.AGA.T...-G..
Scutigera	.T............G.............C.C........A...............G..C.................CC.C
Spirobolus	.T............G.............C.C........A..................C.................CC.C
Thermobius	.T............G.............C.C........A................G.C...A........A....CC.C
Heptagenia	.T............G.............C.C........A................G.C............A....CC.C
Dorocordulia	.T............G.............C.C........A................G.....C........A....CC..
Libellula	.T............G.............C.C........A................G.....C........A....CC..
Mantis	.T............G.............C.C........A................G.C...A........A....CC.C
Tibicen	.T............G.............C.C........A................G.C............A....CC.C
Papilio	.T............G....T........C.C........A...G....-T...T..T..............AC...AC.A
Drosophila	.T............G.............C.C........A................G.....A..TG.GT.T..C.AC..

```
Lepidochiton     CAATGACTCGGCGGGCA-G-CTTCCGGGAAACCAAAGTTTGTGGGTTCCGGGGGAAGTATGGTTGCAAAGCTGAAAC
Loligo           .C.....C.......T......G...........G..CCCT....--------------------------------
Glycera          ..C....C..............................C.T..............................C.....
Haemopis         .T.....C.C...........C..................T...A................................
Lumbricus        .T.....C.C...........C.T..............C.A...A................................
Peripatoides     T......C..............................C.T....................................
Peripatus        TC.....C..............................C.T....................................
Anoplodactylus   .T....................................C.T....................................
Limulus          A..C..................................G.T....................................
Centruroides     A.....................................G.T........A...........................
Mastigoproctus   A.C...................................G.T....................................
Peucetia         A......A..C...........................C.T....................................
Nephila          ACG....G..C......GCT..............G...C.TC..C................................
Callinectes      .C.....C..............................C.T....................................
Balanus          GG.....CTCTTT.....C....A................T....................................
Scutigera        .G.......T............................G.T....................................
Spirobolus       .........T............................A.T....................................
Thermobius       ................G....................C..T....................................
Heptagenia       .G...................................C..T....................................
Dorocordulia     AT...................................C.CT....................................
Libellula        AT...................................C.CT....................................
Mantis           .G..............G....................C..T....................................
Tibicen          .G...................................C..T....................................
Papilio          .T...G...............................A..T...AC...............................
Drosophila       TT...G...TCTCA.TC....................C..T....C...............................
```

continued

Appendix 4—*contd.*

```
"C" Fragment
Lepidochiton     TTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAAACTCACCCGGC
Loligo           ...............................................................................T
Glycera          ................................................................................
Haemopis         ................................................................................
Lumbricus        ................................................................................
Peripatoides     ..................................................................G.............
Peripatus        ..................................................................G.............
Anoplodactylus   ................................................................................
Limulus          ..................................................................G.............
Centruroides     ........................................................................T.......
Mastigoproctus   ...........................................-............................T.......
Peucetia         .................G......T................A..............................T.......
Callinectes      ..................................................................G..C......A...
Balanus          ............................................C.....................G.........A..T
Scutigera        ................................................................................
Spirobolus       ................................................................................
Thermobius       ....................................G................................C......A...
Heptagenia       .....................................................................C......A...
Dorocordulia     ...........................................C.........................C......A...
Libellula        .....................................................................C......A...
Mantis           .....................................................................C......A...
Tibicen          .....................................................................C......A...
Papilio          .....................................................................T......A...
Drosophila       ........................................................................T...A..T
```

```
Lepidochiton     CCGGACACT-GTAAGGATTGACAGATTGAGAGCTCTTTCTTGATTCGGTGGGTGGTGGTGCATGGCCGTTCTTAGTTGGT
Loligo           M.......C.................C..T..........C.......................................
Glycera          ........C.......................................................................
Haemopis         .A......C..CT..................................................................C
Lumbricus        ................................................................................
Peripatoides     ........C......A.............T..................................................
Peripatus        .............................T..................................................
Anoplodactylus   ...........G......................T...........A.................................
Limulus          ...........G...............A..................A.................................
Centruroides     ...........G...............A..................A.................................
Mastigoproctus   ...........G...............A..................A.................................
Peucetia         ........A..G...............A.C................T.................................
Nephila          ......GTGTTATC..T.GA.....C...CG.......T.A..G.TT....A................CAG..G...C..
Callinectes      ..A.....C..G............................C.......................................
Balanus          ..A....TG..G..................................TA................................
Scutigera        ........C..G....................................................................
Spirobolus       ...........G..................................A.................................
Thermobius       ...........G..................................A.................................
Heptagenia       ........C..G..................................A.................................
Dorocordulia     ..A.....C..GT............A.A....................................................
Libellula        ..A.....C..GT............A.A....................................................
Mantis           ..A.....C..G.T..................................................................
Tibicen          ........C..G....................................................................
Papilio          ........C..G...............A.C..................................................
Drosophila       -..A...TA.AGTGT.TAA..........T..........C..A..TA.............................C..
```

continued

Appendix 4—*contd.*

```
Lepidochiton        GGAGCGATTTGTCTGGTTAATTCCGATAACGAACGAGACTC
Loligo              ........CC........G......................
Glycera             .........................................
Haemopis            .........................................
Lumbricus           .........................................
Peripatoides        .........................................
Peripatus           .........................................
Anoplodactylus      .........................................
Limulus             .........................................
Centruroides        .........................................
Mastigoproctus      .........................................
Peucetia            .........................................
Nephila             ..G.T..CC................................
Callinectes         .........................................
Balanus             ....T....................................
Scutigera           .........................................
Spirobolus          .....................................G...
Thermobius          .........................................
Heptagenia          .........................................
Dorocordulia        .........................................
Libellula           .........................................
Mantis              .........................................
Tibicen             .........................................
Papilio             .........................................
Drosophila          ....T....................................
```

Appendix 5

Table A4
Aligned Ubiquitin sequences
(positions 2–42 are primer sequence in all but *Drosophila*)

Lepidochiton	ATGCAGATCTTCGTGAAGACCTTGACCGGAAAGACCATCACCCTTGAAGTCGAGCCTTCAGATACCATCGAGAATGTGAA
Loligo	....A.....T...G....T..C........A.....T..
Glycera	....T..A..G..Y.....A.G...C.................
Haemopis	....T.....G..G.......GT...........A........
Lumbricus	....T.....G..A.....CAGT..C........A........
Peripatoides	....T..A.....T.....A.....C..T.....A.....C..
Peripatus	....T..A..G..G........T..C.....T..A.....C..
Anoplodactylus	....T..G.....T.....A..T..C.....T........C..
Limulus	....T..G.....T.....A..T..C.....T........C..
Centruroides	....T..G.....A.....GAGT..C.....A...........
Mastigoproctus	....TT.G.....T..A...AGT..C..A..T..A........
Peucetia	....TT.A.....T..R..G..T..C.....T...........
Nephila	....AT.A.....T..A..G.....C.....T..A.....T..
Callinectes	....T..A.....G........T..C.....T...........
Balanus	....T........A..A...A.Y..C.....T..R.....A..
Scutigera	....T.G.....T......AGC..C........A..C.....
Spirobolus	....A.....G.....CAGC..C.....W.....C.....
Thermobius	....T.G.....A.....A.....C.....T.....C.....
Heptagenia	....T.G.....G.....C.....C........A..C.....
Dorocordulia	....T.G.....A.....A.....C..G..H........A..
Libellula	....TT.G..G..T.................T........C..
Mantis	....TT.G.....T.................T........T..
Tibicen	....AT.A.....T..A..G.....C.....T........T..
Papilio	....G...........C.....C.....T...........
Drosophila	...........Y........YY....Y..M..............C..G..A.....S..G..C.....T........Y..

continued

Appendix 5—*contd.*

Lepidochiton	AGCTAAGATCCAAGATAARGAAGGAATTCCYCCAGACCAACAGAGTCTGATCTTTGCCGGCAAGCAGCTTGAAGATGGTC
Loligo	............R..C..A..G..C..............G.....GT.A.....C..............G..G.......
Glycera	...............C..A..G..C..C..A........G...C.............T........A.....G.....C.
Haemopis	G........T.....C..G.....C.....T.....T..G.....GT..........T...........G.........A
Lumbricus	G..............C..G..G..C.....A.....T..G...C.............T...........G........C.
Peripatoides	...C...........C..A.....T..C..C.....T..G.....AT..........T..A......T.G........C.
Peripatus	...C.....T.....C..G.....Y.....C.....T..G...C..Y..........T...........G..........
Anoplodactylus	.........T.....C..G..G..C.....A.....T..G.....GT..........T.....................A
Limulus	G........T.....C..G..G..T..C..A..T..T..G...C....C........T..A...................
Centruroides	G..A..A..A........G.....T.....A..T..T..G...C.............T..A........A..........
Mastigoproctus	G..A.....T.....C..A.....T..A..C..T..T........A...........T........A..G........C.
Peucetia	G........T........G.....T.....T..T..T........AT..........T..A...........G.....C.
Nephila	G.....A..T........A.....T..Y..A.....T......C.AT....Y.....T..A......T.G........C.
Callinectes	G..C.....T.....C..G..G..C..C..A..T..T..G...C....C........T..A.................C.
Balanus	G..G.....Y.....C..A.....C..C..A.....T..G...C....C........T..A.....A...........C.
Scutigera	...G.....T..G..C..A..G..C..C..A........G..AC.AT.......C..A...........G........C.
Spirobolus	...C.....T..G..C..A..G..C.....G.....Y..G...M.AT..........T...........G........C.
Thermobius	...C.....T.....C..G.....T..C..A........G...C.A........C..T..T........G........C.
Heptagenia	N.....A..T.....C..A.....T..C..A.....T..G...C.AT....T.....T..A........G........C.
Dorocordulia	...C..A........C..G..R.....C..A..T.....G...C....C........T..A.................C.
Libellula	G..C..A..T........G..G..T..C..A..T..T..G...C....C........T..G.................C.
Mantis	G...........G.....A.....T.....T........G...M.AT..........T...........G........C.
Tibicen	R.....A..T........A.....T.....A.....T......C.AT....T.....T..A......T.G........C.
Papilio	......A.....G..C..G.....C.....C........G.....GT.......Y..............A........A.
Drosophila	G..Y........R..Y..G..G........C.....T..R...C.......Y..C.................G.....A.

```
Lepidochiton     GCACCCTGACTGACTACAATATCCAGAAAGAATCCACTCTCCATTTGGTGCTGAGGCTAAGAGGTGGT
Loligo           ....T..TT.A........C.................C..T..C........T..A..G..G..A...
Glycera          .T..T..C.G.........C...........G.....C..T..C......T..C.A..TC.T......
Haemopis         .A.......GC........C....................T..C.....T.....T..G.........
Lumbricus        .T.......GC........C........G..G.....C..T..C......T..C.T..TC.T......
Peripatoides     ....A..C.G.........C.................A............T..C.T..TC.T..A...
Peripatus        .........G.........C...........G.....C...............C.T..TC........
Anoplodactylus   ....G...TG.........C........G..G.....C..T............C.T..GC........
Limulus          ........T..........C........G........C..T..C.........C.T..GC........
Centruroides     ....T...T..........C...........G.....C.....C.....T.....A..GC.T......
Mastigoproctus   .T.....CT....T.....C..T...........A..A..G...C.T..T..C......C.T......
Peucetia         ....T..C.G.........C.............................T...C.T........A...
Nephila          ....TT..T....T.....C.....A..G..G.....-------------------------------
Callinectes      .....T..T..........C..T.....G..............C.....T..C.....TC........
Balanus          .T.......G.........C........G.....T..C.....C.....C..C.....CC........
Scutigera        .T...T..T....Y.....C........G..G..T.....T..C............T.G.....A..C
Spirobolus       ........T..........C........G..G..T.....A..C.....A...C.T..G.....A..C
Thermobius       .A......T.A..T.....Y.....A..G.....T..A..T..C.........C.T..CC....C..C
Heptagenia       ....AT..T....T.....N........G..G..T..C..T..C...........A..TC.T..C..A
Dorocordulia     ........T.............T.....G.....T..A..............T.....CC.T.....A
Libellula        .....T..T..........C..T.....G.....A.....T..C.....C..C...T.GC........
Mantis           ....A...T.W..T.....C........G..G..A..C..T..C.....A..C.....TC.T..A..W
Tibicen          ....TT..T....T.....C.....A..G..G..A..T..T...C.T..T.....T..----------
Papilio          ....A..CT..........C........G..G..T..C..Y..C........CC.T..GC....A...
Drosophila       ........T.C..T.....C........G..R..S..C..T..C.....S..YC.T..CC.T..W..W
```

Appendix 6. Summary of Character Support (Optimized by CLADOS default with missing data "squeezed out")

Mollusca/Articulata—since this is the basal most split in the cladogram the character polarities cannot be established. There are, however, six morphological features (1–6) which are usually thought of as derived for the molluscs sampled here. None were homoplastic. There were 10 morphological characters usually considered as derived for the articulates (Annelida + Onychophora + Arthropoda; 50, 55, 58–64), with one homoplasy (55). There were 11 unpolarizable molecular differences between these two groups. One of which (a transversion) was from the 18S ribosomal data and the rest (four transitions and six transversions) from ubiquitin. Three ubiquitin transitions and four ubiquitin transversions were homoplastic.

Annelida—five derived morphological features (7–11), none of which were homoplastic, one ribosomal transition (homoplastic), three ubiquitin transitions and two ubiquitin transversions (all homoplastic).

Clitellata—five morphological synapomorphies (12–15, 50; one—50—homoplastic), five ribosomal transitions (three homoplastic), four ribosomal transversions (all homoplastic), three ubiquitin transitions and one transversion (all homoplastic).

Onychophora + Arthropoda—10 morphological synapomorphies (51, 65–73; none homoplastic), four ribosomal transitions (none homoplastic), two transitions (one homoplastic) and one transversion from ubiquitin (homoplastic).

Onychophora—nine morphological synapomorphies (16–23, 54; one homoplastic—54), four ribosomal transitions (all homoplastic), two ribosomal transversions (both homoplastic), three ubiquitin transitions (all homoplastic), two ubiquitin transversions (both homoplastic).

Arthropoda (=Euarthropoda *sensu* Weygoldt and Paulus, 1979)—16 morphological synapomorphies (50, 51, 55, 65, 74–83, 97, 100; one homoplastic—55), eight ribosomal transitions (two homoplastic) and two transversions (neither homoplastic), five ubiquitin transitions (four homoplastic), two ubiquitin transversions (both homoplastic).

Arachnata (Trilobita + Chelicerata)—three morphological synapomorphies (57, 92, 94; one homoplastic—57).

Chelicerata (Pycnogonida + Euchelicerata)—seven morphological synapomorphies (27–29, 50, 52, 95–96; two homoplastic—50, 52), three ribosomal transitions (all homoplastic), three ubiquitin transitions and three transversions (all homoplastic).

Euchelicerata (*Limulus* + Arachnida)—four morphological synapomorphies (33–35, 52), two ribosomal transitions (both homoplastic), four ribosomal transversions (two homoplastic), two ubiquitin transitions and two ubiquitin transversions (all homoplastic).

Arachnida—seven morphological synapomorphies (35–39, 89, 97; one homoplastic—97), three ribosomal transitions (all homoplastic), four ubiquitin transitions (all homoplastic), six transversions (five homoplastic).

Mandibulata (Crustacea + Tracheata)—five morphological synapomorphies (52, 84–87; one homoplastic—52), two ribosomal transitions (none homoplastic) and two transversions (one homoplastic), one ubiquitin transition and one transversion (both homoplastic).

Crustacea—six morphological synapomorphies (24–26, 51, 56–57; two homoplastic—56 and 57), five ribosomal transitions (all homoplastic), three ribosomal transversions (one homoplastic), two ubiquitin transitions and six ubiquitin transversions (all homoplastic).

Tracheata (Myriapoda + Hexapoda)—seven morphological synapomorphies (53–55, 88–89, 97, 99; three homoplastic—54, 55, 99), three ribosomal transitions (all homoplastic, one ribosomal transversion (not homoplastic), five ubiquitin transitions (four homoplastic) and three transversions (all homoplastic).

Myriapoda—two morphological synapomorphies (40 and 52; one homoplastic—52), one ribosomal transition (homoplastic), five ribosomal transversions (four homoplastic), six ubiquitin transitions (all homoplastic) and six transversions (five homoplastic).

Hexapoda—10 morphological synapomorphies (41–49, 90; one homoplastic—90), six ribosomal transitions (four homoplastic) and one transversion (homoplastic), four ubiquitin transitions (all homoplastic).

CLADOS missing data "squeeze" does not allow missing data to be optimized to support nodes. This, in combination with the delayed transformation of characters, biases optimizations against single origins.