

Problematica old and new

Ronald A. Jenner^{1,*} and D. Timothy J. Littlewood²

¹*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK*

²*Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK*

Problematica are taxa that defy robust phylogenetic placement. Traditionally the term was restricted to fossil forms, but it is clear that extant taxa may be just as difficult to place, whether using morphological or molecular (nucleotide, gene or genomic) markers for phylogeny reconstruction. We discuss the kinds and causes of Problematica within the Metazoa, as well as criteria for their recognition and possible solutions. The inclusive set of Problematica changes depending upon the nature and quality of (homologous) data available, the methods of phylogeny reconstruction and the sister taxa inferred by their placement or displacement. We address Problematica in the context of pre-cladistic phylogenetics, numerical morphological cladistics and molecular phylogenetics, and focus on general biological and methodological implications of Problematica, rather than presenting a review of individual taxa. Rather than excluding Problematica from phylogeny reconstruction, as has often been preferred, we conclude that the study of Problematica is crucial for both the resolution of metazoan phylogeny and the proper inference of body plan evolution.

Keywords: Problematica; Metazoa; Bilateria; phylogenomics; animal evolution; character loss

1. PROGRESS AND REMAINING CONTROVERSY

In the proceedings of a previous international symposium held in London two decades ago, Barnes (1985) summed up progress in metazoan evolution in six ‘phylogenetic assumptions’, only one of which was strictly phylogenetic (besides monophyly of Metazoa)—protostomes and deuterostomes ‘continue to be recognized as major lines of animal evolution’. From that starting point, progress in the field can only be labelled as remarkable (Halanych 2004). However, it is difficult to present an unambiguous consensus view of high-level metazoan phylogeny. First, metazoan phylogenetics is thriving, for example with several large-scale projects underway in both Europe and USA, and promoting an ongoing flux of ideas. Preliminary results of new studies that were presented during the current symposium provide a taster of possible new insights, like a potential refutation of the long-standing dichotomy between protostomes and deuterostomes. Second, workers may favour either particular kinds of phylogenetic evidence, and/or modes of phylogenetic analysis (Jenner & Scholtz 2005). The conservative consensus that forms the basis of our discussion (figure 1) will probably become obsolete in the near future, and it holds the middle ground between recent more extreme views that either report only ‘a few minor controversies’ left to clean up (Eernisse & Peterson 2004, p. 204) or depict a virtually unresolved polytomy (Philippe & Telford 2006).

In this paper, we examine progress in high-level animal phylogenetics by focusing on the most challenging cases: the Problematica. We use the term to label extant or fossil taxa that defy robust, unambiguous

phylogenetic placement. Agreeing with Conway Morris (1991, p. 23), we regard Problematica principally as ‘a problem for biologists, not biology’. We review the methodological and biological causes of Problematica in the context of high-level metazoan phylogeny and provide possible strategies for dealing with Problematica. We discuss fossil and extant Problematica from the perspectives of morphological and molecular phylogenetics. A summary of attempts to grapple with Problematica provides valuable insights into the relative abilities of different kinds of data and phylogenetic methods to deal with some of the most challenging problems in all of systematics.

Space limitations dictate three important caveats. First, we address individual Problematica only by way of selected examples. Second, all factors that can influence the outcome of a phylogenetic analysis are relevant for a discussion of Problematica. Evidently, we can do no justice to all relevant data and ideas, however, the general issues that we address apply to all known Problematica in the Metazoa. Third, in line with our own area of least ignorance, we focus on extant Problematica.

2. PROBLEMATICA: CAUSES AND RECOGNITION CRITERIA

Problematica confront phylogeneticists with all the problems that can beset phylogenetic analysis. Problematica arise when we lack unambiguous phylogenetic signal that can relate them to other taxa. In many cases, this is simply the result of not (yet) having enough knowledge of a taxon. This is certainly the case for many fossil Problematica with unfavourable preservation, and it is likewise a factor for many extant high-level metazoan Problematica, for which our knowledge of non-phenotypic characters is still rudimentary.

* Author for correspondence (rj223@bath.ac.uk).

One contribution of 17 to a Discussion Meeting Issue ‘Evolution of the animals: a Linnean tercentenary celebration’.

| | no. of species | NCBI – GenBank | | | | | |
|------------------------|----------------|--------------------|-------------|------------|-----------|---------|-----|
| | | links ^a | nucleotides | | | prots | mt |
| | | | core | ESTs | GSS | | |
| Choanoflagellata | 140 | 39 | 7660 | 7412 | 0 | 129 | 1 |
| Calcarea | 1000 | 123 | 256 | 258 | 0 | 78 | 0 |
| Demospongiae | 9000 | 794 | 2007 | 11 176 | 8 | 998 | 5 |
| Hexactinellida | 10 000 | 26 | 61 | 0 | 0 | 45 | 0 |
| Cnidaria | 9000 | 2139 | 8950 | 230 493 | 2 | 4915 | 25 |
| Ctenophora | 150 | 73 | 112 | 962 | 0 | 61 | 0 |
| Placozoa | 1 | 88 | 340 | 58 | 0 | 206 | 4 |
| Myxozoa | 1300 | 337 | 608 | 765 | 0 | 30 | 0 |
| Acoelomorpha | 329 | 115 | 207 | 846 | 0 | 31 | 0 |
| Platyhelminthes | 55 000 | 3219 | 119 736 | 397 457 | 62 349 | 16 317 | 22 |
| Sipunculida | 320 | 95 | 93 | 0 | 0 | 809 | 0 |
| Nemertea | 7500 | 197 | 401 | 0 | 0 | 189 | 0 |
| Gastrotricha | 450 | 91 | 68 | 0 | 0 | 3 | 0 |
| Cycliophora | 2 | 8 | 292 | 0 | 0 | 247 | 0 |
| Bryozoa | 4500 | 230 | 930 | 0 | 2 | 432 | 1 |
| Entoprocta | 150 | 12 | 16 | 0 | 0 | 2 | 0 |
| Micrognathozoa | 1 | 3 | 119 | 0 | 0 | 2 | 0 |
| Gnathostomulida | 80 | 49 | 111 | 0 | 0 | 25 | 0 |
| Rotifera | 1800 | 203 | 1367 | 290 | 1 | 692 | 0 |
| Acanthocephala | 1000 | 122 | 375 | 0 | 0 | 212 | 1 |
| Mollusca | 70 000 | 8688 | 248 158 | 308 707 | 1464 | 25 379 | 37 |
| Annelida | 15 000 | 1897 | 30 782 | 8462 | 0 | 4852 | 6 |
| Myzostomida | 170 | 55 | 116 | 0 | 0 | 37 | 0 |
| Lobatocerebromorpha | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Brachiopoda | 335 | 180 | 225 | 0 | 0 | 919 | 3 |
| Phoronida | 20 | 11 | 23 | 0 | 0 | 19 | 0 |
| Dicyemida (Mesozoa) | 78 | 13 | 38 | 0 | 0 | 17 | 0 |
| Orthonectida (Mesozoa) | 24 | 3 | 2 | 0 | 0 | 0 | 0 |
| Arthropoda | 1 100 200 | 53 530 | 1 742 231 | 2 215 728 | 376 918 | 324 214 | 129 |
| Pentastomida | 100 | 12 | 16 | 0 | 0 | 130 | 1 |
| Tardigrada | 980 | 153 | 651 | 5235 | 1063 | 66 | 0 |
| Onychophora | 165 | 100 | 238 | 0 | 0 | 211 | 1 |
| Nematoda | >25 000 | 3186 | 83 478 | 774 910 | 683 724 | 84 085 | 22 |
| Nematomorpha | 320 | 38 | 24 | 0 | 0 | 2 | 0 |
| Kinorhyncha | 150 | 14 | 10 | 0 | 0 | 1 | 0 |
| Loricifera | 22 | 4 | 6 | 0 | 0 | 0 | 0 |
| Priapulida | 18 | 13 | 85 | 478 | 0 | 829 | 1 |
| Chaetognatha | 100 | 63 | 333 | 226 | 0 | 246 | 2 |
| Xenoturbellida | 2 | 4 | 18 | 14 162 | 0 | 40 | 1 |
| Echinodermata | 7000 | 915 | 504 455 | 343 299 | 86 326 | 50 413 | 16 |
| Hemichordata | 106 | 28 | 149 | 11 654 | 0 | 176 | 2 |
| Cephalochordata | 29 | 17 | 1054 | 278 076 | 66 720 | 1255 | 6 |
| Tunicata | 2566 | 265 | 83 949 | 943 882 | 1898 | 12 960 | 6 |
| Vertebrata (non-human) | 58 389 | 36 905 | 18 603 467 | 14 937 969 | 6 687 201 | 161 826 | 786 |
| Vertebrata (human) | 1 | — | 3 015 588 | 8 216 212 | 1 115 905 | 397 886 | 2 |

^a refers to taxon links
total number of metazoan species represented on GenBank = 59 748
total number of estimated species > 1.38 million

Figure 1. A conservative estimate of metazoan phylogeny modified after Philippe & Telford (2006), with indications of estimated number of known species and, from the NCBI (GenBank) databases: the number of unique taxon links (an overestimate of the real number of species for which sequence data are known since separate links for different taxonomic levels of the same species are included), the number of nucleotide sequences (core), the number of nucleotide sequences from large-scale expressed sequence tag (ESTs) or genome (GSS) projects, the number of protein (prots) sequences and the number of mitochondrial genomes (mt).

Biological reasons for the absence of sufficient phylogenetic signal can be grouped into three categories.

- (i) Not enough phylogenetic signal has evolved,
- (ii) phylogenetic signal is lost through extinction, and
- (iii) phylogenetic signal is lost or obscured by the evolution of non-phylogenetic signal.

In the first case, if lineage splitting events succeed each other rapidly, there may not be enough time for distinctive features to evolve, which can be used to group descendant species. Although the length of the fuse of the Cambrian explosion is still debated, this has long been considered a distinct possibility for the divergence of the animal phyla.

Extinction may exacerbate the problem of inferring clades on the basis of homoplasy, or erase phylogenetic signal altogether if the organisms are not discovered. For example, reconstruction of the panarthropod stem group revealed that the subventral mouth shared by extant arthropods and onychophorans has evolved convergently (Eriksson & Budd 2000). As is well known, fossils can contribute important phylogenetic signal (Cobbett *et al.* 2007), and in view of the considerable differences between the body plans of extant phyla, extinction must have removed substantial amounts of morphological phylogenetic signal that can only be retrieved by the study of fossils.

The third category groups several causes related to evolutionary change that can erode or obscure phylogenetic signal with the same effects for phylogenetic analysis as extinction of taxa, even when all relevant taxa are included into the analysis. This is especially important when inferring phylogenies with short stems and long terminal branches (Rokas & Carroll 2006), of which the metazoan phylogeny is a prime example. Firstly, if newly evolved lineages have not yet evolved complete intrinsic isolating mechanisms, extensive introgressive hybridization may occur, even of morphologically distinct species (Wiens *et al.* 2006). Although extensive gene exchange between morphologically distinct species is considered rare (Coyne & Orr 2004), this could scramble any original phylogenetic signal (Clarke *et al.* 1996; Chan & Levin 2005). Causes in this category also relate to the power of natural selection or shared constraints to produce convergent evolution, and parallelisms (non-random non-phylogenetic signal) that may lead to the false inference of monophyletic taxa. This can be an important problem for both morphological and molecular phylogenetic analyses (Waagele & Mayer 2007). It also captures problems caused by the evolution of random noise (undirected homoplasy), such as may arise purely by chance, e.g. as with substitutional saturation of sequences.

These causes can affect phylogenetic analyses of both fossil and extant taxa, at any taxonomic level and independent of the type of evidence used. Difficulties generally become greater with increasing age of the divergence events we attempt to reconstruct, and all causes mentioned above have probably confounded attempts to place particular Problematica in the tree of the Metazoa. In the following sections, we pay more detailed attention to specific causes that are of relevance for certain Problematica.

Several criteria can be used to recognize Problematica.

- (i) Number of alternative sister group hypotheses,
- (ii) phylogenetic spread and hierarchical range of alternative sister group hypotheses,
- (iii) controversial homology assessments,
- (iv) absence of phylogenetically informative characters, and
- (v) exclusion from phylogenetic analyses based on explicit datasets.

The first two criteria are the most straightforward for recognizing Problematica, partly because they can only

be applied after phylogenetic analyses. Classic Problematica, such as Chaetognatha, Ectoprocta, and Pogonophora, have long exhibited both a large number of alternative sister group hypotheses and a large phylogenetic spread among these alternatives (covering both Protostomia and Deuterostomia). The phylogenetic spread of alternative hypotheses is positively related to the hierarchical depth across which the alternatives may be distributed. For example, the placement of Pentastomida is problematic only within the Panarthropoda, with a position either within Crustacea or in the arthropod stem group as the two main contending hypotheses (Waloszek *et al.* 2005). In contrast, vetulicolians are problematic on a much larger scale, across a wide phylogenetic spread (Bilateria) and a large hierarchical depth (ranging from being attributed to a separate 'phylum-level' clade to belonging to a subtaxon of Tunicata; Aldridge *et al.* 2007). However, the problematic status of other taxa emerges only after careful study of potential phylogenetic evidence. Tardigrada, for example, are generally considered to be close relatives of the onychophorans and arthropods, together comprising Panarthropoda. In contrast to the latter two taxa, which are often grouped together to the exclusion of tardigrades, tardigrades lack an ostiate heart and nephridia. However, whether these absences are secondary losses due to miniaturization, or primary absences is unclear. Consequently, correct determination of the phylogenetic significance of these characters depends on whether tardigrades are primitively small bodied or secondarily miniaturized. Our current understanding does not clearly favour either hypothesis, leaving panarthropod relationships unresolved. This example also shows that the distinction between Problematica and non-Problematica is not sharp. Other taxa are problematic owing to the lack of or insufficient study of informative characters. Myxozoa, for example, are probably derived cnidarians (Jiménez-Guri *et al.* 2007) that share so few characters with their closest non-parasitic relatives that most textbooks did not even include them in the Metazoa until very recently. Lacking detailed knowledge may also cause Problematica to be excluded from phylogenetic discussions. Species such as not only *Jennaria pulchra*, the lobatocerebrids, *Xenoturbella bocki* (until recently), *Buddenbrockia* and myxozoans but also myzostomids and pentastomids are frequently excluded from morphological phylogenetic analyses. This is not because their phylogenetic position is so well understood.

3. FOSSIL PROBLEMATICA

(a) *The vagaries of preservation, typological thinking and model choice*

All the difficulties that beset phylogenetic analyses of extant taxa also play a role in the systematization of fossils. With fossils, however, several additional factors can cause problems, of which we think three are of particular importance. First, preservational artefacts can lead to formidable problems of interpretation. Although the majority of fossils can be related to extant body plans without much difficulty, 'unusual objects do occur in rocks' (Yochelson 1991, p. 288).

Problematica are particularly common from the fossil record of the Late Neoproterozoic and the earliest Phanerozoic (approx. 575–500 Myr ago) and it is especially these forms that may provide unique clues to the origin and diversification of early animal body plans. Yet, many important taxa found in this time interval defy unambiguous interpretation owing to the limits of preservation, and taphonomic changes of the organism and surrounding sediment. This is clearly illustrated in recent debates over the putative Precambrian animal *Vernanimalcula* (a coelomate bilaterian?), the Cambrian animal *Odontogriphus* (segmented?) and the oldest putative metazoan eggs and embryos (animals or bacteria?) and in the continuing debate about the Ediacaran biota (Dzik 2003; Fedonkin 2003; Bengtson & Budd 2004; Chen *et al.* 2004; Narbonne 2005; Butterfield 2006; Caron *et al.* 2006; Bailey *et al.* 2007; Donoghue 2007).

Budd & Jensen (2000) nominated typological thinking as another factor that may hinder the phylogenetic systematization of fossils, especially in the context of extant taxa. Through a misguided emphasis on differences, fossils have automatically been labelled Problematica if their body plan does not exactly conform to that of a living phylum (see also Briggs *et al.* 1992). Such reasoning is incompatible with proper phylogenetic logic, but it is nevertheless prevalent (Jenner 2006).

A third factor that inescapably affects thinking about fossil Problematica is that fossils can only be interpreted in the light of our knowledge of living species. Consequently, disagreements about the phylogenetic placement of fossil Problematica frequently hinge upon the use of different living species as models, as illustrated by the vetulicolians (Aldridge *et al.* 2007).

(b) Solving fossil Problematica: stem groups, new fossil and new techniques

Yochelson (1991, p. 289) remarked that he could only offer 'a few platitudes' about how 'to do' fossil Problematica. We hope the following suggestions are helpful. In essence, fossils should be treated like any other living taxon. Attempts to systematize fossils will lead to the establishment of stem groups (Budd & Jensen 2000; Conway Morris 2000; Budd 2001, 2003). Although differences between fossils and extant taxa should not be ignored, they should not be interpreted typologically as evidence against affinities (Budd & Jensen 2000; Jenner 2006). Putative stem group taxa are expected to exhibit some but not all of the diagnostic characters of crown groups, and by creating paraphyletic series of stem taxa, we can illustrate the orderly sequential evolution of body plans. This may not be easy of course. If crucial information is not preserved, a fossil may not be reliably placed. In such cases, unless new fossils are found or new techniques reveal new information, ambiguity will endure.

The main reason why fossil Problematica occur frequently in the Late Neoproterozoic and Early Phanerozoic is extinction. These fossils document the early evolution of animal body plans. The older the fossils are, the more they are expected to fall outside the limits of extant body plans (Budd 2003; Valentine 2004). Unless

body plan evolution takes large leaps, failure to systematize fossil Problematica is chiefly the result of not (yet) knowing related taxa that can bridge their morphology with those of the crown group. Hence, most progress is made with fossil Problematica when new specimens are found. Better-preserved fossils and forms with novel character combinations address the problems of taxon and character matrix completeness, allowing unknowns to be substituted with characters. However, this approach relies on much fieldwork and a great deal of luck.

Palaeontological and analytical techniques are constantly being developed that present ways of discerning new characters, or of better resolving existing ones, and of handling existing data. For example, the three-dimensional reconstruction of fossil forms from thin serial sections has achieved remarkable levels of resolution, thanks to refinements in microscopy and computer rendering. This has provided valuable phylogenetic information for a host of taxa, ranging across the Bilateria (Sutton *et al.* 2001a,b, 2005a–c; Thomson *et al.* 2003). X-ray tomographic microscopy and Raman spectroscopy combined with confocal laser scanning microscopy have also yielded images and insights into the biomolecular nature of fossils with unrivalled resolution (Schopf & Kudryavtsev 2005; Donoghue *et al.* 2006; Chen *et al.* 2007).

Other advances will come from improvements in methods of phylogeny reconstruction. Model-based methods of analysis have proven their worth with molecular data, particularly in dealing with long-branch problems in phylogenetic reconstruction. Such methods, although still in their infancy, are now available for the analysis of morphological and fossil data as well (Lewis 2001). This promises the chance to include incomplete taxa, such as fossil Problematica, with morphological and even molecular data from extant taxa using maximum likelihood or Bayesian techniques (Wiens *et al.* 2005b), while at the same time parsimony-based methods are refined to be able to deal efficiently with large amounts of diverse phylogenetic evidence (Wheeler *et al.* 2006).

4. EXTANT PROBLEMATICA

(a) An apparent paradox: weak molecular signal and large amounts of morphological evolution

It is not surprising that Problematica are encountered when metazoan phylogeny is analysed on the basis of extant taxa. First, any comparison between two extant species belonging to different phyla has to bridge in the order of one billion years of independent evolution. This is ample time to erase signs of ancestry, either through extensive modification or loss of characters, and for convergent evolution to obscure phylogenetic signal. It may thus be unsurprising that sessile taxa (ectoprocts, brachiopods, phoronids), very small (possibly miniaturized) taxa (tardigrades, placozoans, *Lobatocerebrum*) and parasitic taxa (pentastomids, myxozoans) have been particularly prominent Problematica. Another consequence is that molecular phylogenies of the Metazoa bear the typical signature of short stems and long terminal branches, providing ample opportunity for long branch attraction (Waegele & Mayer 2007). This has been a problem for the

placement of several taxa, ranging from myxozoans to acoels (Philippe *et al.* 2007). Second, the major metazoan lineages may have radiated very rapidly, potentially allowing for very little phylogenetic signal to evolve. Although it remains disputed whether lack of resolution is a convincing signature of closely spaced cladogenetic events (Giribet 2002; Rokas *et al.* 2005; Rokas & Carroll 2006; Baurain *et al.* 2007; Whitfield & Lockhart 2007), if current molecular clock estimates of metazoan divergence times are approximately accurate (Peterson *et al.* 2004, 2005), the fact remains that the major metazoan lineages diverged over a time span that is significantly shorter than the subsequent independent history of modern phyla (including their stem groups). The appearance in the fossil record of a variety of crown phyla with their distinctive body plans as early as the Cambrian (Budd 2003; Valentine 2004) implies that important morphological traces of ancestry were probably already erased early in metazoan history.

Intriguingly, the relative branch lengths of morphological metazoan phylogenies seemingly contradict the absence of sufficient phylogenetic signal. These typically show a much smaller discrepancy between the length of stems and terminal branches, or even the opposite pattern of relatively longer stems and shorter tips (Zrzavy *et al.* 1998, 2001; Nielsen 2001; Peterson & Eernisse 2001; Brusca & Brusca 2003; Zrzavy 2003). Large amounts of body plan evolution are commonly inferred along almost all stems. This raises interesting issues about the relationship between genetic and phenotypic evolution which we cannot address here. What is pertinent here is the large amount of body plan evolution inferred across a relatively small number of speciation events. For example, depending on the precise topology of the tree, possibly just six or seven nodes separate the body plan of the last common ancestor shared by (at least some) sponges and the remaining animals, and the last common ancestor of the chordates! Unless half a dozen speciation events are really all that are required to evolve from a sponge grade organization to that of a protochordate, we must be missing something. That something is fossils.

Recent studies of the fossil record have yielded important insights that may help explain why extant Problematica are to be expected. First, Wagner (Wagner 2000, 2001b; Wagner *et al.* 2006) drew the important conclusion that during evolutionary history taxa tend to exhaust their character state spaces. This means that as clades age, homoplasies increase in frequency. Not surprisingly, homoplasies are common between the major lineages of animals (Valentine 2004). Our estimates of homoplasy based on morphological phylogenetic studies are probably underestimates, given a widespread problem of character coding (Jenner 2004b).

Distressingly, Wagner (2001b) noted that the inclusion of fossils into a phylogenetic analysis of extant species could reveal a significant amount of previously hidden character change along branches subtending extant taxa. This positive correlation between the amount of character change that is discovered and the number of taxa included is well known by molecular systematists, and is known as the node density effect. However, its effect for morphological phylogenetics and inference of body plan evolution has barely been

acknowledged (Jenner & Wills 2007). Hence, the inclusion of even incomplete fossil taxa has the potential to reveal that synapomorphies of extant taxa may in fact be homoplasies or symplesiomorphies, and their inclusion can improve accuracy of the phylogenetic relationships inferred between living taxa (Wiens *et al.* 2005a). The reconstruction of stem groups is crucial for a complete picture of body plan evolution, and there is ample evidence that phylogenetic inferences based on extant taxa can be misled; for arthropod examples see Eriksson & Budd (2000) and Budd (2001). The amount of character evolution that is missed by a focus on extant taxa is increasingly illustrated by studies that show that rates of morphological character change may be the highest in the early history of a clade, which may go hand in hand both with the general early establishment of morphological disparity in the history of large clades and indications that morphological transformations had larger step sizes early in a clade's history (Valentine 2004; Ruta *et al.* 2006; Erwin 2007). In combination, these insights suggest that by focusing on living taxa only we are missing a lot of character evolution, the recognition of which is crucial to prevent clades being based on homoplasies or symplesiomorphies.

(b) *From the unequal eye to morphological cladistics*

To see all things with equal eye is not within our power: humans, and especially human narrators, always look upon the world with an unequal eye
(O'Hara 1992, p. 140).

Before computers came to assist phylogenetic analysis, Problematica were an inescapable by-product of phylogenetic inference. Without the help of a computer it is impossible to achieve a balanced and unbiased evaluation of large amounts of comparative data for more than a few taxa. Emphasis on different aspects of available evidence as well as the lack of a uniform phylogenetic methodology fostered disagreement between workers. Consequently, from the beginning of our discipline a researcher's central insights were not uncommonly labelled another's 'fata morgana' (Hubrecht 1887, p. 641), and the coordinating theme of one school of zoological thought would deserve to be 'dead and buried' in the opinion of proponents of another (Hyman 1959, p. 750).

The widespread adoption of cladistic reasoning in the second half of the twentieth century increased the promise of reaching a general consensus on metazoan phylogeny. Yet, without the help of computers progress was once again foiled as the huge amounts of conflicting evidence allowed many mutually exclusive conclusions. The computer-assisted morphological cladistic analyses of metazoan phylogeny published over the last decade greatly advanced the objectivity, explicitness and testability of phylogenetic hypotheses. In this period, the field progressed significantly beyond the traditional textbook trees (Adoutte *et al.* 2000), but perhaps the most important insight of this era of fruitful debate was discovering exactly how problematic many taxa and clades were. As reviewed elsewhere (Jenner 2004a,b), differences in the construction of data

matrices, including different strategies of character selection, character coding and scoring, and taxon selection resulted in many incompatible phylogenies. Taxa such as Chaetognatha and Ectoprocta behave like phylogenetic renegades, residing in as many different clades as there are studies, and although other aspects of the phylogenetic backbone seemed more secure (monophyly of Protostomia, Spiralia), total agreement between analyses is absent. Evidently, the phylogenetic signal residing in morphology needs to be supplemented with molecular evidence.

(c) Old Problematica solved and new Problematica revealed with molecules

A new phylogenetic synthesis for the Metazoa (Halanych 2004; figure 1) has emerged largely on the basis of molecular evidence. The backbone of this phylogeny is based on rDNA sequences (18S and 28S), and despite challenges (Rogozin *et al.* 2007) its major aspects are confirmed by increasingly sophisticated phylogenomic analyses based on larger amounts of data, and employing improved model-based analytical methods (Philippe *et al.* 2005; Baurain *et al.* 2007; Irimia *et al.* 2007). However, since for most taxa phylogenomic data are not yet available, workers still rely heavily on sequence data from a few popular loci. Nevertheless, molecular evidence has provisionally solved a number of controversies. For example, Echiura, Pogonophora and Vestimentifera, and possibly Sipunculida and Myzostomida as well, are all parts of Annelida (McHugh 1997; Rouse 2001; Bleidorn *et al.* 2007; Struck *et al.* 2007), and xenoturbellids are deuterostomes with a probable sister group relationship with the Ambulacraria (Bourlat *et al.* 2006). Molecular data have also helped to constrain the number of existing hypotheses for the perennially problematic chaetognaths to being sister group to either Protostomia or Lophotrochozoa (Marlétaz *et al.* 2006; Matus *et al.* 2006). This resolution is intriguing. Peterson & Eernisse (2001) and Eernisse & Peterson (2004) hypothesized that chaetognaths may retain various bilaterian plesiomorphies. It could be that morphological phylogenetic analyses misplaced them based on symplesiomorphies that were erroneously interpreted as synapomorphies. For example, an emphasis on mouth position and cuticle composition would support ecdysozoan affinities, while embryological features are more in line with a deuterostome affinity.

However, in spite of increasingly broad taxon sampling, in places 'overall resolution remains discouraging' (Rousset *et al.* 2007, p. 54), especially within the Lophotrochozoa. Consequently, molecular data have not yet provided a reliable picture of lophotrochozoan phylogeny and the relationships of non-bilaterians in particular.

The combination of molecular and morphological evidence into single analyses has also yielded interesting insights, but differences in the methods and datasets used make it very difficult to judge the relative merits of different studies (Eernisse & Peterson 2004; Glenner *et al.* 2004). Also, considered in isolation, analyses based on morphology versus molecules show varying degrees of conflict depending on the datasets. Some of the more conspicuous differences concern

Gastrotricha, placed in Ecdysozoa (morphology) or Lophotrochozoa (molecules), the monophyly or paraphyly of Cycloneuralia (Nematoida + Scalidophora; morphology and molecules, respectively) and the lack of robust molecular support for the monophyly of morphologically widely accepted taxa such as Mollusca and Gastrotricha and some of their major subtaxa (Giribet *et al.* 2006; Todaro *et al.* 2006).

(d) Guidelines for future progress in metazoan phylogeny

A large literature exists on troubleshooting molecular systematics. Some excellent recent reviews include Gribaldo & Philippe (2002), Sanderson & Shaffer (2002), Delsuc *et al.* (2005), Philippe *et al.* (2005), Boore (2006), Philippe & Telford (2006), Rokas & Carroll (2006), Wiens (2006) and Whitfield & Lockhart (2007). We extract a number of guidelines to be kept in mind to ensure continued progress in understanding rather than mere stochastic change of opinions.

It makes increasingly little sense to label and target particular taxa as Problematica. Their correct placement in the tree of the Metazoa is unlikely to arise through the isolated accumulation of data. Several factors need to be balanced to produce a good phylogenetic analysis—number of taxa, number of characters, quality of data and quality of analytical models. The interaction between these variables determines whether the results of a phylogenetic analysis are informative and reliable, or suffer from stochastic or systematic error. Stochastic error arises as chance correspondences overwhelm true phylogenetic signal when there are not enough informative data. Systematic error results when reconstruction methods are inaccurate and are unable to deal with bias in the raw data, which can have several causes (Philippe & Telford 2006). The common problem of long branch attraction (Anderson & Swofford 2004; Waagele & Mayer 2007) can be a result of both stochastic and systematic errors.

The problem is that in trying to avoid stochastic error by increasing the number of characters in a dataset, the chances of systematic error may increase when the number of taxa sampled is too small or the amount of data across taxa is uneven. So far the molecular data generated for different phyla are wildly uneven (figure 1) owing to the bias towards key taxa that are important as model organisms, or organisms of biomedical or economic importance, or simply because they are the easiest to collect. To avoid systematic error it is therefore important to strive for a better balance in the number of taxa and characters (Philippe & Telford 2006).

(i) Avoiding stochastic error

—Increase the number of characters. In molecular systematics this is the main rationale for doing phylogenomics, based on large amounts of data generated through genome projects, EST projects, or large-scale projects targeting particular genes with degenerate primers (Delsuc *et al.* 2005; Philippe *et al.* 2005; Philippe & Telford 2006; Baurain *et al.* 2007).

(ii) *Avoiding systematic error*

- Develop better models of sequence evolution which can deal with problematic data to prevent systematic error (Delsuc *et al.* 2005; Philippe *et al.* 2005; Rokas & Carroll 2006; Baurain *et al.* 2007),
- move towards less homoplastic characters such as rare genomic changes (Boore 2006; Rokas & Carroll 2006),
- sample more taxa, including at least several species representing a high-level metazoan taxon, which may do more to prevent systematic error than aiming to have whole genome sequences for fewer taxa (Hillis *et al.* 2003),
- recognize and remove problematic data, such as fast evolving taxa or characters, or characters the evolution of which violates phylogenetic model assumptions (Lecointre & Deleporte 2004; Delsuc *et al.* 2005; Philippe *et al.* 2005).

(iii) *Other considerations*

- Care should be taken not to be misled by gene duplication (paralogy), causing gene trees to diverge from the species tree,
- to maximize the power to test the phylogenetic position of a particular taxon, try to include at least all the taxa that have previously been proposed to be its closest relatives,
- if practical, reconstruct a phylogenetic scaffold based on a restricted number of taxa scored for many characters. Additional taxa can then be added sequentially on the basis of smaller number of characters (Wiens 2006). Addition of incompletely known taxa can boost accuracy and confidence. To prevent systematic error it may be better to add a smaller number of characters scored for many taxa, rather than many characters for fewer taxa,
- if there is not enough phylogenetic signal, focus on characters with higher rates of evolution,
- assess data quality as a standard part of any phylogenetic analysis (Waagele & Mayer 2007),
- exploit combined evidence analyses, including fossil data wherever possible (Giribet 2002; Eernisse & Peterson 2004), while recognizing the interpretational difficulties associated with combining molecular exemplar species and inferred morphological ground patterns,
- sample different genes that evolve at different rates to be able to resolve different regions of the tree (Sanderson & Shaffer 2002; Glenner *et al.* 2004; Philippe & Telford 2006),
- boost the amount of descriptive and comparative morphological studies to revise outdated received wisdom, and provide more data crucial for the inference of body plan evolution (Nielsen 2001; Jenner 2006),
- adopt an experimental approach (sensitivity analysis) to phylogenetic analysis to see how results change depending on different assumptions,
- re-evaluate contentious morphological evidence in the light of independent molecular phylogenies, especially to detect cases of unrecognized character loss (Jenner 2004b),
- carefully construct morphological datasets to maximize testing power (Jenner 2004a),

— adopt standardized methods for the presentation, annotation and analysis of molecular data. To this end Leebens-Mack *et al.* (2006) have called for a standard for reporting on phylogenies, the Minimum Information about a Phylogenetic Analysis (MIAPA), in which each component of a phylogenetic analysis (alignment procedures, alignment, sequences, voucher specimens, methods and parameters used, etc.) is outlined using universally accepted criteria. This will facilitate better evaluation and comparison of results of different analyses.

In summary, the recognition of *Problematica* reveals more than the sum of their missing or ambiguous parts. In avoiding fragmentary fossils or extant organisms with combinations of chimaeric or autapomorphic features, and by excluding long branching taxa or heavily biased nucleotide and protein sequences from molecular analyses, we may bring near completeness to data matrices and greater stability to our phylogenetic analyses, but probably at the expense of accuracy and an understanding of the full evolutionary picture. *Problematica* reveal themselves as supremely important; for without their inclusion and accurate placement, other relationships are liable to change. In understanding how to deal with *Problematica*, we realize the limits of systematics and our ability to have faith in our reconstructions of the tree of life.

The authors dedicate this paper to the memories and careers of Reinhard M. Rieger (1943–2006) and Ellis L. Yochelson (1928–2006). R.A.J. gratefully acknowledges the UK Biotechnology and Biological Sciences Research Council for financial support. We thank Matthew Wills, Ciara Ni Dhubbghaill and Sarah Webb for their insightful comments on the manuscript. We are especially indebted to Greg Edgecombe, Graham Budd and Andrew Smith for their ruthlessly perceptive critiques of the final version of this paper.

REFERENCES

Adoutte, A., Balavoine, G., Lartillot, N., Lespinet, O., Prud'homme, B. & de Rosa, R. 2000 The new animal phylogeny: reliability and implications. *Proc. Natl Acad. Sci. USA* **97**, 4453–4456. (doi:10.1073/pnas.97.9.4453)

Aldridge, R. J., Xian-Guang, H., Siveter, D. J. & Gabbott, S. E. 2007 The systematics and phylogenetic relationships of vetulicolians. *Palaeontology* **50**, 131–168. (doi:10.1111/j.1475-4983.2006.00606.x)

Anderson, F. E. & Swofford, D. L. 2004 Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. *Mol. Phylogenet. Evol.* **33**, 440–451. (doi:10.1016/j.ympev.2004.06.015)

Bailey, J. V., Joye, S. B., Kalanetra, K. M., Flood, B. E. & Corsetti, F. A. 2007 Evidence of giant sulphur bacteria in Neoproterozoic phosphorites. *Nature* **445**, 198–201. (doi:10.1038/nature05457)

Barnes, R. D. 1985 Current perspectives on the origins and relationships of lower invertebrates. In *The origins and relationships of lower invertebrates* (eds S. Conway Morris, J. D. George, R. Gibson & H. M. Platt), pp. 361–367. Oxford, UK: Clarendon Press.

Baurain, D., Brinkmann, H. & Philippe, H. 2007 Lack of resolution in the animal phylogeny: closely spaced cladogeneses or undetected systematic errors? *Mol. Biol. Evol.* **24**, 6–9. (doi:10.1093/molbev/msl137)

- Bengtson, S. & Budd, G. 2004 Comment on "Small bilaterian fossils from 40 to 55 million years before the Cambrian". *Science* **306**, 1291a. (doi:10.1126/science.1101338)
- Bleidorn, C., Eeckhaut, I., Podsiadlowski, L., Schult, N., McHugh, D., Halanych, K. M., Milinkovitch, M. C. & Tiedemann, R. 2007 Mitochondrial genome and nuclear sequence data support Myzostomida as part of the annelid radiation. *Mol. Biol. Evol.* **24**, 1690–1701. (doi:10.1093/molbev/msm086)
- Boore, J. L. 2006 The use of genome-level character for phylogenetic reconstruction. *Trends Ecol. Evol.* **21**, 439–446. (doi:10.1016/j.tree.2006.05.009)
- Bourlat, S. J. *et al.* 2006 Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* **444**, 85–88. (doi:10.1038/nature05241)
- Briggs, D. E. G., Fortey, R. A. & Wills, M. A. 1992 Morphological disparity in the Cambrian. *Science* **256**, 1670–1673. (doi:10.1126/science.256.5064.1670)
- Brusca, R. C. & Brusca, G. J. 2003 *Invertebrates*. Sunderland, MA: Sinauer Associates, Inc.
- Budd, G. E. 2001 Why are arthropods segmented? *Evol. Dev.* **3**, 332–342. (doi:10.1046/j.1525-142X.2001.01041.x)
- Budd, G. E. 2003 The Cambrian fossil record and the origin of the phyla. *Integr. Comp. Biol.* **43**, 157–165. (doi:10.1093/icb/43.1.157)
- Budd, G. E. & Jensen, S. 2000 A critical reappraisal of the fossil record of the bilaterian phyla. *Biol. Rev. Camb. Phil. Soc.* **75**, 253–295. (doi:10.1017/S000632310000548X)
- Butterfield, N. J. 2006 Hooking some stem-group "worms": fossil lophotrochozoans in the Burgess Shale. *BioEssays* **28**, 1161–1166. (doi:10.1002/bies.20507)
- Caron, J.-B., Scheltema, A., Schander, C. & Rudkin, D. 2006 A soft-bodied mollusc with radula from the Middle Cambrian Burgess Shale. *Nature* **442**, 159–163. (doi:10.1038/nature04894)
- Chan, K. M. A. & Levin, S. A. 2005 Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* **59**, 720–729.
- Chen, J. Y., Bottjer, D. J., Oliveri, P., Dornbos, S. Q., Gao, F., Ruffins, S., Chi, H., Li, C.-W. & Davidson, E. H. 2004 Small bilaterian fossils from 40 to 55 million years before the Cambrian. *Science* **305**, 218–222. (doi:10.1126/science.1099213)
- Chen, J.-Y., Schopf, J. W., Bottjer, D. J., Zhang, C.-Y., Kudryavtsev, A. B., Tripathi, A. B., Wang, X.-Q., Gao, X. & Yang, Y. 2007 Raman spectra of a Lower Cambrian ctenophore embryo from southwestern Shaanxi, China. *Proc. Natl Acad. Sci. USA* **104**, 6289–6292. (doi:10.1073/pnas.0701246104)
- Clarke, B., Johnson, M. S. & Murray, J. 1996 Clines in the genetic distance between two species of island land snails: how 'molecular leakage' can mislead us about speciation. *Phil. Trans. R. Soc. B* **351**, 773–784. (doi:10.1098/rstb.1996.0072)
- Cobbett, A., Wilkinson, M. & Wills, M. A. 2007 Fossils impact as hard as living taxa in parsimony analyses of morphology. *Syst. Biol.* **56**, 753–766. (doi:10.1080/10635150701627296)
- Conway Morris, S. 1991 Problematic taxa: a problem for biology or biologists? In *The early evolution of Metazoa and the significance of problematic taxa* (eds A. M. Simonetta & S. Conway Morris), pp. 19–24. Cambridge, UK: Cambridge University Press.
- Conway Morris, S. 2000 The Cambrian "explosion": slow-fuse or megatonnage? *Proc. Natl Acad. Sci. USA* **97**, 4426–4429. (doi:10.1073/pnas.97.9.4426)
- Coyne, J. A. & Orr, H. A. 2004 *Speciation*. Sunderland, MA: Sinauer Associates, Inc.
- Delsuc, F., Brinkmann, H. & Philippe, H. 2005 Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* **6**, 361–375. (doi:10.1038/nrg1603)
- Donoghue, P. C. J. 2007 Embryonic identity crisis. *Nature* **445**, 155–156. (doi:10.1038/nature05520)
- Donoghue, P. C. J. *et al.* 2006 Synchrotron X-ray tomographic microscopy of fossil embryos. *Nature* **442**, 680–683. (doi:10.1038/nature04890)
- Dzik, J. 2003 Anatomical information content in the Ediacaran fossils and their possible zoological affinities. *Integr. Comp. Biol.* **43**, 114–126. (doi:10.1093/icb/43.1.114)
- Eernisse, D. J. & Peterson, K. J. 2004 The history of animals. In *Assembling the tree of life* (eds J. Cracraft & M. J. Donoghue), pp. 197–208. Oxford, UK; New York, NY: Oxford University Press.
- Eriksson, B. J. & Budd, G. E. 2000 Onychophoran cephalic nerves and their bearing on our understanding of head segmentation and stem-group evolution of Arthropoda. *Arthropod Struct. Dev.* **29**, 197–209. (doi:10.1016/S1467-8039(00)00027-X)
- Erwin, D. H. 2007 Disparity: morphological pattern and developmental context. *Palaeontology* **50**, 57–73. (doi:10.1111/j.1475-4983.2006.00614.x)
- Fedonkin, M. A. 2003 The origin of the Metazoa in the light of the Proterozoic fossil record. *Paleontol. Res.* **7**, 9–41. (doi:10.2517/prpsj.7.9)
- Giribet, G. 2002 Current advances in the phylogenetic reconstruction of metazoan evolution. A new paradigm for the Cambrian explosion? *Mol. Phylogenet. Evol.* **24**, 345–357. (doi:10.1016/S1055-7903(02)00206-3)
- Giribet, G., Okusu, A., Lindgren, A. R., Huff, S. W., Schrödl, M. & Nishiguchi, M. K. 2006 Evidence for a clade composed of molluscs with serially repeated structures: monoplacophorans are related to chitons. *Proc. Natl Acad. Sci. USA* **103**, 7723–7728. (doi:10.1073/pnas.0602578103)
- Glennier, H., Hansen, A. J., Sørensen, M. V., Ronquist, F., Huelsenbeck, J. P. & Willerslev, E. 2004 Bayesian inference of the metazoan phylogeny: a combined molecular and morphological approach. *Curr. Biol.* **14**, 1644–1649. (doi:10.1016/j.cub.2004.09.027)
- Gribaldo, S. & Philippe, H. 2002 Ancient phylogenetic relationships. *Theor. Popul. Biol.* **61**, 391–408. (doi:10.1006/tpbi.2002.1593)
- Halanych, K. M. 2004 The new view of animal phylogeny. *Annu. Rev. Ecol. Evol. Syst.* **35**, 229–256. (doi:10.1146/annurev.ecolsys.35.112202.130124)
- Hillis, D. M., Pollock, D. D., McGuire, J. A. & Zwickl, D. J. 2003 Is sparse taxon sampling a problem for phylogenetic inference? *Syst. Biol.* **52**, 124–126. (doi:10.1080/10635150390132911)
- Hubrecht, A. A. W. 1887 The relation of the Nemertea to the Vertebrata. *Quart. J. Microscop. Sci.* **27**, 605–644.
- Hyman, L. H. 1959 *The invertebrates*, vol. 5, *smaller coelomate groups*. New York, NY: McGraw-Hill.
- Irimia, M., Maeso, I., Penny, D., Garcia-Fernández, J. & Roy, S. W. 2007 Rare coding sequence changes are consistent with Ecdysozoa, not Coelomata. *Mol. Biol. Evol.* **24**, 1604–1607. (doi:10.1093/molbev/msm105)
- Jenner, R. A. 2004a The scientific status of metazoan cladistics: why current research practice must change. *Zool. Scr.* **33**, 293–310. (doi:10.1111/j.0300-3256.2004.00153.x)
- Jenner, R. A. 2004b Towards a phylogeny of the Metazoa: evaluating alternative phylogenetic positions of Platyhelminthes, Nemertea, and Gnathostomulida, with a critical reappraisal of cladistic characters. *Contrib. Zool.* **73**, 3–163.
- Jenner, R. A. 2006 Unburdening evo-devo: ancestral attractions, model organisms, and basal baloney. *Evol. Dev.* **216**, 385–394. (doi:10.1007/s00427-006-0084-5)

- Jenner, R. A. & Scholtz, G. 2005 Playing another round of metazoan phylogenetics: historical epistemology, sensitivity analysis, and the position of Arthropoda within the Metazoa on the basis of morphology. In *Crustacea and arthropod relationships* (eds S. Koenemann & R. A. Jenner), pp. 355–385. Boca Raton, FL: CRC Press.
- Jenner, R. A. & Wills, M. A. 2007 The choice of model organisms in evo–devo. *Nat. Rev. Genet.* **8**, 311–314. (doi:10.1038/nrg2062)
- Jiménez-Guri, E., Philippe, H., Okamura, B. & Holland, P. W. H. 2007 *Buddenbrockia* is a cnidarian worm. *Science* **317**, 116–118. (doi:10.1126/science.1142024)
- Lecointre, G. & Deleporte, P. 2004 Total evidence requires exclusion of phylogenetically misleading data. *Zool. Scr.* **34**, 101–117. (doi:10.1111/j.1463-6409.2005.00168.x)
- Leebens-Mack, J. *et al.* 2006 Taking the first steps towards a standard for reporting on phylogenies: minimum information about a phylogenetic analysis (MIAPA). *OMICS* **10**, 231–237. (doi:10.1089/omi.2006.10.231)
- Lewis, P. O. 2001 Phylogenetic systematics turns over a new leaf. *Trends Ecol. Evol.* **16**, 30–37. (doi:10.1016/S0169-5347(00)02025-5)
- Marlétaz, F. *et al.* 2006 Chaetognath phylogenomics: a protostome with deuterostome-like development. *Curr. Biol.* **16**, R578–R579. (doi:10.1016/j.cub.2006.07.016)
- Matus, D. Q., Copley, R. R., Dunn, C. W., Hejnal, A., Eccleston, H., Halanych, K. M., Martindale, M. Q. & Telford, M. J. 2006 Broad taxon and gene sampling indicate that chaetognaths are protostomes. *Curr. Biol.* **16**, R575–R576. (doi:10.1016/j.cub.2006.07.017)
- McHugh, D. 1997 Molecular evidence that echiurans and pogonophorans are derived annelids. *Proc. Natl Acad. Sci. USA* **94**, 8006–8009. (doi:10.1073/pnas.94.15.8006)
- Narbonne, G. M. 2005 The Ediacara biota: Neoproterozoic origin of animals and their ecosystems. *Annu. Rev. Earth Planet. Sci.* **33**, 421–442. (doi:10.1146/annurev.earth.33.092203.122519)
- Nielsen, C. 2001 *Animal evolution—interrelationships of the living phyla*. Oxford, UK: Oxford University Press.
- O’Hara, R. J. 1992 Telling the tree: narrative representation and the study of evolutionary history. *Biol. Philos.* **7**, 135–160. (doi:10.1007/BF00129880)
- Peterson, K. J. & Eernisse, D. J. 2001 Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. *Evol. Dev.* **3**, 170–205. (doi:10.1046/j.1525-142x.2001.003003170.x)
- Peterson, K. J., Lyons, J. B., Nowak, K. S., Takacs, C. M., Wargo, M. J. & McPeck, M. A. 2004 Estimating metazoan divergence times with a molecular clock. *Proc. Natl Acad. Sci. USA* **101**, 6536–6541. (doi:10.1073/pnas.0401670101)
- Peterson, K. J., McPeck, M. A. & Evans, D. A. D. 2005 Tempo and mode of early animal evolution: inferences from rocks, Hox, and molecular clocks. *Paleobiology* **31**, 36–55. (doi:10.1666/0094-8373(2005)031[0036:TAMOE]2.0.CO;2)
- Philippe, H. & Telford, M. J. 2006 Large-scale sequencing and the new animal phylogeny. *Trends Ecol. Evol.* **21**, 614–620. (doi:10.1016/j.tree.2006.08.004)
- Philippe, H., Lartillot, N. & Brinkmann, H. 2005 Multi-gene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protozoa. *Mol. Biol. Evol.* **22**, 1246–1253. (doi:10.1093/molbev/msi111)
- Philippe, H., Brinkmann, H., Martinez, P., Riutort, M. & Baguña, J. 2007 Acoel flatworms are not Platyhelminthes: evidence from phylogenomics. *PLoS ONE* **2**, e717. (doi:10.1371/journal.pone.0000717)
- Rogozin, I. B., Wolf, Y. I., Carmel, L. & Koonin, E. V. 2007 Ecdysozoan clade rejected by genome-wide analysis of rare amino acid replacements. *Mol. Biol. Evol.* **24**, 1080–1090. (doi:10.1093/molbev/msm029)
- Rokas, A. & Carroll, S. B. 2006 Bushes in the tree of life. *PLoS Biol.* **4**, e352. (doi:10.1371/journal.pbio.0040352)
- Rokas, A., Krüger, D. & Carroll, S. B. 2005 Animal evolution and the molecular signature of radiations compressed in time. *Science* **310**, 1933–1938. (doi:10.1126/science.1116759)
- Rouse, G. W. 2001 A cladistic analysis of Siboglinidae Caullery, 1914 (Polychaeta, Annelida): formerly the phyla Pogonophora and Vestimentifera. *Zool. J. Linn. Soc.* **132**, 55–80. (doi:10.1006/zjls.2000.0263)
- Rousset, V., Pleijl, F., Rouse, G. W., Erséus, C. & Siddall, M. E. 2007 A molecular phylogeny of annelids. *Cladistics* **23**, 41–63. (doi:10.1111/j.1096-0031.2006.00128.x)
- Ruta, M., Wagner, P. J. & Coates, M. I. 2006 Evolutionary patterns in early tetrapods. I. Rapid initial diversification followed by decrease in rates of character change. *Proc. R. Soc. B* **273**, 2107–2111. (doi:10.1098/rspb.2006.3577)
- Sanderson, M. J. & Shaffer, H. B. 2002 Troubleshooting molecular phylogenetic analyses. *Annu. Rev. Ecol. Evol. Syst.* **33**, 49–72.
- Schopf, J. W. & Kudryavtsev, A. B. 2005 Three-dimensional Raman imagery of precambrian microscopic organisms. *Geobiology* **3**, 1–12. (doi:10.1111/j.1472-4669.2005.00044.x)
- Struck, T. H., Schult, N., Kusen, T., Hickman, E., Bleidorn, C., McHugh, D. & Halanych, K. M. 2007 Annelid phylogeny and the status of Sipuncula and Echiura. *BMC Evol. Biol.* **7**, 57. (doi:10.1186/1471-2148-7-57)
- Sutton, M. D., Briggs, D. E. & Siveter, D. J. 2001a A three-dimensionally preserved fossil polychaete worm from the Silurian of Herefordshire, England. *Proc. R. Soc. B* **268**, 2355–2363. (doi:10.1098/rspb.2001.1788)
- Sutton, M. D., Briggs, D. E. & Siveter, D. J. 2001b An exceptionally preserved vermiform mollusc from the Silurian of England. *Nature* **410**, 461–463. (doi:10.1038/35068549)
- Sutton, M. D., Briggs, D. E. G., Siveter, D. J. & Siveter, D. J. 2005a Fossilized soft tissues in a Silurian platyceratid gastropod. *Proc. R. Soc. B* **273**, 1039–1044. (doi:10.1098/rspb.2005.3403)
- Sutton, M. D., Briggs, D. E. G., Siveter, D. J. & Siveter, D. J. 2005b Silurian brachiopods with soft-tissue preservation. *Nature* **436**, 1013–1015. (doi:10.1038/nature03846)
- Sutton, M. D., Briggs, D. E. G., Siveter, D. J., Siveter, D. J. & Gladwell, D. J. 2005c A starfish with three-dimensionally preserved soft parts from the Silurian of England. *Proc. R. Soc. B* **272**, 1001–1006. (doi:10.1098/rspb.2004.2951)
- Thomson, K. S., Sutton, M. D. & Thomas, B. 2003 A larval Devonian lungfish. *Nature* **426**, 833–834. (doi:10.1038/nature02175)
- Todaro, M. A., Telford, M. J., Lockyer, A. E. & Littlewood, D. T. J. 2006 Interrelationships of the Gastrotricha and their place among the Metazoa inferred from 18S rRNA genes. *Zool. Scr.* **35**, 251–259. (doi:10.1111/j.1463-6409.2006.00228.x)
- Valentine, J. W. 2004 *On the origin of phyla*. Chicago, IL: The University of Chicago Press.
- Waegle, J. W. & Mayer, C. 2007 Visualizing differences in phylogenetic information content of alignments and distinction of three classes of long-branch effects. *BMC Evol. Biol.* **7**, 147. (doi:10.1186/1471-2148-7-147)
- Wagner, P. J. 2000 Exhaustion of morphological character states among fossil taxa. *Evolution* **54**, 365–386.
- Wagner, G. P. 2001a What is the promise of developmental evolution? Part II: a causal explanation of evolutionary innovations may be impossible. *J. Exp. Zool. (Mol. Dev. Evol.)* **291**, 305–309. (doi:10.1002/jez.1130)

- Wagner, P. J. 2001*b* Gastropod phylogenetics: progress, problems, and implications. *J. Paleontol.* **75**, 1128–1140. (doi:10.1666/0022-3360(2001)075<1128:GPPPAI>2.0.CO;2)
- Wagner, P. J., Ruta, M. & Coates, M. I. 2006 Evolutionary patterns in early tetrapods. II. Differing constraints on available character space among clades. *Proc. R. Soc. B* **273**, 2113–2118. (doi:10.1098/rspb.2006.3561)
- Waloszek, D., Repetski, J. E. & Maas, A. 2005 A new Late Cambrian pentastomid and a review of the relationships of this parasitic group. *Trans. R. Soc. Edin.* **96**, 163–176.
- Wheeler, W. *et al.* 2006 *Dynamic homology and phylogenetic systematics: a unified approach using POY*. New York, NY: American Museum of Natural History.
- Whitfield, J. B. & Lockhart, P. J. 2007 Deciphering ancient rapid radiations. *Trends Ecol. Evol.* **22**, 258–265. (doi:10.1016/j.tree.2007.01.012)
- Wiens, J. J. 2006 Missing data and the design of phylogenetic analyses. *J. Biomed. Inform.* **39**, 34–42. (doi:10.1016/j.jbi.2005.04.001)
- Wiens, J. J., Bonett, R. M. & Chippindale, P. T. 2005*a* Ontogeny discombobulates phylogeny: paedomorphosis and higher-level salamander relationships. *Syst. Biol.* **54**, 91–110. (doi:10.1080/10635150590906037)
- Wiens, J. J., Fetzner Jr, J. W., Parkinson, C. L. & Reeder, T. W. 2005*b* Hylid frog phylogeny and sampling strategies for speciose clades. *Syst. Biol.* **54**, 719–748. (doi:10.1080/10635150500234625)
- Wiens, J. J., Engstrom, T. N. & Chippindale, P. T. 2006 Rapid diversification, incomplete isolation, and the “speciation clock” in North American salamanders (genus *Plethodon*): testing the hybrid swarm hypothesis of rapid radiation. *Evolution* **60**, 2585–2603.
- Yochelson, E. L. 1991 *Problematica/Incertae Sedis*. In *The early evolution of Metazoa and the significance of problematic taxa. Proc. Int. Symp. held at the University of Camerino 27–31 March 1989* (eds A. M. Simonetta & S. Conway Morris), pp. 287–296. Cambridge, UK: Cambridge University Press.
- Zrzavy, J. 2003 Gastrotricha and metazoan phylogeny. *Zool. Scr.* **32**, 61–81. (doi:10.1046/j.1463-6409.2003.00104.x)
- Zrzavy, J., Mihulka, S., Kepka, P., Bezdek, A. & Tietz, D. 1998 Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. *Cladistics* **14**, 249–285.
- Zrzavy, J., Hypsa, V. & Tietz, D. F. 2001 Myzostomida are not annelids. Molecular and morphological support for a clade of animals with anterior sperm flagella. *Cladistics* **17**, 170–198.