Current Biology
Review

Embracing Uncertainty in Reconstructing Early Animal Evolution

Nicole King[1] and Antonis Rokas[2]
[1]Howard Hughes Medical Institute and Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA
[2]Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA
Correspondence: nking@berkeley.edu (N.K.), antonis.rokas@vanderbilt.edu (A.R.)
http://dx.doi.org/10.1016/j.cub.2017.08.054

The origin of animals, one of the major transitions in evolution, remains mysterious. Many key aspects of animal evolution can be reconstructed by comparing living species within a robust phylogenetic framework. However, uncertainty remains regarding the evolutionary relationships between two ancient animal lineages — sponges and ctenophores — and the remaining animal phyla. Comparative morphology and some phylogenomic analyses support the view that sponges represent the sister lineage to the rest of the animals, while other phylogenomic analyses support ctenophores, a phylum of carnivorous, gelatinous marine organisms, as the sister lineage. Here, we explore why different studies yield different answers and discuss the implications of the two alternative hypotheses for understanding the origin of animals. Reconstruction of ancient evolutionary radiations is devilishly difficult and will likely require broader sampling of sponge and ctenophore genomes, improved analytical strategies and critical analyses of the phylogenetic distribution and molecular mechanisms underlying apparently conserved traits. Rather than staking out positions in favor of the ctenophores-sister or the sponges-sister hypothesis, we submit that research programs aimed at understanding the biology of the first animals should instead embrace the uncertainty surrounding early animal evolution in their experimental designs.

> "Science works on the frontier between knowledge and ignorance. We're not afraid to admit what we don't know. There's no shame in that. The only shame is to pretend that we have all the answers."
>
> —Neil deGrasse Tyson, Cosmos: A Spacetime Odyssey

Introduction

Knowledge of phylogenetic relationships is critical for understanding the evolution of genes, gene regulatory pathways and morphological traits, as well as for generating and testing evolutionary hypotheses about life's major transitions. To generate phylogenies, gene or protein sequence alignments can be interrogated to separate phylogenetic signal (sequence changes that reflect the evolutionary relationships among species) from noise. In the post-genomics era, the analysis of hundreds or thousands of genes using powerful models and computational algorithms has dramatically improved the accuracy with which the evolutionary relationships among living organisms can be deduced, but certain key branches have remained recalcitrant and sparked controversy [1–3].

The biggest controversy in the phylogeny of animals (metazoans) concerns the evolutionary relationships among three lineages that originated near the base of the animal tree. The first lineage comprises the vast majority of animals and includes placozoans (an enigmatic group of multicellular organisms with a very simple organization), cnidarians (such as jellyfish and sea anemones) and all bilaterally symmetrical animals (including humans, fruit flies and worms). The remaining two lineages — and the major actors in the controversy — are sponges and ctenophores, a marine phylum of gelatinous animals that bear distinctive 'combs' of cilia, earning them the moniker 'comb jellies' (Figure 1). The debate centers on which of these two, sponges or ctenophores, is the sister lineage (Box 1) to the rest of the animals. The uncertainty stemming from this controversy has major implications for understanding the origin of animal multicellularity and development, for deciphering the biology of early animals [4,5] and for unraveling the evolution of many animal traits, such as muscles or nervous systems [6–8].

Here, we review the historical background of the sponge–ctenophore controversy, summarize where it stands now and discuss its implications for understanding and studying the evolution of key animal traits. As short branches at the base of ancient evolutionary radiations are challenging to resolve, we argue that research aimed at deciphering the cellular and molecular foundations of animal multicellularity and development should embrace the uncertainty surrounding the early evolution of animals.

A Brief Historical Perspective

Pre-molecular era efforts to reconstruct evolutionary relationships among animal phyla were largely based on their cellular and morphological characteristics [9,10]. In those phylogenies, sponges were invariably placed as the sister branch to the rest of the animals, and ctenophores were thought to represent either the sister lineage to cnidarians [9] or to bilaterians (Box 1) [10]. In fact, even the notion that sponges are animals was debated early on [11,12], with later leading opinions arguing that sponges should be confined to Parazoa (Box 1), animals of the "cellular grade of construction," leaving the rest of animals in the "tissue grade of construction" Eumetazoa (Box 1) [13]. Our modern classification of sponges as animals is based on phylogenomics, comparative genomics, and findings of conserved processes during embryogenesis [14,15]. Sponges, like the rest of animals,

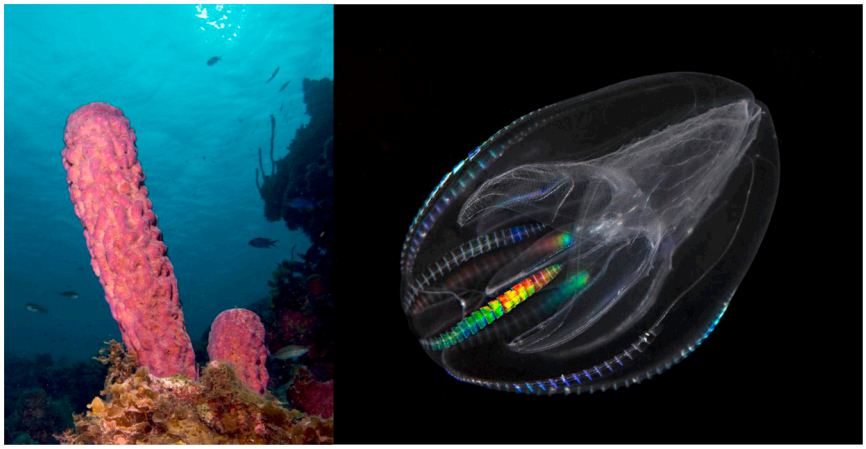

Figure 1. Sponges and ctenophores, the two animal phyla at the center of the controversy.
Left: The stove-pipe sponge *Aplysina archeri* (Photo: Nick Hobgood, Wikimedia Commons CC BY-SA 3.0). Right: the ctenophore *Mnemiopsis leidyi* (Photo: © Stefan Siebert)

produce differentiated sperm and eggs, have epithelia, contain a suite of animal-specific genes, develop through a clonal process of serial cell division and exhibit conserved developmental gene expression patterns [14–17]. Sponge collar cells, which bear a single flagellum surrounded by a collar of microvilli, resemble the cell morphology of the choanoflagellates [18,19] (but see [20]), the closest living relatives of animals [21], which was interpreted as further support for the notion that sponges are the sister lineage to the rest of animals.

Despite eliciting early interest from zoologists and embryologists [13], ctenophores escaped the attention of most cell and molecular biologists until relatively recently. The unusual challenges of working with ctenophores meant that they were excluded from most early molecular phylogenies. Perhaps most importantly, early examinations of their morphological characteristics, such as their diffuse nervous system and musculature [22], inspired some to infer a close relationship to the bilaterians [13], while others argued for a sister group relationship with cnidarians based on the similarities of the two phyla in adult morphology [9].

Early molecular phylogenies based on a single or a few genes tended to place sponges as the sister lineage to the rest of animals [23,24] and rejected the close affinity of ctenophores with cnidarians, instead placing ctenophores as the sister to a clade comprised of cnidarians, placozoans and bilaterians [24–26]. In hindsight, these early molecular efforts were underpowered [24,25,27–29], and their results weakly supported and highly varied. For example, some early analyses favored now obsolete scenarios such as animal polyphyly caused by the grouping of cnidarians with ciliates and fungi [30] or the inference of a clade comprised by sponges, ctenophores, cnidarians, and placozoans that was sister to that of bilaterians [31]. At the dawn of the 21st century, it was unclear whether the observed volatility in animal relationships was due to the use of small amounts of molecular sequence data, the poor fit of phylogenetic models to the sequence data, or the genuine lack of phylogenetic signal that might be expected from the early phase of the animal evolutionary radiation.

The Sponge–Ctenophore Controversy in the Phylogenomic Era

Over the last decade, remarkable advances in DNA sequencing technologies have led to the sequencing of the first sponge [15] and ctenophore [7,32] genomes as well as to large amounts of transcriptome data [32–38]. This increase in data by orders of magnitude, coupled with considerable developments in computational phylogenetics, culminated in faster software and enabled researchers to use new and more sophisticated strategies for phylogenomic inference. Consequently, molecular phylogenies of animal lineages were no longer based on just a handful of genes, but on hundreds of them. An early example of using phylogenomics to address long-standing questions about the animal tree came in 2008, with the first study suggesting that ctenophores — not sponges — were the sister lineage to the rest of the animals [33]. Since then, more than a dozen conflicting phylogenomic analyses have offered support for the ctenophores-sister hypothesis [2,3,7,32,38–40], the sponges-sister hypothesis [34–37,41,42] or, much more rarely, neither [43] (Figure 2).

Why is it that, despite a decade of ever-increasing amounts of data and phylogenomic analyses, we have yet to reach consensus on the relative position of sponges and ctenophores in the animal tree of life? In general, several different biological and analytical factors can conspire to make a particular branching pattern recalcitrant to resolution (Box 2). Although several such factors are likely to be at play, the controversy can be whittled down to two key issues: model selection and analytical strategy. The first concerns how one models sequence evolution in phylogenomic inference. Briefly, the standard site-homogeneous models (Box 1) of protein evolution assume that all amino acid sites in a given protein sequence alignment have evolved under the same substitution process, whereas site-heterogeneous models (Box 1) allow each site in the protein sequence-alignment to have its own substitution process. Consequently, site-heterogeneous models are better descriptors of biological sequence evolution than site-homogeneous models [44,45], but are computationally much more costly [46]. In the context of phylogenomic analyses, where data matrices contain hundreds to thousands of protein sequence alignments, researchers typically resort to one of two strategies: they either use a site-homogeneous model for each gene in the data matrix (this partitioning improves the fit of models to the data) or a site-heterogeneous model across the entire data matrix.

In the case of the sponge–ctenophore controversy, analyses employing site-homogeneous models of protein evolution with partitioning (a different model for each gene) tend to recover ctenophores as the sister branch to all animals [2,3,7,32,33,38–40] (but see [35,42]). In contrast, analyses employing site-heterogeneous models (and in particular the CAT model [47]) typically recover sponges as the sister branch ([34–37,41,42], but see [38,40,48]). At this time, there is no

Box 1. Glossary.

Bilateria: The monophyletic group (clade) of animal phyla with a bilateral axis of symmetry; this clade includes all extant animal phyla except sponges, ctenophores, cnidarians, and placozoans.
Eumetazoa: A hypothetical sub-kingdom of animals (i.e. metazoans) that includes all phyla that exhibit a "tissue grade of construction" [13]; it includes all extant animal phyla, except sponges and placozoans (which belong to the sub-kingdom Parazoa).
Hybridization and introgression: Hybridization occurs when two organisms from different, typically closely related, species mate and produce offspring. When the hybrids mate back with members of one of the parent species, genes (or genetic regions) from the other parent species can cross the species barrier through introgression. Hybridization and introgression can lead to the evolutionary history of genes differing from the history of their species, complicating inference of phylogenetic relationships from gene sequence data.
Incomplete lineage sorting of ancestral polymorphisms: The retention of two or more alleles from an ancestral population in descendant species following successive speciation events. Incomplete lineage sorting is usually followed by random allele fixation in the descendant species, which can result in gene histories differing from the history of their species.
Pan-animal homologies: Homologous traits found in all or nearly all extant animal phyla. Pan-animal homologies include animal synapomorphies (shared, derived traits) as well as more ancient, pre-animal traits that have been conserved in all animals.
Parazoa: A now-refuted, paraphyletic sub-kingdom of animals comprised of sponges and placozoans to the exclusion of the rest of the animal phyla [13].
Site-heterogeneous models of sequence evolution: These models allow different amino acid positions in a protein alignment to have their own substitution models [47]; they are inspired by the observation that certain positions of protein sequence alignments tolerate substitutions between only a specific subset of the 20 amino acids. For example, consider a position in a highly conserved protein that can only tolerate one of the two negatively charged amino acids (aspartic acid, D, or glutamic acid, E) for the protein to retain its function. If D and E are functionally equivalent, we expect that multiple substitutions from D to E and *vice versa* will occur after hundreds of millions of years of evolution. Thus, the similarities and differences between animal proteins on the basis of the observed amino acid at this position will not accurately reflect evolutionary affinity. Furthermore, because such positions will lack phylogenetic information but give the appearance of doing so, they contribute to phylogenetic error. As site-heterogeneous models are specifically tailored to individual amino acid positions, they can in principle reduce the negative impact of such sites on phylogenetic inference much better than site-homogeneous sites [41] (but see [46]).
Site-homogeneous models of sequence evolution: The standard models of protein sequence evolution; these models are constructed from empirically derived amino acid substitution matrices. As these models assume that all amino acid positions in a given protein sequence alignment have evolved under the same substitution process, all positions in the alignment use the same amino acid exchange rate matrix. Different site-homogeneous models may be used for different gene alignments, or data partitions, within a concatenated data matrix.
Sister lineage: A lineage that is the closest relative of another lineage and *vice versa*. In the context of the animal controversy, the ctenophore-sister hypothesis proposes that ctenophores are the closest relatives to the rest of the animals (and *vice versa*), whereas the sponges-sister hypothesis proposes that sponges are the closest relatives to the rest of the animals (and *vice versa*).
Urmetazoan: The last common ancestor of all animals (i.e., metazoans).

consensus as to which of these two strategies is more likely to yield an accurate phylogeny.

The second issue (or set of issues) concerns choices made during assembly of the phylogenomic data matrix, including strategies for identifying orthologous proteins, the level of tolerance for missing orthologs from certain species in the protein sequence alignment, the choice of non-animal organisms to root the phylogeny and rules for excluding taxa whose protein sequences show unusual characteristics (e.g., unusually high evolutionary rate) [7,35,37,38,40–42]. Here, it is less clear how these choices influence the recovery of either sponges or ctenophores as the sister lineage of all other animals. Although the strategies and models employed by the latest studies represent the field's state-of-the-art, the absence of independent types of data for testing the validity of either of the two alternative phylogenetic hypotheses makes it hard to determine an optimal strategy. Thus, we argue that we cannot yet confidently infer which of the two hypotheses — sponges-sister or ctenophores-sister — is more likely to be correct.

How can we gain greater confidence in our reconstructions of the earliest stages of animal diversification? Taking a cue from well-established branches of the tree of life, we should expect that the 'correct' resolution should be robustly supported by independent sources of data, experimental designs and methods. Much remains to be done. To date, phylogenomicists have had only one sponge genome [15], two ctenophore genomes [7,32] and a handful of transcriptomes [32–38] from either at their disposal; clearly more genomes from diverse sponges, ctenophores and animal outgroups are needed. Sequencing more ctenophore genomes may, for example, uncover one or more new lineages that have evolved more slowly, which would be valuable for investigating whether the ctenophores-sister placement is an artifact stemming from the high evolutionary rates of the two known ctenophore genomes. Likewise, genomic sequencing of additional outgroup taxa may reveal the extent to which the current selection of taxa skews analyses toward either sponges-sister or ctenophores-sister in phylogenomic analyses (Box 2).

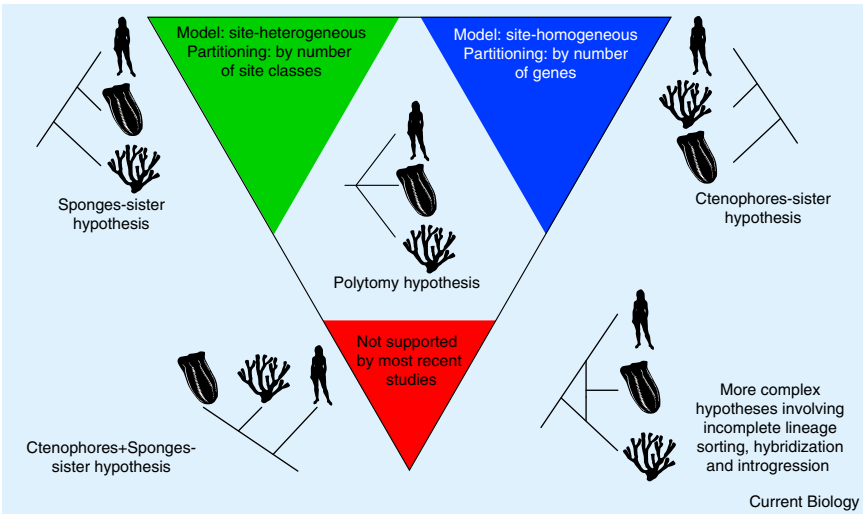

Figure 2. Different models of amino acid evolution favor different scenarios for early animal evolution.

Hypotheses for the evolutionary relationships among sponges, ctenophores and all other animals (represented by human). Three alternative hypotheses are shown on each of the corners of the triangle: sponges as the sister lineage to the rest of the animals (top left) [34–37,41,42], ctenophores as the sister lineage (top right) [2,3,7,32,33,38–40], or a clade of sponges and ctenophores as the sister lineage to the rest of the animals [2,3]. The triangles summarize some of the key characteristics of phylogenomic studies supporting or rejecting each hypothesis (e.g., application of site-homogeneous models tends to favor the ctenophores-sister hypothesis, whereas application of the CAT site-heterogeneous model tends to favor the sponges-sister hypothesis). The polytomy hypothesis, under which all three lineages originated at the same time, is displayed inside the triangle. The inset depicts more complex phylogenetic scenarios, such as hybridization, introgression, and incomplete lineage sorting of ancestral polymorphisms (Box 1). Although evidence in support of either polytomy or a more complex scenario is lacking, had it occurred, it remains questionable whether it would be detectable with existing data and methods. (Sponge image: Mali'o Kodis, photograph by Derek Keats (http://www.flickr.com/photos/dkeats/) Attribution 3.0 Unported (CC BY 3.0) https://creativecommons.org/licenses/by/3.0/).

Additional genomes may also catalyze phylogenomic method development, either by making the branch easier to resolve or by facilitating a clearer demarcation of the strengths and weaknesses of the different strategies in data-rich, taxon-rich data matrices. More genomes could also facilitate the discovery of independent characters, such as rare genomic changes, that might shed light on early animal diversification. So far, the only type of rare genomic change analyzed has been the presence or absence of genes (gene content) across animal phyla, with one study favoring ctenophores-sister [7] and a reanalysis favoring sponges-sister [42]. While a step in the right direction, gene content can be even more sensitive than linear sequence data to variation in organismal lifestyle [48], raising concerns about its utility for resolving ancient radiations, such as that of animals.

Finally, our desire for a neatly bifurcating tree may mask the true history of early branching animal phyla. It is possible that more complex phylogenetic scenarios involving hybridization, introgression or incomplete lineage sorting (Box 1) were part of the early diversification of animals (Figure 2), even though these processes cannot be robustly detected using existing data and methods. This would imply that the evolutionary history of traits encoded by genes that experienced any of these processes may be genuinely different from that implied by the species phylogeny [49], regardless of whether sponges or ctenophores are the sister lineage.

Implications for Reconstructing the Urmetazoan

While reconstructing the relationships among animal phyla is a formidable challenge, it is essential for understanding the origins of modern animals and the evolutionary processes that gave rise to their diversity. Notwithstanding the current ambiguities regarding sponges and ctenophores, the robustness of other parts of the animal and eukaryotic phylogeny mean that much can yet be inferred about the biology of the last common ancestor of all animals, the Urmetazoan (Box 1).

Many traits can be assigned to the Urmetazoan with near certainty, regardless of whether sponges or ctenophores are the sister lineage (Figure 3). Among these are the pan-animal homologies (Box 1), homologous traits found in nearly all animals, and by inference, in their last common ancestor, the Urmetazoan. These Urmetazoan traits include traits predicted to have evolved along the animal stem lineage, such as obligate, clonal multicellularity, spermatogenesis and oogenesis (i.e., oogamy), animal-specific developmental signaling pathways (e.g., Wnt and TGF-β), epithelia, and, potentially, Vasa- and Piwi-driven regulation of stem cell multipotency [4,50–54]. Moreover, the Urmetazoan also contained several more ancient traits that are today conserved in animals and in some of their closest relatives, including phagotrophy, genes involved in animal cell signaling, cell adhesion and transcriptional regulation (e.g., Brachyury, cadherins, integrins, and receptor tyrosine kinases), as well as cells bearing a single apical flagellum/cilium (e.g., found today on animal sperm and epithelial cells) [4,55–58].

Several other traits, while absent from either sponges or ctenophores, can also be traced back to the Urmetazoan because they are conserved in diverse animals and in one or more of their protozoan relatives. For example, the conservation of the cadherin-repeat-containing gene *hedgling* in choanoflagellates, sponges, and cnidarians suggests that it was present in the Urmetazoan, despite its absence from the genomes of ctenophores, placozoans and bilaterians [59,60]. Similarly, the collar cells found in choanoflagellates, sponges, cnidarians and many bilaterians are likely to have appeared in the Urmetazoan, despite their absence from ctenophores and most ecdysozoans [19].

Reconstructing the ancestry of other traits can be more challenging. If a trait is absent from non-animals and also not present in either sponges or ctenophores, inferences about its ancestry are contingent upon and await the resolution of the controversy (Figure 3). For some of these traits, ancestral reconstruction is further complicated by different interpretations of their homology relationships. A case in point is the neuron, a cell type present in ctenophores and absent in placozoans and sponges (but see [61]). Historically, neurons in ctenophores, cnidarians and bilaterians have been inferred to be homologous, with their homology inspiring some zoologists to place these lineages in a clade

Box 2. Why are some evolutionary relationships so controversial?

Although many parts of the tree of life have been robustly and reproducibly resolved using many independent types of data and approaches, others have proven more challenging. The most heavily debated phylogenetic controversies center around short branches at the base of ancient evolutionary radiations [1]. This is largely because the resolution of such short, deep branches is highly susceptible to analytical artifacts. As the taxa compared are only distantly related, orthologous gene sequences are highly divergent (or altogether absent in some taxa), reducing the accuracy of orthology inference as well as of multiple sequence alignment of these putative orthologous sequences, while also increasing the amount of sequence data missing. Distantly related taxa also differ from each other in ways that influence evolutionary substitution rates (e.g., generation time, population size, mutation rate and GC content), reducing the fit between models of sequence evolution and the data at hand. Poor fit between the model of evolution and the sequence data being analyzed can lead to long branch attraction, a phenomenon in which taxa whose sequence data have experienced the largest amounts of change are artifactually grouped together, irrespective of their true evolutionary relationships [74].

Variation in evolutionary rates across taxa also means that the same set of genes may be fast-evolving in some taxa but slowly-evolving in others, which can also lead to long branch attraction. In the context of the animal phylogeny, the branches leading to sponges and ctenophores, which are some of the longest, are particularly prone to long branch attraction. Sampling of additional sponge and ctenophore taxa could potentially help break these long branches and ameliorate long branch attraction [75], a strategy employed in all recent phylogenomic studies on the controversy. However, while all major sponge lineages appear to be ancient (>500 million years old), the last common ancestor of extant ctenophores is thought to be much younger [26], making long branch attraction amelioration strategies based on taxon sampling potentially less effective.

Finally, short branches at the base of evolutionary radiations, including the one at the center of the controversy about animal origins, are hotspots for lineage sorting of ancestral polymorphisms [76] as well as for hybridization and introgression (Box 1) [77], each of which can produce gene histories that differ from those of their species. Such events are commonplace in the animal phylogeny [49,76,77]. Although evidence of lineage sorting, introgression or hybridization in the early history of animals is lacking, had it occurred, it remains questionable whether it would be detectable with existing data and methods (Figure 2).

called 'Neuralia' [54]. Not only do neurons from ctenophores, cnidarians and bilaterians produce a conserved set of diagnostic neuropeptides, but the genomes of ctenophores and cnidarians encode homologs of diverse proteins that have been shown in bilaterians to be required for neuronal fate, patterning and the formation of synapses [7,32,62–66]. Nonetheless, ctenophore genomes appear to lack a number of bilaterian neuronal genes (e.g., neuroligin), and other neuronal genes are not expressed in a neuron-specific manner [32,62]. Those who emphasize the similarities between neurons from ctenophores and other animals infer that they have a shared ancestry and are homologous, while some others, who have focused on the differences, argue for independent origins.

Whether sponges or ctenophores are the sister group to all animals also has implications for the presence or absence of neurons in the Urmetazoan. Under the sponges-sister hypothesis, we might reasonably infer that the Urmetazoan lacked neurons, and that neurons subsequently evolved in the stem lineage leading to ctenophores and all other animals. (It is not possible to rule out an Urmetazoan origin of neurons, only for them to be lost in sponges.) Under the ctenophore-sister hypothesis, one must explain the absence of neurons from sponges and placozoans and take account of the fact that some features of the ctenophore nervous system resemble those of cnidarians and bilaterians, while others are quite different. Note that under both of these scenarios, placozoans are presumed to have lost neurons (Figure 3A). One recent hypothesis focuses on the differences and suggests that neurons in ctenophores, cnidarians and bilaterians might have evolved convergently rather than being homologous cell types that evolved through descent from a common ancestor [32,62]. A seemingly more likely scenario, regardless of whether ctenophores or sponges are the sister to all other animals, is that the last common ancestor of ctenophores and cnidarians/bilaterians may have had a rudimentary nervous system that provided the genetic and cellular building blocks, including neurons, of modern nervous systems. This ancestral nervous system may have then been independently elaborated upon in the ctenophore and cnidarian–bilaterian lineages [66–68], yielding divergent nervous systems in the extant organisms. According to this scenario, if ctenophores are the sister lineage, sponges would have lost the ancestral nervous system, akin to the more recent and incontrovertible losses of neurons in placozoans and the parasitic myxozoans [69].

This picture of a rudimentary ancestral nervous system being elaborated in some lineages and lost in others parallels our understanding of the evolution of other complex traits, such as body appendages, the heart and eyes [70]. For example, the eyes of molluscs, arthropods and vertebrates were thought to have originated independently until the discovery that they all develop under the control of a conserved master regulatory gene, *Pax6*, and the subsequent inference that they each are derived from an ancestral photoreceptor system [71,72]. To unravel the evolutionary history of complex traits, such as the nervous system, we must not only resolve the animal phylogeny, but also test our assumptions about trait homology at multiple levels (i.e., genetic, cellular ultrastructure, and system function [73]) by trying to understand the molecular and cellular mechanisms that give rise to these traits in diverse animals.

Science is an iterative and evolutionary process. We now understand that sponges are animals, and that ciliates are not. Moreover, we are able to make ever more specific inferences about the cellular and molecular biology of organisms that lived

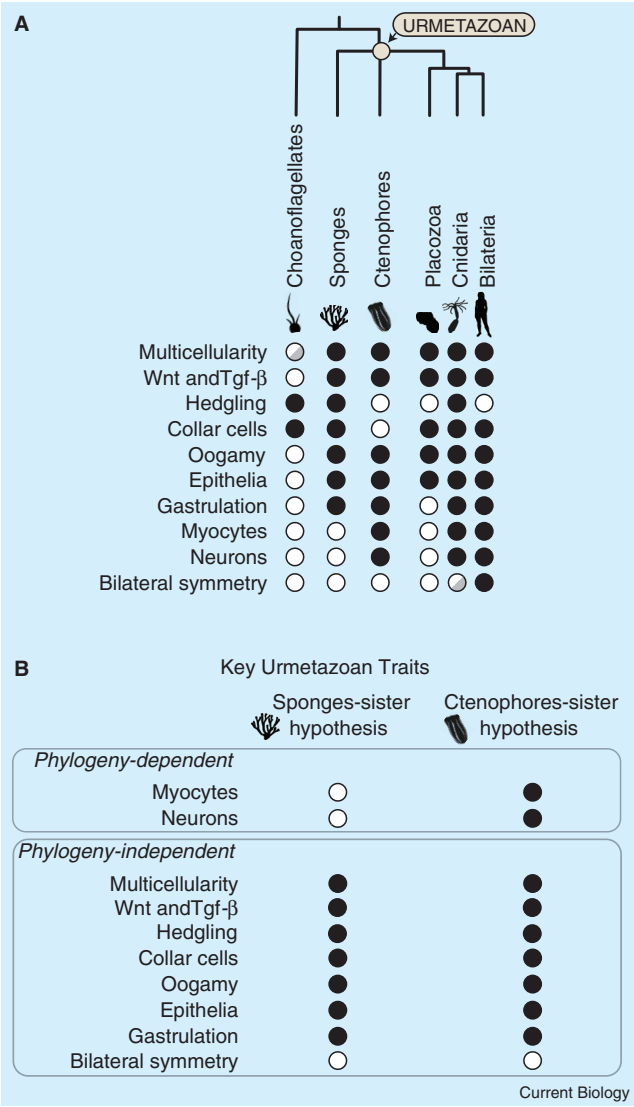

Figure 3. Reconstructing the Urmetazoan.
Much can be inferred about the Urmetazoan, despite the ongoing phylogenetic controversy. (A) Distribution of illustrative traits in diverse animals and their closest living relatives, the choanoflagellates. The base of the animal tree is depicted as a polytomy to emphasize current uncertainty about the relative placement of sponges and ctenophores. Each trait is indicated as being either present in a clade (black circles), not detected in a clade (white circles), or as having been detected in an intermediate form in a subset of taxa within the lineage (half-gray circles). For example, some choanoflagellates produce an intermediate multicellular form, a 'colony,' in which all cells have apparently equivalent morphology. How traits are coded (i.e., present vs. absent for intermediate or convergent traits) influences inferences about their presence or absence in the Urmetazoan. (B) Controversies regarding the phylogenetic relationships among sponges, ctenophores, and other animals shape inferences about the biology of the Urmetazoan. Inferences about the presence (black circles) or apparent absence (white circles) of myocytes and neurons in the Urmetazoan are 'phylogeny-dependent' and are contingent on whether ctenophores or sponges are the sister group to all other animals. In contrast, the presence or absence from the Urmetazoan of 'phylogeny-independent' traits can be reasonably inferred, regardless of the branch order of sponges and ctenophores.

and died over 600 million years ago, leaving hardly a trace in the fossil record. We can confidently infer that they produced eggs and sperm, became multicellular through serial cell division, initiated the production of many differentiated cell types and tissues during gastrulation and are likely to have fed on bacteria using specialized collar cells. But our window through time to the Urmetazoan remains obscured by uncertainty regarding whether sponges or ctenophores are the sister lineage to all other animals. While we concur with the sentiment in the opening quote that "we're not afraid to admit what we don't know" and we revel in the insights that are currently possible regarding the dawn of animal life (Figure 3), we eagerly anticipate future breakthroughs that will allow us to move from ignorance to knowledge in our quest to reconstruct the origin of animals.

ACKNOWLEDGEMENTS

We thank Thibaut Brunet, Pawel Burkhardt, Casey Dunn, Babcock Hall, Joe Ryan, Xing-Xing Shen, Monika Sigg and Max Telford for constructive feedback on earlier drafts of this manuscript. Research in the King laboratory is supported by the Howard Hughes Medical Institute, the National Institutes of Health (R01GM099533), and the Gordon and Betty Moore Foundation. Research in the Rokas laboratory is supported by the National Science Foundation (DEB-1442113), the Burroughs Wellcome Fund, and the March of Dimes Prematurity Research Center Ohio Collaborative.

REFERENCES

1. Rokas, A., and Carroll, S.B. (2006). Bushes in the tree of life. PLoS Biol. *4*, e352.
2. Shen, X.X., Hittinger, C.T., and Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nature Ecol. Evol. *1*, 0126.
3. Arcila, D., Ortí, G., Vari, R., Armbruster, J.W., Stiassny, M.L.J., Ko, K.D., Sabaj, M.H., Lundberg, J., Revell, L.J., and Betancur-R., R. (2017). Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. Nature Ecol. Evol. *1*, 0020.
4. Richter, D.J., and King, N. (2013). The genomic and cellular foundations of animal origins. Annu. Rev. Genet. *47*, 509–537.
5. Schierwater, B., Holland, P.W.H., Miller, D.J., Stadler, P.F., Wiegmann, B.M., Wörheide, G., Wray, G.A., and DeSalle, R. (2016). Never ending analysis of a century old evolutionary debate: "Unringing" the Urmetazoon bell. Front. Ecol. Evol. *4*, 5.
6. Dohrmann, M., and Worheide, G. (2013). Novel scenarios of early animal evolution–is it time to rewrite textbooks? Int.egr. Comp. Biol. *53*, 503–511.
7. Ryan, J.F., Pang, K., Schnitzler, C.E., Nguyen, A.D., Moreland, R.T., Simmons, D.K., Koch, B.J., Francis, W.R., Havlak, P., Program, N.C.S., *et al*. (2013). The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. Science *342*, 1242592.
8. Dunn, C.W., Leys, S.P., and Haddock, S.H. (2015). The hidden biology of sponges and ctenophores. Trends Ecol. Evol. *30*, 282–291.
9. Brusca, R.C., and Brusca, G.J. (1990). Invertebrates (Sunderland, Mass: Sinauer).
10. Nielsen, C. (2001). Animal Evolution, 2nd Edition (Oxford: Oxford University Press).
11. Saville-Kent, W. (1880-1882). A Manual of the Infusoria, *Volume 1-3* (London: David Bogue).
12. Haeckel, E. (1874). Memoirs: The gastraea-theory, the phylogenetic classification of the animal kingdom and the homology of the germ-lamellæ. J. Cell Sci., s2–14, 223-247.
13. Hyman, L.H. (1940). The Invertebrates: Protozoa through Ctenophora, *Volume I*, 1st Edition (New York: McGraw-Hill).

14. Leys, S.P., and Degnan, B.M. (2002). Embryogenesis and metamorphosis in a haplosclerid demosponge: gastrulation and transdifferentiation of larval ciliated cells to choanocytes. Invert. Biol. *121*, 171–189.

15. Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M.E., Mitros, T., Richards, G.S., Conaco, C., Dacre, M., Hellsten, U., *et al*. (2010). The *Amphimedon queenslandica* genome and the evolution of animal complexity. Nature *466*, 720–726.

16. Leininger, S., Adamski, M., Bergum, B., Guder, C., Liu, J., Laplante, M., Brate, J., Hoffmann, F., Fortunato, S., Jordal, S., *et al*. (2014). Developmental gene expression provides clues to relationships between sponge and eumetazoan body plans. Nat. Commun *5*, 3905.

17. Nakanishi, N., Sogabe, S., and Degnan, B.M. (2014). Evolutionary origin of gastrulation: insights from sponge development. BMC Biol. *12*, 26.

18. James-Clark, H. (1867). IV.—Conclusive proofs of the animality of the ciliate sponges, and of their affinities with the Infusoria flagellata. J. Nat. Hist. *19*, 13–18.

19. Brunet, T., and King, N. (2017). The origin of animal multicellularity and cell differentiation. bioRxiv. http://dx.doi.org/10.1101/161695.

20. Mah, J.L., Christensen-Dalsgaard, K.K., and Leys, S.P. (2014). Choanoflagellate and choanocyte collar-flagellar systems and the assumption of homology. Evol. Dev. *16*, 25–37.

21. King, N., Westbrook, M.J., Young, S.L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., *et al*. (2008). The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. Nature *451*, 783–788.

22. Harbison, H.R. (1985). On the classification and evolution of the Ctenophora. In The Origins and Relationships of Lower Invertebrates, S. Conway Morris, J.D. George, R. Gibson, and H.M. Platt, eds. (Oxford: Oxford University Press), pp. 78–100.

23. Wainright, P.O., Hinkle, G., Sogin, M.L., and Stickel, S.K. (1993). Monophyletic origins of the metazoa: an evolutionary link with fungi. Science *260*, 340–342.

24. Collins, A.G. (1998). Evaluating multiple alternative hypotheses for the origin of Bilateria: an analysis of 18S rRNA molecular evidence. Proc. Natl. Acad. Sci. USA *95*, 15458–15463.

25. Medina, M., Collins, A.G., Silberman, J.D., and Sogin, M.L. (2001). Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. Proc. Natl. Acad. Sci. USA *98*, 9707–9712.

26. Podar, M., Haddock, S.H., Sogin, M.L., and Harbison, G.R. (2001). A molecular phylogenetic framework for the phylum Ctenophora using 18S rRNA genes. Mol. Phylogenet. Evol. *21*, 218–230.

27. Rodrigo, A.G., Bergquist, P.R., Bergquist, P.L., and Reeves, P.R. (1994). Are sponges animals? An investigation into the vagaries of phylogenetic inference. In Sponges in Time and Space, R.W.M. van Soest, T.M.G. van Kempen, and J.C. Braekman, eds. (Rotterdam: Balkema), pp. 47–54.

28. Abouheif, E., Zardoya, R., and Meyer, A. (1998). Limitations of metazoan 18S rRNA sequence data: implications for reconstructing a phylogeny of the animal kingdom and inferring the reality of the Cambrian explosion. J. Mol. Evol. *47*, 394–405.

29. Rokas, A., King, N., Finnerty, J., and Carroll, S.B. (2003). Conflicting phylogenetic signals at the base of the metazoan tree. Evol. Dev. *5*, 346–359.

30. Field, K.G., Olsen, G.J., Lane, D.J., Giovannoni, S.J., Ghiselin, M.T., Raff, E.C., Pace, N.R., and Raff, R.A. (1988). Molecular phylogeny of the animal kingdom. Science *239*, 748–753.

31. Philippe, H., Chenuil, A., and Adoutte, A. (1994). Can the Cambrian explosion be inferred through molecular phylogeny? Development Suppl. 15–25.

32. Moroz, L.L., Kocot, K.M., Citarella, M.R., Dosung, S., Norekian, T.P., Povolotskaya, I.S., Grigorenko, A.P., Dailey, C., Berezikov, E., Buckley, K.M., *et al*. (2014). The ctenophore genome and the evolutionary origins of neural systems. Nature *510*, 109–114.

33. Dunn, C.W., Hejnol, A., Matus, D.Q., Pang, K., Browne, W.E., Smith, S.A., Seaver, E., Rouse, G.W., Obst, M., Edgecombe, G.D., *et al*. (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. Nature *452*, 745–749.

34. Philippe, H., Derelle, R., Lopez, P., Pick, K., Borchiellini, C., Boury-Esnault, N., Vacelet, J., Renard, E., Houliston, E., Queinnec, E., *et al*. (2009). Phylogenomics revives traditional views on deep animal relationships. Curr. Biol. *19*, 706–712.

35. Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D.J., Di Franco, A., Roure, B., Satoh, N., Queinnec, E., Ereskovsky, A., *et al*. (2017). A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. Curr. Biol. *27*, 958–967.

36. Pick, K.S., Philippe, H., Schreiber, F., Erpenbeck, D., Jackson, D.J., Wrede, P., Wiens, M., Alie, A., Morgenstern, B., Manuel, M., *et al*. (2010). Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. Mol. Biol. Evol. *27*, 1983–1987.

37. Nosenko, T., Schreiber, F., Adamska, M., Adamski, M., Eitel, M., Hammel, J., Maldonado, M., Muller, W.E., Nickel, M., Schierwater, B., *et al*. (2013). Deep metazoan phylogeny: when different genes tell different stories. Mol. Phylogenet. Evol. *67*, 223–233.

38. Whelan, N.V., Kocot, K.M., Moroz, L.L., and Halanych, K.M. (2015). Error, signal, and the placement of Ctenophora sister to all other animals. Proc. Natl. Acad. Sci. USA *112*, 5773–5778.

39. Hejnol, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G.W., Edgecombe, G.D., Martinez, P., Baguna, J., Bailly, X., Jondelius, U., *et al*. (2009). Assessing the root of bilaterian animals with scalable phylogenomic methods. Proc. R. Soc. S. B *276*, 4261–4270.

40. Borowiec, M.L., Lee, E.K., Chiu, J.C., and Plachetzki, D.C. (2015). Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. BMC Genomics *16*, 987.

41. Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T., Manuel, M., Worheide, G., and Baurain, D. (2011). Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. *9*, e1000602.

42. Pisani, D., Pett, W., Dohrmann, M., Feuda, R., Rota-Stabelli, O., Philippe, H., Lartillot, N., and Worheide, G. (2015). Genomic data do not support comb jellies as the sister group to all other animals. Proc. Natl. Acad. Sci. USA *112*, 15402–15407.

43. Schierwater, B., Eitel, M., Jakob, W., Osigus, H.J., Hadrys, H., Dellaporta, S.L., Kolokotronis, S.O., and Desalle, R. (2009). Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoon" hypothesis. PLoS Biol. *7*, e20.

44. Halpern, A.L., and Bruno, W.J. (1998). Evolutionary distances for protein-coding sequences: modeling site-specific residue frequencies. Mol. Biol. Evol. *15*, 910–917.

45. Lartillot, N., Brinkmann, H., and Philippe, H. (2007). Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. BMC Evol. Biol. *7* (*Suppl 1*), S4.

46. Whelan, N.V., and Halanych, K.M. (2017). Who let the CAT out of the bag? Accurately dealing with substitutional heterogeneity in phylogenomic analyses. Syst. Biol. *66*, 232–255.

47. Lartillot, N., and Philippe, H. (2004). A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. *21*, 1095–1109.

48. Chang, E.S., Neuhof, M., Rubinstein, N.D., Diamant, A., Philippe, H., Huchon, D., and Cartwright, P. (2015). Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. Proc. Natl. Acad. Sci. USA *112*, 14912–14917.

49. Hahn, M.W., and Nakhleh, L. (2016). Irrational exuberance for resolved species trees. Evolution *70*, 7–17.

50. Babonis, L.S., and Martindale, M.Q. (2017). Phylogenetic evidence for the modular evolution of metazoan signalling pathways. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, http://dx.doi.org/10.1098/rstb.2015.0477.

51. Alie, A., Hayashi, T., Sugimura, I., Manuel, M., Sugano, W., Mano, A., Satoh, N., Agata, K., and Funayama, N. (2015). The ancestral gene repertoire of animal stem cells. Proc. Natl. Acad. Sci. USA *112*, E7093–7100.

52. Alie, A., Leclere, L., Jager, M., Dayraud, C., Chang, P., Le Guyader, H., Queinnec, E., and Manuel, M. (2011). Somatic stem cells express Piwi and Vasa genes in an adult ctenophore: ancient association of "germline genes" with stemness. Dev. Biol. *350*, 183–197.

53. Reitzel, A.M., Pang, K., and Martindale, M.Q. (2016). Developmental expression of "germline"- and "sex determination"-related genes in the ctenophore *Mnemiopsis leidyi*. Evodevo *7*, 17.

54. Nielsen, C. (2008). Six major steps in animal evolution: are we derived sponge larvae? Evol. Dev. *10*, 241–257.

55. Abedin, M., and King, N. (2008). The premetazoan ancestry of cadherins. Science *319*, 946–948.

56. Sebe-Pedros, A., Ariza-Cosano, A., Weirauch, M.T., Leininger, S., Yang, A., Torruella, G., Adamski, M., Adamska, M., Hughes, T.R., Gomez-Skarmeta, J.L., *et al*. (2013). Early evolution of the T-box transcription factor family. Proc. Natl. Acad. Sci. USA *110*, 16050–16055.

57. Sebe-Pedros, A., Roger, A.J., Lang, F.B., King, N., and Ruiz-Trillo, I. (2010). Ancient origin of the integrin-mediated adhesion and signaling machinery. Proc. Natl. Acad. Sci. USA *107*, 10142–10147.

58. Suga, H., Chen, Z., de Mendoza, A., Sebe-Pedros, A., Brown, M.W., Kramer, E., Carr, M., Kerner, P., Vervoort, M., Sanchez-Pons, N., *et al*. (2013). The *Capsaspora* genome reveals a complex unicellular prehistory of animals. Nat. Commun. *4*, 2325.

59. Adamska, M., Matus, D.Q., Adamski, M., Green, K., Rokhsar, D.S., Martindale, M.Q., and Degnan, B.M. (2007). The evolutionary origin of hedgehog proteins. Curr. Biol. *17*, R836–837.

60. Nichols, S.A., Roberts, B.W., Richter, D.J., Fairclough, S.R., and King, N. (2012). Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/beta-catenin complex. Proc. Natl. Acad. Sci. USA *109*, 13046–13051.

61. Leys, S.P. (2015). Elements of a 'nervous system' in sponges. J. Exp. Biol. *218*, 581–591.

62. Moroz, L.L., and Kohn, A.B. (2015). Unbiased view of synaptic and neuronal gene complement in ctenophores: Are there pan-neuronal and pan-synaptic genes across Metazoa? Integr. Comp. Biol. *55*, 1028–1049.

63. Burkhardt, P., Gronborg, M., McDonald, K., Sulur, T., Wang, Q., and King, N. (2014). Evolutionary insights into premetazoan functions of the neuronal protein homer. Mol. Biol. Evol. *31*, 2342–2355.

64. Jager, M., Chiori, R., Alie, A., Dayraud, C., Queinnec, E., and Manuel, M. (2011). New insights on ctenophore neural anatomy: immunofluorescence study in *Pleurobrachia pileus* (Muller, 1776). J. Exp. Zool. B Mol. Dev. Evol. *316B*, 171–187.

65. Simmons, D.K., Pang, K., and Martindale, M.Q. (2012). Lim homeobox genes in the Ctenophore *Mnemiopsis leidyi*: the evolution of neural cell type specification. Evodevo *3*, 2.

66. Ryan, J.F. (2014). Did the ctenophore nervous system evolve independently? Zoology *117*, 225–226.

67. Jekely, G., Paps, J., and Nielsen, C. (2015). The phylogenetic position of ctenophores and the origin(s) of nervous systems. Evodevo *6*, 1.

68. Liebeskind, B.J., Hillis, D.M., and Zakon, H.H. (2015). Convergence of ion channel genome content in early animal evolution. Proc. Natl. Acad. Sci. USA *112*, E846–851.

69. Ryan, J.F., and Chiodin, M. (2015). Where is my mind? How sponges and placozoans may have lost neural cell types. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, http://dx.doi.org/10.1098/rstb.2015.0059.

70. Carroll, S.B., Grenier, J.K., and Weatherbee, S.D. (2005). From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design, 2nd Edition (Malden, MA: Blackwell Pub.).

71. Halder, G., Callaerts, P., and Gehring, W.J. (1995). New perspectives on eye evolution. Curr. Opin. Genet. Dev. *5*, 602–609.

72. Randel, N., and Jekely, G. (2016). Phototaxis and the origin of visual eyes. Philos. Trans. R. Soc. Lond. B Biol. Sci. *371*, 20150042.

73. Abouheif, E., Akam, M., Dickinson, W.J., Holland, P.W., Meyer, A., Patel, N.H., Raff, R.A., Roth, V.L., and Wray, G.A. (1997). Homology and developmental genes. Trends Genet. *13*, 432–433.

74. Anderson, F.E., and Swofford, D.L. (2004). Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. Mol. Phylogenet. Evol. *33*, 440–451.

75. Graybeal, A. (1998). Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. *47*, 9–17.

76. Degnan, J.H., and Rosenberg, N.A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. *24*, 332–340.

77. Nakhleh, L. (2013). Computational approaches to species phylogeny inference and gene tree reconciliation. Trends Ecol. Evol. *28*, 719–728.