Acta Zoologica (Stockholm), Vol. 73, No. 5, pp. 383–389, 1992
Printed in Great Britain

0001–7272/92$5.00+ .00
Pergamon Press Ltd

Computerized Analysis of Crustacean Relationships

George D. F. Wilson
Division of Invertebrates, Australian Museum, P.O. Box A285, Sydney South, NSW 2000, Australia

Abstract

Wilson, G. D. F. 1992. Computerized analysis of crustacean relationships.—*Acta Zoologica* (Stockholm) 73: 383–389.

The strengths and weaknesses of phylogenetic analysis using computers are reviewed from the viewpoint of understanding crustacean evolution. Computerized methods require the explicit presentation of characters and character state homologies. New techniques allow investigators to design evolutionary models into a character data matrix, or to use evolutionary models that make minimal a priori assumptions. The computer analysis relieves the investigator from the highly repetitious testing of trees, allows the concentration on the character state data, and provides objective methods for comparing trees, primarily their length. These are regarded as the strengths of computerized methods. The weaknesses of these methods include the relatively inscrutable nature of the character data matrix compared with the overall 'gestalt' of resulting trees, the difficulties of defining discrete homologies within the Crustacea, especially for counts of segmentation, the lack of clear intermediate character states in some multistate segmental characters, and the inability to define evolutionary polarity. These difficulties may be overcome by analysing the data using the minimal assumption models of character evolution, and by a recognition that the trees are a result of the input data, and therefore the data should be criticized, rather than the trees themselves. A 'consensus' character data set, including most extant major groups of the Crustacea as well as several key fossils, was assembled and revised by the participants in the workshop. An artificial taxon, 'ur-crustacean characters', was introduced to root the tree. Three observations may be made from parsimony analyses using several weighting and tree rooting methods. (1) The currently accepted large scale phylogeny and classification of the Crustacea is not corroborated. (2) The number of supposed plesiomorphic traits possessed by a taxon is not a good index for early derivation in crustacean evolution. (3) The taxon Maxillopoda is not supported by the arrangement of any of the trees.

George D. F. Wilson, Division of Invertebrates, Australian Museum, P.O. Box A285, Sydney South, NSW 2000, Australia.

Introduction

The advent of effective computer programs for the estimation of phylogenies has revolutionized our ability to objectively study these problems in crustacean taxa. Recently, several objective studies of crustacean phylogeny using cladistic methods (Schram 1986; Brusca & Brusca 1990) derived solutions differing from more traditional classifications. Simple verbal argumentation is no longer sufficient for demonstrating a particular evolutionary scenario for the Crustacea. Workers can now meet on the common intellectual ground provided by objective techniques of phylogenetic estimation. This paper discusses the strengths and weaknesses of the cladistic approach as applied to crustacean relationships. As a test case, a new analysis of the Crustacea, based on data provided by the members of the workshop (see list of participants and methods below), is presented. Although this analysis is limited at the outset by not being provided with a detailed character analysis, it provides additional hypotheses to test in future inquiry.

A fundamental goal for our workshop is to estimate the phylogeny of all Crustacea. Determination of ancestral crustacean character states, a traditional pursuit in previous papers (e.g. Hessler & Newman 1975) is only a secondary objective. A collection of ancestral features was deduced for this analysis, although this practice is largely unnecessary if outgroups are included to provide a rooting of the taxa of interest. Once the phylogeny is understood reasonably well, the ancestor might be profitably appraised. A focus on ancestors during the analysis may lead to circularity: if the ancestral characters are assumed, the analysis runs the risk of being forced along paths that may omit more parsimonious and reasonable solutions. Moreover, using deduced ancestors of selected subsets of taxa within the entire group as a means of grouping taxa into subsolutions also greatly limits possible outcomes in the search for the best tree. Therefore, the use of all taxa and all characters simultaneously in one analysis has the highest probability of obtaining an optimal solution.

Why should computer-assisted methods be used in such a simultaneous analysis? This question has been discussed in the literature (see Felsenstein 1982; Swofford & Olsen 1990), so little detail will be presented here. In short, homoplasy and the vast number of trees to be examined control our choice of method. Our analysis of character state homologies may contain errors, or homoplasy. If no conflicting characters were present in a data set, phylogenetic descent perhaps could be deduced without

recourse to computerized algorithms, as in Hennig's (1964) method. Most real data sets will contain much homoplasy, including misinterpreted homologies via convergence, parallel evolution or reversals. Characters also may be used at an inappropriate level; for example if a feature changes many times over the evolutionary expanse subsumed by the taxa in an analysis, the feature may be largely informationless for the problem at hand.

If homoplasy exists, our problem becomes how to find the best solution for a phylogeny that has the least number of changes, or said another way, the fewest hypotheses of character evolution. In existing algorithms, the shortest tree cannot be calculated from the character distributions. Rather a tree must be assembled and then tested against existing shortest trees. If a test tree is shortest, it is retained; if not, it is discarded. Computers are excellent at this highly repetitive task. In large data sets, this type of search is nearly impossible to do by hand, primarily because the number of potential trees goes up as a polynomial function of the number of taxa (Felsenstein 1978). For the number of taxa in the analysis below (17), approximately $2.91898783962511 \times 10^{17}$ possible fully bifurcating, rooted trees exist that must be evaluated if an exhaustive search is made.

The strengths of computer-assisted phylogenetic analysis are the overall objectivity enforced on actual estimation of the phylogenies, the standardization of data representation used for analysis, and the freedom from the algorithms (the computer does the work), so that the focus of research can be on the characters and homologies. Because the same methods are used by all phylogenetic programs, the same data set analysed under the same presumed models of character evolution will produce the same results, regardless of who is doing the analysis. The trees have an objective index for comparison—their relative length in the number of character changes required. Furthermore, because the data must be specified for the analysis, the investigator must present a detailed justification for homologies of all characters used in the analysis and codify this into a data matrix for analysis. Such detailed accounting will allow any scientist to evaluate decisions relating to character evolution. This focus on the characters themselves is the real strength of the computerized methods: the relative freedom from algorithms.

The weaknesses of computerized analysis of phylogenies originate not from the methods themselves, but from the data used. When a particular 'computer tree' is viewed with disfavour, the real disagreement should be placed on the data used to derive the tree. Unfortunately, the tree itself comes under criticism and the data are less likely to be evaluated because one cannot easily grasp the 'gestalt' of a large matrix of character states. Nevertheless, if a critic of a particular phylogeny finds fault with a particular tree topology, it is because she or he has special knowledge of character state distributions that imply relationships not seen in the viewed tree. The tree may be rejected out of hand, while the proper course is to see if the special knowledge held by the critic is actually present in the data set. If not, then these data can be added, and the analysis re-run. The outcome from such a critical test should lead towards a better understanding of the taxa being evaluated.

Crustacea, by virtue of their possession of a rigid exoskeleton, provide many external characters for use. The data set in Table 2 represents a subset of the possible features that one could use. Within the Crustacea, however, the greatest problem is the determination of segment homologies, both for body somites and for podomeres. Boxshall & Huys (see earlier remarks by Boxshall in this issue) have elaborated on this problem for the Copepoda, but it is a general problem for all crustacean phylogenetic studies. When we score a group of taxa with (for example) 11 post-cephalic somites, we are assuming that these somites are indeed homologous, and that the path by which the possessors of this character state came to this condition is identical, i.e. shared ancestry. The 11 somites could be obtained, say, by loss of segments in the middle of the body, or by increment from a previous shorter condition. Scoring both situations as homologous would inject important homoplasy into the data analysis. This form of homoplasy may be even more common in the limb configuration, where the segments can be lost or gained somewhat more easily anywhere along their length.

A difficulty with computer analysis of phylogenies is the relative complexity of the decisions to be made concerning allowable character transitions. The easiest and most effective way to estimate phylogenies is to make as few assumptions (or decisions) as possible about how the characters evolved. Analytical constraints on character or taxon evolution can be used when the evidence for a particular path is strong. In most instances of characters in the crustacean analysis below, however, this is not the case. Consequently, I have chosen the simplest course by using unordered multistate characters in a reversible parsimony analysis of the crustacean data. To provide some understanding for this approach, I discuss the various options available. These are reviewed elsewhere (Felsenstein 1982; Swofford & Olsen 1990), so I only touch on these issues briefly.

An ordered 'Wagner' (Kluge & Farris 1969; Farris 1970) or unordered 'Fitch' (Fitch 1971) parsimony analysis allows characters to reverse direction in a search for the shortest tree. Reversals in character states may seem unsatisfactory to many systematists, while parallelisms and convergence are common features of evolutionary descent. Reversals should be interpreted as an attempt of the algorithm to find the most parsimonious path to represent changes in character states—not as a representation of the real evolutionary path. The most parsimonious path of character evolution using a reversal might signify a place where we have classified non-homologous character states as homologous, thereby creating a homoplasy. (Homologous character states can be defined as a character state originating in a single ancestral species, and perpetuated unchanged to descendent taxa sharing this state.) For example, consider the simple case of the gain and then loss of a seta in a lineage. Under morphological examination, we cannot tell the difference between the state of the original absence and the state after the loss. Without supporting data, we would be forced to conclude that both states were homologous, while in fact they are not: that is, the loss state is descended from a state where the seta was present and therefore not homologous to the ancestral absence state.

A reversible parsimony evolutionary model with supporting data will correctly identify that a gain and a loss has taken place. When we estimate the real evolutionary derivation of characters which reverse, we should reexamine specimens and data for the causes of these anomalous transitions.

A phylogenetic analysis also may allow multistate characters to change from one state to any other state in one step. This is the Fitch (unordered) parsimony method, called 'unordered' or 'nonadditive' characters in various lexicons, although these terms are not necessarily equivalent. By not specifying the presumed character evolution, the analysis can find the shortest trees, after which the character transitions can be determined after the analysis. Other character evolution models may be used at the expense of simplicity and maximum parsimony. The character transitions can be ordered in a linear sequence in the Wagner parsimony model. If convergence or parallelisms are expected, then perhaps a 'Camin–Sokal' parsimony method (1965), which does not allow reversals and forces convergence and parallelisms, is appropriate. Other possible evolutionary models, e.g. Dollo (Farris 1977) or polymorphism (Felsenstein 1979), may also apply. Recently, techniques have become available that subsume many models into a single generalized model, wherein the cost of each character state transition can be assigned independently (Swofford 1990; Swofford & Olsen 1990). Generalized parsimony may be most useful for molecular characters, although the construction of a specific model for morphological evolution is possible. Any of these models, however, are always more restrictive than Fitch parsimony, both by limiting the possibilities of character transitions as well as forcing character state transitions into models that may or may not be appropriate.

Character transitions from an analysis using fully reversible methods, however, can be reinterpreted to determine the most parsimonious transitions a posteriori, rather than a priori. If constraints are enforced on character evolution, either by specifying the path by which characters can evolve or by allowing only certain transitions, then the researcher is left with testing a myriad possibilities in different combinations of transitions. Consequently, an unordered parsimony analysis has the best chance of finding a result closest to the true phylogeny. This philosophy has been employed in the following estimates of the crustacean phylogeny.

In the Crustacea, a general trend of decrease in body segment number is thought to correspond to the sequence of taxon derivation; i.e. more 'primitive' taxa have more body somites. I suspect that this might be an idealized concept based on presumed ancestors (an annelid ancestry), rather than one based on the evidence. The fossil evidence does not provide much guidance because the taxa with the most body segments are those alive today with little or no fossil record (branchiopods, remipeds, cephalocarids), while the ostracodes (appearing in the Cambrian) and the extinct Orsten fauna (Walossek & Müller 1990), have limited somite counts. The simplest crustacean, the nauplius larva, has only three pairs of limbs and no body segments. The homology of segmentation is especially a problem for the Ostracoda, which have few distinct body segments and an uncertain arrangement of the post-mandibular limbs. If we assume the ostracodes got to this state by substantial reduction, then they are highly advanced in body construction. If the crustacean ancestors had only a few body segments, then the Ostracoda evolved from an ancestral state without many changes to their current form. By constraining our phylogenetic solutions to a strictly ordered analysis of segmentation, we may arrive at answers that only reflect our preconceived (and possibly wrong) notions of crustacean phylogeny.

Other difficulties with the crustacean data set include the lack of clear intermediates for some character states. This observation is probably a reflection of the vast time gap between when the taxa arose and now. An especially interesting view of crustaceans is provided by the Orsten fauna (Walossek & Müller 1990) where despite the fundamental differences of head and limb structure observed in many of these animals, some general aspects of construction recur in many of the taxa. These animals are a window into the early times in crustacean evolution.

Materials and Methods

A crustacean character data set was assembled by using a list of 'ur-crustacean' character states. These states were accumulated by an *ad hoc* committee (Boxshall, Briggs & Newman). The list was then entered into a spreadsheet program, and alternative states were added to the list. A scoring for the Copepoda was provided by Boxshall as an example. This list with the example was then distributed to the workshop attendees for contributions to the data set. Contributions were made by Boxshall, Briggs, Fryer, Grygier, Hessler, Newman, Rolfe, Wägele, Walossek, Watling and Wilson. The taxa scored were: 'ur-crustacean characters', Branchiopoda, Branchiura, *Bredocaris*, *Canadaspis*, Cephalocarida, Copepoda, Eumalacostaca, *Lepidocaris*, Leptostraca, *Martinssonia*, Mystacocarida, *Nahecaris*, Ostracoda, Remipedia, Skaracarida, Tantulocarida, and Thecostraca. *Martinssonia*, however, was not used in most analyses because insufficient characters were included to identify this taxon as a member of the stem Crustacea group (Walossek & Müller 1990). These contributions were entered into the same data spreadsheet with several new characters suggested by the participants. The data were corrected where a character state attributed to a taxon was found to be in disagreement with the literature or workshop participants' findings; some differences were due to typographical errors. The assembled data set was submitted to the participants for further correction and addition. Further corrections and changes were provided by Bergström, Briggs, Fryer, Grygier, Hessler, and Rolfe, as well as by Ann Cohen (Los Angeles County Museum) for ostracodes. During this analysis, several characters were found to be uninformative or confusing to some of the workshop participants; these characters were eliminated from the analysis. Moreover, some scorings for 'ur-crustacean characters' were downgraded to unknown ('?'), or were expanded to allow for more than one possibility. The list of characters and their states is in Table 1 and the actual data matrix used in the analysis is in Table 2. Space and the source of the data do not allow a detailed discussion of each of these characters. Ordinarily I would make such a presentation, but I can only report the data as provided by the workshop participants.

The character state data, at the various stages of editing, were submitted to the following PAUP (version 3.0g or 1; Swofford 1990) procedures. Multistate taxa were interpreted as having polymorphisms, and the outgroup was set to the hypothetical ancestor 'ur-crustacean'. Two weighting algorithms were tried. The first weighting method is referred to herein as 'normalized weighting', i.e. all characters weighted inversely to the number of character states, using a base weight of 1000. Therefore, all two state characters received a weight of 1000, and multistate characters received a weight adjusted so that all steps in each character would have a summed effective weight over all steps equivalent to all other characters. In the second weighting method, all character weights were initially set to one, weighting each step of each character equally. This method will give the multistate characters more weight in a parsimony analysis. After the most parsimonious trees were found, the data were reweighted once or several times using a successive weighting algorithm (Farris 1989). The weights were based on the rescaled consistency index of the best fits of each character among all trees found. This

Table 1. *Crustacean evolution workshop phylogenetic analysis character state list. The presumed ancestral state is shown in curly brackets*

1...Antennules: 0, uniramous; 1, biramous; {0}.
2...Compound eyes: 0, stalked; 1, sessile; {?}.
3...Eye type: 0, malacostracan; 1, entomostracan; {?}.
4...Nauplius eye: 0, 3-cup; 2, 4-cup; 3, 2-cup; {?}.
5...Antenna: 0, biramous; 1, uniramous; {0}.
6.....Naupliar process: 0, present; 1, absent; {0}.
7.....Exopod: 0, flagellate; 1, non-flagellate; {0}.
8.....Segment counts: 0, 20+; 1, 19–17; 2, 14–13; 3, 10–9; 4, 3–4; 6, 2–3; 7, 1; 8, absent {1}.
9.....Exopod seta position: 0, inner; 1, outer; {0}.
10.....Endopod: 0, 5-seg; 1, 4-seg; 2, 3-seg; 3, 2-seg; {0}.
11.....Antennal gland: 0, present; 1, absent {0}.
12...Labrum: 0, present; 1, absent; {0}.
13...Mandible: 0, biramous; 1, uniramous {0}.
14.....Coxal gnathobase: 0, present; 1, absent; {0}.
15.....Endopod: 0, 3-segs; 1, 2-segs; 2, absent; {0}.
16.....Exopod: 0, flagelliform; 1, non-flagelliform; 2, absent; {0}.
17.....Segment number: 0, 12–11; 1, 7–6; 2, 5; 3, 1; 4, absent; {0}.
18...Maxillula: 0, biramous; 1, uniramous; {0}.
19.....Protopodal endites: 0, present; 1, absent; {0}.
20.....Endopod: 0, 4-segs; 1, 3-segs; 2, absent; {?}.
21.....Exopod: 0, flagelliform; 1, 2-segmented paddle; 2, 1-segmented paddle; 3, rudimentary; 4, absent; {1}.
22.....Epipodite: 0, absent; 1, present; {0}.
23...Maxilla: 0, biramous; 1, uniramous; 2, lobate; {0}.
24.....Maxilliary gland: 0, present; 1, absent; {0}.
25.....Protopodal endites: 0, 8–6; 1, 5–4; 2, 3–1; 3, 0; {1}.
26.....Endopod: 0, 6 segments; 1, 4 segments; 2, 1 segment; {?}.
27.....Exopod: 0, flagelliform; 1, 2-segments; 2, unsegmented; 3, absent; {2}.
28...Thoracopod 1: 0, biramous; 1, uniramous; 2, triramous; {0}.
29.....Protopod endites: 0, present; 1, absent; {0}.
30.....Exopod segments: 0, 3; 1, 2; 2, unsegmented; 4, absent; {2}.
31.....Endopod segments: 0, 7; 1, 6; 2, 5; 3, 4; 5, 3; 6, 2; 7, 1; {0 or 1}.
32...Remaining thoracopods: 0, biramous; 1, uniramous; 2, triramous; {0}.
33.....Protopod endites: 0, present; 1, absent; {0}.
34.....Exopod segments: 0, flagelliform; 1, 3; 2, 2; 3, unsegmented; 4, absent; {3}.
35.....Endopod segments: 0, 7; 1, 6; 2, 5; 3, 4; 4, 3; 5, 2; 6, 1; {1}.
36...Trunk limbs (post-maxilliary) on female: 0, 40–16; 1, 15–12; 2, 9; 3, 8; 4, 7; 5, 6; 6, 5; 7, 4; 8, 3; 9, 1; {0}.
37...Trunk limbs on male: 0, 40–16; 1, 15–12; 2, 9; 3, 8; 4, 7; 5, 6; 6, 5; 7, 4; 8, 3; 9, 1; {0}.
38...Trunk somites including telson: 0, 15+; 1, 12; 2, 11; 3, 8; 4, 5; {0}.
39...Limb size posteriorly: 0, decreasing; 1, not decreasing; {0}.
40...Limb position: 0, ventrally; 1, laterally {0}.
41...Anus: 0, posterior; 1, ventral; {0}.
42...Furcal rami: 0, present; 1, absent; {0}.
43...Furcal rami articulation: 0, articulated basally; 1, fused basally {0}.
44...Male 7th trunk limb: 0, not modified; 1, partially fused; 2, fused; {0}.
45...Sperm: 0, flagellated; 1, modified; {0}.
46...Trunk somite pleurae: 0, present; 1, absent; {?}.
47...Pleural adductor muscle: 0, present; 1, absent; {0}.
48...Dorsal cephalic shield: 0, no extension; 1, not fused; 2, fused; {0}.
49...Cephalic segments: 0, 5; 1, 4; 2, 6; {0}.
50...Dorsal organ: 0, present; 1, absent; {?}.
51...Development: 0, anamorphic; 1, metamorphic; 2, abbreviated; {0}.
52...Naupliar limbs: 0, 3; 1, 4; {0}.
53...Pleopods: 0, absent; 1, present; {?}.
54...Carapace adductor muscle: 0, present; 1, absent; {0}.
55...Telson posteriorly elongated: 0, present; 1, absent; {0}.
56...Carapace: 0, univalved; 1, bivalved; {?}.
57...Rostral Plate: 0, fused to carapace; 1, articulated at carapace margin; {?}.
58...Segments in Pleotelson: 0, no pleopods on abdomen; 1, seven segments; 2, six segments; {?}.
59...Non-locomotory limbs on thorax: 0, absent; 1, present; {?}.

decreases the number of characters that potentially could be ignored.

The resulting data set was too large and too homoplasous for the use of an effective branch and bound procedure (Hendy & Penny 1982), so several heuristic procedures were employed instead. A randomized taxon addition sequence with 10 iterations was effective at finding the shortest trees. Minimal trees were found regardless of whether the 'tree bisection–reconnection' or the 'subtree pruning–reconnection' methods were used (see Swofford 1990) with the randomized addition sequence.

Results of the Phylogenetic Analyses

Weighting algorithms proved to be unnecessary for the rooted analysis, or they produced misleading results. Only one minimal tree was found in the equal weighted characters analysis (Fig. 1), so successive weighting was unnecessary. This tree had a length of 242 steps, a consistency index (CI) of 0.702, a homoplasy index (HI) of 0.533, a retention index (RI: see Farris 1989) of 0.493, and a rescaled consistency index (RC) of 0.346. With normalized weighting, the decrease in the segmental characters' impact caused the Eumalacostraca to group with the Copepoda, clearly an incorrect result. Most synapomorphies of the Eumalacostraca and the Leptostraca are in multistate characters, so decreasing the strength of these characters left little to define the unique position of the Eumalacostraca. Other parts of the tree varied little in most analyses. In one combination of taxa, i.e. excluding the 'hypothetical ur-crustacean', produced similar results, although with less resolution. The unweighted analysis produced 24 trees each with a minimal length of 231 steps. These trees were used for successive weighting to increase the resolution. The characters were reweighted using the maximum value of their rescaled consistency indices, and then the analysis was run again. Only one run with the rescaled weights was necessary before a stable result was obtained. In this case, three trees were found (length 232 steps, CI = 0.716, HI = 0.517, RI = 0.496, RC = 0.355). The unrooted strict consensus of these three trees (Fig. 2) is reasonably well resolved, but suggests a radically different classification for the Ostracoda. Without the 'ur-crustacean' to root the tree, the exact form of evolutionary descent comes under discussion. Consequently, three different trees, each rooted differently, were generated from the Nelson tree to provide examples (Fig. 3) of how the unrooted tree (Fig. 2) might be directed.

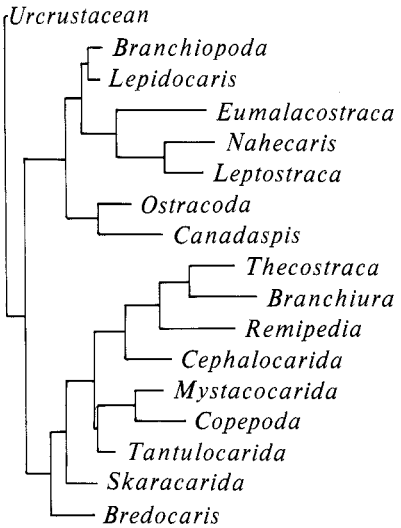

Fig. 1. Single tree resulting from an unweighted, unordered character analysis of the data in Table 2. This tree had a length of 242 steps, CI = 0.702, HI = 0.533, RI = 0.493, and RC = 0.346.

Table 2. *Input data matrix. Uninformative/excluded characters, and deleted taxa not included. Alternative states for each taxon in same column below row with taxon name*

Characters (10's digit)							1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5
Characters	1	2	5	6	7	8	0	1	5	6	7	0	1	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	6	7	8	9	0	1	3	4	5	6	7	8	9
Taxa	0	?	0	0	0	1	0	?	0	0	0	2	1	1	?	2	0	0	2	0	0	0	3	0	0	0	0	0	0	0	0	0	?	0	0	0	?	0	0	?	0	?	?	?	?
Ur-crustacean													2							2					2																				
Branchiopoda	0	0 1	0	1	?	1 to 5	?	2	2	2	4	0	4	2	3	3	0	0	2	0	0	0	3	0	0 to 7	0 to 7	0 to 4	0 1	0	0	1	0	0	?	1	1	0 1	0	0	0 1 2	0	0	1	0 1	?	0	0
Eumalacostraca	1	0 1	0	0	0	1	0	0	0	1	4	5	4	2	3	2	0	0	1	4	0	0	0 2	4	1	1	0	1	0	1	0	0	0	0	1	0 2	0	?	1	1	1	1	0	0 1	2	0
Lepidocaris	0	?	0	1	0	4	2	?	2	2	4	4	4	0	3	2	0	0	2	8	0	0	3	8	1	1	0	0	0	0	0	0	0	?	?	0	0	?	0	0	1	1	0	0	0	0
Ostracoda	0	0 1	0	1	0 1	3 5 6	1 2	0	0	1	3	4	2	2 3	?	2 3	0	0 1	?	5	1	1	4	0 4 5 6	8	8	2	0 1	0	0	1 1	1	0	?	1	1	0	0	?	0	0	0	1	1	0	0	0
Thecostraca	0	0 1	2	0	?	?	?	1	2	2	3	5	4	2	0	3	0	1	1	7	0	1	2	6	4	4	2	1	0	0	0	0	1 2	0	0	1	0	?	1	0	0	1	0 1	0	0	0	0
Nahecaris	1	0	0	?	0	0	2	?	0	2	4	?	?	?	?	?	0	?	?	4	0	?	?	4	1	1	0	0	0	1	0	0	0	1	1	1	?	?	?	1	0	0	1	1	1	1
Brcdocaris	0	1	0	0	0	1	2	?	0	0	0	3	3	0	2	2	0	2	2	6	0	0	3	6	4	4	8	0	0	0	0	1	0	1	?	0	0	0	1	0	?	1	0	0	0	0
Canadaspis	0	0	1	?	0	2	?	?	2	2	3	1	2	2	0	2	0	0	2	1	0	0	3	1	3	3	0	1	0	0	1	?	0	1	1	1	1	?	?	0	0	1	1	0	0	0
Mystacocarida	0	?	0	0	0	3	1	?	0	0	1	3	4	0	2	3	0	0	1	6	2	?	4	8	6	6	2	1	0	0	0	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0
Skaracarida	0	?	0	0	0	2	2	0	0	0	0	3	1	1	2	0	0	0	2	3	1	2	?	?	9	9	1	?	0	0	0	0	0	0	?	0	0	0	?	0	1	1	0	0	0	0
Leptostraca	1	0	1	1	?	?	?	0	0	2	4	5	4	1	3	2	2	0	2	4	2	0	3	4	1	1	0	0	0	0	0	0	0	0 1	1	2	0	?	2	1	0	1	0	1	1	1
Cephalocarida	0	?	0	0	0	1	3	2	2	2	4	2	2	1	1	1	2	0	1	3	2	0	2	3	2	2	0	0	0	0	0	0	0	3	0 1	0	0	0	?	0	0	0	1	0	0	0	0
Remipedia	1	?	0	1	1	6	2	?	2	2	4	3	4	1	2	0	1	0	3	4	0	1	1	5	0	0	0	0	1	0	0	0	0	0	?	2	2	1	?	0	?	1	0	0	0	0
Tantulocarida	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	0	0	1	7	0	0	2	7	5	4	1	0	0	0	0	0	0	2	0	?	0	0	1	1	0	1	1	0	0	0	0
Branchiura	0	1	0	0	1	6	2	1	1	2	4	5	4	3	2	3	0	1	1	3	0	1	2	8	7	7	4	1	1	0	0	1	0	1	?	2	2	0	?	2	0	?	1	0	0	0	0
Copepoda	0	?	0	0	0	3	1	0	1	0	2	4	2	1	2	?	1	0	3	3	0	1	1	6	5	6	2	1	0	0	0	0	0	?	1	2	2	1	0	0	1	1	0	0	0	0

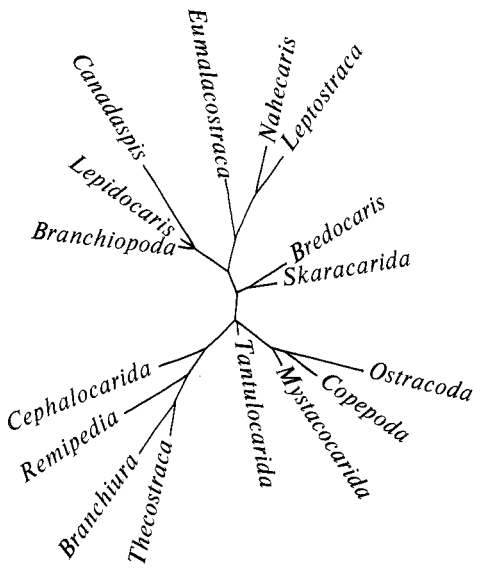

Fig. 2. Unrooted strict consensus of three trees resulting from a single iteration of successive weighting with 'ur-crustacean' deleted. The three trees had a length 232 steps, CI = 0.716, HI = 0.517, RI = 0.496, and RC = 0.355.

Discussion

These trees contain several important results. Firstly, the large scale arrangements of the Crustacea (see Brusca & Brusca 1990) are not corroborated by these results. These phylogenies recognize potentially non-monophyletic taxa (either paraphyletic or polyphyletic) such as the Maxillopoda, and therefore may be disadvantaged at the outset. Computerized methods now available can permit successful estimation of phylogenies for large numbers of taxa. Consequently, any taxon whose monophyly is in doubt can be evaluated via all of its components, rather than as a single entity. This criticism can be aimed at this report because a few of the taxa, particularly the Branchiopoda, subsume considerable diversity and are potentially non-monophyletic.

Secondly, the possession of many plesiomorphic traits (e.g. some Branchiopods and the Cephalocarida) is not a good index of early derivation or relative primitiveness. Despite many discussions in the literature and in this workshop, remipeds, cephalocarids, or branchiopods do not appear in this analysis to be the earliest derived of the extant living Crustacea. The Ostracoda, after deleting extinct taxa, may get that honour. This result, however, is somewhat meaningless, because all crustacean taxa living today have been evolving and changing for at least the last 600 million years and therefore are all quite derived compared with the fossil taxa. The tree in Fig. 1 is satisfactory in that most fossil taxa branch off early.

Thirdly, the concept of the Maxillopoda is not supported in any of the trees. The discussions at the workshop suggested that few (if any) synapomorphies could unambiguously define the group, despite the interest in maintaining the Maxillopoda. In Figs 1–3, the Thecostraca, the Branchiura, and the Remipedia cluster together with the Cephalocarida, while the Copepoda and the Mystacocarida cluster with the Tantulocarida. In the unrooted tree in Fig. 2, the two main groups of the Maxillopoda are separated by the intervening Remipedia and Cephalocarida. Unless the Maxillopoda is redefined to contain Remipedia and Cephalocarida, it will remain a poorly corroborated concept under the hypothesis of crustacean phylogeny presented in this report.

At this juncture, I hasten to point out that the tree in Fig. 1 is a working hypothesis, perhaps best regarded as a target for further research. Undoubtedly, readers will find some or all of the tree at odds with their knowledge of the Crustacea. In my defence, I can only offer that a necessary part of the analysis, the character analysis, is missing from this paper. As mentioned in the introduction, automated methods do not give correct or reasonable results if non-homologous characters and character states are mixed. Most of the data used here were collected quickly over a few days, and then corrected over the intervening months. This analysis is dependent on the impressive and collective wisdom of the participants of the workshop; I do not doubt the providers of these data. And yet, our knowledge of homologies spanning 600–700 million years of evolution in such a diverse array of taxa remains limited in many respects. Therefore, the weaknesses of this study can only help to highlight the general need for future morphological and biochemical studies.

Acknowledgements

I am grateful for the invitation from Professors Eric Dahl and Jarl-Ove Strömberg to participate in this workshop. The workshop participants should be recognized for their contribution to this report via providing and checking of the basic data which were used to generate the new hypothesis of crustacean phylogeny. Drs Ann Cohen and Fred Schram (Los Angeles County Museum) provided corrections and additions to the data matrix. Mr Ricardo Martinez (Scripps Institution of Oceanography) and Dr Pat Hutchings (Australian Museum) kindly allowed this inveterate PC user to run PAUP 3.0 on their Macintosh computers.

References

Brusca, R. C. & Brusca, G. 1990. *Invertebrates*. Sinauer Associates, Sunderland, Massachusetts.

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

Farris, J. S. 1970. Methods for computing Wagner trees.—*Systematic Zoology* 19: 83–92.

Farris, J. S. 1977. Phylogenetic analysis under Dollo's Law.—*Systematic Zoology* 26: 77–88.

Farris, J. S. 1989. The retention index and the rescaled consistency index.—*Cladistics* 5: 417–419.

Felsenstein, J. 1978. The number of evolutionary trees.—*Systematic Zoology* 27: 27–33.

Felsenstein, J. 1979. Alternative methods of phylogenetic inference and their interrelationship.—*Systematic Zoology* 28: 49–62.

Felsenstein, J. 1982. Numerical methods for inferring evolutionary trees.—*Quarterly Review of Biology* 57: 379–404.

Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specified tree topology.—*Systematic Zoology* 20: 406–416.

Hendy, M. D. & Penny, D. 1982. Branch and bound algorithms to determine minimal evolutionary trees.—*Mathematical Biosciences* 59: 277–290.

Hennig, W. 1966. *Phylogenetic systematics*. University of Illinois Press, Urbana.

Hessler, R. R. & Newman, W. A. 1975. A trilobitomorph origin for

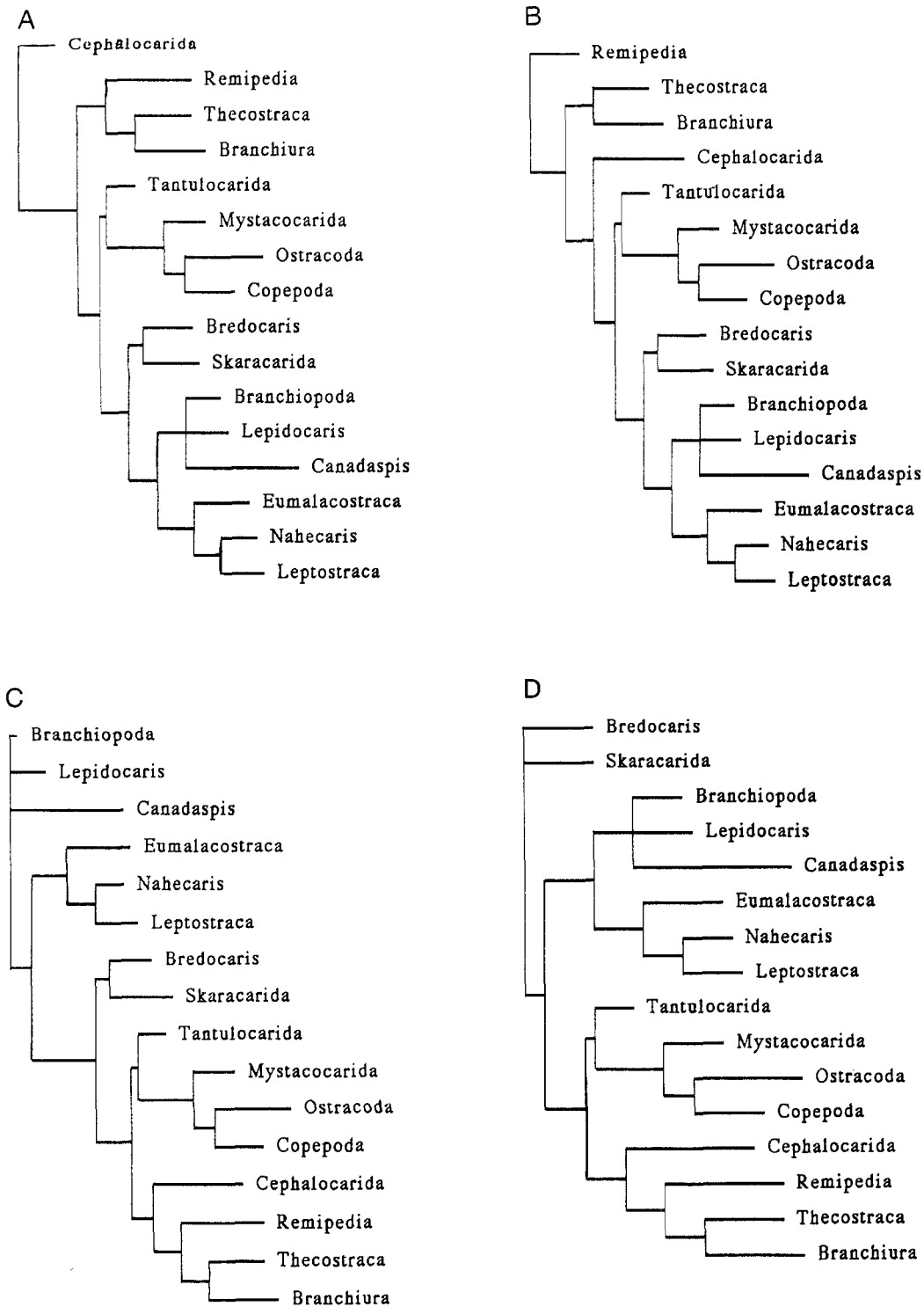

Fig. 3. Same strict consensus tree as in Fig. 2, but re-rooted with several taxa hypothesized by different authors to be earliest derived among Crustacea.—*A*, Cephalocarida;—*B*, Remipedia;—*C*, Branchiopoda, plus one tree (*D*) rooted on the Cambrian fossil crustaceans *Bredocaris Skara*.

the Crustacea.—*Fossils and Strata* 4: 437–459.
Kluge, A. G. & Farris, J. S. 1969. Quantitative phyletics and the evolution of anurans.—*Systematic Zoology* 18: 1–32.
Schram, F. R. 1986. *Crustacea*. Oxford University Press, New York.
Swofford, D. L. 1990. PAUP: phylogenetic analysis using parsimony, Version 3.0. Illinois Natural History Survey, Champaign, Illinois.
Swofford, D. L. & Olsen, G. J. 1990. Phylogeny reconstruction. In D. M. Hillis & C. Moritz (eds): *Molecular systematics*, pp. 411–501. Sinauer Associates, Sunderland, Massachusetts.
Walossek, D. & Müller, K. J. 1990. Upper Cambrian stem-lineage crustaceans and their bearing upon the monophyletic origin of Crustacea and the position of *Agnostus*.—*Lethaia* 23: 409–427.