Renaissance taxonomy: integrative evolutionary analyses in the classification of Scyphozoa

Michael N Dawson

Coral Reef Research Foundation, Koror, Palau. Present address: Section of Evolution and Ecology, Division of Biological Sciences, Storer Hall, One Shields Avenue, University of California, Davis, CA 95616, USA.
E-mail: mndawson@ucdavis.edu

New tools and techniques invigorate taxonomy through discovery and description of new organisms. New editions of the International Code on Zoological Nomenclature explicitly attempt to accommodate scientific advances, the problems they bring, and individual expertise in documenting the diversity of animal life. Yet, in practice, one of the most important biological and technical advances of the last quarter century—the democratization of large scale sequencing of DNA—remains at the fringe of metazoan taxonomy, where it keeps remarkable company with evolution. I discuss a more inclusive approach to taxonomy, primarily in the context of differentiating and describing species and subspecies of Scyphozoa. Global concern regarding biodiversity has rejuvenated efforts to discover and describe species; it may yet stimulate a renaissance if biological classification adopts a total evidence approach.

REVIEW

‘We are entering a renaissance of species discovery with the use of new tools and technology’
(Raskoff & Matsumoto, 2004)

Specifically, Raskoff & Matsumoto (2004) were referring to advances in submersible and remotely operated vehicle technology that have allowed researchers to dive to greater depths and discover new families of scyphomedusae (Matsumoto et al., 2003; Raskoff & Matsumoto, 2004). Their message though, has a broader context. Deep-ocean exploration is just one of many advances in biological oceanography starting with the Challenger Expedition—and more recently including SCUBA diving, blue-water diving and zooplankton ethology (Hamner et al., 1975; Hamner, 1985), new culturing techniques (Greve, 1975; Kaebelkein et al., 2002), and environmental genetics (particularly of bacteria; e.g. Béjá et al., 2002; Fuhrman et al., 2002)—that have increased our knowledge and understanding of marine biodiversity through discovery and observation. Other advances have been co-opted or developed in parallel for the more pedestrian but essential activity of differentiating and describing the newly discovered biodiversity, including each new edition of the International Code of Zoological Nomenclature (e.g. ICZN, 1999), electron microscopy (e.g. Östman & Hydman, 1997), molecular phylogenetics (particularly of bacteria; Woese et al., 1990; Acinas et al., 2004), and online publication of key historical documents (e.g. Kramp, 1961; Russell, 1970, see www.mba.ac.uk/nmbl/publications/pub_lists/publications.html). Many innovations have been incorporated quickly and seamlessly into modern taxonomy, but some have not. For example, electron microscopy was used in 1940—less than a decade after invention of the transmission electron microscope and before its commercial development—to establish the particulate nature of bacteriophages (Summers, 2000) and electron microscopy has since been assimilated without controversy into the repertoire of metazoan taxonomists (Sperling, 2003) becoming an essential component of many species descriptions (e.g. Moreira et al., 2003). In contrast, molecular genetic techniques—including protein electrophoresis, which was developed in the late 1930s (Tiselius, 1937), and DNA sequencing (e.g. Sanger et al., 1977) which was made broadly accessible after invention of the Polymerase Chain Reaction (e.g. Sakai et al., 1985)—remain extraneous to metazoan species descriptions.

The disparate histories of electron microscopy and molecular genetics present an interesting contrast because both intuitively fulfill the same role in taxonomy, i.e. they are ways of acquiring additional taxon-specific information when traditional morphological
approaches fail to identify or differentiate taxa. They are similar in other ways, too. Both require particular expertise (oneself, a colleague, or a specialist facility), both require expensive machinery and reagents, and neither is applicable in the field. So, why, in contrast to electron microscopy, has molecular genetics not been assimilated into the taxonomic toolkit?

Why not molecular genetics?

The debate over whether molecular genetics should or should not be used to describe species has been re-ignited by the proposal of DNA barcoding—using a single short DNA sequence to catalogue all species on Earth (Hebert et al., 2003a). Protagonists for DNA barcoding advocate replacing morphology-based with sequence-based classification on the grounds that barcoding: (1) is still a powerful identifier if done on parts of, rather than whole, organisms; (2) is practicable, and easily comparable, across all life history stages; (3) can identify cryptic species; (4) provides a precise, digital, discrete, description; and (5) makes species identification possible by non-specialists unfamiliar with the intricacies of morphology (e.g. Hebert et al., 2003a; Stoeckle et al., 2004).

Critics of DNA barcoding dispute claims by Hebert et al. (2003a) that identification problems attributable to phenotypic plasticity, cryptic species, and misdiagnoses are sufficiently common to warrant transition to sequence-based classification (Will & Rubinoff, 2004). They also highlight the heterogeneous rates of molecular evolution that exist between different segments of genomes and among taxa, such that any molecular approach would have to consider many rather than a single gene (see also Hebert et al., 2003a,b; Tautz et al., 2003), and they lament the loss of breadth of information that would inevitably result from a sole focus on DNA sequencing (Will & Rubinoff, 2004). Will & Rubinoff (2004) are almost diametrically opposed to Hebert et al. (2003a) in their belief that morphology is sufficient in most cases and that molecular data should be used only as a last resort.

Both arguments have merit. For example, an unequivocal digital species description (Hebert et al., 2003a) is practically preferable to a ‘vivid ... poem’ (Winston, 1999, p. 201–202). Molecular analyses demonstrate that cryptic species, or at least species that are exceedingly difficult to differentiate reliably using only morphological criteria, and misdiagnoses are at least as common as morphologically robust identifications in some groups of organisms, including Scyphozoa (e.g. Dawson, 2004). In contrast, the very slow rate of molecular evolution of COI is a serious drawback for any proposed DNA barcode of Anthozoa (Hebert et al., 2003b), and our information base will be restricted if non-molecular approaches are neglected. This last point highlights the fact that several proposed benefits of DNA barcoding—speeding documentation of biodiversity and ensuring accurate rapid identifications in non-taxonomic work, promoting technological advances that will speed and increase accuracy of field identifications, contributing directly to building a complete tree of life, and increasing the profile and value of museums’ collections (Stoeckle et al., 2004)—could equally be tangible benefits of renewed effort and funding focused on morphological taxonomy. Indeed, one—hastening the advent of an online encyclopedia with a web-page for every species of plant and animal (Stoeckle et al., 2004)—would be positively enriched by a multitude of datatypes.

Fortunately, DNA barcoding is just one aspect of a broader, longer-standing (e.g. Thorpe et al., 1978), and more moderate debate over ‘molecular taxonomy’. Protagonists on both sides of this debate recognize benefits in integrating molecular genetic data into species descriptions, although they may differ on the approach to and extent of integration. For example, modern taxonomic definitions of ‘diagnosis’ and ‘description’ do not exclude or recommend any type of data, deferring that decision instead to specialists in different taxa (ICZN, 1999; Seberg et al., 2003). In turn, Tautz et al. (2003) note that ‘a DNA-based system must be firmly anchored within the knowledge, concepts, techniques and infrastructure of traditional taxonomy.’ So, the real issue concerning molecular genetic data in species descriptions is, like deep sea research (Malakoff, 2002), not whether to do it, but how to do it.

How to integrate molecular techniques?

The most obvious and controversial (after DNA barcoding) approach is also the most facile; molecular data should be routinely assimilated into species diagnoses and descriptions. Both advocates and critics of DNA taxonomy support this position (Seberg et al., 2003). Logically, this is simply good scientific practice; scientific publications should report all the data necessary to support the conclusions made. Increasingly, species (re)descriptions will define units that were delimited initially by molecular analyses whether or not subsequent re-analysis of morphology is able to disentangle historical confusion borne of, for example, morphological homoplasy. Giving morphological data pre-eminence in species descriptions at the expense of molecular data that were critical to distinguishing taxa is disingenuous and, more importantly,
simply invites the same mistakes to be made again. For example, even though ‘morphological … [n]umerical taxonomy [and] effective statistical packages … offer … a renaissance of morphological studies’ (Guerro et al., 2000) such approaches also demonstrate that traditional, morphology-based, scyphozoan systematics has been unstable for the simple reason that morphological variation does not always reflect species boundaries (Dawson, 2003).

It might be argued that the diagnosis and description are not the appropriate place for molecular genetic data and that data other than morphological data should be, and already are, accommodated in the taxonomic discussion (e.g. ‘Remarks’ in this journal; Winston, 1999; e.g. Matsumoto et al., 2003). Such data usually include ecology, distribution, and behaviour and other such information, sometimes including genetics, on the taxon. Most of these are not specific properties contained within a type specimen and are, therefore, in many ways accessory information, but molecular genetic (e.g. DNA sequence) differences, like the morphological features discovered through electron microscopy, are actual, recoverable, properties of type specimens (particularly those preserved in ethanol, or otherwise appropriately treated). There is, therefore, no obvious reason to treat molecular genetic and electron microscopic characters differently in species descriptions. Both can, and should, be included in the diagnosis and description, particularly when molecular data are more definitive than morphological data. Interestingly, this argument can even be extended to behaviours if they can be preserved appropriately, for which precedent exists in the description of ichnotaxa (ICZN, 1999).

Several frameworks for integrating multiple datatypes exist. One, emerging from dissatisfaction surrounding solely molecular definitions of Evolutionarily Significant Units (ESUs; Moritz, 1994), recognizes the continuity and complexity of the evolutionary process and proposes a scheme to guide taxonomic interpretation and conservation decisions (e.g. Crandall et al., 2000). This scheme is based on the idea of ‘exchangeability’ or ecological equivalence and was originally described as the ‘cohesion species concept’ (Templeton, 1989). Although there are problems with any species concept, in this case possibly the number and complexity of factors that can be considered, exchangeability has received renewed favour in the marine literature as a kind of ecological species concept (Knowlton, 2000). It implicitly acknowledges that no one datatype will be sufficient to correctly identify all species in all given circumstances and it therefore promotes something of a ‘total evidence’ approach (sensu Carnap, 1950; see Kluge, 2004).

Total evidence is not a simple topic. Although its essence is intuitive—if multiple independent lines of evidence support a particular conclusion they will overwhelm random variation leading to the correct phylogenetic conclusion and, moreover, the more data that support the conclusion (and reject competing hypotheses) then the greater the confidence in the final result—it’s application can be challenging. Fortunately, 15 years of discussion in the phylogenetic community has resulted in philosophical and methodological advances that provide explicit (although debated) frameworks and tools for integrated analyses of molecular, morphological, and other datatypes (e.g. Kluge, 1989, 2004; Bull et al., 1993; Jenner, 2004; Lecointre & Deleporte, 2004).

Total evidence is valuable because it can incorporate morphological and molecular and other datatypes providing the large amount of (appropriately distributed) data necessary to reconstruct robust phylogenies (Jenner, 2004). Many solely morphological studies of scyphomedusae will likely unavoidably lack sufficient data for robust statistical conclusions because the simple structure of scyphomedusae yields insufficient characters to generate a compelling signal of historical relationships between even a moderate number of species (e.g. Gershwin & Collins, 2002; see Dawson, 2003). Total evidence also can integrate datatypes as diverse as molecules of extant and morphology of extinct taxa. Furthermore, assuming simultaneous (character congruence) analyses as opposed to sequential (taxonomic congruence) analyses of multiple datasets, the total evidence framework is most conceptually in tune with evidence of interaction between morphology and behaviour (and molecules) during evolution, as in the evolution of morphology, swimming behaviour, and vertical migrations in Mastigias (Dawson & Hamner, 2003; Dawson, 2003a; see also Peichel et al., 2001; Shapiro et al., 2004).

The application of phylogenetic methods to resolve these difficult issues will have the added benefit of advancing the true ‘systema naturae’ of classification, that based on evolutionary relationships (e.g. Darwin, 1859). The central importance of phylogeny is recognized in the preface to the 4th edition of the ICZN, ‘conventional Linnaean hierarchy will not be able to survive alone: it will … coexist with the ideas and terminology of phylogenetic (cladistic) systematics’.

1. For \( n \) species, \( 2(n-1) \) synapomorphies are required to create a fully dichotomous rooted phylogeny; for all branches in this tree to receive over 90% bootstrap support c. \( 26(n-1) \) synapomorphies are required, i.e. three synapomorphies per branch uncompromised by homoplasies or down-weighting (Dawson, 2003).
(Minelli & Kraus, 1999), and in development of the Phylocode, which has been designed to be used concurrently with rank-based codes, such as the ICZN, with minimal disruption of existing nomenclature (Cantino & de Queiroz, 2004). The phylogenetic approach is not new to scyphozoan taxonomy—Mayer (1910) favoured the ‘indication of affinities and the discovery of relationships’—but advances in data gathering and analytical techniques put us in an unprecedented position to apply this approach in an objective and robust manner.

How much sequence divergence represents a species-level difference?

The question of how much sequence divergence is prescriptive of species-level divergences inevitably arises in discussions such as this (e.g. Hebert et al., 2003a; Acinas et al., 2004). In a total evidence framework, the question is redundant because other attributes always need to be considered.

The question can be recast more usefully: how much sequence divergence separates species that historically have been reliably differentiated using morphologic criteria? Rephrasing the question in this way has two benefits. It explicitly delimits the taxa being discussed, which circumvents the issue of rate heterogeneity for predictive species descriptions, and it allows future measurements to change the answer. In scyphozoans, to date, mitochondrial cytochrome c oxidase subunit I (COI) sequence, the ‘barcode’ proposed by Hebert et al. (2003a), shows divergences between Aurelia aurita, A. labiata, and A. limbata of ≥18% (Dawson & Jacobs, 2001) and divergences between Cassiopea andromeda and C. frondosa of ≥19% (Holland et al., 2004). Nuclear internal transcribed spacer one (ITS1) sequence divergence between A. aurita, A. labiata, and A. limbata is ≥16% (reanalysis of Dawson & Jacobs, 2001). Mitochondrial 16S sequence divergence between A. aurita, A. labiata, and A. limbata is ≥11% (Sroth et al., 2002). Thus, divergences of similar magnitude in COI and ITS1 that have been measured within other morphospecies of Aurelia and Cassiopea (e.g. sensu Kramp [1961] or Hummelinck [1968]) have been interpreted as strong evidence of ‘cryptic’ species (e.g. Dawson & Jacobs, 2001; Dawson, 2004; Holland et al., 2004). On reconsideration, these cryptic species often appear to correspond morphologically to varieties or species recognized by earlier workers (e.g. Mayer, 1910; see Dawson, 2003). Intraspecific genetic variation in these cases is generally ~2% or less (re-analyses of Dawson & Jacobs, 2001; Sroth et al., 2002; Dawson, 2004; Holland et al., 2004), with a few notable exceptions. Clades of Aurelia labiata north versus south of the Strait of Juan de Fuca, Aurelia aurita in the eastern and western North Atlantic, and Cassiopea mosaicus west versus north of Cape Howe are each separated by approximately 4% sequence divergence in either COI or ITS1 (but not in both; Dawson & Jacobs, 2001; Dawson, 2005b). In each case, slight morphological differences also distinguish the molecular clades (Gershwin, 2001; Dawson, 2005c). These clades, like those that are distinguished by large morphological and only slight molecular differences and which are geographically isolated and very recently or relatively recently diverged, clearly represent ESUs and have been interpreted as distinct subspecies (e.g. Dawson, 2005a,c).

Subspecies are an often under-appreciated taxonomic rank (Randall, 1998). Problems perceived in working with subspecies include that they are transient entities, that boundaries of many marine species are poorly known (and that subspecies therefore cannot be known), and that they afford less attention than species in taxonomic, funding, and conservation circles. These factors no doubt contribute to the dearth of taxonomic papers describing new sub-species (3%; most of them in well-known taxa such as birds and butterflies) compared to those describing new species (33%) or new genera (8%; Winston, 1999) although, logically (as well as empirically [see above]), subspecies must be abundant2. However, these considerations are not unique to subspecies. Species and orders, e.g. Conulatae, are also transient (just on different time-scales), infra-specific taxa can be recognized and defined if sufficient contextual information exists even though species may not be well circumscribed, and subspecies still attract more attention than varieties, forms, populations, ESUs, or other management units. Moreover, the foundational role of subspecies in evolution of higher taxa means they have particular relevance in terms of the evolutionary process and that they should be clearly identifiable entities below the species level in a phylogenetic framework. Indeed, as an indication of their importance, it is pertinent to note that Ernst Mayr, one of the most influential systematists and evolutionary biologists of the 20th Century, described 445 subspecies of birds from collections at the American Museum of Natural History (Pennisi, 2004).

---

2. Assuming cladogenesis predominates over anagenesis, and barring saltatory evolution such as polyploidy, nearly all species must evolve from subspecies but all subspecies do not necessarily become species. So, assuming that evolution is an ongoing process and that the mean longevity of sub-species is not dramatically less than one-tenth the mean longevity of species (~2–7 million years; Koch, 1980; Baumiller, 1993), there should be proportionally more sub-species being described.
The way ahead?

Inevitably, each person has particular aptitude dealing with certain kinds of data which, in a total evidence framework, are all pertinent to understanding relationships between taxa and their classification. But no one can do it alone (as is apparent from the limited scope of several recent publications in this area: Dawson & Jacobs, 2001; Dawson, 2003; Dawson, 2004). Accurate taxonomy is the foundation for comparative biology, biodiversity studies, and successful conservation and, to be completed quickly and robustly, will require international consortia of expertise. Describing the global biodiversity of Scyphozoa is just the kind of 'big science' currently favoured by biological funding agencies. Scyphozoan taxonomy finds itself at a time of great opportunity to meet this challenge. Technological advances allow new waters to be charted (Raskoff & Matsumoto, 2004). Molecular phylogenetic techniques permit species boundaries to be identified with relatively high statistical confidence (e.g. Dawson, 2003). Modern morphological approaches then enable detailed comparisons of populations, subspecies, and species (e.g. Dawson, 2003, 2005a). Invaluable expertise exists in established approaches to taxonomy accompanied by a deep appreciation and knowledge of the ICZN and historical literature. Thus we have an unprecedented opportunity to integrate different datatypes in holistic approaches that allow us to chart patterns of morphological and life history evolution (Collins, 2002; Gershwin & Collins, 2002; Marques & Collins, 2004), geographic patterns of speciation (Dawson, 2005b), ecology (Dawson & Martin, 2001; Schroth et al., 2002), and so on, greatly enriching our understanding of, and ability to describe, the biodiversity of Scyphozoa.

Discussions with I. Martin, H. Mianzan, T. Trnski, and G. Wilson encouraged me to express and clarify these ideas. A. Collins provided a thoughtful review that improved the final manuscript. This work would not have been possible without the help of the many people who have contributed in numerous ways to the studies out of which this discussion grew.

REFERENCES


*Journal of the Marine Biological Association of the United Kingdom* (2005)


