CHAPTER 1

Molecular Systematics and Arthropods

Ward Wheeler

Abstract

Nucleic acids have offered a wealth of novel information to systematic analyses of arthropod taxa. Recent molecular accounts of the systematics of extant arthropods have been marred, however, by three defects. The first of these is the lack of attention to potentially important groups such as pycnogonids and tardigrades, second is the nonuse of existing information, and third is the nongenerality of assumptions. Molecular and morphological data sets created for arthropod phylogeny have been treated *de novo* with little if any attempt to integrate existing work in these "novel" hypotheses, and these results are dependent on specific, unexamined assumptions. Only through the integration of research efforts, intelligent sampling, and the probing of our assumptions can coherent hypotheses be erected and tested.

The past decade has presented us with nearly annual molecular reanalyses of Arthropoda. Although there has been some consensus among these investigations, results have been highly dependent on the specifics of the study (Field et al. 1988; Abele et al. 1989; Lake 1990; Turbeville at al. 1991; Ballard et al. 1992; Wheeler et al. 1993; Boore et al. 1995; Friedrich and Tautz 1995; Garey et al. 1996; Giribet et al. 1996).

The initial molecular forays were limited to nuclear small ribosomal subunit (18S rRNA) sequences — usually RNA based. Field et al. (1988) were primarily interested in metazoan relationships; however, within their sample were several representatives of the major extant arthropod lineages. The scheme of relationships they proposed has a monophyletic Arthropoda with the basal dichotomy between myriapods (*Spirobolus*) and (chelicerates + (crustaceans + hexapods)) represented by *Limu-*

lus, *Artemia*, and *Drosophila* (fig. 1.1a). The authors note that this is at variance with traditional ideas of arthropod evolution and ascribe this difference to "fast-clock" taxa. In adding a pentastomid sequence to the Field et al. data, Abele et al. (1989) supported a monophyletic arthropod clade with Mandibulata (Crustacea + Myriapoda + Hexapoda) appearing for the first time (fig. 1.1b). Lake (1990) reanalyzed these data deriving a novel result — that the arthropods were neither monophyletic nor polyphyletic, but paraphyletic with respect to roundworms and mollusks (fig. 1.1c). This truly original idea had no previous complement in morphological or developmental work. Interestingly, when Lake's tree is rerooted such that arthropods are monophyletic, the internal arrangements are highly consistent with previous

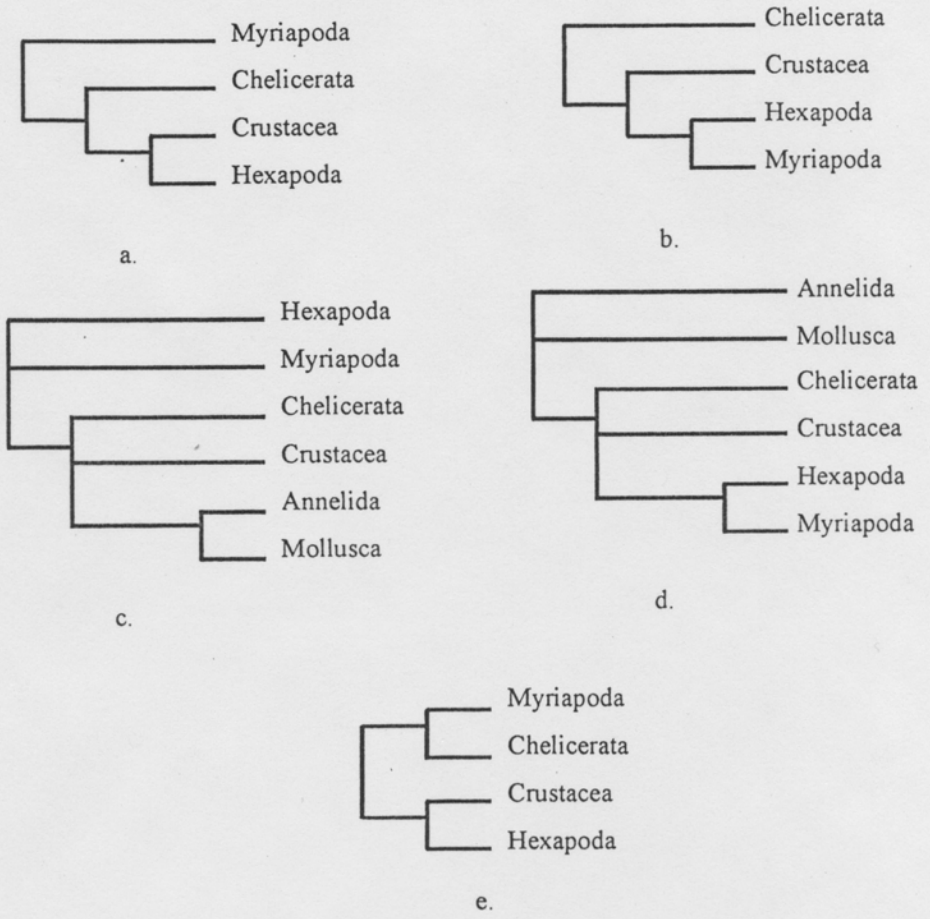

FIGURE 1.1
Small subunit ribosomal RNA-based arthropod phylogenies: *(a)* after Field et al. (1988); *(b)* after Abele et al. (1989); *(c)* after Lake (1990); *(d)* Lake (1990) rerooted with arthropods monophyletic; and *(e)* after Turbeville et al. (1991).

non-molecular analyses (Snodgrass 1938; Weygoldt 1986) (fig. 1.1d). Turbeville et al. (1991) added two new rRNA sequences (and a single DNA-based sequence) and jostled the remainder, straining to remove "fast-clock" taxa. They removed *Drosophila* and *Artemia* from analysis, substituting *Tenebrio* and *Procambarus*. They also supported a crustacean/hexapod clade with the myriapods now sister to the chelicerates (fig 1.1e).

In a departure from nuclear rRNA, Ballard et al. (1992) sequenced the 12S rDNA. This molecule is the mitochondrial version of the nuclear 18S sequenced so frequently. Ballard et al. displayed a tree for forty arthropod (and related) taxa, with good resolution for all the major lineages. The tree presented by the authors does not support a monophyletic Arthropoda, placing Myriapoda outside (Onychophora + (Chelicerata + (Crustacea + Hexapoda))) (fig 1.2a). Unfortunately, the authors mentioned but did not present the most parsimonious cladograms for the taxa (seven steps shorter out of 1418), and the data are considerably less informative (Carpenter et al. in press) (fig. 1.2b). Although their presentation lacked transparency, Ballard et al. (1992) laudably did sequence the crucial Onychophora.

Polyubiquitin was added to the mix by Wheeler et al. (1993). In addition to extending molecular analysis to a new locus, the authors attempted to explicitly integrate the character-based morphological data generated by other investigators. Although the ubiquitin sequences on their own seemed to offer little, they did aid in the resolution of arthropod groups when combined with the 18S rDNA data.

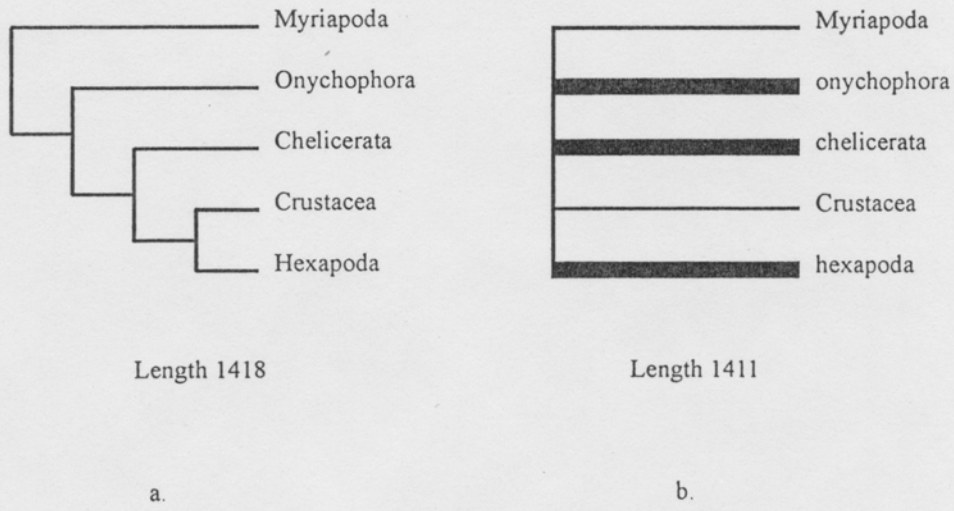

FIGURE 1.2

Mitochondrial small subunit-based arthropod phylogenies: (*a*) presented by Ballard et al. (1992) at length 1418 steps; (*b*) strict consensus cladogram (from 54 trees) at length 1411 steps reported but not presented. The taxa in lower case (and with thicker lines) are not monophyletic in the most parsimonious reconstructions.

Interestingly, the molecular data alone linked Crustacea with Hexapoda within Mandibulata. The main conclusions of this study came from the total evidence (Kluge 1989) analysis with morphological and molecular data yielding a phylogeny identical to that proposed by Snodgrass (1938) (fig. 1.3). Although presented (graphically) by the authors, this point was elegantly restated by Kraus and Kraus (1994).

The entirety of molecular data to this point has been sequence data. Boore et al. (1995) presented gene order data. Through studying the variation in the order of tRNA and other genes in the mitochondrial genome of arthropods and their relatives (including Onychophora), Boore et al. presented seven new characters that support a monophyletic Arthropoda and Mandibulata.

Friedrich and Tautz (1995) presented a somewhat retrograde analysis of arthropod taxa. The authors added a new genetic system to the analysis, the large ribosomal (28S) subunit. Their analysis included data from both the 18S and 28S loci, but did not include morphological data or analysis of crucial taxa such as the Onychophora. Additionally, the criteria for the exclusion of data and taxa seem ad hoc (removal of "fast-clock" taxa and "hard-to-align" areas, etc.). Although Manton's (e.g., 1964, 1973) Uniramia hypothesis cannot be directly tested with these data (because onychophorans were neglected), the authors proposed groupings of Hexapoda + Crustacea and Chelicerata + Myriapoda. This topology is identical to that shown in figure 1.1e.

The glaring hole in all of these analyses is the absence of the Tardigrada. This taxon is potentially the sister taxon of arthropods to the exclusion of Onychophora (reviewed in Brusca and Brusca 1990) but had never been sequenced. This was

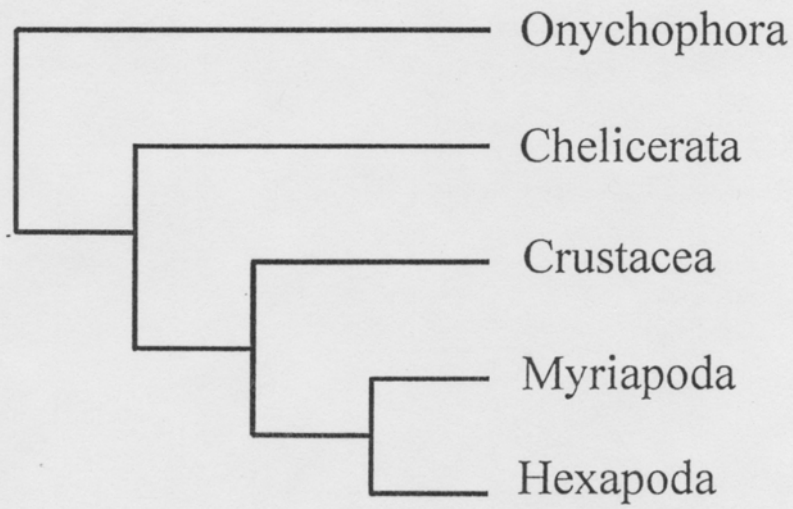

FIGURE 1.3
Total evidence phylogeny of Arthropoda supported by Wheeler et al. (1993). This is largely equivalent to the phylogenetic tree of Snodgrass (1938).

due, no doubt, to the technical difficulties in gathering pure samples of these near-microscopic creatures. Giribet et al. (1996) determined the 18S rDNA sequence for *Macrobiotus hufelandi*, one of the larger tardigrades. Their analysis confirmed the placement of the Tardigrada as sister to the arthropods, but could not place them relative to the Onychophora (no data). In line with other molecular work, the myriapod representative was sister to the arachnids, with crustaceans and hexapods interdigitated (fig. 1.4).

Garey et al. (1996) also analyzed tardigrade sequence (the same genus as Giribet et al. 1996) in their metazoan level sample. Their analysis did not contain any myriapods, pycnogonids, or Onychophora. A distance analysis (bootstrapped Neighbor-Joining) "supports" a tardigrade-arthropod grouping, and the parsimony analysis they performed shows a similar "bootstrap" support of 50%. It is difficult to interpret the analysis of Garey et al. since the authors do not present any of the most parsimonious cladograms or their consensus. Although laudable for the inclusion of tardigrade sequence, the impact of Garey et al. (1996) is unclear given their depauperate sampling and analytical fog.

Common Themes

Although the molecular analyses to date have been polymorphous in the extreme, several common themes surface. The first is that with few exceptions investigators have generated data independently. Although integrating these studies, by adding data to an ever-growing pool of information, seems an obvious step, this was rarely

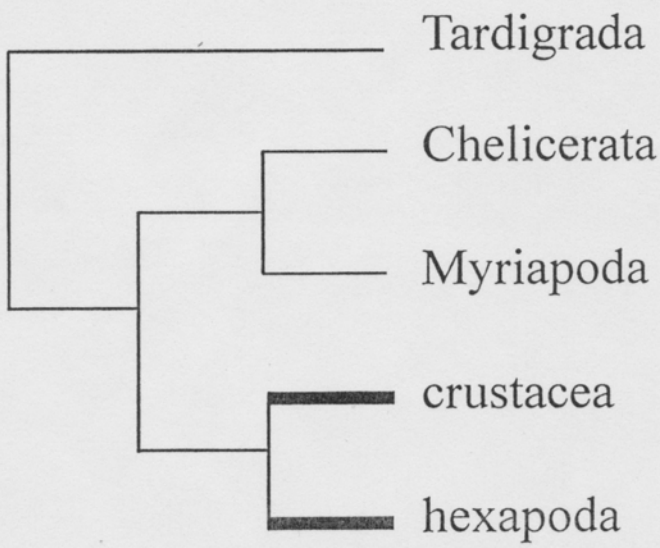

FIGURE 1.4
Small subunit ribosomal DNA-based arthropod phylogeny of Giribet et al. (1996). The taxa in lower case (and with thicker lines) are not monophyletic.

done. Phylogenetically, there are three groupings that are common to almost all the molecular work. They are: (1) monophyletic Arthropoda; (2) chelicerates + myriapods; and (3) crustaceans + hexapods. The second and third results are at great variance with morphological studies, which usually support Crustacea + (Hexapoda + Myriapoda).

Yet Another Reanalysis

To address the problems mentioned above and to summarize current information on the extant arthropods, an analysis was performed including 18S and 28S rDNA, ubiquitin, and morphological information. This analysis examined thirty taxa (table 1.1), including both Onychophora and Tardigrada. Where sequences were not available (e.g., tardigrade 28S and ubiquitin), they were treated as missing data. The inclusion of all data sources was attempted with the caveat that 50% of the taxa must have been analyzed for a given data set to be included. For this reason the Ballard et al. (1992) data were not included — there was almost no overlap between that and other analyses. An exception to this were the Boore et al. (1995) data. These data were coded somewhat generously (see table 1.2) to surpass the 50% criterion. The nonsequence data (table 1.2; appendix 1.1) consisted of 117 discrete characters. They are based on the 100 characters Wheeler et al. (1993) garnered from the literature, to which were added the 7 gene arrangement characters of Boore et al. (1995), 1 feature of the Crustacea from Boxshall (1996), and 9 characters relating to myriapod-hexapod relationships from Kraus and Kraus (1994) and not included in the previous codings.

The sequence data are derived from several previous studies augmented by several generated for this analysis (table 1.1). There were approximately 1100 bases of the 18S rDNA (defined by the 5' and 3' end points of Wheeler et al. 1993), 350 bases of 28S rDNA, and 228 bases of ubiquitin. The entire contiguous 18S and 28S sequences were employed (as defined by primers, not the complete loci: Whiting et al. [1997]) with the exception of three regions. These areas (the E21–2 to E21–4 region of the 18S and two in the "D3" region of 28S: Hendriks et al. 1988; Gutell and Fox 1988) contained large inserts that were unique to some terminal taxa. They correspond to unique secondary-structure "loops" that had no homologue in other taxa. They are, in effect, autapomorphies for terminal taxa.

These data were analyzed separately and in combination. In the molecular and the combined (total evidence) analyses, two analytical parameters were varied: insertion-deletion cost and transition-transversion ratio (as in Wheeler 1995). When nonsequence data were included, they were weighted equal to the insertion-deletion cost, and when the transition-transversion ratio was set other than unity, the insertion-deletion cost was set according to the cost of transversions. In total, nine combinations of analysis parameters were employed for each of the molecules separately, the molecular data as an ensemble, and in combination with the nonsequence

TABLE 1.1
Taxa Used in the Study

Higher Group Taxon		18S rDNA	28S rDNA	Ubiquitin
Mollusca				
Cephalopoda	*Loligo pealei*	Wheeler	ND	Wheeler
Polyplacophora	*Lepidochiton cavernae*	Wheeler	ND	Wheeler
Annelida				
Polycheata	*Glycera* sp.	Wheeler	ND	Wheeler
Oligocheata	*Lumbricus terrestris*	Wheeler	ND	Wheeler
	Tubifex sp.	Friedrich	Friedrich	ND
Hirudinea	*Haemopis marmorata*	Wheeler	ND	Wheeler
Onychophora				
Peripatoidae	*Peripatus trinitatis*	Wheeler	ND	Wheeler
Peripatopsidae	*Peripatoides novozealandia*	Wheeler	ND	Wheeler
Tardigrada				
	Macrobiotus hufelandi	Giribet	ND	ND
Chelicerata				
Pycnogonida	*Anoplodactylus portus*	Wheeler	Here	Wheeler
Xiphosura	*Limulus polyphemus*	Wheeler	Here	Wheeler
Scorpiones	*Centruroides hentzii*	Wheeler	Here	Wheeler
Uropygi	*Mastigoproctus giganteus*	Wheeler	Here	Wheeler
Araneae	*Peucetia viridans*	Wheeler	Here	Wheeler
Crustacea				
Cirrepedia	*Balanus* sp.	Wheeler	Here	Wheeler
Malacostraca	*Callinectes* sp.	Wheeler	Here	Wheeler
Phyllopoda	*Artemia salina*	Nelles	Friedrich	ND
Myriapoda				
Chilopoda	*Scutigera coleoptrata*	Wheeler	Here	Wheeler
	Lithobius forficatus	Friedrich	Friedrich	ND
Diplopoda	*Spirobolus* sp.	Wheeler	Here	Wheeler
	Polyxenus lagurus	Friedrich	Friedrich	ND
	Megaphyllum sp.	Friedrich	Friedrich	ND
Hexapoda				
Collembola	*Pseudachorutes* sp.	Friedrich	Friedrich	ND

TABLE 1.1 (*continued*)

Higher Group Taxon		18S rDNA	28S rDNA	Ubiquitin
Archeognatha	*Petrobius brevistylis*	Friedrich	Friedrich	ND
Odonata	*Libellula pulchella*	Wheeler	Whiting	Wheeler
Dictyoptera	*Mantis religiosa*	Wheeler	Whiting	Wheeler
Heteroptera	*Saldula pallipes*	Wheeler	Whiting	ND
Lepidoptera	*Papilio* sp.	Wheeler	Whiting	Wheeler
Diptera	*Drosophila melanogaster*	Tautz	Tautz	Lee
Coleoptera	*Tenebrio molitor*	Hendriks	Whiting	ND

Giribet = Giribet et al. (1996); Hendriks = Hendriks et al. (1988); Friedrich = Friedrich and Tautz (1995); Nelles = Nelles et al. (1984); Lee = Lee et al. (1988); Tautz = Tautz et al. (1988); Wheeler = Wheeler et al. (1993); Whiting = Whiting et al. (1997); ND = no data; Here = this study.

data (morphology and other character data), for a total of forty-five phylogenetic analyses. These analyses were performed to assess the effect of variation in unmeasurable factors (such as insertion-deletion cost) on systematic conclusions. Phylogenetic analysis was performed using the direct character optimization method of Wheeler (1996), which directly diagnoses phylogenetic cost without an intervening multiple-alignment step. In each case, eleven addition sequences were employed (one based on proximity to the outgroup taxa and ten random) with TBR-type branch swapping using the program MALIGN, version 2.7 (Wheeler and Gladstein 1992, 1994). The values reported are all parsimony tree lengths weighted by the relative costs of gaps, transversions, transitions, and nonsequence character transformation costs.

In each case, the character incongruence values (Mickevich and Farris 1981) were calculated. This value was used to choose among these various analyses in order to present those schemes of relationship that best represent all the data.

Results

The morphological data yielded two cladograms at 141 steps with a consistency index of 0.88 and a retention index of 0.97. The first cladogram shows a monophyletic Myriapoda and the second a paraphyletic Myriapoda with a sister group relationship between the diplopods and hexapods to the exclusion of the centipedes (as in Kraus and Kraus 1994). The strict consensus of the two leads to a trichotomy among Chilopoda, Diplopoda, and Hexapoda (fig. 1.5).

For the combined analyses, the phylogenetic tree lengths (weighted parsimony costs) are shown in table 1.3. For each of the three insertion-deletion costs and three transition-transversion ratios, five phylogenetic reconstructions were performed. These consisted of total evidence including nonsequence data (morphological and other) and sequence data, molecular sequence information (18S, 28S, and ubiquitin), and 18S rDNA, 28S rDNA, and ubiquitin independently. Character-based incongruity was calculated for the entire data, among the sequence data sets, and between

TABLE 1.2

Nonsequence character matrix

Lepidochiton	1111110000	0000000000	0000?00000	0?00?000?0	0000000000
	0000000000	0000000000	0000000000	0000000?00	0000?????0
	0000110???	???????			
Loligo	1111110000	0000000000	0000?00000	0?00?000?0	0000000000
	0000000000	0000000000	0000000000	0000000?00	0000?????0
	0000110???	???????			
Glycera	0000001111	1000000000	0000?00000	0?00?000?0	0000000001
	0000100111	1111000000	0000000000	0000000?00	0000?????0
	0000110???	???????			
Haemopis	0000001111	1111100000	0000?00000	0?00?000?0	0000000000
	0000100111	1111000000	0000000000	0000000?00	0000?????0
	0000110???	???????			
Lumbricus	0000001111	1111100000	0000?00000	0?00?000?0	0000000000
	0000100111	1111000000	0000000000	0000000?00	0000?????0
	0000110???	???????			
Tubifex	0000001111	1111100000	0000?00000	0?00?000?0	0000000000
	0000100111	1111000000	0000000000	0000000?00	0000?????0
	0000110???	???????			
Peripatoides	0000000000	0000011111	1110000000	0?00?000?0	0000000001
	10?1100111	1111111111	1110000000	0000000?00	0000??00?0
	000000001?	???????			
Peripatus	0000000000	0000011111	1110000000	0?00?000?0	0000000001
	10?1100111	1111111111	1110000000	0000000?00	0000??00?0
	000000001?	???????			
Macrobiotus	0000000000	0000000000	0000?00000	0?00?000?0	0000000000
	1000?00??1	111?11111?	???100?00?	1000000?01	000000000?
	???????0??	???????			
Anoplodactylus	0000000000	0000000000	0000001111	1100000000	0000000000
	0100000111	1111211111	1111111111	1110000?00	0?01001001
	???????000	0000000			
Limulus	0000000000	0000000000	0000001110	0011100000	0000000000
	2200011111	1111211111	1111111111	1110000?00	1111001101
	1111000000	0000000			
Centruroides	0000000000	0000000000	0000001110	0011211110	0000000000
	2200001111	1111211111	1111111111	1110000?10	0101000001
	111100?000	0000000			
Mastigoproctus	00000000000	0000000000	0000011100	0112111100	0000000002
	200001111	1111211111	1111111111	1110000?10	0101000001
	111100?000	0000000			
Peucetia	0000000000	0000000000	0000001110	0011211110	0000000000
	2200001111	1111211111	1111111111	1110000?10	0101000001
	111100?000	0000000			
Artemia	0000000000	0000000000	0001110000	0000000000	0000000002
	3100011111	1111211111	1111111111	1111111000	0000111001
	1111001100	0000000			

TABLE 1.2 (*continued*)

Callinectes	0000000000	0000000000	0001110000	0000000000	0000000002
	3100011111	1111211111	1111111111	1111111000	0000111001
	1111001100	0000000			
Balanus	0000000000	0000000000	0001110000	0000000000	0000000002
	3100011111	1111211111	1111111111	1111111000	0000111001
	111100?100	0000000			
Scutigera	0000000000	0000000000	0000200000	0?000000?1	0000000002
	2011100111	1111211111	1111111111	1111111121	0000110011
	1111001010	0111000			
Lithobius	0000000000	0000000000	0000200000	0?000000?1	0000000002
	2011100111	1111211111	1111111111	1111111121	0000110011
	1111001010	0111000			
Spirobolus	0000000000	0000000000	0000200000	0?000000?1	0000000002
	2011100111	1111211111	1111111111	1111111120	0000110011
	1111001011	1000111			
Polyxenus	0000000000	0000000000	0000200000	0?000000?1	0000000002
	2011100111	1111211111	1111111111	1111111120	0000110011
	1111001011	1000111			
Megaphyllum	0000000000	0000000000	0000200000	0?000000?1	0000000002
	2011100111	1111211111	1111111111	1111111120	0000110011
	1111001011	1000111			
Pseudachorutes	0000000000	0000000000	0000200000	0000000000	1111111112
	2111100111	1111211111	1111111111	1111111121	0000110011
	1111001011	1000000			
Petrobius	0000000000	0000000000	0000200000	0000000000	1111111112
	2111100111	1111211111	1111111111	1111111121	0000110011
	1111001011	1000000			
Saldula	0000000000	0000000000	0000200000	0000000000	1111111112
	2111100111	1111211111	1111111111	1111111121	0000110011
	1111001011	1000000			
Tenebrio	0000000000	0000000000	0000200000	0000000000	1111111112
	2111100111	1111211111	1111111111	1111111121	0000110011
	1111001011	1000000			
Libellula	0000000000	0000000000	0000200000	0000000000	1111111112
	2111100111	1111211111	1111111111	1111111121	0000110011
	1111001011	1000000			
Mantis	0000000000	0000000000	0000200000	0000000000	1111111112
	2111100111	1111211111	1111111111	1111111121	0000110011
	1111001011	1000000			
Papilio	0000000000	0000000000	0000200000	0000000000	1111111112
	2111100111	1111211111	1111111111	1111111121	0000110011
	1111001011	1000000			
Drosophila	0000000000	0000000000	0000200000	0000000000	1111111112
	2111100111	1111211111	1111111111	1111111121	0000110011
	1111001011	1000000			

Characters 25, 50, 52, and 89 are nonadditive, the remaining multistates are ordered.

the nonsequence and sequence data (table 1.3). In no case was more than one most parsimonious result found. Hence, the polytomies in phylogenetic results are due to zero-branch-lengths rather than to incongruence. This is not as unusual as it sounds. Since the tree lengths are so long, there is ample opportunity for topologies to differ (however slightly) in length. Phylogenetic conclusions are fairly robust with most variation in the disposition of the pycnogonid *Anoplodactylus* (figs. 1.6–1.8). These achieved minima at two points (fig. 1.9). The overall and sequence-data incongruence were minimized (12.9% and 13.9%) when the insertion-deletion cost was four times that of a base substitution and transitions and transversions were equally weighted. The discordance between the nonsequence and sequence data (more or less morphological and molecular) was minimized when the insertion-deletion cost was twice that of base substitutions (0.996% — again with equal weighting of transversions and transitions). These results are in accord with those of Wheeler (1995), which is not terribly surprising given the similarity of data and problem.

Discussion and Conclusions

The basic phylogenetic conclusions of these analyses are that Arthropoda, Mandibulata, Tracheata, and Myriapoda are all monophyletic. The data are less sanguine about the placement of the pycnogonids with the other chelicerates, but this appears to be more due to lack of evidence than to conflict among characters. Given

TABLE 1.3
Phylogenetic Tree Statistics

	Insertion-Deletion Cost Ratio		
	2:1	4:1	8:1
Transversion-Transition Cost Ratio			
1:1	5221/4605/2762 374/672/564/ 16.3%/17.3%/0.996%	5910/4669/2992/ 356/672/1128 12.9%/13.9%/1.91%	7145/4531/2991/ 118/672/2256 15.5%/15.1%/5.01%
2:1	4040/3388/1998/ 432/478/564 14.1%/14.2%/2.18%	4832/3506/2174/ 249/478/1128 16.6%/17.3%/4.10%	5963/3465/2127/ 86/478/2256 17.0%/22.3%/4.06%
4:1	6673/5454/3158/ 709/751/1128 13.9%/15.3%/1.36%	8114/5335/3438/ 352/751/2256 16.2%/14.9%/6.45%	11132/6430/3961/ 217/751/4512 15.2%/23.3%/1.71%

In each cell: Total Evidence/Sequence Data/18S rDNA/28S rDNA/Ubiquitin/Nonsequence tree lengths/Mickevich-Farris (1981) Character Incongruence Overall/Among Sequence Data/Nonsequence vs. Sequence

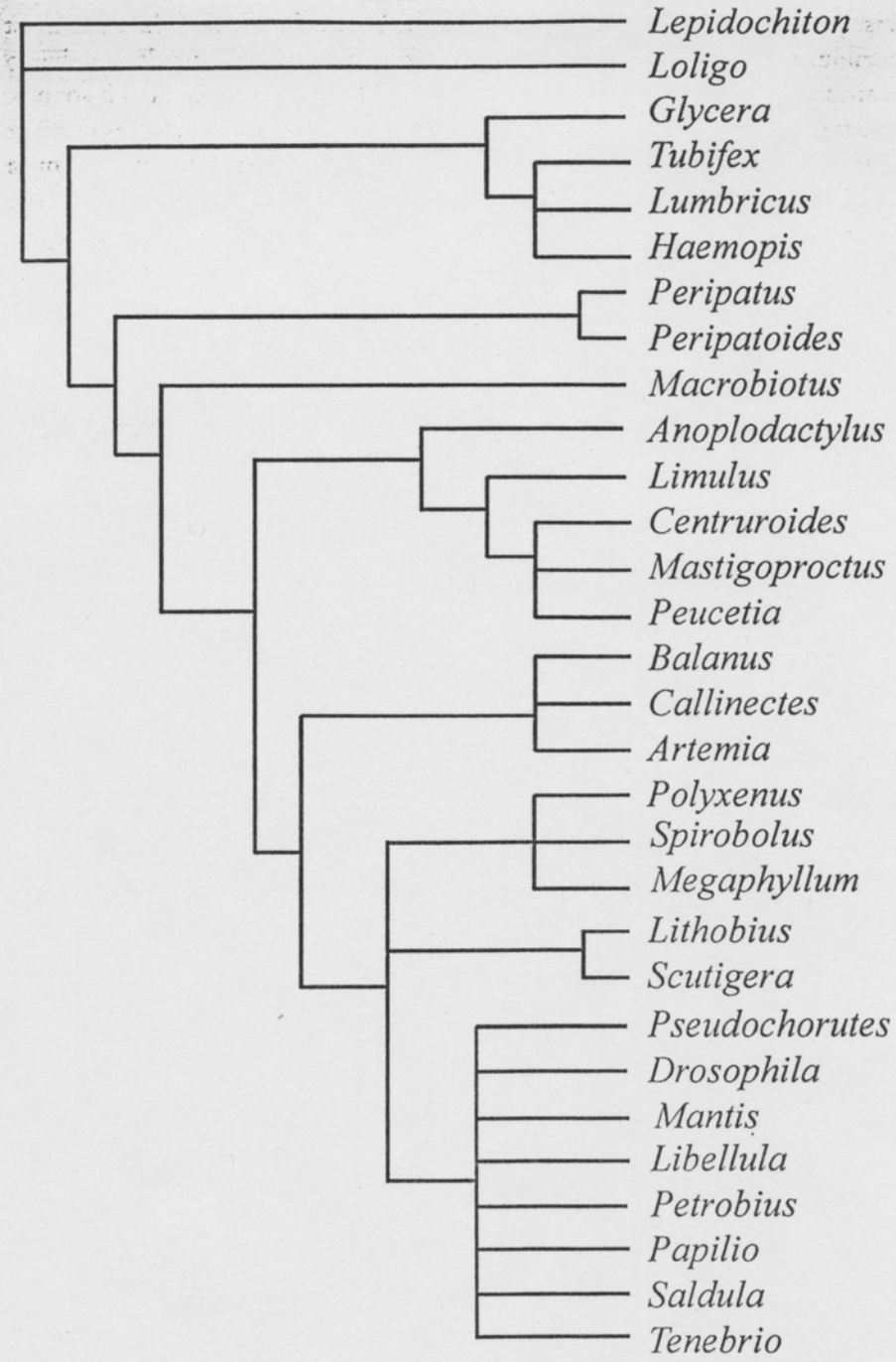

FIGURE 1.5

Nonsequence (morphology and other character data such as gene rearrangements) data-based phylogeny of the arthropod taxa in this study. There were two trees of length 141 with CI of 0.88 and RI of 0.97. This is the strict consensus of the two.

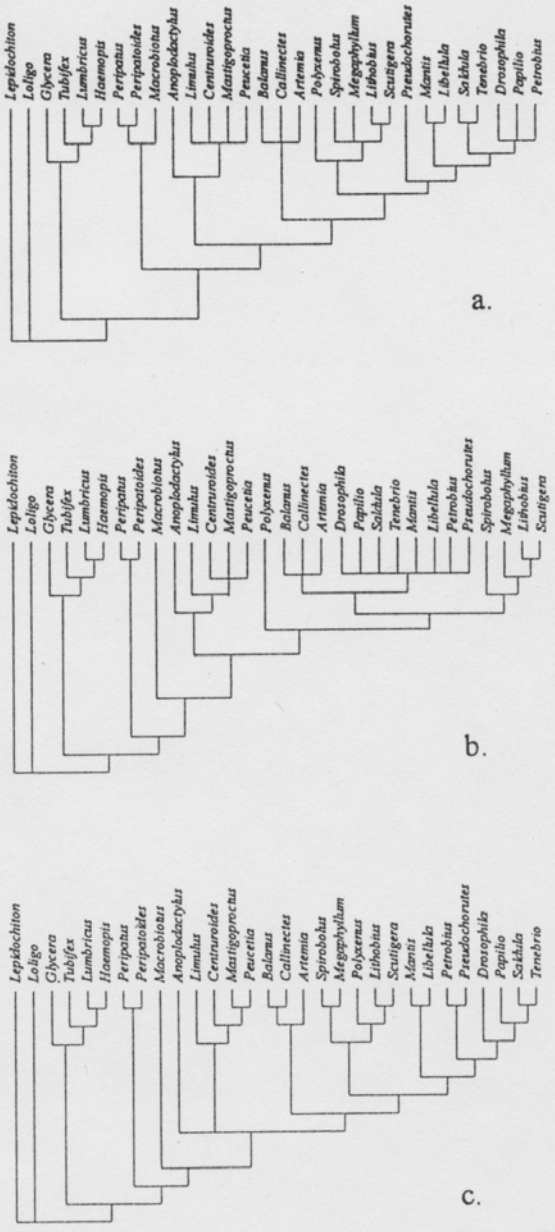

FIGURE 1.6

Total-evidence-derived phylogenies of arthropod relationships. These three cladograms and those in figures 1.7 and 1.8 are derived from nonsequence (morphological and other), small subunit, large subunit, and ubiquitin sequence data. Parsimonious reconstructions of phylogeny were derived via the method of Wheeler (1996) varying analysis parameters. These three cladograms were derived using an insertion-deletion cost (gap ratio) of twice that of a base substitution: *(a)* transversions weighted four times transitions; *(b)* transversions weighted twice transitions; *(c)* equal weighting of transversions and transitions. In each case, the nonsequence characters were assigned a weight equal to that of a sequence gap.

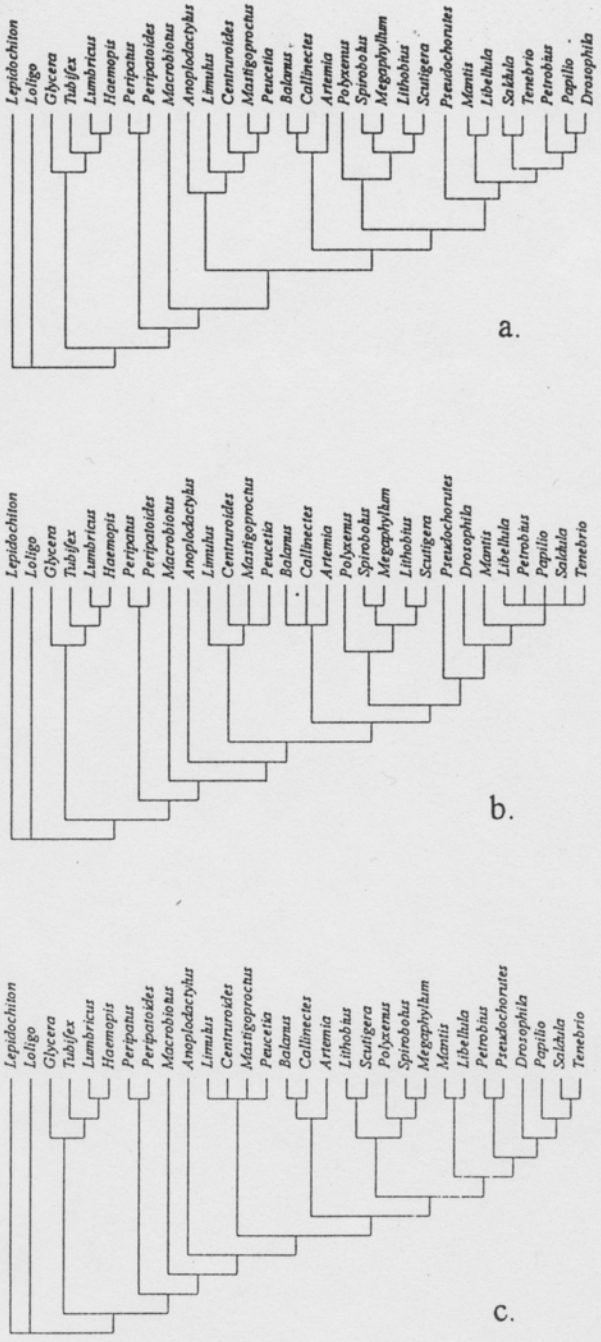

FIGURE 1.7

Total evidence cladograms derived using an insertion-deletion cost of four times the substitution cost (cf. figs. 1.6 and 1.8): *(a)* transversions weighted four times transitions; *(b)* transversions weighted twice transitions; *(c)* equal weighting of transversions and transitions.

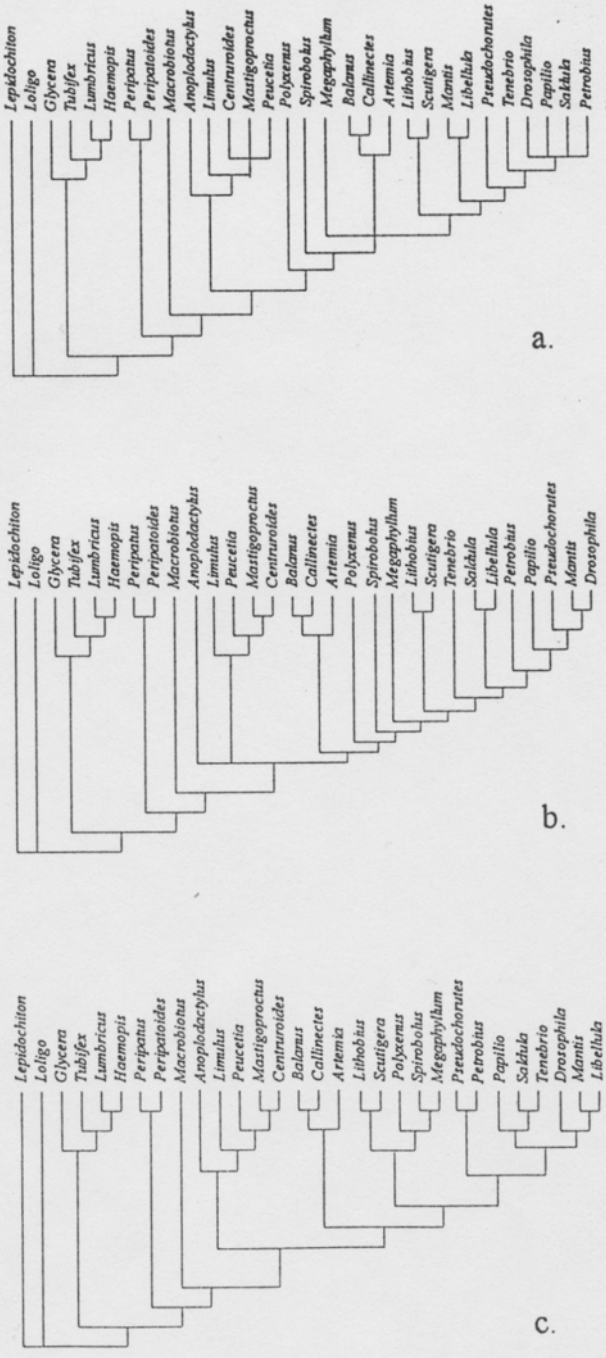

FIGURE 1.8

Total evidence cladograms derived using an insertion-deletion cost of eight times the substitution cost (cf. figs. 1.6 and 1.7): *(a)* transversions weighted four times transitions; *(b)* transversions weighted twice transitions; *(c)* equal weighting of transversions and transitions.

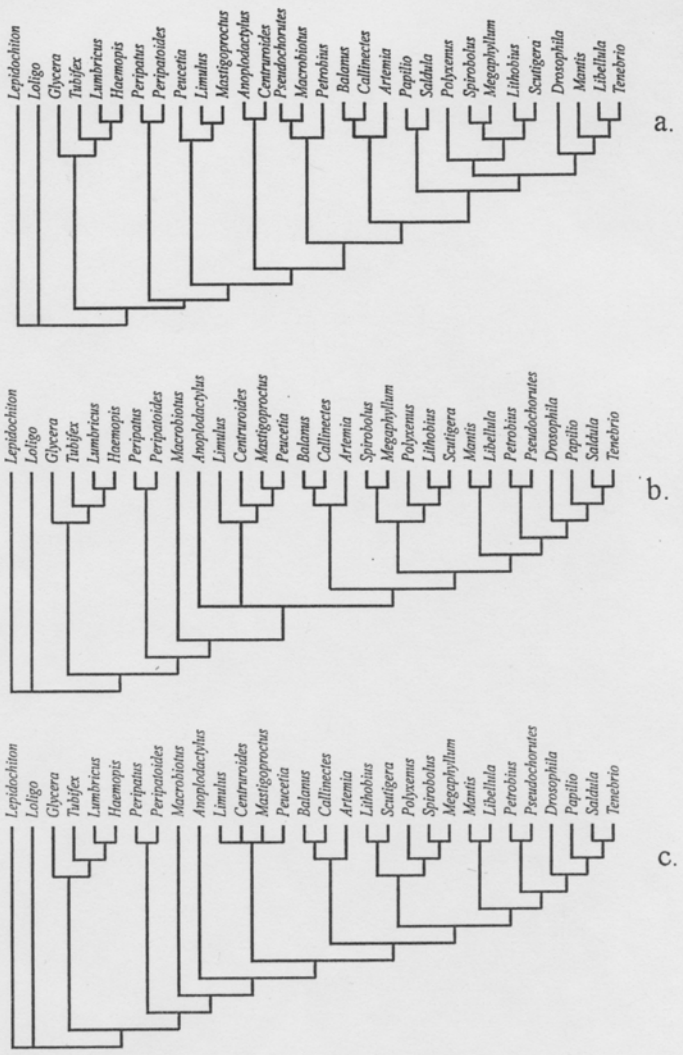

FIGURE 1.9

Arthropod phylogenies with minimal character-based incongruence as assayed by the Mickevich and Farris metric (1981): *(a)* optimal sequence-only cladogram; *(b)* optimal for nonsequence (morphological and other) versus sequence data; *(c)* optimal for all data.

the inherent uncertainties of weighting character information versus molecular sequence data, and transitions versus transversions versus insertion-deletion events, it is comforting to see the relative consistency among the results. In all cases (employing total evidence), the arthropods form a monophyletic group. In eight cases (of nine), Mandibulata is supported, and seven cases support Tracheata. Interestingly, the Crustacea + Hexapoda clade is present only with a transversion-to-transition cost ratio of two to one (the same analysis conditions and result of Friedrich and Tautz [1995]); under any other parameter regime this clade is not supported. This result highlights the importance of examining the robustness of phylogenetic conclusions for variation in these parameter values, which are essentially arbitrary assumptions (Wheeler 1995). Friedrich and Tautz reported a result (verified here) that is based on a particular (but not general) set of analysis conditions. It is dangerous to make strong phylogenetic claims without explicit analysis of sensitivity to parameter values.

Although the weight of nonsequence (e.g., morphological) data was kept the same as the insertion-deletion cost, the contribution of this weighted change information varied from approximately 10% of the total length of the cladogram (at an insertion-deletion cost ratio of 2:1 and transversion to transitions at equality) to over 40% where insertion-deletion events cost was eight times that of transversions and that of transversions four times that of transitions. This variation is due to the ability of the sequences to favor base substitution over insertion-deletion events while the nonsequence changes remain constant in type and number. Over this range, the combined data yielded remarkably homogeneous results. The main areas of inconsistency are in the placement of the pycnogonid *Anoplodactylus* and the status of the Myriapoda.

Roughly half the analyses support a monophyletic Chelicerata (*sensu* Weygoldt and Paulus 1979). Among the others, the pycnogonids are placed either agnostically, in an unresolved polytomy at the base of the euarthropods, or as sister to other euarthropods — a paraphyletic Chelicerata. In two-thirds of the cases examined, the myriapods are monophyletic (and in the least incongruous). Other situations show the myriapods as paraphyletic, but not in the sense of Kraus and Kraus (1994). These two areas may benefit disproportionately from increased sampling. The diversity of pycnogonids is not well circumscribed and major myriapod lineages (Pauropoda and Symphyla) are not sampled in these analyses.

The overall message of this analysis is simple. First, phylogenetically, these data support the Tardigrada as the sister taxon to the Euarthropoda to the exclusion of the Onychophora. Within the arthropods, Euchelicerata, Mandibulata, and Tracheata (= Atelocerata) are all generally supported. The placement of the pycnogonids is suggestive of the morphologically defined Chelicerata, but the evidence here is not strong. Similarly, some evidence supports the monophyly of the myriapods, but this is far from unequivocal.

Second, the combination of data — sequence and other — is certainly achievable and these results should be persuasive as to the importance of total evidence. Total

evidence does not mean that more data will not be uncovered, but that responsible investigators must include all evidence within their grasp. Is it satisfactory to base conclusions on a minority of information? Is it scientific? And what of fossils (although none are included here)? Without the simultaneous analysis of sequence and anatomical information, there will be two worlds of systematics — molecular data from the extant taxa and anatomical from the extinct and extant. Synthesis will be impossible.

The third message concerns the robustness of phylogenetic conclusions. The results discussed above are generally supported. They are reasonably independent of perturbations in the tested assumptions (of course, the examinations performed here are simplistic — but salutary nonetheless). This need not have been true. Combining data and varying their relative importance could have caused chaotic, incoherent results. This may be true in other phylogenetic situations, but only by examining multiple parameter values (insertion-deletion costs and transition-transversion ratios) can this be determined.

Arthropods are too large and diverse a group to be allied based on a single shot in the dark, whether that be due to taxonomic, empirical, or epistemological myopia.

ACKNOWLEDGMENTS

I would like to thank Norman Platnick, Amy Litt, and Michael Whiting for reading drafts 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. *Molecular Biology and Evolution* 6:685–691.

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

Ballard, J. W. O., G. J. Olsen, D. P. Faith, W. A. Odgers, D. M. Rowell, and P. W. Atkinson. 1992. Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods. *Science* 258:1345–1348.

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

Boore, J. L., T. M. Collins, D. Stanton, L. L. Daehler, and W. M. Brown. 1995. Deducing the pattern of arthropod phylogeny from mitochondrial rearrangements. *Nature* 376:163–165.

Boudreaux, H. B. 1979. *Arthropod Phylogeny with Special Reference to Insects*. New York: John Wiley.

Boxshall, G. 1996. Limb segmentation patterns in copepods and other major crustacean groups. In *International Symposium on the Relationships of Major Arthropod Groups: Programme and Abstracts*, p. 5. London.

Brusca, R. C. and G. L. Brusca. 1990. *Invertebrates*. Sunderland, Mass.: Sinauer.

Carpenter, J. M., P. A. Goloboff, and J. S. Farris. In press. Randomization and confusion: A reply to Trueman. *Cladistics*.

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. *American Museum Novitates* 2543:1–41.

Field, K. G., G. J. Olsen, D. J. Lane, S. J. Giovannoni, M. T. Ghiselin, E. C. Raff, N. R. Pace, and R. A. Raff. 1988. Molecular phylogeny of the animal kingdom. *Science* 239:748–753.

Friedrich, M. and D. Tautz. 1995. Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of the myriapods. *Nature* 376:165–167.

Garey, J. R., M. Krotec, D. R. Nelson, and J. Brooks. 1996. Molecular analysis supports a tardigrade-arthropod association. *Invertebrate Biology* 112:79–88.

Giribet, G., S. Carranza, J. Baguñà, M. Riutort, and C. Ribera. 1996. First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Molecular Biology and Evolution* 13:76–84.

Gutell, R. and G. Fox. 1988. A compilation of large subunit RNA sequences presented in a structural format. *Nucleic Acids Research* 16 (supplement): r175–r269.

Hendriks, L., R. De Baere, C. Van Broeckhoven, and R. De Wachter. 1988. Primary and secondary structure of the 18S ribosomal RNA of the insect species *Tenebrio molitor*. *Federation of European Biochemical Societies* 232:115–120.

Hennig, W. 1981. *Insect Phylogeny*. New York: John Wiley.

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

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

King, P. E. 1973. *Pycnogonids*. London: Hutchinson.

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

Kraus, O. and M. Kraus. 1994. Phylogenetic system of Tracheata (Mandibulata): on "Myriapoda" - Insecta interrelationships, phylogenetic age, and primary ecological niches. *Verhandlungen des naturwissenschaftlichen Vereins in Hamburg* (NF) 34:5–31.

Kristensen, N. P. 1975. The phylogeny of hexapod "orders": A critical review of recent accounts. *Zeitschrift für zoologische Systematik und Evolutionsforschung* 13:1–44.

Lake, J. A. 1990. Origin of the Metazoa. *Proceedings of the National Academy of Sciences USA* 87:763–766.

Lee, H. J., A. Simon, and J. T. Lis. 1988. Structure and expression of ubiquitin genes in *Drosophila melanogaster*. *Molecular and Cellular Biology* 8:4727–4735.

Manton, S. M. 1964. Mandibular mechanisms and the evolution of arthropods. *Philosophical Transactions of the Royal Society of London*, series B, 247:1–183.

Manton, S. M. 1973. Arthropod phylogeny: A modern synthesis. *Journal of Zoology* 171:111–130.

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

Mickevich, M. F. and J. S. Farris. 1981. The implications of congruence in *Menidia*. *Systematic Zoology* 30:351–370.

Nelles, L., B.-L. Fang, G. Volckaert, A. Vandenberghe, and R. De Wachter. 1984. Nucleotide sequence of a crustacean 18s ribosomal RNA gene and secondary structure of eukaryotic small subunit RNAs. *Nucleic Acids Research* 14:2345–2364.

Paulus, H. F. 1979. Eye structure and the monophyly of Arthropoda. In A. P. Gupta, ed., *Arthropod Phylogeny*, pp. 299–383. New York: Van Nostrand Reinhold.

Pocock, R. I. 1893. On the classification of the tracheate Arthropoda. *Zoologischer Anzeiger* 16:271–275.

Schram, F. R. 1978. Arthropods: A convergent phenomenon. *Fieldiana, Geology* 39:61–108.

Schram, F. R. 1986. *Crustacea*. New York: Oxford University Press.

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

Snodgrass, R. E. 1938. Evolution of the Annelida, Onychophora, and Arthropoda. *Smithsonian Miscellaneous Collections* 97:1–159.

Snodgrass, R. E. 1952. *A Textbook of Arthropod Anatomy*. Ithaca: Cornell University Press.

Tautz, D., J. M. Hancock, D. A. Webb, C. Tautz, and G. A. Dover. 1988. Complete sequence of the rRNA genes in *Drosophila melanogaster*. *Molecular Biology and Evolution* 5:366–376.

Tiegs, O. W. and S. M. Manton. 1958. The evolution of the Arthropoda. *Biological Reviews of the Cambridge Philosophical Society* 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. *Molecular Biology and Evolution* 8:669–702.

Weygoldt, P. 1986. Arthropod interrelationships: The phylogenetic-systematic approach. *Zeitschrift für zoologische Systematik und Evolutionsforschung* 24:19–35.

Weygoldt, P. and H. F. Paulus. 1979. Untersuchungen zur Morphologie, Taxonomie, und Phylogenie der Chelicerata, Teil 2: Cladogramme und die Entfaltung der Chelicerata. *Zeitschrift für zoologische Systematik und Evolutionsforschung* 17:117–200.

Wheeler, W. C. 1995. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Systematic Biology* 44:321–332.

Wheeler, W. C. 1996. Optimization Alignment: The end of multiple sequence alignment in phylogenetics? *Cladistics 12:1–9*.

Wheeler, W. C., P. Cartwright, and C. Y. Hayashi. 1993. Arthropod phylogeny: A combined approach. *Cladistics* 9:1–39.

Wheeler, W. C. and D. S. Gladstein. 1992. *MALIGN version 2.7: Program and Documentation*. New York.

Wheeler, W. C. and D. S. Gladstein. 1994. MALIGN: A multiple sequence alignment program. *Journal of Heredity* 85:417.

Whiting, M. F., J. C. Carpenter, Q. D. Wheeler, and W. C. Wheeler. 1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology* 46:1–68.

APPENDIX 1.1

Morphological characters. Literature sources refer to all characters listed since previous citation.

1. Reduction of coelom and development of an open hemocoelic 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 (Hyman 1967; Brusca and Brusca 1990)
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 (Boudreaux 1979)
12. Clitellum
13. Hermaphroditism
14. Direct development without intervening larval stages
15. Cerebral ganglion moved into anteriormost trunk segment (Jamieson 1988; Brusca and Brusca 1990)
16. Suppression of external segmentation
17. Oblique muscle layer in body wall
18. Subcutaneous hemal channels
19. Oral papillae
20. Body papillae and scales
21. Slime glands
22. Non-migratory gastrulation
23. Lobopods with pads and claws (Boudreaux 1979; Brusca and Brusca 1990)
24. Two pairs of antennae
25. Third head metamere: 0 = unspecialized appendages; 1 = biramous antenna; 2 = intercalary (appendages absent) (Kraus and Kraus 1994)
26. Nauplius or egg-nauplius stage in ontogeny (Schram 1986)
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 and Paulus 1979; Weygoldt 1986)
30. Opisthosoma reduced
31. Proboscis (King 1973; Weygoldt and Paulus 1979)
32. Inverse retina in four median eyes (Paulus 1979)
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); enclosed to form book lungs (2)
36. Extraintestinal digestion
37. Five simple lateral eyes
38. Slit sensillae (Weygoldt and Paulus 1979)
39. Vitreous body present in median eyes (Paulus 1979) Characters 36 and 37 are questioned by Shultz (1990).
40. Palps on first and second maxillae absent (Brusca and Brusca 1990) Other features mentioned as myriapod synapomorphies may be more broadly distributed (such as the organ of Tömösváry) or have unclear homology relationships (stink glands).
41. Thorax divided into three segments, each with a pair of limbs

42. Locomotory limbs six-segmented
43. Abdomen with twelve 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 (Kristensen 1975; Boudreaux 1979; Paulus 1979; Hennig 1981)
50. Antennae and palps: 0 = nothing; 1 = prostomial/pre-oral palps; 2 = post-oral antennae
51. Lateral eyes: 0 = absent; 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 unclear due to the structural dissimilarities between their tracheae and those in myriapods and hexapods. This feature also occurs in some araneamorph 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-stomodaeum-midgut-mesoderm-posterior) v. anterior-midgut-mesoderm-stomodaeum-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 pre-oral acron (prostomium) and the non-metameric 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 = no cilia except in sperm (although some sensillae may be cilia-derived) (Paulus 1979)
66. Elongated dorsal gonads
67. Development of ventrolateral appendages
68. Reduction of coelom-hemocoel 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 (Boudreaux 1979; Weygoldt 1986; Brusca and Brusca 1990)
73. Rhabdomeric retinular structure in eye facets (Paulus 1979)
74. Nephridia in, at most, first four cephalic and first two post cephalic segments (Weygoldt 1986)
75. Hard exoskeleton
76. Articulating jointed appendage with arthrodial membrane
77. Fully segmented sclerites
78. Cephalic (at least anterior) ecdysis glands
79. Cephalon with one pair of pre-oral and four pair post-oral appendages (Weygoldt 1986; Brusca and Brusca 1990)
80. All muscles striated
81. Suppression of all circular body wall muscle
82. Similar intersegmental 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 4 and 5 (Brusca and Brusca 1990)
88. Loss of mandibular palpus (inactive recoded in crustacean area)
89. Malpighian tubules: 0 = absent; 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 and Paulus 1979; Weygoldt 1986)
94. Embryos with four gangliar post-oral segments
95. Separation of cephalic and locomotory functions onto different tagma
96. Coxal endites on appendage of second somite (primitively on all) (Boudreaux 1979)
97. Anterior (0) v. 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 pre-oral limb-bearing segment in larva (if not adult)
101. $tRNA^{C}$ is between $tRNA^{W}$ and $tRNA^{Y}$
102. $tRNA^{Y}$ is inverted with respect to $tRNA^{W}$
103. $tRNA^{M}$ is between $tRNA^{Q}$ and ND2
104. $tRNA^{S(AGN)}$ is between $tRNA^{N}$ and $tRNA^{E}$
105. l-rRNA / $tRNA^{L(CUN)}$ / $tRNA^{L(UUR)}$ / ND1
106. COI / $tRNA^{L(UUR)}$ / COII

107. l-rRNA / $tRNA^{L(CUN)}$ / ND1 (Boore et al. 1995—somewhat charitably coded)
108. Unique crustacean limb segmentation pattern based on a shared coxa-basis muscular arrangement (Boxshall 1996)
109. Tarsal claws (paired?) (Hennig 1981)
110. Coxal vesicles, styli
111. Maxillary plate (mouth cavity bordered by second maxillae)
112. Appendages of first post-cephalic segment transformed into maxillipedes
113. Specialization in the ventral border of the mouth cavity
114. Stemmata
115. Diplosegments
116. Antennae with four sensory cones in distal segment
117. First post-cephalic segment into collum (Kraus and Kraus 1994)

Characters 25, 50, 52, and 89 are nonadditive, the remaining multistates are ordered.