

Beyond congruence: evidential integration and inferring the best evolutionary scenario

Arsham Nejad Kourki; School of Biological Sciences, University of Bristol; Life Sciences Building, 24 Tyndall Ave, Bristol, United Kingdom, BS8 1TQ

Integration of evidence from diverse fields is required for inferring the best evolutionary scenario

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 collectively transformed our understanding of the evolutionary history of animals. Moreover, the diversity of methods and models within molecular phylogenetics has resulted in significant disagreement among molecular phylogenies as well as between these and traditional phylogenies. Here, I argue that tackling this multifaceted problem lies in integrating evidence to infer the best evolutionary scenario. I begin with an overview of recent developments in early metazoan phylogenetics, followed by a discussion of key conceptual issues in phylogenetics revolving around phylogenetic evidence and theory. I then argue that integration of different kinds of evidence is necessary for arriving at the best evolutionary scenario rather than the best-fitting cladogram. Finally, I discuss the prospects of this view in stimulating interdisciplinary cross-talk in early metazoan research and beyond.

Keywords: evidential integration; causal explanation; early animal evolution; phylogenetics; macroevolution; evolutionary scenario; cross-disciplinary research

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 [1-9], often calibrated by fossils [10, 11]. 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 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 subjectivity—has given way to uncertainties around incongruence [12], matching models to huge quantities of data [12-14], and criteria for model choice [15, 16]. Molecular approaches are not flawless.

The primary implication for phylogenetics is that molecular approaches, and thus the resulting phylogenies, cannot be deemed categorically superior to traditional ones. This is especially important because, unsurprisingly, 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 [3, 4, 6, 15, 17]; the placement of Placozoa as either sister to Eumetazoa or belonging to crown-group Eumetazoa [4, 17, 18]; the placement of Xenacoelomorpha as sister to

Nephrozoa or deep within Bilateria [15, 19, 20]; and the relationship between arthropods and annelids [21, 22]. Of these, only the last can be considered “resolved” in any meaningful way (Fig. 1).

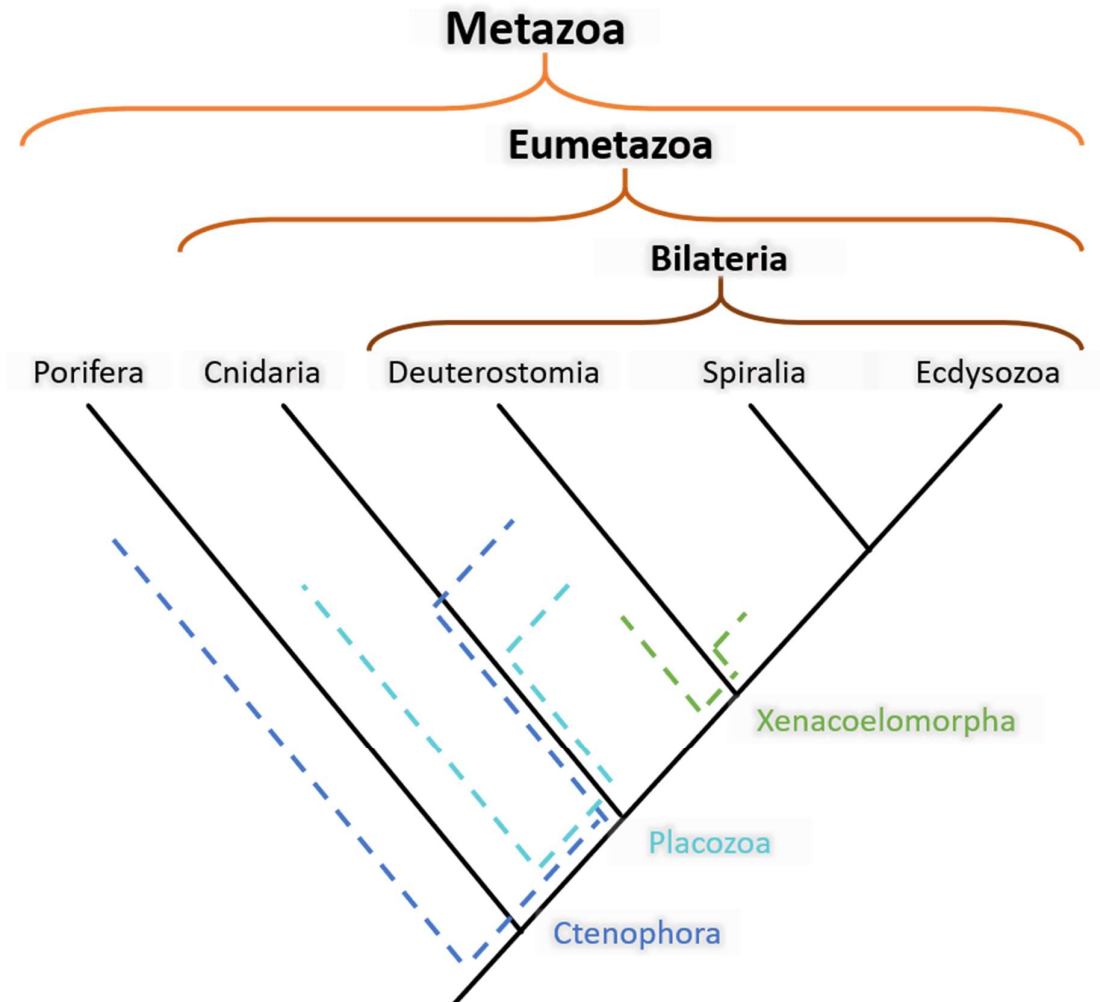


Figure 1: Unresolved relationships between major animal clades. The ctenophore lineage (navy blue) is shown either as sister to all other animals or to Cnidaria (a third possibility not shown here is sister to Bilateria). The placozoan lineage (cyan) is shown either as sister to Eumetazoa or to Cnidaria. Xenacoelomorpha (olive green) is shown either as sister to the rest of Bilateria (=Nephrozoa) or within Deuterostomia.

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. Next, I attempt to provide a solution to the problem of disagreement more generally, 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). This line of argument calls for multidisciplinary research in phylogenetics by providing a theoretical framework centred around evidential integration.

From homologues to cladograms

Elusive homology

Homology is an elusive concept [23], and homologues (i.e. homologous characters) can be equally hard to identify [24]. The two closely-related concepts have a long history stretching back to the late 18th century and are, at their core, concerned with *sameness* [25]. There are nevertheless several existing definitions of homology [23, 26, 27] 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 [28].

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 [29] or the convergent evolution of “limbs” in cephalopods, vertebrates, and arthropods [30, 31]. Examples of homologue identification abound, ranging from the unequivocally homologous such as the vertebrate forelimb or the insect hindwing 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 [28, 32] and Remane [33] (also see [34, 35]). 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; in such cases, the alternative approach, namely identifying and utilising homologous gene families, is potentially more appropriate [36].

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 [13] 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 [37], Annelida [38], and the recovery of Lophotrochozoa at the expense of Articulata [39]—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 [25, 27, 34, 35, 40-43], 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 [44-48]. 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 and needs to be addressed in future work/is the topic of the next chapter; however, it is something to which I will briefly return later on 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 [49-52]. The abundance of morphologically relatively simple bilaterian phyla [53], 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 [54-56] 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. 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 on them 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 [3, 4, 17] and xenacoelomorphs [15, 20]. 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 [3], 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 [53] as well as numerous molecular phylogenies published since [4, 6, 17]. 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 [57], and the xenoturbellids, previously included withing bivalves [58] and later Deuterostomia [59]. Since its discovery as a clade through molecular phylogenies [60] and subsequent corroboration by the identification of morphological synapomorphies for the clade [53], 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 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 [16]. 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 [15]. 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 [61]). 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 (Figure 2).

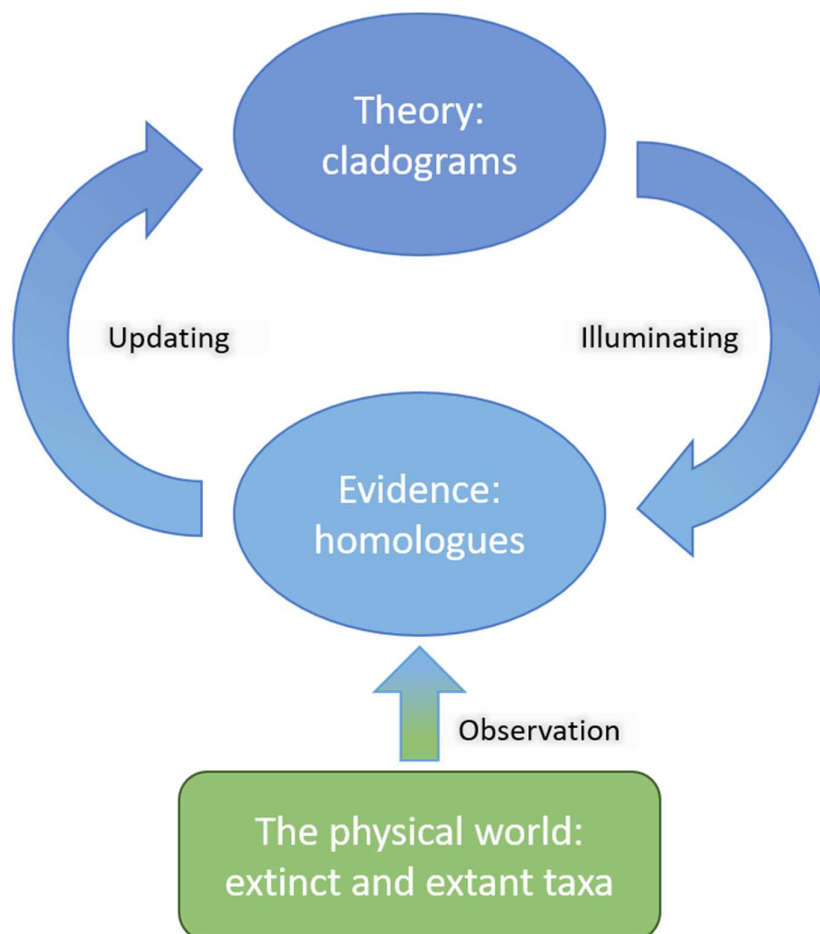


Figure 2: The cyclical relationship between theory, evidence, and observations in phylogenetics. Observations *become* evidence in light of theories, which in turn update the theories by favouring hypotheses over another. In phylogenetics, these broadly amount to cladograms, homologues, and physico-chemical features of extinct and extant taxa.

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 [23, 24].

¹ Though the distinction matters here, some authors in fact use the terms “theory” and “hypothesis” interchangeably.

² Roughly speaking, a semaphoronts 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.

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 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 sophisticated 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 irreflexive 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 [cs]. 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. 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 [62, 63] or alternatively as sister to Bilateria within Triploblastica [53]. Moreover, the Coelenterata hypothesis has been further corroborated more recently in a palaeontological study [64], 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 [28, 65, 66]; 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.

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.

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 [40, 44, 67-70]. 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)⁴. 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-world, historical events of speciation and extinction. 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. 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 absent of such characters in ancestral species.

Palaeoecology and biogeography can be equally informative for constructing evolutionary scenarios, as has been discussed elsewhere in detail [23, 24]. 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 [71, 72]. 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 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 planktrotrophy were in fact not

⁴ 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 is not crucial.

yet present in the pre-Cambrian (when the last common bilaterian ancestor would have lived), but appeared later on in the late Cambrian [73].

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).

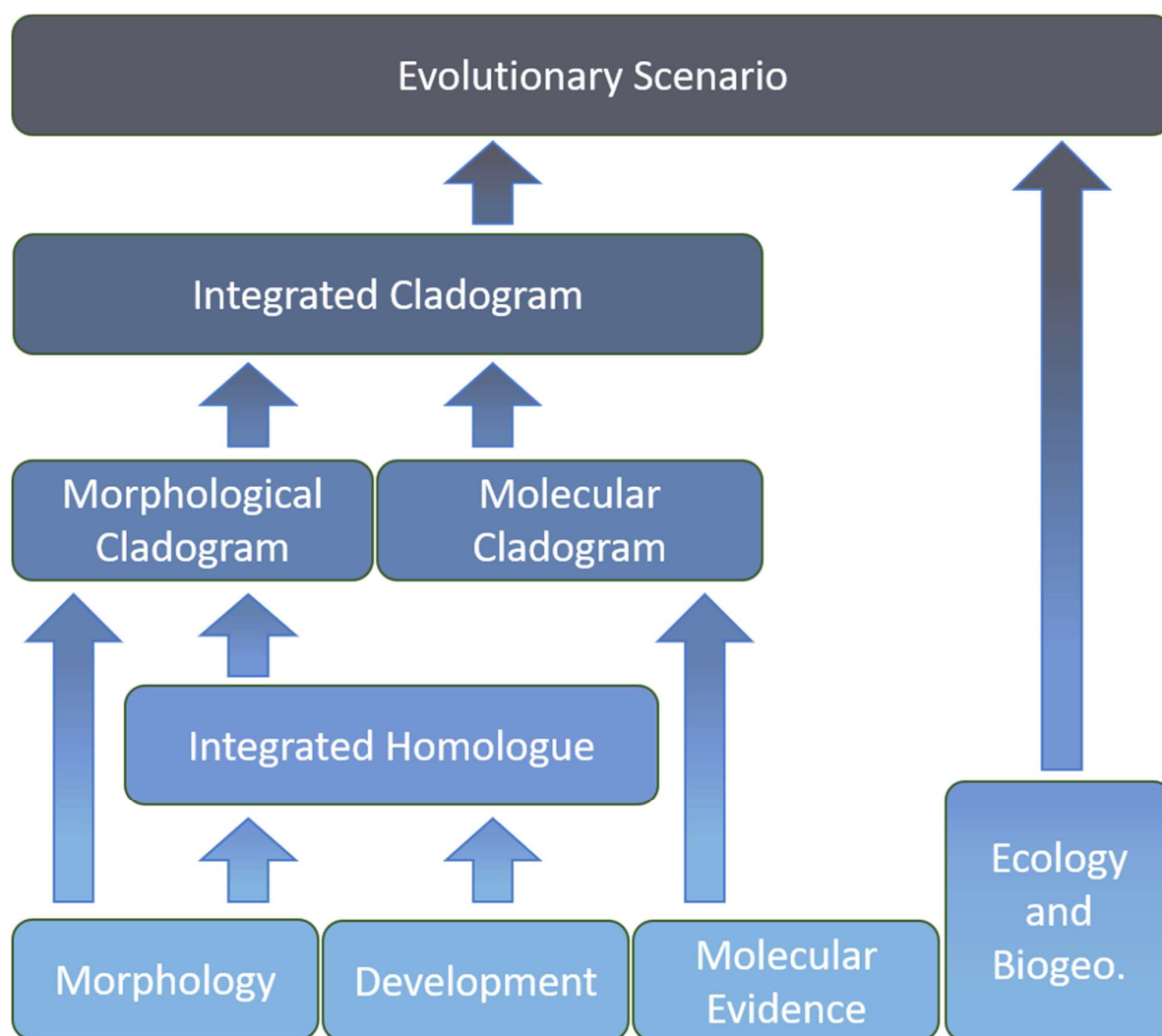


Figure 3: Evidential integration at several levels leading to construction of an Evolutionary Scenario—a multifaceted causal explanation encompassing several lines of evidence. Downward arrows (illumination of evidence by theory) are omitted for the sake of simplicity.

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. Causal explanations are considered in mainstream philosophy of science to be the most paradigmatic scientific explanations [61], the key reason here being that causal explanations *aim to directly reflect real-world series of events*, in this case the 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.

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 just aren't 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 [74] is an encouraging example 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 counter-intuitive 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.

Acknowledgements

This research was carried out as part of my PhD project at the School of Biological Sciences, University of Bristol, partially funded by the university's fee waiver scholarship for international students. I would especially like to thank my 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 (Bristol Palaeobiology) for a helpful discussion on phylogenomics, and Professor Samir Okasha for offering invaluable philosophical comments on the manuscript.

References

1. Paps, J., *What makes an animal? The molecular quest for the origin of the animal kingdom*. Integrative and comparative biology, 2018. **58**(4): p. 654-665.
2. Telford, M.J. and D.T.J. Littlewood, *Animal evolution: genomes, fossils, and trees*. 2009: Oxford University Press.
3. Dunn, C.W., et al., *Broad phylogenomic sampling improves resolution of the animal tree of life*. Nature, 2008. **452**(7188): p. 745-749.
4. Simion, P., et al., *A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals*. Current Biology, 2017. **27**(7): p. 958-967.

5. Laumer, C.E., et al., *Revisiting metazoan phylogeny with genomic sampling of all phyla*. Proceedings of the royal society B, 2019. **286**(1906): p. 20190831.
6. Feuda, R., et al., *Improved modeling of compositional heterogeneity supports sponges as sister to all other animals*. Current Biology, 2017. **27**(24): p. 3864-3870. e4.
7. Rota-Stabelli, O., et al., *A congruent solution to arthropod phylogeny: phylogenomics, microRNAs and morphology support monophyletic Mandibulata*. Proceedings of the Royal Society B: Biological Sciences, 2011. **278**(1703): p. 298-306.
8. Lozano-Fernandez, J., et al., *Pancrustacean evolution illuminated by taxon-rich genomic-scale data sets with an expanded remipede sampling*. Genome biology and evolution, 2019. **11**(8): p. 2055-2070.
9. Williams, T.A., et al., *Phylogenomics provides robust support for a two-domains tree of life*. Nature ecology & evolution, 2020. **4**(1): p. 138-147.
10. Tihelka, E., et al., *Integrated phylogenomic and fossil evidence of stick and leaf insects (Phasmatodea) reveal a Permian–Triassic co-origination with insectivores*. Royal Society open science, 2020. **7**(11): p. 201689.
11. Donoghue, P.C. and M.J. Benton, *Rocks and clocks: calibrating the Tree of Life using fossils and molecules*. Trends in ecology & evolution, 2007. **22**(8): p. 424-431.
12. Jeffroy, O., et al., *Phylogenomics: the beginning of incongruence?* TRENDS in Genetics, 2006. **22**(4): p. 225-231.
13. Cai, C., et al., *Data curation and modeling of compositional heterogeneity in insect phylogenomics: a case study of the phylogeny of Dytiscoidea (Coleoptera: Adephaga)*. Molecular phylogenetics and evolution, 2020. **147**: p. 106782.
14. Nagy, L.G., et al., *Novel phylogenetic methods are needed for understanding gene function in the era of mega-scale genome sequencing*. Nucleic acids research, 2020. **48**(5): p. 2209-2219.
15. Kapli, P. and M.J. Telford, *Topology-dependent asymmetry in systematic errors affects phylogenetic placement of Ctenophora and Xenacoelomorpha*. Science advances, 2020. **6**(50): p. eabc5162.
16. Kapli, P., T. Flouri, and M.J. Telford, *Systematic errors in phylogenetic trees*. Current Biology, 2021. **31**(2): p. R59-R64.
17. Pisani, D., et al., *Genomic data do not support comb jellies as the sister group to all other animals*. Proceedings of the National Academy of Sciences, 2015. **112**(50): p. 15402-15407.
18. Laumer, C.E., et al., *Support for a clade of Placozoa and Cnidaria in genes with minimal compositional bias*. Elife, 2018. **7**: p. e36278.
19. Kapli, P., et al., *Lack of support for Deuterostomia prompts reinterpretation of the first Bilateria*. Science Advances, 2021. **7**(12): p. eabe2741.
20. Brauchle, M., et al., *Xenacoelomorpha Survey Reveals That All 11 Animal Homeobox Gene Classes Were Present in the First Bilaterians*. Genome Biology and Evolution, 2018. **10**(9): p. 2205-2217.
21. Eernisse, D.J., J.S. Albert, and F.E. Anderson, *Annelida and Arthropoda are not sister taxa: a phylogenetic analysis of spiralian metazoan morphology*. Systematic Biology, 1992. **41**(3): p. 305-330.
22. Telford, M.J., G.E. Budd, and H. Philippe, *Phylogenomic insights into animal evolution*. Current Biology, 2015. **25**(19): p. R876-R887.
23. Hall, B.K., *Homology: The hierarchical basis of comparative biology*. 2012: Academic Press.
24. Williams, D.M. and M.C. Ebach, *Cladistics*. 2020: Cambridge University Press.
25. Ramsey, G. and A. Siebels Peterson, *Sameness in biology*. Philosophy of Science, 2012. **79**(2): p. 255-275.
26. Van Valen, L.M., *Homology and causes*. Journal of Morphology, 1982. **173**(3): p. 305-312.
27. Novick, A., *The fine structure of 'homology'*. Biology & Philosophy, 2018. **33**(1-2): p. 6.

28. Patterson, C., *Homology in classical and molecular biology*. Molecular biology and evolution, 1988. **5**(6): p. 603-625.
29. Martín-Durán, J.M., et al., *Convergent evolution of bilaterian nerve cords*. Nature, 2018. **553**(7686): p. 45-50.
30. Tarazona, O.A., et al., *Evolution of limb development in cephalopod mollusks*. Elife, 2019. **8**: p. e43828.
31. Prpic, N.-M., *Limb Development: A lesson in homology*. Elife, 2019. **8**: p. e48335.
32. Patterson, C., *Morphological characters and homology*. Problems of phylogenetic reconstruction, 1982: p. 21-74.
33. Remane, A., *Die Grundlagen des natürlichen Systems, der vergleichenden Anatomie und der Phylogenetik: theoretische Morphologie und Systematik*. 1956.
34. Griffiths, P.E., *The phenomena of homology*. Biology & Philosophy, 2007. **22**(5): p. 643-658.
35. DiFrisco, J., A.C. Love, and G.P. Wagner, *Character identity mechanisms: a conceptual model for comparative-mechanistic biology*. Biology & Philosophy, 2020. **35**(4): p. 1-32.
36. Natsidis, P., et al., *Systematic errors in orthology inference: a bug or a feature for evolutionary analyses?* bioRxiv, 2020.
37. Poustka, A.J., et al., *New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships*. Scientific Reports (Nature Publisher Group), 2020. **10**(1).
38. Struck, T.H., et al., *Phylogenomic analyses unravel annelid evolution*. Nature, 2011. **471**(7336): p. 95-98.
39. Kocot, K.M., et al., *Phylogenomics of Lophotrochozoa with consideration of systematic error*. Systematic Biology, 2017. **66**(2): p. 256-282.
40. Wagner, G.P., *Homology, genes, and evolutionary innovation*. 2018: princeton university press.
41. Wagner, G.P., *The developmental genetics of homology*. Nature Reviews Genetics, 2007. **8**(6): p. 473-479.
42. DiFrisco, J. and J. Jaeger, *Homology of process: developmental dynamics in comparative biology*. Interface Focus, 2021. **11**(3): p. 20210007.
43. Brigandt, I., *Typology now: homology and developmental constraints explain evolvability*. Biology & Philosophy, 2007. **22**(5): p. 709-725.
44. Genikhovich, G. and U. Technau, *On the evolution of bilaterality*. Development, 2017. **144**(19): p. 3392-3404.
45. Steinmetz, P.R., et al., *Gut-like ectodermal tissue in a sea anemone challenges germ layer homology*. Nature ecology & evolution, 2017. **1**(10): p. 1535-1542.
46. Nielsen, C., T. Brunet, and D. Arendt, *Evolution of the bilaterian mouth and anus*. Nature ecology & evolution, 2018. **2**(9): p. 1358-1376.
47. Arendt, D., et al., *Gastric pouches and the mucociliary sole: setting the stage for nervous system evolution*. Philosophical Transactions of the Royal Society B: Biological Sciences, 2015. **370**(1684): p. 20150286.
48. He, S., et al., *An axial Hox code controls tissue segmentation and body patterning in Nematostella vectensis*. Science, 2018. **361**(6409): p. 1377-1380.
49. Balavoine, G. and A. Adoutte, *The segmented Urbilateria: a testable scenario*. Integrative and Comparative Biology, 2003. **43**(1): p. 137-147.
50. Northcutt, R.G., *Cladistic analysis reveals brainless urbilateria*. Brain, behavior and evolution, 2010. **76**(1): p. 1.
51. Hejnol, A. and M.Q. Martindale, *Acoel development supports a simple planula-like urbilaterian*. Philosophical Transactions of the Royal Society B: Biological Sciences, 2008. **363**(1496): p. 1493-1501.
52. Jacobs, D.K., et al., *Terminal addition, the Cambrian radiation and the Phanerozoic evolution of bilaterian form*. Evolution & development, 2005. **7**(6): p. 498-514.

53. Nielsen, C., *Animal evolution: interrelationships of the living phyla*. 2012: Oxford University Press on Demand.
54. Chen, H., et al., *A Cambrian crown annelid reconciles phylogenomics and the fossil record*. Nature, 2020: p. 1-4.
55. Vinther, J., et al., *Ancestral morphology of crown-group molluscs revealed by a new Ordovician stem aculiferan*. Nature, 2017. **542**(7642): p. 471-474.
56. Budd, G.E. and S. Jensen, *The origin of the animals and a 'Savannah' hypothesis for early bilaterian evolution*. Biological Reviews, 2017. **92**(1): p. 446-473.
57. Rohde, K., N. Watson, and L.G. CANNON, *Ultrastructure of epidermal cilia of Pseudactinoposthia sp. (Platyhelminthes, Acoela); implications for the phylogenetic status of the Xenoturbellida and Acoelomorpha*. Journal of submicroscopic cytology, 1988. **20**(4): p. 759-767.
58. Norén, M. and U. Jondelius, *Xenoturbella's molluscan relatives* ". Nature, 1997. **390**(6655): p. 31-32.
59. Nielsen, C., *After all: Xenoturbella is an acoelomorph!* Evolution & development, 2010. **12**(3): p. 241-243.
60. Philippe, H., et al., *Acoel flatworms are not platyhelminthes: evidence from phylogenomics*. PLoS one, 2007. **2**(8): p. e717.
61. Lipton, P., *Evidence to the best explanation*. 2004, London and New York: Routledge.
62. Schierwater, B., et al., *Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoon" hypothesis*. PLoS Biol, 2009. **7**(1): p. e1000020.
63. Philippe, H., et al., *Phylogenomics revives traditional views on deep animal relationships*. Current Biology, 2009. **19**(8): p. 706-712.
64. Zhao, Y., et al., *Cambrian sessile, suspension feeding stem-group ctenophores and evolution of the comb jelly body plan*. Current Biology, 2019. **29**(7): p. 1112-1125. e2.
65. Rieppel, O., *The philosophy of total evidence and its relevance for phylogenetic inference*. Papeis Avulsos de Zoologia, 2005. **45**(8): p. 77-89.
66. Rieppel, O., *'Total evidence' in phylogenetic systematics*. Biology & Philosophy, 2009. **24**(5): p. 607-622.
67. Genikhovich, G., et al., *Axis patterning by BMPs: cnidarian network reveals evolutionary constraints*. Cell reports, 2015. **10**(10): p. 1646-1654.
68. Arendt, D., et al., *Evolution of neuronal types and families*. Current opinion in neurobiology, 2019. **56**: p. 144-152.
69. Arendt, D., et al., *The origin and evolution of cell types*. Nature Reviews Genetics, 2016. **17**(12): p. 744-757.
70. Martín-Durán, J.M. and A. Hejnol, *A developmental perspective on the evolution of the nervous system*. Developmental biology, 2019.
71. Nielsen, C., *Early animal evolution: a morphologist's view*. Royal Society open science, 2019. **6**(7): p. 190638.
72. Peterson, K.J., R.A. Cameron, and E.H. Davidson, *Bilaterian origins: significance of new experimental observations*. Developmental biology, 2000. **219**(1): p. 1-17.
73. Peterson, K.J., *Macroevolutionary interplay between planktic larvae and benthic predators*. Geology, 2005. **33**(12): p. 929-932.
74. Beavan, A.J., et al., *Performance of a priori and a posteriori calibration strategies in divergence time estimation*. Genome biology and evolution, 2020. **12**(7): p. 1087-1098.