

# Origin of the Metazoa

(ribosomal RNA/phylogenetic trees/Arthropoda/Annelida/Mollusca)

JAMES A. LAKE

Molecular Biology Institute and Department of Biology, University of California, Los Angeles, Los Angeles, CA 90024

Communicated by Everett C. Olson, November 3, 1989

**ABSTRACT** The origin of the multicellular animals has been investigated by rate invariant analysis of 18S rRNA sequences. These analyses indicate that (i) the Metazoa is a monophyletic taxon; (ii) the Deuterostomia is a monophyletic taxon; (iii) the Annelida–Mollusca lineage is the sister group of an arthropod subgroup; and (iv) the last common ancestor of the Annelida–Mollusca lineage is most parsimoniously derived from a segmented, hemocoelic ancestor with an open circulatory system.

Molecular sequence data can provide new opportunities to understand metazoan phylogenetic relationships and complement the extensive data obtained by functional, developmental, structural, and paleontological analyses. The set of metazoan rRNA sequences recently obtained by Field *et al.* (1) is a promising source of phylogenetic information. Their trees, however, have been criticized because they are mutually contradictory and inconsistent with traditional morphological characters (2) and because they are inconsistent with data supporting a monophyletic origin of the Cnidaria and the Eumetazoa (3, 4). In this paper, I derive a tree that supports a monophyletic origin for the Metazoa using a method less sensitive to artifacts of tree reconstruction. This tree also indicates that the Deuterostomia are monophyletic and that the Annelida–Mollusca lineage is the sister group of an arthropod subgroup. Furthermore, this tree is generally consistent with morphological data and the fossil record.

## MATERIALS AND METHODS

Unrooted trees were determined using the rate-independent technique of evolutionary parsimony, which estimates the length and statistical significance of the central branch of four taxon trees (5). The multi-taxon, unrooted tree was constructed using a modified neighborliness procedure (6). In this method, all four taxon trees that were supported by evolutionary parsimony at a statistically significant level (5%) were calculated. If two taxa were consistently juxtaposed in all of the four taxon trees containing both taxa, then they were combined into a new taxon. When a new taxon was created, all four taxon trees were again calculated and the process was iterated until the tree was determined. At each step, parsimony scores (7), both transition and transversion counts (8), were also computed for the four taxon trees. These solutions coincided with the evolutionary parsimony ones, except when unequal rate effects obtained—i.e., as judged by equal or nearly equal parsimony counts for two alternative trees. In practice, bifurcations could not always be resolved, and it was sometimes necessary to combine more than two taxa into a new taxon. The Mollusca–Brachiopoda–Sipuncula grouping is an example of an unresolved pentafurcation. Branch lengths were determined by operator metrics and represent numbers of transversion

differences rather than substitutions (9). rRNA sequences were used exactly as aligned by Field *et al.* (1) and are available on request from them.

## RESULTS AND DISCUSSION

**The Metazoa Is Monophyletic.** The interpretation of molecular sequence data from rapidly radiating groups, such as the Metazoa, is neither simple nor direct. Even the most extensively employed algorithms, such as the parsimony (7) and distance matrix procedures (10, 11), can fail when rates of nucleotide substitution are unequal in juxtaposed branches and when substitutions are frequent. When this happens, incorrectly reconstructed four taxon trees characteristically have their long branches on one side and their short branches on the other (12).

A newly created algorithm, evolutionary parsimony (5), is less biased by unequal rates (see refs. 13–16) and is used to derive the trees in this paper. This algorithm works by subtracting artifactual background counts (that are generated by peripheral tree branches of unequal length) from parsimony counts to obtain a topology that is not affected by unequal rates.

Field *et al.* (1) concluded that the Metazoa are polyphyletic, using outgroups from diverse taxa, evolving at different characteristic rates. That tree, including the published branch lengths, is labeled “polyphyletic metazoan origins” and is shown at the top right of Fig. 1. In this unrooted representation, the coelenterates and the yeast, maize, and a ciliate are adjacent and both groups are represented by short branches, whereas the coelomates and the cellular slime mold are both represented by branches two to three times longer. This is the type of pattern that would be produced if unequal rates were biasing the tree. This tree contrasts with the traditional tree (labeled “monophyletic metazoan origins”). In the traditional tree, the single-celled organisms, yeast, and cellular slime mold are juxtaposed.

The results of evolutionary parsimony analysis of the sequences used in the Field *et al.* (1) paper are shown in the significance plot at the left of Fig. 1. Neither the monophyletic nor the polyphyletic origins are supported. Since simulations have shown that trees are most reliably calculated when the most slowly evolving tip species are used to construct them, I then used the eight coelomates with the shortest peripheral branches and yeast (the shortest branched single-celled organism) to recalculate the significance of each of the topologies (for taxa, see Fig. 1). With this precaution, the results now support the monophyletic origins tree and lend no support to the alternative polyphyletic origins tree. Patterson (17), in a thoughtful parsimony analysis of the Field *et al.* (1) data, has independently concluded the Metazoa are monophyletic. Some molecular characters supporting the monophyletic tree are listed elsewhere (4).

**Topology and Rooting of the rRNA Eumetazoa Tree.** The topology of the rRNA tree derived by evolutionary parsimony is shown in Fig. 2A with possible roots indicated by letters (see ref. 8). When the cnidarian sequences were used

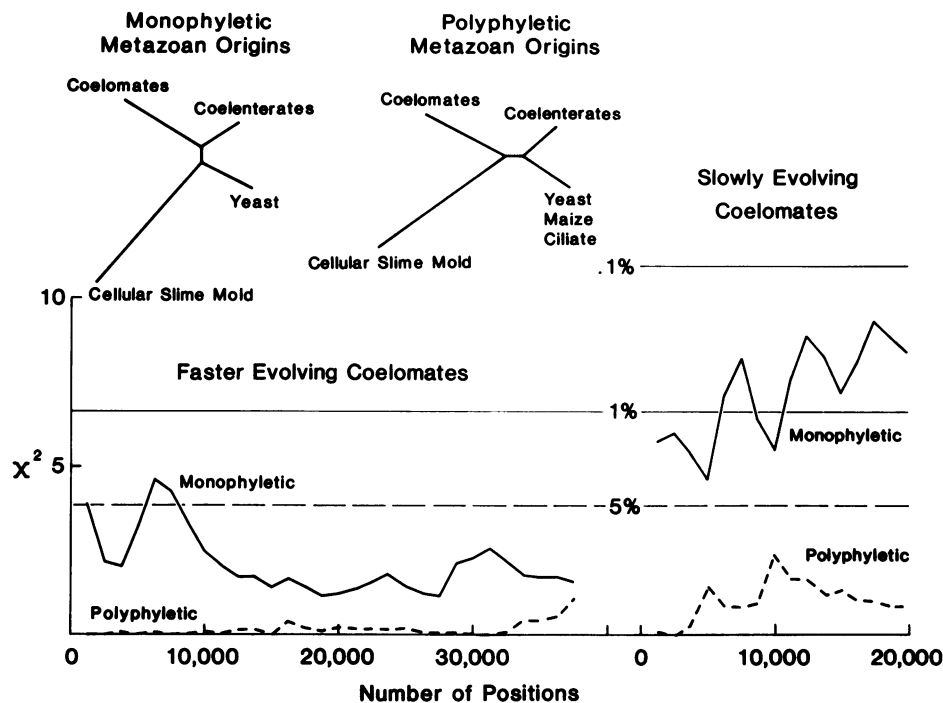


FIG. 1. A test, by evolutionary parsimony, of the polyphyletic and monophyletic origins of the Metazoa. The unrooted trees representing each theory are illustrated at the top. No significant support is provided for either theory when the (faster evolving) coelomates used in ref. 1 (starfish, earthworm, human, brine shrimp, and planarian) are included. This is shown at the lower left. When the most slowly evolving coelomates and the most slowly evolving single-celled organism, yeast, are used to calculate the tree, however, the monophyletic origins tree is supported to the exclusion of the others (including a third alternative, not illustrated) as shown by the significance plot on the right. Correlations among the pooled trees were estimated from the data and are included in the  $\chi^2$  test as described (5). The slowly evolving coelomates (as defined in ref. 1) are Echinodermata (starfish, brittle star, sea urchin, crinoid), Arthropoda (horseshoe crab), Annelida (polychaete), Brachiopoda (lamp shell), and Mollusca (chiton).

as an outgroup, only the roots in Fig. 2A labeled a–e were significantly supported (at the 5% level) in four taxon tests. Each of these is explicitly shown as a separate tree in Fig. 2B. The cnidarians are at the base of the tree in all five figures but are shown in only “a.”

Of the five possible trees, that shown in “a” is closest to the conventional view for the origin of the coelomates. This rooting corresponds approximately to the protostome–deuterostome dichotomy. Traditionally the deuterostomes are thought to represent an early branching of the coelomates

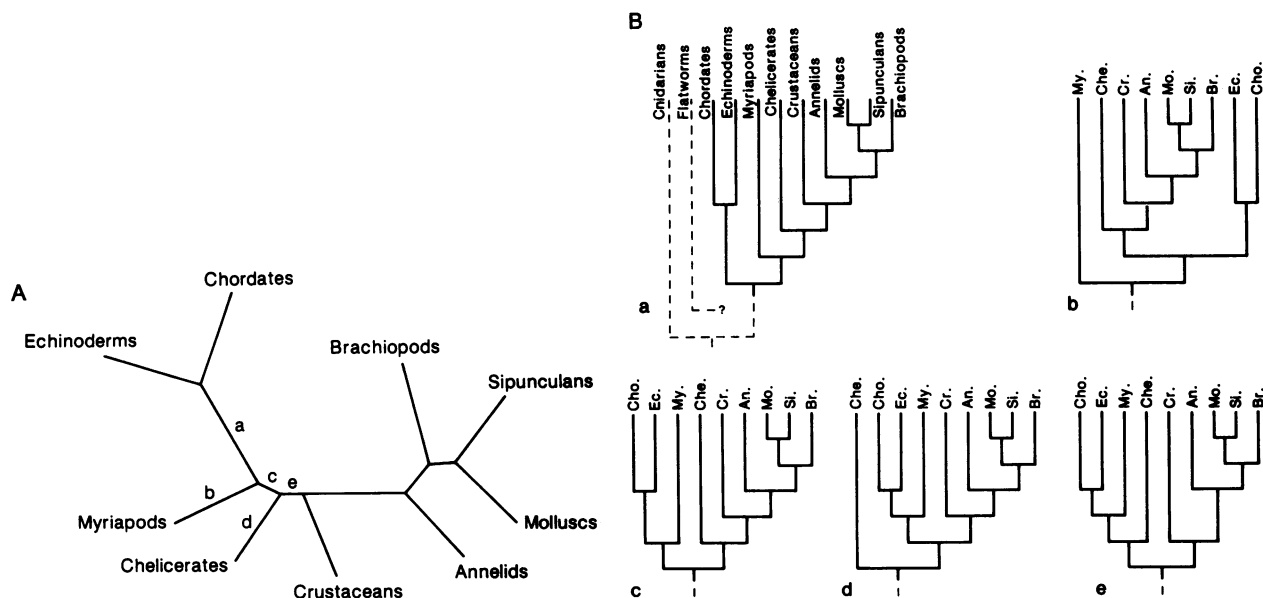


FIG. 2. Schematic illustration of the metazoan 18S rRNA tree. (A) Possible rootings of the unrooted tree are indicated by the letters a–e. Using cnidarians as an outgroup, roots at any other positions were inconsistent with evolutionary parsimony analysis. (B) The rooted trees corresponding to each of these five alternatives are illustrated. The cnidarians are shown as a dashed line in the first drawing of the set. The “?” on the flatworm lineage indicates that it was less useful as an outgroup than were the Cnidarians. (The position shown was significant at only the 7% level.)

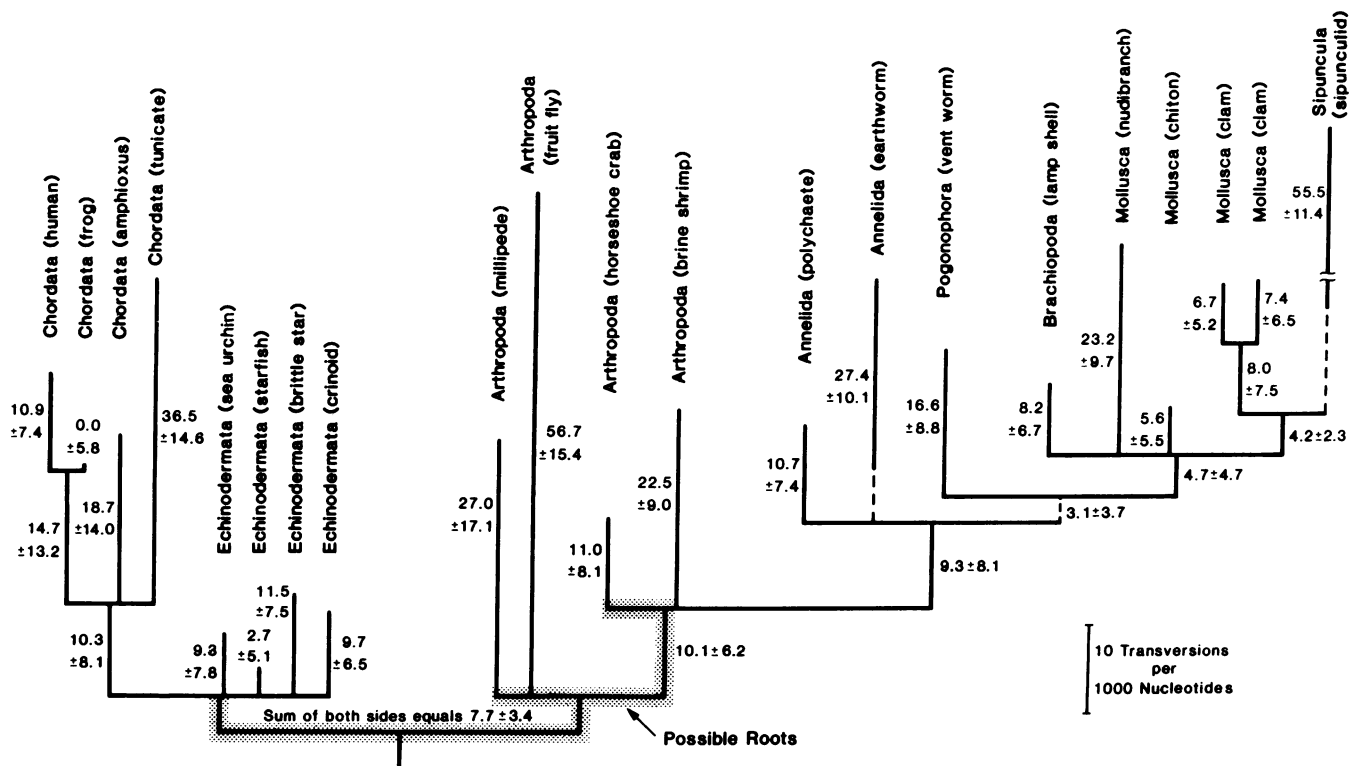


FIG. 3. The rooted metazoan evolutionary tree calculated from 18S rRNA sequences. In the conventional rooting (the "a" root), the left half of the tree includes the Deuterostomia—namely, the Chordata—and Echinodermata. The right half of the tree includes the Arthropoda, Annelida, Pogonophora, Brachiopoda, and Mollusca. Since a unique root was not determined, all rootings that are consistent with the data are shaded. The topology in Fig. 2 differs slightly from that here, since additional, slowly evolving crustaceans (21) were used to position the Annelida–Mollusca branch in Fig. 2. Distances, calculated by operator metrics, are given in transversions per 1000 nucleotides and the  $\pm$  figures represent 1.96 standard deviations. Internal and peripheral branches that could not be positioned within the 5% confidence criterion are indicated by dashed lines.

because they differ from the protostomes so extensively (18, 19). In practice, both differ so greatly from the cnidarians that there are few morphological features that can help one select the root (20). Although the "a" root is used for Fig. 3, the sequence data do not make a significant distinction between any of the five possible rootings. This rooting differs from that presented at the Nobel Symposium on the Hierarchy of Life (22) since additional, slowly evolving crustacean sequences have been used in the present analysis. That preliminary report did not attempt to relate the phylogenetic tree to the distribution of morphological characters and the fossil record.

**Annelid–Mollusc–Brachiopod–Sipunculid–Pogonophoran Relationships.** The branching order of the Annelida–Mollusca portion of the tree is the same for all five allowed roots; hence its interpretation is relatively direct. The upper right half of the tree in Fig. 3 consists of annelids, a pogonophoran, a brachiopod, molluscs, and a sipunculid. Synapomorphies of some traditional morphological characters supporting these relationships are described below.

The Pogonophora occupy a position in the RNA tree intermediate between the Annelida and Mollusca. Morphological analyses have placed them with the deuterostomes, on a separate line intermediate between deuterostomes and protostomes (ref. 23; reviewed in ref. 24), or most recently near the annelids. Synapomorphies of the Pogonophora and the Annelida, in agreement with the RNA tree, include terminal segmented opisthosoma, with setae similar to those of annelids (25); coelomic compartments in each of the body divisions (26); septa between coelomic compartments of the opisthosoma; and a closed vascular system (19).

Relationships within the Mollusca–Brachiopoda–Sipuncula grouping are not uniquely resolved and are presented as

a pentafurcation of the tree. The Brachiopoda have been considered as both candidate deuterostomes and protostomes (19). Possible synapomorphies of the Brachiopoda with the Annelida include chitinous setae in the adults and larvae of most brachiopods and a mouth probably derived from near the blastopore (23). Synapomorphies of the Sipuncula with the Mollusca (and with the Brachiopoda) include an open circulatory system and a nonsegmented body plan. The Annelida, Mollusca, and Sipuncula share a ventral nerve cord that runs the body length, a similar body wall construction, and a trochophore larval stage (19).

Previously, Field *et al.* (1) found a clustering of the Annelida, Pogonophora, Brachiopoda, Mollusca, and Sipuncula. Their tree differs from the one presented here, in the detailed placement of taxa within the group, however. In general, it places short branches with short branches (e.g., a chiton with a polychaete and both with a brachiopod) and long branches with long branches (e.g., sipunculid with nudibranch). This suggests that placement of individual taxa within this tree may have been biased by unequal rates effects (12).

**The Last Common Ancestor of the Annelida–Mollusca Lineage May Have Had a Hemocoel, an Open Circulatory System, and Segmentation.** The unrooted tree shown in Fig. 2 positions the arthropods between the Annelida–Mollusca grouping and the root of the tree. Hence, those character states jointly shared among the myriapods, chelicerates, and crustaceans will determine, most parsimoniously, the state of the last common ancestor of the Annelida–Mollusca grouping.

The nature of the coelom and the segmentation of the ancestral mollusc–annelid are of fundamental theoretical interest (27–29). Given the character states of the Arthropoda (30–32), then most parsimoniously, the ancestral annelid–mollusc is predicted to have a hemocoel with an open

circulatory system and to be segmented. By segmented, I mean that their body walls are divided antero-posteriorly into distinctive sections, with blocs of circular muscles that are usually interrupted at intersegmental junctions and with blocs of longitudinal muscles commonly restricted to one or a few segments (33). According to this interpretation, the eucoelomic annelids would be derived from a segmented ancestor with a hemocoel.

Recent interpretations of the fossil record of the late Proterozoic support this prediction of the RNA tree. As Valentine has noted (34), although horizontal traces are not uncommon, there is little evidence of late Precambrian burrowing (35–37) that is attributable to animals with eucoelomic, annelidan body plans. Hence, he interprets the late Precambrian segmented organisms, best known in South Australia and in roughly correlative strata near the White Sea in the U.S.S.R., as hemocoelic, but not coelomic. Thus there is, indeed, evidence of segmented, hemocoelic forms that antedate the first known traces of annelidan fossils.

**The Deuterostomia Is Monophyletic.** The deuterostomes form a monophyletic taxon in the rRNA-derived tree for all of the five possible rooted trees. Their monophyly is well supported by morphological and developmental data. Synapomorphies of the Deuterostomia include their defining character—namely, the mode of formation of the mouth, as well as the method of formation of the coelom and the mesoderm (19, 23). Field *et al.* (1) did not observe this grouping, perhaps because substitution rates are much faster for the chordates than for the echinoderms.

Unlike the deuterostomes, the relationships of the arthropod groups depend strongly upon the position of the root. Whatever the choice of the root, provided the tree topology is correct, the Arthropoda will be a paraphyletic taxon. In fact, no choice of a root can make them monophyletic. This is unsettling and suggests that revisions based on molecular sequence data and traditional morphological characters may be forthcoming. The phylogenetic coherence of the Arthropoda has been vigorously debated in the past (30, 31), and this is certain to continue in the future.

**Conclusions and Some Qualifications.** Some features of the tree are more strongly supported than are others. For example, the paraphyly of the Arthropoda is probably not as certain as the distances on the tree would indicate. Even though the branch that defines the deuterostomes is  $7.7 \pm 3.4$  transversion units long, its position is referenced to a relatively rapidly evolving sequence (a millipede) and to a rapidly evolving sequence (the fruit fly). Clearly, additional sequences with lower substitution rates will be needed from the Myriapoda to resolve the origins of the deuterostomes. Other conclusions, such as the origin of the annelid–mollusc line within the Arthropoda are probably more secure. In this instance, three additional sequences of slowly evolving crustaceans (21), not in the Field *et al.* data set (1), were used to construct Fig. 2.

In summary, those features of the tree that seem best supported are (i) the Metazoa is a monophyletic group; (ii) the Deuterostomia is monophyletic; (iii) the Annelida–Mollusca line is the sister group of an arthropod subgroup; and (iv) the Annelida–Mollusca lineage is most parsimoniously derived from a segmented, hemocoelic ancestor with an open circulatory system. The analysis of molecular sequences, yet to be obtained, promises that new, deep phylogenetic relationships among the Metazoa may soon be discovered.

I thank Jack Farmer for pointing out the correspondence between this tree and the Precambrian fossil record, Bruce Runnegar for numerous discussions of functional morphology and the fossil record, and Jim Morin and Michael Greenfield for discussions of invertebrate systematics. This research was supported by grants from the National Science Foundation and the Sloan Foundation.

- Field, K. G., Olsen, G. J., Lane, D. J., Giovannoni, S. J., Ghiselin, M. T., Raff, E. C., Pace, N. R. & Raff, R. A. (1988) *Science* **239**, 748–753.
- Nielsen, C. (1989) *Science* **243**, 548.
- Walker, W. F. (1989) *Science* **243**, 548–549.
- Bode, H. R. & Steele, R. E. (1989) *Science* **243**, 549.
- Lake, J. A. (1987) *Mol. Biol. Evol.* **4**, 167–191.
- Fitch, W. (1981) *J. Mol. Evol.* **18**, 30–37.
- Fitch, W. (1977) *Am. Nat.* **111**, 223–257.
- Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. (1982) *J. Mol. Evol.* **18**, 225–239.
- Lake, J. A. (1987) *J. Mol. Evol.* **26**, 59–73.
- Jukes, T. H. & Cantor, C. R. (1969) in *Mammalian Protein Metabolism*, ed. Monro, H. N. (Academic, New York), pp. 21–131.
- Doolittle, R. F. (1986) *Of Urfs and Orfs* (University Science Books, Mill Valley, CA).
- Felsenstein, J. (1978) *Syst. Zool.* **27**, 401–410.
- Cavender, J. (1989) *Mol. Biol. Evol.* **6**, 301–316.
- Holmquist, R., Miyamoto, M. M. & Goodman, M. (1988) *Mol. Biol. Evol.* **5**, 201–216.
- Felsenstein, J. (1988) *Nature (London)* **335**, 118.
- Penny, D. (1988) *Nature (London)* **331**, 111–112.
- Patterson, C. (1989) in *The Hierarchy of Life*, proceedings from Nobel Symposium 70, eds. Fernholm, B., Bremer, K. & Jornvall, H. (Elsevier, Amsterdam), pp. 471–488.
- Hyman, L. H. (1951) *The Invertebrates: Echinodermata* (McGraw-Hill, New York), Vol. 4.
- Barnes, R. D. (1987) *Invertebrate Zoology* (Saunders, Philadelphia).
- Bergstrom, J. (1986) *Zool. Scr.* **15**, 189–200.
- Abele, L. G., Kim, W. & Felgenhauer, B. E. (1989) *Mol. Biol. Evol.* **6**, 689–691.
- Lake, J. (1989) in *The Hierarchy of Life*, proceedings from Nobel Symposium 70, eds. Fernholm, B., Bremer, K. & Jornvall, H. (Elsevier, Amsterdam), pp. 273–278.
- Hyman, L. H. (1951) *The Invertebrates: The Smaller Coelomate Groups* (McGraw-Hill, New York), Vol. 5.
- Cutler, E. B. (1975) *Syst. Zool.* **24**, 512–513.
- George, J. D. & Southward, E. C. (1973) *J. Mar. Biol. Assoc. U.K.* **53**, 403–424.
- Flügel, H. J. & Langhoff, I. (1983) *Sarsia* **68**, 131–138.
- Runnegar, B. & Pojeta, J. (1985) in *The Mollusca* (Academic, New York), Vol. 10, pp. 1–57.
- Vagvolgyi, J. (1967) *Syst. Zool.* **16**, 153–168.
- Wingstrand, K. G. (1985) *Galathea Report* (Brill/Scandinavian Science Press, Leiden and Copenhagen).
- Boudreaux, H. B. (1979) *Arthropod Phylogeny with Special Reference to Insects* (Wiley, New York).
- Manton, S. M. (1977) *The Arthropoda* (Clarendon, Oxford, U.K.).
- Tiegs, O. W. & Manton, S. M. (1958) *Biol. Rev.* **33**, 255–337.
- Clark, R. B. (1964) *Dynamics in Metazoan Evolution* (Clarendon, Oxford, U.K.).
- Valentine, J. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2272–2275.
- Fedonkin, M. A. (1985) *Philos. Trans. R. Soc. London Ser. B* **311**, 27–45.
- Fedonkin, M. A. (1985) in *Vendian System*, eds. Sokolov, B. S. & Fedonkin, M. A. (Nauka, Moscow).
- Cope, J. C. W. (1985) *Philos. Trans. R. Soc. London Ser. B* **311**, 45.