Abstract
Molecular methods have revolutionised virtually every area of biology, and metazoan phylogenetics is no exception: molecular phylogenies, molecular clocks, comparative phylogenomics, and developmental genetics have generated a plethora of molecular data spanning numerous taxa and collectively transformed our understanding of the evolutionary history of animals, often corroborating but at times opposing results of more traditional approaches. Moreover, the diversity of methods and models within molecular phylogenetics has resulted in significant disagreement among molecular phylogenies as well as between these and earlier phylogenies. How should this broad and multifaceted problem be tackled? I argue that the answer lies in integrating evidence to infer the best evolutionary scenario. I begin with an overview of recent development in early metazoan phylogenetics, followed by a discussion of key conceptual issues in phylogenetics revolving around phylogenetic evidence, theory, methodology, and interrelations thereof. I then argue that the integration of different kinds of evidence (e.g. molecular, morphological, ecological) is necessary for arriving at the best evolutionary scenario (causal explanation) rather than merely the best-fitting cladogram (statistical explanation). Finally, I discuss the prospects of this view in stimulating interdisciplinary cross-talk in early metazoan research and beyond, and challenges that need to be overcome.

Keywords Evidential integration · Triangulation · Early animal evolution · Phylogenetics · Macroevolution · Inference to the best explanation
Introduction

Phylogenetics, or genealogical systematics, has been transformed in recent decades by the advent of molecular approaches to classification relying on similarities in the genetic material of living taxa; starting with analyses on small sets of genes, they have come a long way to the large-scale phylogenomics of recent years (Dunn et al. 2008; Feuda et al. 2017; Giribet 2016; Kocot et al. 2011, 2017; Laumer et al. 2019; Lozano-Fernandez et al. 2019; Paps 2018; Poustka et al. 2020; Rota-Stabelli et al. 2011; Simion et al. 2017; Telford and Littlewood 2009; Williams et al. 2020), often calibrated by fossils (Donoghue and Benton 2007; Tihelka et al. 2020). The impact of molecular approaches to phylogenetics cannot be understated: we now have access to an unprecedented wealth of data from a wide variety of organisms independent of—and almost overwhelmingly more abundant than—morphological data, and a panoply of sophisticated analytical tools enabling their interpretation. Nevertheless, initial optimism about molecular data being less prone to long-standing issues concerning more traditional approaches—most notably false homologies, relatively small amounts of data, and non-neutrality compared to molecular data—has given way to uncertainties around incongruence (Jeffroy et al. 2006), matching models to huge quantities of data (Cai et al. 2020; Jeffroy et al. 2006; Nagy et al. 2020), and criteria for model choice (Kapli et al. 2021a; Kapli and Telford 2020). Molecular approaches are not flawless.

The primary implication for phylogenetics is that molecular approaches and the resulting phylogenies cannot be deemed categorically superior to traditional ones. This is especially important because phylogenies based on molecular characters are often not in full agreement with those based on morphological characters, as well as not being in full agreement with each other (the same holds for morphological phylogenies). A few notable cases concern early metazoan evolution: the placement of Ctenophora as either the sister group to all other animals or as a eumetazoan taxon (Dunn et al. 2008; Feuda et al. 2017; Kapli and Telford 2020; Pisani et al. 2015; Simion et al. 2017); the placement of Placozoa as either sister to Eumetazoa or belonging to crown-group Eumetazoa (Laumer et al. 2018; Pisani et al. 2015; Simion et al. 2017); the placement of Xenacoelomorpha as sister to Nephrozoa or deep within Bilateria (Brauchle et al. 2018; Kapli, Natsidis, et al., 2021; Kapli and Telford 2020; Ruiz-Trillo et al. 2004); and the relationship between arthropods and annelids (Eernisse et al. 1992; Telford et al. 2015). Of these, only the last is meaningfully resolved (Fig. 1).

This in turn gives rise to a number of theoretical and philosophical questions: what are the main underlying reasons for such disagreements? Should we expect full agreement in the first place, and what reasons might we have for or against this? In which approach or phylogeny should we have more confidence, and what criteria can we have for making such judgments? The aim of this paper is to offer answers to these questions. I begin with a number of key theoretical issues including homologue identification in morphological and molecular phylogenetics, methodological artifacts, and model-data fit, followed by a model of the relationship between homologues and cladograms in order to put the aforementioned issues in perspective and provide part of the answer to the overarching question: evidence of any kind strictly
requires a model that fits if it is to be informative. Next, I attempt to provide another general solution to the problem of disagreement, namely meta-congruence through the integration of different kinds of evidence—molecular and morphological as well as developmental, ecological, and biogeographical—leading to inferring the best evolutionary scenario (a causal explanation) rather than the most likely cladogram (a statistical explanation). Crucially, different kinds of evidence ought to be considered together because they compensate for one another’s weaknesses. This line of argument calls for multidisciplinary research in phylogenetics by providing a theoretical framework centred around evidential integration.

Before delving into the discussion, it is worth clarifying the core argument of this paper. Standardly, phylogeneticists start with (1) inferring a phylogenetic tree, nowadays typically from molecular data, and use the inferred tree to (2) interpret morphological, genetic, or developmental evolution, and finally (3) speculate on causality. One way to read the argument of this paper is as follows: (1) gives a partial picture...
of evolution and therefore (2) is also necessary. This is somewhat trivial, and most researchers in the field would immediately agree. My aim is rather to argue that (1) would benefit from incorporation of data of the kind usually included in step (2), and this would make inferences about (3) more accurate.

**From homologues to cladograms**

**Elusive homology**

Homology is an elusive concept (Hall 2012), and homologues (homologous characters) can be equally hard to identify (Williams and Ebach 2020). The two closely-related concepts have a long history stretching back to the late 18th century and are, at their core, concerned with *sameness* (Ramsey and Siebels Peterson 2012). There are nevertheless several existing definitions of homology (Hall 2012; Novick 2018; Van Valen 1982) pertaining to varying aspect of it and aimed at explaining overlapping though not coextensive sets of phenomena (e.g. serial homology, developmental homology, etc.). For the purposes of the present discussion, I shall focus on the phylogenetic aspect of homology and homologues as characters which are the *same* (not identical) due to common ancestry, on whose basis we infer genealogical relationships between taxa—in the same spirit as Patterson’s concept of homologues as synapomorphies defining clades (Patterson 1988).

Inferring cladistic (=genealogical) relationships between taxa starts with the identification of homologues. A classic and fairly intuitive example is the vertebrate forelimb; it comes in a variety of forms (arms, wings, flippers, etc.), performing a variety of functions (handling, flying, swimming, etc.), but nonetheless remains the *same* organ—the vertebrate forelimb. Unsurprisingly, however, not all homologues are so easily discerned from analogues—characters exhibiting similarity due to reasons such as convergent evolution or the co-option of gene regulatory networks. Two notable examples from recent literature on early bilaterian evolution include the convergent evolution of bilaterian nerve cords (Martín-Durán et al. 2018) or the convergent evolution of “limbs” in cephalopods, vertebrates, and arthropods (Prpic 2019; Tarazona et al. 2019). Examples of homologue identification abound, ranging from the unequivocally homologous such as the vertebrate forelimb or the insect hind-wing to the contentious, such as bilaterian metamerism or the tripartite coelom; often with serious implications for understanding the evolution of key characters of taxa of interest, and subsequently their interrelations: correctly identifying homologues—i.e. discerning them from analogues—forms the basis of inferring phylogenies.

The difficulty involved in correctly identifying homologues (at least morphological ones), almost invariably requiring extensive expertise acquired over years of research, has led to generalised criteria for doing so—most notably those of Patterson (Patterson 1982, 1988) and Remane (Remane 1956) (also see (DiFrisco et al. 2020; Griffiths 2007)). These tend to touch upon key general features of homologues, including but by no means exclusive to relative complexity, similarity unattributable to convergence, phylogenetic traceability, and congruence with other characters.
Molecular homology

Regarding molecular characters, there are two main approaches to inferring homology and its utilisation in phylogenetics. The first, more commonly used approach is orthology inference: homologous genes are categorised primarily into orthologues and paralogues, the former resulting from speciation events, the latter from duplication events; therefore, only orthologues are informative about relationships between taxa (which also result from speciation events), which means phylogenetic analyses based on individual genes should be, and indeed has been, restricted to orthologues. The main potential issue with this approach is that orthologue identification algorithms lose their accuracy when dealing with clades with high rates of evolution. Furthermore, once orthologues have been identified by their overall sequence similarity, there still remains a question about the exact homology relationships between their sites and regions, which is often dealt with using multiple sequence alignment; which, once again, can lead to inaccurate results if the model used is not well-suited for the data (e.g. if sequences are under high evolutionary rates). In such cases, the alternative approach, namely identifying and utilising homologous gene families, is potentially more appropriate (Natsidis et al. 2020).

The issue just outlined is one among a few concerning the pertinence of the choice of methodological approaches to molecular phylogenetics. Others include using all available data or applying filters to them (Cai et al. 2020) and choosing the best model for clades of interest, which are discussed below under Two controversies and a common error. Before proceeding, it is worth pointing out the invaluable contributions of molecular phylogenetics; notable examples include resolving relationships within Mollusca (Poustka et al. 2020), Annelida (Struck et al. 2011), and the recovery of Lophotrochozoa at the expense of Articulata (Kocot et al. 2017)—cases that morphological approaches had long struggled to resolve. The primary aim of this paper is to advocate utilising different kinds of evidence when they are most informative, and molecular approaches have shown to be highly so in cases mentioned above, as well as corroborating innumerable traditional phylogenies. With this in mind, I will go on to stress specific cases where other sources of evidence can be used to complement molecular approaches in phylogenetics.

Integrative homology

Another such issue concerns the identification of homologues in light of different kinds of evidence—e.g. genetic, developmental, or morphological. Though the literature on homology and especially its developmental basis has made great progress in recent years in elucidating the relationship between genetic, developmental, and morphological homology (Brigandt 2007; DiFrisco and Jaeger 2021; DiFrisco et al. 2020; Griffiths 2007; Novick 2018; Ramsey and Siebels Peterson 2012; Wagner 2007, 2018), there is as of yet no general theoretical framework elucidating the identification of homologues on the basis of different kinds of evidence. One example where this would be illuminating concerns the evolution of the eumetazoan body plan: this has for many decades been, and remains to be, a matter of debate, and a hotly-debated topic in recent years due to methodological advances in developmental biology. More
specifically, morphological and developmental genetic evidence disagree in this case over the ancestral eumetazoan body plan and the homology relations between the body axes of placozoans, cnidarians, deuterostomes, and protostomes (Arendt et al. 2015; Genikhovich and Technau 2017; He et al. 2018; Nielsen et al. 2018; Steinmetz et al. 2017). While this specific case more directly pertains to inferring the characters of ancestral organisms rather than inferring tree topologies, the issue nevertheless remains that there is as of yet no principled method for integrating different kinds of evidence to identify homologues. A full discussion on how this problem could be overcome is beyond the scope of this paper; however, I will briefly return to it in this paper under Evidential Integration and meta-congruence.

Methodological artefacts and systematic errors

An inescapable feature of any scientific endeavour is the partial dependence of results on the method used to obtain those results, rather than complete dependence on the data alone. The inescapability is a direct outcome of the scientific method: data is most often collected in order to test specific hypotheses, informed by specific theories, and subjected to analytic methods also informed by specific theories and bearing specific sets of idealising assumptions. Phylogenetics is no exception to this general rule. I will now draw attention to two outstanding problems in animal phylogenetics to stress the importance of the role that methodological commitments play in obtaining phylogenetic results—i.e. cladograms and reconstructions of ancestral taxa.

Urbilateria: complex or simple?

The first concerns the differential weighting of a certain class of homologous characters in morphological analyses of major metazoan phyla, namely those relating to morphological complexity and especially metamerism/segmentation. It has long been a matter of debate whether the last common bilaterian ancestor (= Urbilateria) was a complex segmented organism with an array of organs and organ systems (e.g. eyes, hearts, central nervous systems), or a relatively simple animal lacking most if not all such features, very much akin to living xenacoelomorph worms (Balavoine and Adoutte 2003; Hejnol and Martindale 2008; Jacobs et al. 2005; Northcutt 2010). The abundance of morphologically relatively simple bilaterian phyla (Nielsen 2012), as well as the plausible—yet contentious (see below)—phylogenetic placement of the xenacoelomorph worms as the sister group of all other living bilaterians (= Nephrozoa) together lend support to the “simple Urbilateria” hypothesis. The “complex Urbilateria” hypothesis, on the other hand, gains support from the complexity and shared developmental features of aforementioned traits as well as the contentiousness of the placement of Xenacoelomorpha and a growing body of evidence from the early Cambrian fossil record, including a number of fossils with quite complex morphology inferred to be stem-group (Budd and Jensen 2017; Chen et al. 2020; Vinther et al. 2017) members of several bilaterian phyla, whose crown-group members tend to be morphologically simple.

The key issue is thus whether it is more plausible that Urbilateria was a complex organism and numerous bilaterian phyla subsequently lost multiple traits, or that it
was a simple organism and the shared complex traits are the result of convergent evolution. To put it more simply, the question is one of assigning weights to certain characters: should they be given equal weight as other characters, or should they be given more (or possibly less) weight than average. Differential weighting of characters might be a plausible option in circumstances where there is independent reason to assume that those characters are more (or less) likely to evolve independently (due to relative complexity, for example), particularly favoured (or disfavoured) by ecological factors at the time of phylogenetic divergence (e.g. widespread evolution of heavy mineralisation across bilaterians in the early Cambrian), a robust (or volatile) underlying developmental basis, or the presence of stem-group fossils with particular morphological features. While a detailed discussion on the relative importance of these factors for determining character weighting—in this particular case or any other—cannot fit in a few short paragraphs, it should nevertheless be apparent by now that methodological commitments regarding which of these factors to take into account, and how to do so, do in fact make a difference to the phylogenetic result obtained—in this case the relative complexity of the last common bilaterian ancestor.

**Two controversies, a common issue**

Though the preceding example primarily concerned morphological characters, it is in fact at its core quite similar to a broadly encompassing issue in molecular phylogenetics: choosing the appropriate character evolution model for the taxon under investigation. Generally speaking, molecular phylogenetic characters—primarily nucleotides, but also amino acids—have clade-specific rates of evolution (e.g. nucleotide substitution). Furthermore, this clade-specificity pertains not only to individual nucleotides or amino acids, but also to particular regions of genes and even particular genes and genomic regions. Rate heterogeneity by site is in fact usually incorporated into modern molecular phylogenies; less so with rate heterogeneity by clade, and especially less so with rate heterogeneity across various sites across different clades. Discovering precisely which nucleotides or genes in which clade or organism are more or less subject to evolution is therefore a major empirical challenge, without which molecular phylogenetic inferences would suffer from an inevitable lack of precision. Once again, empirical knowledge is needed in order to determine how to utilise phylogenetically relevant data; in the preceding example this concerned assigning differential weights to characters, in the present one it concerns assigning differential rates of evolution.

Assigning differential rates of evolution in molecular phylogenetics standardly takes the form of the employment of a diverse array of character evolution models. The sheer enormity of the literature on these models precludes the feasibility or even utility of an informative discussion in this paper. Instead, I will focus on a specific recent example pertaining to two relatively well-known controversies in the literature on early animal phylogenomics: those concerning the placement of ctenophores (Dunn et al. 2008; Pisaní et al. 2015; Simion et al. 2017) and xenacoelomorphs (Brauchle et al. 2018; Kapli and Telford 2020). As it turns out, they have more in common than previously anticipated: they are both strongly affected by systematic errors.
The phylogenetic placement of ctenophores is perhaps more appropriately dubbed a controversy: since its inception (Dunn et al. 2008), it has featured in popular science articles as well as academic ones; in the latter it has raised time and again the possibility of ctenophores (commonly known as comb-jellies) being the sister clade to all other animals, in sharp contrast to virtually every morphology-based phylogeny (Nielsen 2012) as well as numerous molecular phylogenies published since (Feuda et al. 2017; Pisani et al. 2015; Simion et al. 2017). The controversial nature of the “ctenophore-first” hypothesis is at least in part fuelled by its shocking implications: that a whole suite of traits otherwise specific to eumetazoans, including muscle cells, nerve cells, a quasi-radial body plan, and a digestive cavity (to name a few) have either evolved independently in the eumetazoan and ctenophore lineages, or have been present in the last common metazoan ancestor and subsequently lost in the poriferan (=sponge) lineage. Both are highly implausible evolutionary scenarios given our understanding of metazoan morphological evolution.

The xenacoelomorphs are a relatively recently discovered clade comprised by acoelomorphs, which were previously included within the phylum Platyhelminthes (Rohde et al. 1988), and the xenoturbellids, previously included within bivalves (Norén and Jondelius 1997) and later Deuterostomia (Nielsen 2010). Since its discovery as a clade through molecular phylogenies (Philippe et al. 2007) and subsequent corroboration by the identification of morphological synapomorphies for the clade (Nielsen 2012), the exact phylogenetic placement of Xenacoelomorpha within the broader metazoan tree has been a matter of contention; the two leading hypotheses being their placement as the sister group to all other bilaterians (the Nephrozoa hypothesis), or as a member of the deuterostome crown-group (the Xenambulacraria hypothesis); these alternative hypotheses have strong implications for the evolution of key bilaterian characters, as briefly mentioned above.

The situation is in fact quite similar to the ctenophore controversy in that it also involves two hypotheses, one of which takes the clade of interest as the sister group to the rest of a much larger clade (Metazoa and Nephrozoa) while the other takes it to belong deep within the larger clade; and that both have implications for the evolution of key characters pertaining to the clades involved. But there is also another similarity which potentially explains the others: both clades have long branch lengths—in other words, both exhibit unusually high rates of molecular as well as morphological evolution. The importance of this point lies in the long-recognised fact that taxa with long branches in cladograms tend to artificially cluster together and get “pushed down” towards the base of the tree (Kapli, Flouri, et al., 2021). The importance of this artefactual effect in resolving the phylogenetic placement of ctenophores and xenacoelomorphs is discussed in relation to the aforementioned issue of model choice in a recent paper by Kapli and Telford (Kapli and Telford 2020). A very brief summary is as follows: a combination of empirical data and simulation shows that the ctenophore-first and Nephrozoa hypotheses, unlike the sponge-first and Xenambulacraria hypotheses, are strongly supported by analyses affected by systematic errors, suggesting that the support found in empirical studies for the former two hypotheses can be explained by the models being used not appropriately accommodating the data. The key feature of the data here is in both cases the long branch leading to each respective taxon: site-homogenous models of molecular evolution systematically underestimate
branch length, thereby asymmetrically favouring the ctenophore-first and Nephrozoa hypotheses even in simulations where these are not the correct topologies—in contrast to site-heterogenous models, which do not in fact overestimate branch lengths. Thus, in addition to the amount of data present and the filtering methods used on the data, it also seriously matters what model is used to analyse the data, in this case a site-homogenous vs. a site-heterogenous model.

Evidence and theory in phylogenetics

We have therefore seen, through two interrelated examples from modern metazoan phylogenetics, how methodological commitments such as model choice or selective consideration of different lines of evidence can affect the outcome of any phylogenetic study: inappropriate models and insufficient or biased data can give rise to methodological artefacts and inaccurate results. I will now attempt to incorporate this into a more general model of the relationship between evidence and theory in phylogenetic analysis.

As briefly mentioned above, scientific enquiry involves two distinct but interconnected parts: evidence and theory. Evidence is produced when observations (=raw data) are utilised in order to favour one hypothesis over others, either negatively or positively (i.e. evidence for or evidence against a specific hypothesis; see (Lipton 2004)). A theory is produced, as is widely accepted, when a specific hypothesis is favoured over its rivals by specific evidence. Thus, the production of theories and evidence is a dynamic process which involves observation, hypothesis formation, and their synergistic utilisation. As more observations are made, more hypotheses are produced to explain them, and more observations are in turn made to test the hypotheses¹. So does science grow (Fig. 2).

What are the observations, hypotheses, evidence, and theories of the science of phylogenetics? I suggest here that observations can be taken simply as the physical attributes of organisms—anything ranging from their genomic sequences, patterns of gene expression in each semaphoront², distribution of cell types and tissues across the body, or the presence/absence or configuration of organs or organ systems in their body. The hypotheses can be taken as putative homologies: these are hypotheses of relationships between parts of organisms (putative homologues). Putative homologies result from the recognition of similarities not easily explained in terms of their adaptive value in a specific ecological setting³; in other words, they are hypotheses of genealogical relationships between characters (Hall 2012; Williams and Ebach 2020).

The process in phylogenetics that transforms observations and hypotheses into evidence and theories, respectively, is the application of a method of analysis (e.g. parsimony, maximum likelihood, or Bayesian methods) to the data, along with some

¹ Though the distinction matters here, some authors in fact use the terms “theory” and “hypothesis” interchangeably.
² Roughly speaking, a semaphoront is a specific region of a developing organisms with a specific developmental trajectory.
³ A classic example is the presence of wings in both bats and birds: both are needed for flight.
idealising assumptions (e.g. substitution rate models), to arrive at one or a few tree topologies. The trees thus obtained can be characterised as the theories of phylogenetics, and the (secondary) homology relations supporting the trees as the evidence for them. Both, of course, are subject to revision when new observations are made or existing methodological approaches are improved: these are indeed two areas where molecular phylogenetic approaches, most notably phylogenomics, has made unprecedented progress in recent years with a growing number of sequenced genomes of microbial as well as macroscopic organisms available for an increasingly sophisti-
cated set of methodological approaches, resolving phylogenies of problematic taxa that traditional approaches have struggled with for decades (as mentioned above).

Thus, the characterisation made here identifies phylogenetic trees with theories, homologies with evidence, and putative homologies as hypotheses, and physical attributes of taxa as observations, as well as the vital role of methodological approaches in the scientific process of phylogenetics puts into perspective the importance of not only the quality of data and their relation to hypotheses, but also that of finding the right methodological approach in analysing the data. Biased or insufficient data as well as inappropriate methods to deal with existing data can both lead to results reflective of the history of life on Earth. Nevertheless, integration of different kinds of evidence remains an outstanding concern. In the next part of the paper, I will attempt to extend the characterisation made here to incorporate this issue and argue why evidential integration is highly relevant to the reconstruction of the history of life on Earth.

**Inferring the best evolutionary scenario.**

**Evidential integration and meta-congruence**

As with any other science, specific questions in phylogenetics are generally best answered by providing specific lines of evidence, either obtained through the same or different methodological approaches, or of the same or different kinds. Thus, questions about patterns of gene and genome evolutionary lineages are generally best answered by molecular phylogenetic approaches; likewise, questions about patterns of morphological evolution are best answered by comparative or functional morphology. But this is an oversimplification of phylogenetics as it overlooks a crucial fact: evidence of any kind is *inevitably imperfect*: as seen above, for example, while molecular approaches are typically more powerful than morphological ones in resolving phylogenies, they nevertheless occasionally (albeit significantly) give rise to problematic cases through susceptibility to systematic errors, such as when certain clades undergo high rates of evolution. More or less the same can be said about morphological evidence.

However, phylogenetics is not restricted to questions of the sorts mentioned above. Rather, it is often questions of a more overarching sort, namely those concerned with phylogenetic history—i.e. genealogical relationships between taxa. Research into the two sorts of questions generally complement and inform each other, as more fine-tuned knowledge of the phylogenies of characters (= homologues) is crucial to constructing better phylogenies of taxa, which in turn facilitates inferences around character evolution (as touched upon in the previous section). And it is precisely in the context of these overarching questions that the value of evidential integration—i.e. the utilisation of different kinds of evidence—is revealed: as such shortcomings are inevitable for any line of evidence, a reliable course of action is often to rely on a different line of evidence, either of the same or a different kind, to *compensate for such shortcomings* with respect to a question of the overarching sort. In other words, it is precisely because questions of this sort are *overarching*—in the sense that they
inform and are informed by evidence pertaining to questions of the more specific sort—that evidence of different kinds can and arguably should be utilised in concordance with each other to provide answers to them.

A notable example concerns (once again) the phylogenetic placement of ctenophores. As discussed above, phylogenomic studies have resulted in a rift dividing two camps on this issue, already suggesting that molecular evidence might not be best for tackling this part of animal phylogeny. Furthermore, a recent study highlights how the hypothesis favoured by one camp (the ctenophore-first hypothesis) is likely supported due mainly to systematic errors in analysis related to high rates of ctenophore evolution, and that this indirectly favours the other (sponge-sister) hypothesis [citation]. Interestingly, this is concordant with a long tradition in comparative morphology placing Ctenophora within Eumetazoa as the sister group to Cnidaria within the clade Coelenterata (Philippe et al. 2009; Schierwater et al. 2009) or alternatively as sister to Bilateria within Triploblastica (Nielsen 2012). Moreover, the Coelenterata hypothesis has been further corroborated more recently in a palaeontological study (Zhao et al. 2019), which traces the ctenophore lineage through a series of early Cambrian fossils down to stem-group cnidarians. As the fossil record is the only more or less direct way of investigating evolutionary history, the existence of such an account of the emergence of ctenophores from stem-group coelenterates ought to tilt the balance towards a theory that places ctenophores within Eumetazoa—as opposed to being placed as the sister group to the rest of animals—even in the absence of the revelation of systematic errors.

Thus, the utilisation of different lines of evidence in this compensatory capacity is what I refer to as evidential integration. This is a special version of utilising different lines of evidence more generally whether of the same or different kinds, the former of which lies at the core of congruence in phylogenetics (Patterson 1988; Rieppel 2005, 2009); i.e. the support gained for a particular tree topology through congruence (=agreement) between characters supporting that topology, in turn also supporting the homologies of congruent characters (these here being the different “lines” of evidence). However, there is more to evidential integration than mere congruence: by utilising different kinds of evidence, it is possible to answer questions beyond simply what tree topology is best supported by the evidence presented; thus, as congruence results from the utilisation of different lines of evidence of the same kind, the result of evidential integration in the sense described here can be termed meta-congruence as it goes beyond mere congruence\(^4\).

Two types of questions where meta-congruence matters particularly besides inferring phylogenies are those surrounding (1) the inference of character homologies through the integration of different kinds of evidence (e.g. developmental and morphological), as well as (2) reconstructing evolutionary history informed by palaeontology, ecology, and biogeography. I will now briefly discuss how meta-congruence pertains to these two types of question in turn.

\(^4\)The account of evidential integration presented here is in the same spirit as that of triangulation (Munafò and Smith 2018): both are ultimately concerned with compensatory utilisation of different kinds of evidence using appropriate methods to infer the best causal explanation rather than the statistically best-supported result. However, the present account is distinct in that it is explicitly hierarchical.
Integrated homology inferences

The discussion presented thus far ought to have demonstrated the importance of the credibility of homology hypotheses: once again, these form the basis of any phylogenetic study. Regardless, the homology of numerous morphological characters of interest to macroevolutionary research is not inferred exclusively through comparative morphology, as the field of developmental genetics (broadly construed) has made significant progress in elucidating the molecular mechanisms underlying development at the cellular as well as tissue and organ levels; notable examples from animals include the development of animal limbs, eyes, cell types, nervous systems, and body plans (Arendt et al. 2016, 2019; Genikhovich et al. 2015; Genikhovich and Technau 2017; Martín-Durán and Hejnol 2019; Wagner 2018). One of the core tenets of developmental genetics is that differences in spatiotemporal patterns of gene expression are directly responsible for and therefore capable of explaining differences in morphological characters, which naturally leads to the conclusion that they should be informative for inferences of morphological homology.

Nevertheless, not everything matches up. Evidence from developmental genetics and from morphology disagree about as often as they corroborate each other on the homology of morphological characters, making it unclear how they should be weighed against each other in specific cases and whether there are any conditions under which one should be favoured over the other, nor is there any set of general principles or theoretical framework that can guide inferences of homology in light of both kinds of evidence—at least as of yet. This demonstrates the importance of meta-congruence not just at the level of inferring trees (as discussed above) but also at the more fundamental level of inferring putative homologues, as meta-congruence provides at least one general clue as to how evidential integration at this level can be achieved; namely, to utilise one source of evidence where the other is relatively inconclusive.

Evolutionary scenarios and cladograms

I will now explain how reconstructing evolutionary history goes beyond phylogeny reconstruction. The science of biological classification has for decades been largely a statistical science providing statistical explanations, concerned with variants on the following question: What is the most likely tree topology given the data presented? The trees obtained thus are, roughly speaking, referred to as either cladograms (if they lack a temporal element) or phylogenetic trees (if they contain one)\(^5\). The presence of a temporal element in the definition of a phylogenetic tree as opposed to a cladogram is indicative of the fact that the distinction between cladograms and phylogenetic trees is a bit subtler than this rough distinction. Cladograms *really are just statistical explanations*, whereas phylogenetic trees are supposed to be more reflective of actual evolutionary history: after all, the aim of phylogenetics is to reveal the *actual genealogical relationships* between taxa, which in turn result from real-

---

\(^5\) Alternatively, they are distinguished by whether degree of phenotypic distance is represented in branch lengths (phylogenetic tree) or not (cladogram). This supplements the present argument, but it is not crucial.
world, historical events of speciation and extinction. By “statistical explanation” here I mean an explanation that consists entirely of a nested hierarchy of similarities inferred from numerical data, and is therefore devoid of any causal links between the taxa under study; this stands in contrast with phylogenetic trees (and more so with evolutionary scenarios—see next paragraph), which do at least speculate on genealogical relationships between taxa by virtue of their reliance on inferred temporal or phenotypic distances between taxa. The genealogical relationships between taxa, as well as patterns of character evolution, form causally-connected series of events; thus, the explanations of these evolutionary series of events that are provided by phylogenetics or other branches of macroevolutionary science comprise, in a strong sense causal explanations in contrast to statistical ones. Therefore, on a spectrum ranging from purely statistical explanations to complete causal explanations, cladograms lie very close to the statistical end and phylogenetic trees lie somewhat further towards the causal end.

We have already seen how evidential integration can be useful in inferring the most likely phylogenetic tree through compensation by one source of evidence where another tends to be inconclusive. We have also seen how different kinds of evidence can, on their own, illuminate other aspects of macroevolutionary biology such as large-scale patterns of genomic or morphological evolution. Not only are these evolutionary patterns closely associated with phylogenetic trees, but they can also be integrated with phylogenies, alongside other key sources of evidence such as biogeography and ecology, to arrive at evolutionary scenarios, thereby informing phylogenies as well. An evolutionary scenario is here taken to refer to an overarching explanation consisting of the combination of phylogenetic trees in the sense described above and accurate reconstruction of the evolution of traits and the patterns they form—which consists of causally-related series of evolutionary events and processes.

Earlier, I used the example of the reconstruction of the stages of ctenophore evolution from stem-coelenterate ancestors, which demonstrates the importance of palaeontology in reconstructing phylogenies even in the face of evidence (of another kind) to the contrary. This is in fact an example of an evolutionary scenario as it elucidates not just genealogical relations between taxa, but also the series of the evolution of key characters belonging to these taxa—i.e. information about the presence or absence of such characters in ancestral species. It also highlights the importance of informing phylogeny inferences in light of the likelihood of one evolutionary scenario vs. another, where one is more likely because it better incorporates certain patterns of morphological evolution.

Palaeoecology and biogeography can be equally informative for constructing evolutionary scenarios, as has been discussed elsewhere in detail (Hall 2012; Williams and Ebach 2020). Here, I will use the evolution of planktotrophic larvae in animals as an example to demonstrate this point. A fairly popular theory of the origins and early evolution of Bilateria states that the earliest bilaterians were either microscopic free-swimming organisms feeding on plankton, or at least took this form at some point in their life-cycle (Nielsen 2019; Peterson et al. 2000). Both versions of the theory rest primarily on the widespread presence of planktotrophic larvae in modern bilaterians, as well as similarities between morphological traits belonging to such larvae, most
Beyond congruence: evidential integration and inferring the best…

notably ciliary bands used for swimming and food capture, leading to the conclusion that these structures are homologous across bilaterians possessing them and therefore possessed by the last bilaterian common ancestor. The issue here is therefore whether this is a correct inference of homology, or that these have independently evolved and are similar due to convergent evolution and possibly in part due to co-opted developmental mechanisms present in the last common ancestor. This latter hypothesis is in fact best supported by palaeoecological evidence: the ecological conditions for planktotrophy were in fact not yet present in the pre-Cambrian (when the last common bilaterian ancestor would have lived), but appeared later on in the late Cambrian (Peterson 2005).

We have therefore seen through various examples how evidential integration and meta-congruence can be of tremendous service at the two levels of phylogenetic inference—namely the identification of putative homologues and inferring phylogenies—as well as more overarching macroevolutionary problems revolving around the reconstruction of the evolutionary history of life on Earth (in which phylogenetics plays an irreplaceable part). This can be conceptualised as an extension of the model of phylogenetic inference described in the preceding part of this paper: here, phylogenetic inference is contextualised within the overarching aim of macroevolutionary biology, with evidential integration playing a significant role at each step of the process (Fig. 3).

With respect to the statistical-causal spectrum, evolutionary scenarios—the final result of evidential integration at multiple levels—are far closer to the causal end than phylogenetic trees, as they aim to reflect various aspects of several interrelated
and multifaceted series of real-world events in the immeasurably complex history of life on Earth. This ought to provide a strong argument for the utilisation of different kinds of evidence when approaching macroevolutionary biology and phylogenetics as is incorporated in this science, as well as being a powerful vindication of the causal notion of explanations described above.

**Relative weight and context-sensitivity of evidence**

In the preceding paragraphs, we have seen how the weight ought to be given to different kinds of evidence can vary in relation to other kinds in a context-dependence way. Based on this, we can arrive at a sketch of a principle way evidential integration can be operationalised. To start with, I emphasise that specific evidence always answers specific questions. The example at the focal point of this paper concerns molecular evidence, which only gives rise to molecular phylogenies. Crucially, these are no more than the cladograms that fit best with the molecular evidence itself—rather than being the best explanations of all available evidence. It is widely accepted nowadays that molecular evidence, in large part due to the sheer amount of data contained in it, leads to the best cladograms simpliciter. Here, I have tried to argue that, at least in certain contexts, molecular evidence can lead to results that contradict not only the bulk of existing morphological evidence, but also results from other carefully-conducted analyses based on essentially the same data. Such disagreements, of course, come down to the particular model used or the way in which data is “trimmed”, a detailed discussion of which is well beyond the present scope. Regardless, the fact remains that in certain contexts molecular evidence can deteriorate in overall value for phylogeny inference. In such cases, if morphological evidence (for instance) offers a more certainty, or a more coherent explanation overall, it ought to regain some of its weight in leading to the best phylogenetic tree. In other words, where one kind of evidence is less conclusive, another can fill a compensatory role. Seemingly, it just happens that molecular evidence tends to be more conclusive in many cases, less so in a few contentious ones.

This general principle of course holds with respect to integration of other kinds of evidence: while a phylogeny based on molecular or morphological evidence might suggest that, say, evolutionary scenario (a) is more likely than evolutionary scenario (b), palaeontological evidence might suggest otherwise—for example by showing that scenario (a) is incompatible with the order of appearance of certain species in the fossil record. In such a case, at least some weight ought to be detracted from the evidence supporting scenario (a). Alternatively, palaeontological evidence could point strongly towards a period when a diverse array of complex organisms were likely to thrive, followed by a period of widespread extinction. In such cases, it would be similarly permissible to opt for an evolutionary scenario that can account for this evidence, even if the best inferred cladogram does not support this scenario and only the second- or third-best cladogram does. In other words, palaeontological evidence gains a relatively greater weight in cases where it makes a certain scenario particularly difficult to refute, even if it is not best supported by available phylogenies.

To summarise, the choice of the best evolutionary scenario lies not just in simply considering all sources of evidence, but more vitally giving different kinds of
evidence different weights depending on (1) the number of independent data points they rely on (e.g. molecular evidence relies on enormous amounts of data, giving it a general advantage), (2) how conclusive they are in terms of strongly favouring one scenario over others (the ctenophore controversy is an example where molecular evidence is not conclusive in this sense), and (3) how difficult it is to refute a scenario supported by this evidence (e.g. order of appearance of species in fossil record). Ultimately, the best evolutionary scenario must take these differential weightings into account in a manner that does not render it internally inconsistent or incoherent.

Conclusions and prospects

To reiterate, the aim of this paper has been to stress the utility in phylogenetics (and beyond) of different kinds of evidence in compensating the inevitable shortcomings of any kind of evidence on its own; some data are just not best for answering some questions. This is obviously not a novel claim. Most scientists presumably know the value of evidential integration and use it on a regular basis. It is nevertheless a point worth stressing, at least in phylogenetics and macroevolutionary biology where it can be all too easy to underestimate the value of traditional approaches and their associated evidence—notably comparative morphology, palaeontology, biogeography, and ecology (though recent advances in integrating molecular and palaeontological evidence in calibrated molecular clocks (Beavan et al. 2020) are encouraging examples to the contrary).

Besides articulating this position for better conceptual clarity, this paper has also attempted to provide a sketch of a theoretical framework within which integrative inferences can be better understood and evaluated, and at least one general principle for integration—namely evidential compensation. The path to a sensu stricto explanatory macroevolutionary biology is through meta-congruence.

This, of course, does not mean integration at every possible level: it is counterintuitive and plausibly counter-productive to create a data matrix consisting of genomic and morphological characters, as the informative value of the latter is likely to be swamped by the sheer scale of the former. Thus, further theoretical research is needed to elucidate more precisely under what conditions and how evidential integration would be beneficial, in conjunction with empirical research utilising it (explicitly or implicitly). A notable outstanding challenge in this area with theoretical as well as empirical facets is, as already mentioned, integrated homology inference, especially regarding morphological and developmental evidence, which I hope will be addressed in future research. A potentially viable programme for future research will therefore include (1) careful consideration of phylogenies produced by molecular evidence and particularly avoiding overreliance on molecular phylogenies that are especially plagued with uncertainty, and in such cases (2) giving more weight to phylogenies based on other kinds of characters—notably morphological ones—while keeping inferences of morphological homology informed by more certain molecular as well as developmental evidence; and (3) situating phylogenies within broader macroevolutionary science, e.g. appealing to biogeography or palaeoecology in deciding between rival phylogenies or in giving weight to character losses vs. gains; (4) ultimately aiming for multifaceted, overarching explanations—evolutionary scenarios.
Acknowledgements  I would like to thank my former supervisor Dr Jakob Vinther (Bristol Palaeobiology) for instilling some of the early ideas in this paper and giving the initial motivation, Dr Jordi Paps Montserrat and Miss Maria Eleonora Rossi (Bristol Palaeobiology) for helpful discussions on phylogenomics, Professor Samir Okasha (Bristol Philosophy) for offering invaluable philosophical comments on the manuscript, and two anonymous reviewers for helping me improve the paper even further.

Funding:  This research was undertaken as part of my PhD project at the University of Bristol, partially funded by the university’s fee waiver scholarship for international students.

Data Availability  N/A.

Code Availability  N/A.

Conflicts of interest/Competing interests  N/A.

Ethics approval  N/A.

Consent to participate  N/A.

Consent for publication  N/A.

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Arendt D, Benito-Gutierrez E, Brunet T, Marlow H (2015) Gastric pouches and the mucociliary sole: setting the stage for nervous system evolution. Philosophical Trans Royal Soc B: Biol Sci 370(1684):20150286

Arendt D, Musser JM, Baker CV, Bergman A, Cepko C, Erwin DH, Laubichler MD (2016) The origin and evolution of cell types. Nat Rev Genet 17(12):744–757

Arendt D, Bertucci PY, Achim K, Musser JM (2019) Evolution of neuronal types and families. Curr Opin Neurobiol 56:144–152

Balavoine G, Adoutte A (2003) The segmented Urbilateria: a testable scenario. Integr Comp Biol 43(1):137–147

Beavan AJ, Donoghue PC, Beaumont MA, Pisani D (2020) Performance of a priori and a posteriori calibration strategies in divergence time estimation. Genome Biol Evol 12(7):1087–1098

Brauchle M, Bilican A, Eyer C, Bailly X, Martinez P, Ladurner P, Sprecher SG (2018) Xenacoelomorpha Survey Reveals That All 11 Animal Homeobox Gene Classes Were Present in the First Bilaterians. Genome Biol Evol 10(9):2205–2217. doi:https://doi.org/10.1093/gbe/evy170

Brigandt I (2007) Typology now: homology and developmental constraints explain evolvability. Biology & Philosophy 22(5):709–725

Budd GE, Jensen S (2017) The origin of the animals and a ‘Savannah’ hypothesis for early bilaterian evolution. Biol Rev 92(1):446–473
Beyond congruence: evidential integration and inferring the best...

Cai C, Tihelka E, Pisani D, Donoghue PC (2020) Data curation and modeling of compositional heterogeneity in insect phylogenomics: a case study of the phylogeny of Dytiscoidae (Coleoptera: Adephaga). Mol Phylogenet Evol 147:106782

Chen H, Parry LA, Vinther J, Zhai D, Hou X, Ma X (2020) A Cambrian crown annelid reconciles phylogenomics and the fossil record. Nature, 1–4

DiFrisco J, Love AC, Wagner GP (2020) Character identity mechanisms: a conceptual model for comparative-mechanistic biology. Biology & Philosophy 35(4):1–32

DiFrisco J, Jaeger J (2021) Homology of process: developmental dynamics in comparative biology. Interface Focus 11(3):20210007

Donoghue PC, Benton MJ (2007) Rocks and clocks: calibrating the Tree of Life using fossils and molecules. Trends Ecol Evol 22(8):424–431

Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Edgecombe GD (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452(7188):745–749

Eernisse DJ, Albert JS, Anderson FE (1992) Annelida and Arthropoda are not sister taxa: a phylogenetic analysis of spiralian metazoan morphology. Syst Biol 41(3):305–330

Feuda R, Dohrmann M, Pett W, Philippe H, Rota-Stabelli O, Lartillot N, Pisani D (2017) Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. Curr Biol 27(24):3864–3870e3864

Genikhovich G, Fried P, Prünster MM, Schinko JB, Gilles AF, Fredman D, Technau U (2015) Axis patterning by BMPs: cnidarian network reveals evolutionary constraints. Cell Rep 10(10):1646–1654

Genikhovich G, Technau U (2017) On the evolution of bilaterality. Development 144(19):3392–3404

Giribet G (2016) New animal phylogeny: future challenges for animal phylogeny in the age of phylogenomics. Organisms Divers Evol 16(2):419–426

Griffiths PE (2007) The phenomena of homology. Biology & Philosophy 22(5):643–658

Hall BK (2012) Homology: The hierarchial basis of comparative biology. Academic Press

He S, Del Viso F, Chen C-Y, Ikmi A, Kroesen AE, Gibson MC (2018) An axial Hox code controls tissue segmentation and body patterning in Nematostella vectensis. Science 361(6409):1377–1380

Hejnol A, Martindale MQ (2008) Acoel development supports a simple planula-like urbilateral. Philosophical Trans Royal Soc B: Biol Sci 363(1496):1493–1501

Jacobs DK, Hughes NC, Fitz-Gibbon ST, Winchell CJ (2005) Terminal addition, the Cambrian radiation and the Phanerozoic evolution of bilaterian form. Evol Dev 7(6):498–514

Jeffroy O, Brinkmann H, Delsuc F, Philippe H (2006) Phylogenomics: the beginning of incongruence? Trends Genet 22(4):225–231

Kapli P, Telford MJ (2020) Topology-dependent asymmetry in systematic errors affects phylogenetic placement of Ctenophora and Xenacoelomorpha. Sci Adv 6(50):eabc5162

Kapli P, Flouri T, Telford MJ (2021a) Systematic errors in phylogenetic trees. Curr Biol 31(2):R59–R64

Kapli P, Natsisidis P, Leite DJ, Fursman M, Jeffrie N, Rahman IA, Telford MJ (2021b) Lack of support for Deuterostomia prompts reinterpretation of the first Bilateria. Sci Adv 7(12):eabc2741

Kocot KM, Cannon JT, Todt C, Citarella MR, Kohn AB, Meyer A, Lieb B (2011) Phylogenomics reveals deep molluscan relationships. Nature 477(7365):452–456

Kocot KM, Struck TH, Merkel J, Waits DS, Todt C, Brannock PM, Lieb B (2017) Phylogenomics of Lophotrochozoa with consideration of systematic error. Syst Biol 66(2):256–282

Laumer CE, Gruber-Vodicka H, Hadfield MG, Pease VB, Riesgo A, Marioni JC, Giribet G (2018) Support for a clade of Placozoa and Cnidaria in genes with minimal compositional bias. Elife 7:e36278

Laumer CE, Fernández R, Lemer S, Combsorch D, Kocot KM, Riesgo A, Giribet G (2019) Revisiting metazoan phylogeny with genomic sampling of all phyla. Proceedings of the Royal Society B, 286(1906), 20190831

Lipton P (2004) Evidence to the best explanation. Routledge, In: London and New York

Lozano-Fernandez J, Giacomelli M, Fleming JF, Chen A, Vinther J, Thomsen PF, Iliffe TM (2019) Pancrustacean evolution illuminated by taxon-rich genomic-scale data sets with an expanded remipede sampling. Genome Biol Evol 11(8):2055–2070

Martín-Durán JM, Pang K, Barve A, Lê HS, Furu A, Cannon JT, Hejnol A (2018) Convergent evolution of bilaterian nerve cords. Nature 553(7686):45–50

Martín-Durán JM, Hejnol A (2019) A developmental perspective on the evolution of the nervous system. Developmental biology

Munafò MR, Smith GD (2018) Robust research needs many lines of evidence. In: Nature Publishing Group
Nagy LG, Merényi Z, Hegedüs B, Bálint B (2020) Novel phylogenetic methods are needed for understanding gene function in the era of mega-scale genome sequencing. Nucleic Acids Res 48(5):2209–2219

Natsidis P, Kapli P, Schiffer PH, Telford MJ (2020) Systematic errors in orthology inference: a bug or a feature for evolutionary analyses? bioRxiv

Nielsen C (2010) After all: Xenoturbella is an acoelomorph! Evol Dev 12(3):241–243

Nielsen C (2012) Animal evolution: interrelationships of the living phyla. Oxford University Press on Demand

Nielsen C, Brunet T, Arendt D (2018) Evolution of the bilaterian mouth and anus. Nat Ecol Evol 2(9):1358–1376

Nielsen C (2019) Early animal evolution: a morphologist’s view. Royal Soc open Sci 6(7):190638

Norén M, Jondelius U (1997) Xenoturbella’s molluscan relatives”. Nature 390(6655):31–32

Northcutt RG (2010) Cladistic analysis reveals brainless urbilateria. Brain Behav Evol 76(1):1

Novick A (2018) The fine structure of ‘homology’. Biology & Philosophy 33(1–2):6

Paps J (2018) What makes an animal? The molecular quest for the origin of the animal kingdom. Integr Comp Biol 58(4):654–665

Patterson C (1982) Morphological characters and homology. Problems of phylogenetic reconstruction, 21–74

Patterson C (1988) Homology in classical and molecular biology. Mol Biol Evol 5(6):603–625

Peterson KJ, Cameron RA, Davidson EH (2000) Bilateral origins: significance of new experimental observations. Dev Biol 219(1):1–17

Peterson KJ (2005) Macroevolutionary interplay between planktic larvae and benthic predators. Geology 33(12):929–932

Philippe H, Brinkmann H, Martinez P, Riutort M, Baguna J (2007) Acoel flatworms are not platyhelminthes: evidence from phylogenomics. PLoS one, 2(8), e717

Philippe H, Derelle R, Lopez P, Pick K, Borchielini C, Boury-Ensault N, Quéinnec E (2009) Phylogenomics revives traditional views on deep animal relationships. Curr Biol 19(8):706–712

Pisani D, Pett W, Dohrmann M, Feuda R, Rota-Stabelli O, Philippe H, Wörheide G (2015) Genomic data do not support comb jellies as the sister group to all other animals. Proceedings of the National Academy of Sciences, 112(50), 15402–15407

Pousta AJ, Isabella S, Halanych KM, Schrödl M (2020) New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships. Scientific Reports (Nature Publisher Group), 10(1)

Prpic N-M (2019) Limb Development: A lesson in homology. Elife 8:e48335

Ramsey G, Siebels Peterson A (2012) Sameness in biology. Philos Sci 79(2):255–275

Remane A (1956) Die Grundlagen des natürlichen Systems, der vergleichenden Anatomie und der Phylogenetic: theoretische Morphologie und Systematik

Rieppel O (2005) The philosophy of total evidence and its relevance for phylogenetic inference. Papeis Avulsos de Zoologia 45(8):77–89

Rieppel O (2009) ‘Total evidence’in phylogenetic systematics. Biology & Philosophy 24(5):607–622

Rohde K, Watson N, CANNON LG (1988) Ultrastructure of epidermal cilia of Pseudactinoposthia sp.(Platyhelminthes, Acoela); implications for the phylogenetic status of the Xenoturbellida and Acoelomorpha. J submicroscopic Cytol 20(4):759–767

Rota-Stabelli O, Campbell L, Brinkmann H, Edgecombe GD, Longhorn SJ, Peterson KJ, Telford MJ (2011) A congruent solution to arthropod phylogeny: phylogenomics, microRNAs and morphology support monophyletic Mandibulata. Proceedings of the Royal Society B: Biological Sciences, 278(1703), 298–306

Ruiz-Trillo I, Riutort M, Fourcade HM, Baguñà J, Boore JL (2004) Mitochondrial genome data support the basal position of Acoelomorpha and the polyphyly of the Platyhelminthes. Mol Phylogenet Evol 33(2):321–332

Schierwater B, Eitel M, Jakob W, Osigus H-J, Hadrys H, Dellaporta SL, DeSalle R (2009) Concatenated analysis sheds light on early metazoan evolution and fuels a modern “urmetazoon” hypothesis.PLoS Biol, 7(1), e1000020

Simion P, Philippe H, Baurain D, Jager M, Richter DJ, Di Franco A, Ereskovsky A (2017) A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. Curr Biol 27(7):958–967

Steinmetz PR, Aman A, Kraus JE, Technau U (2017) Gut-like ectodermal tissue in a sea anemone challenges germ layer homology. Nat Ecol Evol 1(10):1535–1542

Struck TH, Paul C, Hill N, Hartmann S, Hösel C, Kube M, Purschke G (2011) Phylogenomic analyses unravel amniid evolution. Nature 471(7336):95–98
Beyond congruence: evidential integration and inferring the best...

Tarazona OA, Lopez DH, Slota LA, Cohn MJ (2019) Evolution of limb development in cephalopod mollusks. Elife 8:e43828

Telford MJ, Littlewood DTJ (2009) Animal evolution: genomes, fossils, and trees. Oxford University Press

Telford MJ, Budd GE, Philippe H (2015) Phylogenomic insights into animal evolution. Curr Biol 25(19):R876–R887

Tihelka E, Cai C, Giacomelli M, Pisani D, Donoghue PC (2020) Integrated phylogenomic and fossil evidence of stick and leaf insects (Phasmatodea) reveal a Permian–Triassic co-origination with insectivores. Royal Soc open Sci 7(11):201689

Van Valen LM (1982) Homology and causes. J Morphol 173(3):305–312

Vinther J, Parry L, Briggs DE, Van Roy P (2017) Ancestral morphology of crown-group molluscs revealed by a new Ordovician stem aculiferan. Nature 542(7642):471–474

Wagner GP (2007) The developmental genetics of homology. Nat Rev Genet 8(6):473–479

Wagner GP (2018) Homology, genes, and evolutionary innovation. princeton university press

Williams DM, Ebach MC (2020) Cladistics. Cambridge University Press

Williams TA, Cox CJ, Foster PG, Szöllősi GJ, Embley TM (2020) Phylogenomics provides robust support for a two-domains tree of life. Nat Ecol Evol 4(1):138–147

Zhao Y, Vinther J, Parry LA, Wei F, Green E, Pisani D, Cong P (2019) Cambrian sessile, suspension feeding stem-group ctenophores and evolution of the comb jelly body plan. Curr Biol 29(7):1112–1125e1112

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.