**Morphology should not be forgotten in the era of genomics—a phylogenetic perspective**

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1016/j.jcz.2015.01.003</td>
</tr>
<tr>
<td>Accessed</td>
<td>January 16, 2018 4:37:04 AM EST</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:34609598">http://nrs.harvard.edu/urn-3:HUL.InstRepos:34609598</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:34609598">http://nrs.harvard.edu/urn-3:HUL.InstRepos:34609598</a> terms-of-use#OAP</td>
</tr>
</tbody>
</table>
Morphology should not be forgotten in the era of genomics—a phylogenetic perspective

Gonzalo Giribet

*Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA*
ABSTRACT

Morphology has traditionally played a pivotal role in animal phylogeny since the first evolutionary biologists began to decipher the animal tree of life. In recent times, however, morphological characters have lost prominence in inferring deep relationships among animals due to fundamental issues with homology statements and the choice of higher taxa as terminals, but most importantly, due to declining costs of generating molecular data. As genomes and transcriptomes become widely available, the burden of evidence for morphology in overcoming molecular tree topologies has become ever heavier. In addition, collection and coding of morphological characters must accelerate and become less subjective. These needs have led to developments in the field of ontology, the use of a controlled and formalized morphological vocabulary. Finally, the role of fossils for dating molecular phylogenies is discussed. While molecules seem to have displaced morphology for phylogenetic inference, the need to incorporate fossils as explicit terminals, including dating of molecular-driven trees, together with new methods for high-throughput morphological data acquisition and annotation, may rescue morphology to become again an important player in phylogenetics. The reasons for maintaining a research program in morphology are thus countless, as ultimately, a zoologist's interest is to understand form and function, ecology and evolution, as well as all other aspects that may explain how our favorite organisms live, behave and evolve.

Keywords: phylogenetics, phylogenomics, tip dating, node dating, total evidence dating
1. The historical role of morphology in systematics

Since the beginnings of systematics as the science for inferring organismal relationships, morphology has played a prominent role in elucidating the tree of life (e.g., Haeckel, 1866). With the advent of cladistics (phylogenetic systematics) (Hennig, 1950, 1966) and numerical methodologies (mostly parsimony-based methods), the prevalence of explicit morphological data sets exploded in the 1980s and 1990s with the aim of testing animal classification systems at various taxonomic levels. With respect to the interrelationships of the animal phyla, several morphological data matrices were assembled (e.g., Brusca and Brusca, 2003; Eernisse et al., 1992; Jenner, 2004; Jenner and Scholtz, 2005; Meglitsch and Schram, 1991; Nielsen et al., 1996; Schram, 1991; Sørensen et al., 2000; Zrzavý et al., 1998), some becoming the basis for subsequent combined analyses of morphological and molecular data (e.g., Giribet et al., 2000; Glenner et al., 2004; Peterson and Eernisse, 2001; Zrzavý et al., 1998).

The development of these matrices was however hampered by two methodological impediments and few advances have been made since the early to mid-2000s. First, these matrices often used higher taxa as terminals, a practice eloquently critiqued by Prendini (2001), among other things by not allowing such higher taxa to be non-monophyletic. For example, it is now broadly accepted that Acanthocephala nests within Rotifera (e.g., Garey et al., 1996; Sørensen and Giribet, 2006; Wey-Fabrizius et al., 2014), something that could never be discovered by coding Rotifera and Acanthocephala as terminals of their own—at most one could place them as sister taxa. Second, many characters coded as absent/present assumed homology of non-comparable conditions (this is independent from other issues with applying a/p coding; see for example Pleijel, 1995), as critiqued by Jenner (2002). For
example a character state such as presence of spiral cleavage bears a homology statement (Hejnol, 2010), but its absence coding assumes that all other types of cleavage are homologous—something that obviously is not the case. Many other issues with coding characters and recycling data sets have also been exposed (e.g., Jenner, 2001).

While the use of species as terminals is a logical, albeit time consuming, solution to the first problem—and may require coding multiple species with disparate morphologies per phylum—resolving issues of homology can be complicated in many cases, especially when dealing with complex characters such as egg cleavage patterns or segmentation (Scholtz, 2002). Large morphological data matrices based on the exemplar approach advocated by Prendini (2001) are occasionally assembled for many animal groups (e.g., Bieler et al., 2014; Lee et al., 2013). However, the exemplar approach has yet to be applied to broad metazoan relationships, and few efforts seem to be currently directed that way.

2. **Molecular systematics in the era of genomics**

While morphological data sets played a pioneering role in elucidating the animal phylogenetic tree, molecules entered the game relatively early (Field et al., 1988), and included data from nearly all the animal phyla just a decade later (e.g., Giribet et al., 2000; Zrzavý et al., 1998). But these phylogenies relied mostly on sequences from one or a few genes obtained through Sanger sequencing of PCR (polymerase chain reaction)-amplified fragments, even after the first invertebrate genomes were rapidly being sequenced and assembled (Adams et al., 2000; Consortium, 1998). Sanger-based sequencing approaches were however limited in the number of genes that could be amplified and they often lacked signal in many deep nodes. Soon, with the availability of the first animal genomes, several
data sets used information from 100 (Blair et al., 2002) to 3,000 gene sequences (Dopazo et al., 2004). These data sets, severely limited in taxonomic scope, questioned what had become recent “dogmas” of molecular phylogenetics, such as the clade Ecdysozoa (Philip et al., 2005; Wolf et al., 2004). These studies however omitted most of the diversity of invertebrates, save for a few model organisms for which genomes were available.

While these few genomes were available for a handful of model organisms, generating whole genome data remained unreachable to the average systematist. The first attempts to analyze large data sets using broad animal sampling efforts thus combined the previous genomic data with a rapidly evolving set of multigene data sets derived from expressed sequence tags (ESTs) (Delsuc et al., 2006; Delsuc et al., 2005; Lartillot and Philippe, 2008; Philippe et al., 2007; Philippe et al., 2005; Philippe and Telford, 2006). Still, these studies mostly relied on publicly available data until Dunn et al. (2008) published the first comprehensive analysis of animal relationships using transcriptomics, soon expanded to include a few additional missing species and key lineages like acoelomorphs and placozoans (Hejnol et al., 2009). Many other analyses followed, especially re-evaluating the base of the animal tree (e.g., Lartillot and Philippe, 2008; Moroz et al., 2014; Philippe et al., 2009; Pick et al., 2010; Ryan et al., 2013), or some of its largest clades (e.g., Andrade et al., 2014; Dordel et al., 2010; González et al., 2015; Kocot et al., 2011; Lemer et al., 2015; Meusemann et al., 2010; Nesnidal et al., 2013; Philippe et al., 2011; Rehm et al., 2011; Rehm et al., 2014; Sharma et al., 2014; Smith et al., 2011; Struck et al., 2014; von Reumont et al., 2012; Weigert et al., 2014; Zapata et al., 2014). The newer analyses transitioned from small and sparsely populated EST data matrices, generated by Sanger sequencing, to more complete ones using genomes or transcriptomes obtained by 454 pyrosequencing, or Illumina sequencing, among other techniques. These high-throughput sequencing techniques have exploded, with an associated decrease in sequencing costs, shifting the
paradigm of nucleic acid sequencing, and resulting in a low cost per base and a throughput on the gigabase order of magnitude. Some phylogenetic studies therefore now include thousands of genes—mostly protein-coding genes are used in phylogenomic studies—and dozens of taxa, and larger data sets will soon be available (see several global initiatives in Kumar et al., 2012; Lopez et al., 2014; Robinson et al., 2011). One such data set recently explored the phylogeny of insects with an unprecedented amount of data (Misof et al., 2014). The first multi-genome phylogeny of an animal group appeared more than half a decade ago (Clark et al., 2007), and dozens of genomes can now be obtained for non-model invertebrate species, as recently shown by the 101 genomes of the monarch butterfly Danaus plexippus (Zhan et al., 2014). It is thus unavoidable that rapidly decreasing costs of sequencing genomes coupled with analytical developments will result in sufficient data to thoroughly examine many key nodes in the animal tree.

Whereas the available amounts of molecular data scale up faster than requisite technological developments (i.e., computation power, algorithmic efficiency), the gradual expansion of morphological data is far from paralleling the growth of molecular data acquisition, in spite of more efficient data collection. This has led some to question whether one should collect morphological data at all for the purpose of inferring phylogenies (Scotland et al., 2003).

3. Why is it important to maintain a research program in morphology?

The reasons for maintaining a research program in morphology are countless, as ultimately, a zoologist’s interest is to understand form and function, ecology and evolution, as well as all other aspects that may explain how our favorite organisms live, behave and
evolve (see Richter and Wirkner, 2014). Put plainly, one’s interest in animal phylogeny is
not the phylogeny per se, but to understand how the organisms evolve and adapt to their
environment, and how this reflects upon their morphology, ecology and behavior. And of
course, morphology will always be required for recognizing new taxa representing novel
body plans (e.g., Funch and Kristensen, 1995; Just et al., 2014; Kristensen, 1983; Kristensen
and Funch, 2000), “testing” molecular hypotheses, and to understand the evolution of body
plans (evo-devo). Therefore, rephrasing the title of this section, we may ask, “Why is it
important to maintain a research program in morphology for estimating phylogenies?”,
especially when entire genomes can now be collected for hundreds of species. It is thus
unquestionable that large molecular data sets derived from next-generation sequencing
techniques are becoming available and the preferred way to infer the relationships of many
animal groups, including most phyla. It is also evident that cladistic matrices to infer deep
metazoan relationships are comparatively stagnant, given the amount of effort and
expertise required to code properly a matrix using an exemplar approach and produce
sound hypotheses of homology for animals that share few macroscopic features, such as
between for example the two "C. elegans", Crella, the sponge, and Caenorhabditis, the
nematode. There are obviously some sound alternatives to the exemplar approach
advocated here, such as combining actual observations from different exemplars for
analytical purposes, especially to minimize amounts of missing data, but still this strategy
does not seem to scale up with molecular data. Nevertheless, molecular hypotheses face the
insurmountable limitation of necessarily omitting vast amounts of biodiversity—those that
are extinct but still provide sound explanations of a group’s phylogeny (Edgecombe, 2010).
Fossils constitute a key element for understanding the phylogenies of many animal
groups (Donoghue et al., 1989; Edgecombe, 2010; Wiens, 2004), from placental mammals
(O’Leary et al., 2013) to early panarthropods (Legg et al., 2013). In some cases, fossils in fact
increase congruence between morphological and molecular phylogenies. The role of morphology in coding fossil taxa for phylogenetic inference is thus undeniable and such practice becomes a strong advocate (one of many) for maintaining a healthy program in morphology. However, the proportion of phylogenies produced using fossils as terminals is just a tiny fraction of the total number of phylogenies generated every year, as they require expertise in more complex data than molecular phylogenies, and access to material often not broadly available for study.

Another aspect worth exploring, but seldom touched upon in the literature, is the information content of morphology and molecules, or even in some cases, that of the different character systems. Some early studies compared measures such as the consistency index of molecular and morphological data (Sanderson and Donoghue, 1989), but to my knowledge, few invertebrate studies have focused on discerning the information content of different systems. A seminal study by Harper et al. (2000) analyzed 43 anatomical and shell characters in anomalodesmatan bivalves, and then re-analyzed 18 characters which could be scored from shell characters alone—as many of these could also be scored for fossils. They then showed conflict between the shell-based tree and the cladistic hypothesis based on all the evidence (Harper et al., 2000). More recently, Bieler et al. (2014) analyzed broader bivalve relationships using a combined approach with multiple molecular markers and a specimen-based morphological approach of 210 characters (for a total of 22,680 cells with 16.6% of missing data), and then assessed the phylogenetic signal of multiple character systems. From the 210 characters, 99 had significant phylogenetic signal (when compared to the molecular and total evidence results), but a few character systems, including sperm ultrastructure and shell gross morphology and microstructure contained above average signal, while other character systems, such as alimentary and gill characters, are less informative than the average. Studies such as these can direct morphological
research or even assess whether fossils would contribute with significant signal—as in this case, two of the best character systems are available in the fossils. Analyses of phylogenetic signal in morphological characters are however scarce, or in some cases, limited to just a few characters to assess their possible value for inferring phylogenies (e.g., Gainett et al., 2014). Further and rigorous assessment of the phylogenetic signal in morphological data is certainly needed (see Wiens, 2004).

Another aspect to consider when relying on morphological data for inferring phylogenetic relationships is the time of data collection and interpretation, an aspect that has been properly outlined by just a few authors (e.g., Ramírez et al., 2007). Discussion has emerged in the community about “high throughput morphology”, and novel terms such as “next-generation histology”, or “phenomics”, have appeared, often trailing on the newest tendencies in molecular data generation. The use of automated methods for generating the raw data that will be transformed into the comparative morphological data is becoming more widespread, at least for technologies such as micro-computed tomography and magnetic resonance imaging (e.g., Ziegler et al., 2011), leading to accelerated acquisition of anatomical data (Ziegler, 2012; Ziegler et al., 2008; Ziegler and Menze, 2013). A major impediment to the phylogenetic analysis of such data is the so-called linguistic problem of morphology, or the translation of the anatomical raw data into anatomical descriptions, and these into morphological data matrices (Vogt et al., 2010). A solution to this problem can be provided by the development of ontologies—the use of a controlled and formalized morphological vocabulary (e.g., Ramírez et al., 2007; Richter et al., 2010; Vogt et al., 2010; Vogt et al., 2012; Vogt et al., 2013), but the practice and application of ontologies to distantly related groups is difficult. Modernizing data collection, storage, display and annotation are, together with the training of morphological experts who can interpret the characters, key aspects for returning morphological data to a more prominent position in
inferring phylogenetic patterns. However, there will always be a limiting step, character conceptualization, that will require expert assessment and that will limit automated procedures.

4. Morphology, fossils, dating, and phylogenomics

Dating trees has become common in systematics (Donoghue and Benton, 2007), as a time-tree allows not only discussing patterns, but also tempo, of evolution, and thus allows for testing specific hypotheses. The hypotheses often pertain to biogeography, and increasingly to comparative methods evaluating diversification rates and alternative models for evolutionary dynamics through time. Comparative methods are thus on the rise (Harmon et al., 2008; Rabosky and Alfaro, 2010) and require both precise dating for testing key innovations (Slater et al., 2012), where morphological or behavioral innovations cause differential diversification rates (e.g., Dimitrov et al., 2012) and allow testing for adaptive radiations, where niches are defined on the basis of morphology (e.g., Gillespie, 2004; Losos and Ricklefs, 2009). It is thus ironic that fossils have become formidable allies of molecules wherever inference of divergence times is required, while molecules have relegated morphological characters of living organisms to an almost marginal role in phylogenetics.

For the most part, researchers use fossils to calibrate molecular trees, providing information to constrain certain nodes (Donoghue and Benton, 2007). Some even explore issues of uncertainty of calibrations and multiple models (e.g., Mao et al., 2012; Warnock et al., 2012), but often assume that the taxonomy reflects the correct phylogenetic position of the fossils, something difficult to justify without proper phylogenetic assessment. It is for the latter reason that a minority of workers has opted to treat fossils as actual terminals
(the so-called “tip dating” or “total evidence dating” approach), using morphological data to infer fossil placement—as commonly done by paleontologists—while the molecular data guides the dating (Murienne et al., 2010; Pyron, 2011; Sharma and Giribet, 2014; Wood et al., 2013). Tip dating addresses on the one side the aspect of phylogenetic uncertainty of fossils (see also Parham et al., 2012), but also the concern that the date estimates are more accurate when fossils appear in a specific phylogenetic position, and not just as approximate constraints to nodes (Pyron, 2011), especially when taxon sampling is limited. This has been illustrated in a large arachnid data set with an exemplar fossil record (Sharma and Giribet, 2014), and has now been tested with ancient DNA from fossil humans (Rieux et al., 2014). The latter study, using a comprehensive data set of 350 ancient and modern human complete mitochondrial DNA genomes, shows that for the same data set, estimates based on individual dated tips are far more consistent with each other than those based on node calibration, and should thus be considered as more reliable. This result thus provides the scientific justification for preferring tip dating (using fossils as terminals) over node dating (using fossils to constrain nodes), encouraging the use of tip dating whenever possible—which may not always be possible.

There may also be other reasons for preferring tip versus node dating, these having to do with the idiosyncrasies of the fossil record. Some groups (e.g., molluscs, arthropods, brachiopods) fossilize well, while others (e.g., platyhelminths, nemerteans, rotifers) do not. In addition, some geological periods have more rock available to us or more favorable conditions for fossilization, so in some cases, a group that has missed a fossilization window (e.g., in the Carboniferous) may not appear again until many millions of years later. To provide an empirical example, the fossil record of Opiliones (harvestman arachnids) is exceptional in many respects, with the first Devonian fossils (Dunlop et al., 2003) probably close to the first appearance of this ancient arachnid order, with multiple lineages
preserved as three-dimensional fossils in the Carboniferous (Garwood et al., 2011; Garwood et al., 2014), and several other lineages known from the Mesozoic (see Dunlop, 2010 for a recent review). The phylogeny of Opiliones is well understood (e.g., Giribet et al., 2002; Giribet et al., 2010; Hedin et al., 2012a; Sharma et al., 2014; Shultz, 1998; Shultz and Regier, 2001), and while stem-group Eupnoi and Dyspnoi are known from the Paleozoic (Garwood et al., 2011), the first fossil Cyphophthalmi is only from Cretaceous amber (Poinar, 2008), between 45 and more than 300 million years after the diversification of the group (following Sharma and Giribet, 2014), and its exact phylogenetic position is unclear. In the case of Laniatores—the largest Opiliones suborder, which separated from its sister group around the Devonian—the oldest known fossil is from Eocene Baltic amber (Ubick and Dunlop, 2005). Therefore, while for some clades the first fossil occurs near their time of origin and thus may provide good estimates of the age of a clade, in other cases the first fossil occurrences may be closer to the tip than to the base of a clade. If the latter are used for dating a clade, the excessively young fossil may produce a “push towards the present” effect for the whole clade, and thus grossly underestimate the age of the clade in question (Fig. 1).

It is clear that this may oversimplify the way node dating is done, as this fossil would generally be treated as some sort of minimum age, but still, in the absence of much older fossils in nearby clades, this effect should be noticeable. However, if the fossil were included as a morphological terminal in a phylogeny instead of used for constraining a clade, its position within the phylogeny should reflect its more derived position, and thus engender a more precise dating of the whole clade. Of course one could argue that a more precise node than “Cyphophthalmi” or “Laniatores” should be chosen when using those groups, but the Burmese cyphophthalmid fossil was described in the wrong family and the phylogenetic position of the Baltic fossils is difficult to ascertain without a comprehensive phylogenetic
analysis of Laniatores, which was entirely lacking at the time of its description. One could also argue that the taxonomy should be fixed first, but in the real world, molecular systematists tend to follow the taxonomy from the taxonomists describing the fossils and such fossils may often not be precisely positioned.

Another common problem of using fossils to date nodes is that sometimes such nodes are misrepresented, because most authors "recycling" fossil papers for their analyses without critical phylogenetic assessment tend to take conservative approaches. For example, Hedin et al. (2012b) used the age estimate of two members of closely related families of Eupnoi as the age for the entire suborder (see Sharma and Giribet, 2014), resulting again in a push towards the present effect, but in this case for a different reason. Had they coded the fossils as terminals, the two species represented molecularly would have not been placed with the fossil they used for dating the node. A clear demonstration of these was recently provided in two analyses of fossil Opiliones (Garwood et al., 2014; Sharma and Giribet, 2014).

Tip dating however requires on the one hand morphological expertise, and also algorithmic developments, as currently no software allows analyzing simultaneously complex genomic-scale data (i.e., thousands of loci) and morphological data in a probabilistic framework. To my knowledge, no software additionally enables efficient dating using morphological and phylogenomic data matrices (see Sharma and Giribet, 2014). Due to the complexity of nucleotide and/or amino acid transition matrices, assigning accurate models to simultaneous analyses of morphological and molecular sequence data is inherently difficult, and recapitulates an older problem of how different data classes should be weighted in total evidence analyses (Bull et al., 1993; Giribet, 2010; Kluge, 1989).
5. Final Remarks

Morphological data have traditionally played a key role in inferring animal phylogeny but recently have been replaced by molecular data due to multiple factors, perhaps most prominently due to the efficiency and low cost of sequence data acquisition and the development of analytical methods to deal with a vast amount of characters with a small number of states. This is unfortunately leading to a large asymmetry in the amount of studies using either type of data, while morphological data can still provide an important framework to assess molecular trees—congruence between the two data sources generate confidence in the results, while incongruence often leads to interesting aspects of body plan evolution, including convergence. Several developments, including high throughput morphological data collection and the establishment of clear ontologies are now required to restore the place of morphology as a main player in elucidating phylogenetic relationships. These will also aid in ameliorating previous problems with higher taxa as terminals and poorly formulated homology statements. But most important is the need of using fossils to improve phylogenetic inference and to help generate time-trees. In the latter arena, it is beginning to become evident that tip dating (total evidence dating) may be philosophically preferable to other forms of dating using fossils to calibrate accepted nodes. As refined bioinformatic tools become available, it should be possible to analyze genomic and morphological data in conjunction and to use these analyses for providing the most accurate estimate of the timed phylogenies of a group. Molecular data may easily provide the skeleton of the animal tree of life; morphology will still be required to test such skeleton, to place key fossils, and to provide a sound temporal framework for such animal tree.
Acknowledgments

This invited paper derived from the Presidential Lecture at III International Congress on Invertebrate Morphology, and I am indebted to the organizers and the editors of this volume for their invitation, especially to Gerhard Scholtz, Thomas Stach and Carsten Lüter. Some of the ideas outlined here derive from many discussions and conversations about total evidence, morphology, dating and genomics, I have had with many close colleagues through the years, especially with Greg Edgecombe, Rosa Fernández, Gustavo Hormiga, Prashant Sharma, Lars Vogt, and Ward Wheeler. Prashant Sharma, Rosa Fernández, Stefan Richter and an anonymous reviewer provided comments and criticism that helped to considerably improve this review.
References


transcriptome data reveal syndermatan relationships and suggest the evolution of endoparasitism in Acanthocephala via an epizoic stage. PLoS One 9, e88618.


Fig. 1. Effect of a young fossil on the phylogeny of an old lineage, represented in this case by an amber fossil for a clade with an old history (A). Supposing that the fossil could be assigned to the red clade, based on morphology, it could be used for tip dating (B) or for constraining a node (C). In tip dating, the clade could accommodate for extensive cladogenesis occurring prior to the fossil, if the fossil were positioned in a derived position in the phylogeny of the group. In node dating, the fossil may cause a “push towards the present” effect. See text for discussion.