

The Systematist

Newsletter of the Systematics Association
Number 31 2010

www.systass.org

ISSN 1744-5701

In Memoriam

- Christopher Humphries
1947-2009



2009 Biennial Reports

- Asking the right questions
- Flowers on the tree of life
- Evolution of plant-pollinator relationships
- The importance of being small

Book reviews

- Major transitions in vertebrate evolution
- The science of describing
- All creatures
- Biodiversity databases
- Darwin: a reader's guide



Other Reports

- Late Triassic Terrestrial Biotas and the Rise of Dinosaurs
- Xth Symposium of the International Organization of Plant Biosystematists
- SRF Report

Inspirations

- Olaf Bininda-Emonds



information about the medicinal properties of plants were increasingly de-emphasized. At the same time an independent literature took root that catered to the blossoming population of amateur lovers of plants and flowers. Florilegia became the coffee table books of the seventeenth century. They were picture books with detailed copperplate engravings or etchings of plants and flowers. Engraved pictures could show much more details than the woodcuts previously used. The emphasis in florilegia came heavily to rest on the illustrations, so that the text withered down to the bare essentials of practical information of interest to gardeners, leaving out scholarly controversies and details about the places and times of the collected species.

The final period discussed in Ogilvie's book is from 1590 to 1620, and this was the period of the systematizers. The number of newly described plants became so large that naturalists were now forced to concern themselves with efficiently organizing their data. They needed to be able to classify the known species of plants (the number increased from hundreds in the mid sixteenth century to thousands a few decades later), but also accommodate the unstoppable flood of new species. Consequently, simply adopting a kind of folk taxonomy, such as the Theophrastean categories of trees, shrubs, bushes, and herbs, was no longer sufficient. It is here that we encounter the very foundations of modern scientific systematics. The work of Andrea Cesalpino looms large as the forerunner of the later Linnaean system, and the work of Adam Zaluzianski is noteworthy as a first attempt to introduce classification as a pedagogical or practical tool.

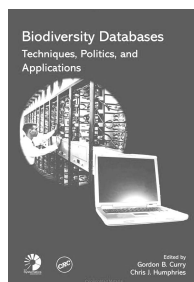
In short, Ogilvie's study shows how with a minimum of intermediate stepping stones the ancient knowledge of plants and

animals as recorded in the works of Aristotle and Theophrastus (BC), and Pliny and Dioscorides (AD) came to form the foundational canon of Renaissance natural history. In the hands of the humanistically educated Renaissance naturalists these works formed the basis of a monumental change in the scholarly attitude towards nature's variety, replacing an earlier natural philosophical approach to animals and plants that was aimed at distilling universal truths, with a new and explicitly descriptive tradition that charged itself with producing a complete inventory of all nature's particulars. It was when the number of known and described species started to outstrip the individual scholar's capacity to keep them all apart, that we find the cradle of modern systematics. Ogilvie's book is so rich in detail that it deserves prolonged study after an initial reading. Because of its almost exclusive focus on botany, the reader is left with enough food for thought, for example how developments in the nascent discipline of zoology mirrored or differed from those observed in botany, and how the new focus on accurate and comprehensive description in natural history dovetailed with similar concerns in the visual arts.

Ronald A. Jenner

Department of Zoology, The Natural History Museum, London, UK
<r.jenner@nhm.ac.uk>

Biodiversity Databases: Techniques, Politics, and Applications



Curry GB, Humphries CJ (eds.) 2007. Systematics Association Special Volume 73. CRC Press, Taylor & Francis Group,

Boca Raton. ISBN: 9780415332903 (Hardcover). US\$89.95.

This concise volume of ten contributed chapters is based on papers presented during the 4th Biennial Conference of the Systematics Association held in Dublin from August 18-22, 2003 (see Humphries 2003). The field of biodiversity informatics is moving forward at a fast pace, facing on one side a vast amount of heterogeneous legacy data and on the other side a somewhat amorphous, but increasingly demanding, user community. To make it all work we need potent ideas and technologies at the center, along with a "new spirit of collaboration" (Holsinger et al. 2007). For those who struggle to keep up with the latest in cybertaxonomy, a compilation of papers by leading experts is certainly welcome. Nevertheless, the scope of *Biodiversity Databases* is more modest in comparison to relevant predecessors such as Fortuner (1993) or the excellent series *Taxonomy for the Twenty-First Century* (*Phil. Trans. Royal Soc. B* 359: 559-739). Most chapters emphasize conceptual and technical improvements within the European Network of Biodiversity Information (ENBI), where more than half of all described species and type specimens are located. I will first offer a brief overview of each chapter [1, 2, etc.], and then close with general comments.

Lane & Edwards [1] give a three-page summary of the mission of the Global Biodiversity Information Facility (GBIF). Los & Hof [2] describe how this mission is being implemented at the level of ENBI. The number of relevant projects (and corresponding acronyms!) is impressive, spanning from species name checklists, initiatives to virtually access specimen records and molecular sequences, to standardized information resources on European plant communities. The authors then list activities aimed at improving the Network in

terms of database integration, analytical tools, and web-based platforms for users. More than 20 relevant WWW resources are cited at the end, a useful convention that was adopted in several other chapters.

Berendsohn & Geoffroy [3] review a salient proposal to improve semantic integration between databases using the taxonomic concept approach (see Kennedy et al. 2006; Franz et al. 2008). The idea to employ taxonomic concepts – names used explicitly and exclusively in the context of a

empirical comparison of two approaches that use morphometric information to produce automated species identifications, viz. linear discriminant analysis and artificial neural networks. Both methods are tested on a set of foraminiferal species. Although the former performed slightly better, the latter is more generally applicable and less time consuming, leading to a nuanced call for exploiting these complementary properties to maximize the potential for automation. The authors finish with a wide ranging defense of

distributions, and other relevant data. Jones [7] introduces the notion of a biodiversity Grid – a distributed infrastructure that facilitates high-performance computing to model climate change, future species distributions, and similar data- and algorithm-intensive analyses (see also Jones et al. 2006). In addition to outlining the architectural needs for such a Grid, the author stresses the importance of establishing a stable biodiversity ontology in order to run workflows involving heterogeneous datasets. The BiodiversityWorld project has made progress towards these goals.

Triebel et al. [8] provide an update on LIAS, a global information system for lichenized and non-lichenized ascomycetes. Originally based on the DELTA format (<http://delta-intkey.com/>), LIAS now has a networked database and on-line portal offering taxonomic and descriptive information on more than 2700 taxa, identification keys, and regional checklists, among other services. An active community of specialists is involved in the expansion and integration of LIAS within the GBIF network. White [9] analyzes common challenges that occur when biodiversity databases are merged to yield a more comprehensive information resource. Most examples are taken from the 20+ year history of ILDIS, the International Legume Database and Information Service. Of particular interest are situations where regional databases represent different assessments with regards to the validity and taxonomic inclusiveness of species names (*pro parte* synonymy, etc.). The Litchi Project aims to merge checklists while applying a set of rules to identify and resolve taxonomic conflict. "Intelligent" species links (generated via prior searches or registration) can connect multiple resources with similar content. The author advocates maintaining a system of cross-mappings that

I recommend *Biodiversity Databases* to anyone who is looking for a good entry point into the field of biodiversity informatics

particular circumscription (Berendsohn 1995) – is one of the most significant conceptual contributions to systematics made so far by developers of taxonomic databases. The authors explain how this approach permits assessments of taxonomic stability as multiple succeeding classifications are imported into a single repository and its elements cross-linked via concept relationships. They argue that taxonomic concepts represent a valid solution to the problem of integrating evolving taxonomic perspectives without imposing a consensus view. Projects focusing on algae, mosses and vascular plants underscore the feasibility of the concept approach. Scoble & Berendsohn [4] recount the historical trajectory and challenges of networking specimen-level information generated by natural history collections in Europe. Issues of technical implementation, scalability, information quality, intellectual ownership, and general operability across 30 or more countries are addressed with laudable realism. Support from the European Commission has played an important role in developing the database network.

MacLeod et al. [5] present an

automated identification projects (see also Gaston & O'Neill 2004), arguing that such services are needed to reinvigorate morphological systematics. But surely there remain some less than ideal parallels to traditional DNA barcoding, in the sense that we are limited methodologically to one or a few sets of features to identify species when others may reflect more immediately upon particular historical speciation events. There are trade-offs between the practicality for automation and maximizing the diversity and diagnostic power of features selected to achieve reliable species recognition.

Curry & Connor [6] examine new technologies capable of extracting taxonomic descriptions from historical publications and parsing them into a structured digital format based on markup language tags. Their proof-of-concept study involves a monograph of brachiopods from which they scanned pages with high accuracy using optical character recognition software. The development of a suitable XML parser program adds semantic structure and allows selective querying of names, synonyms, authors, descriptions,

resemble concept relationships (see [3]).

In the final and longest chapter of the volume, Andersen et al. [10] develop an innovative approach using WORLDMAP (Williams 2001) to prioritize areas for conservation on Borneo based on distributional information (> 5000 records) of endemic rattan palm species. Their analysis reflects great attention to detail and scholarship. Application of the criterion of complementarity indicates that 26 grid cells out of a total of 1087 are required to represent all taxa at least once. Congruence with less extensive datasets of birds and butterflies appears to solidify the original recommendations for areas of high conservation priority.

Even though *Biodiversity Databases* no longer reflects on the very latest developments (such as globally unique identifiers; see Pyle et al. 2008), it succeeds in touching on many prevalent themes in the field. The exclusive spotlight on Europe is readily justified given that more than 50% of all recognized species and type specimens and over 4000 collections are potentially involved. As stated on numerous occasions, GBIF and incentives stemming from the Framework Programme of the European Commission have helped forge Europe's leadership role in biodiversity informatics. The fruits of such high-level support are fully on display in the volume, as large-scale networks of high-quality biodiversity data are starting to become reality. One must also applaud the balanced line-up of papers emphasizing new concepts (taxonomic concepts, workflows), issues of implementation (databases, networking), and empirical applications (morphometrics, conservation). As different factions (e.g. phylogeneticists versus conservationists) within biodiversity informatics pursue their primary agendas, it is easy to move away from the vision of a globally

coordinated platform for information about species.

On the other hand, in my estimation *Biodiversity Databases* would have benefited from a more confrontational and critical perspective towards some of the greatest challenges we are experiencing at this juncture. Too often the reader gets little more than a rundown of original goals, then progress towards them as implemented in a series of projects, then more goals. The message one might take from this promotional style of presentation is that simply plugging ahead will take care of things. But there are *major* obstacles to be overcome on the path towards reliable and precise integration of biodiversity information. The sociopolitical challenges involved in merging data and factional interests (including financial ones) are not adequately stated and scrutinized. An unresolved tension runs through the volume between producing traditional checklists with species names versus integrating different regional or temporal taxonomic perspectives which would require a community-wide commitment to the taxonomic concept approach. Is it time for a paradigm change? Other relevant issues are left unmentioned, for example the discouraging fact that none of the most popular – and populated – biodiversity databases are designed to readily incorporate new phylogenetic insights. Do we need a new ontology translating trees into database fields? In short, the vision of a global and semantically concordant network needs strong ideas and passionate advocates in order to succeed. In that sense I was disappointed by the sober tone adopted by many authors. I also missed a critical synthesis and outlook at the end of the volume.

On the technical side, the chapters are generally well presented, illustrated, and indexed. An unfortunate exception is Figure 7.1 which should show a Grid

architecture instead of repeating Figure 6.7. I recommend *Biodiversity Databases* to anyone who is looking for a good entry point into the field of biodiversity informatics, with the qualification that the reality of data integration might be more "lively" than some chapters let on.

References

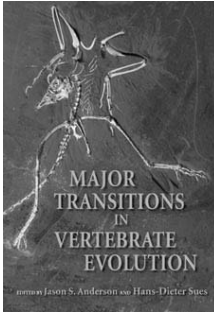
- Berendsohn WG. 1995. The concept of "potential taxa" in databases. *Taxon* 44: 207-212.
- Fortuner R. 1993. *Advances in Computer Methods for Systematic Biology. Artificial Intelligence, Databases, Computer Vision*. Baltimore: John Hopkins University Press, Baltimore.
- Franz NM, Peet RK, Weakley AS. On the use of taxonomic concepts in support of biodiversity research and taxonomy; pp. 63-86. In Wheeler QD. (ed.), *The New Taxonomy. Systematics Association Special Volume Series 74*. Boca Raton: Taylor & Francis, 63-86.
- Gaston KJ, O'Neill MA. 2004. Automated species identification: why not? *Philosophical Transactions of the Royal Society of London B* 359: 655-667.
- Holsinger E, Kleinhans M, Wheeler Q. 2007. Planet Bob. <http://www.planetbob.asu.edu/>
- Humphries CJ. 2003. Biennial meeting – a presidential review. *The Systematist* 22: 10-14.
- Jones MB, Schildhauer MP, Reichman J, Bowers S. 2006. The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annual Review of Ecology, Evolution and Systematics* 37: 519-544.
- Kennedy J, Hyam R, Kukla R, Paterson T. 2006. A standard data model representation for taxonomic information. *OMICS* 102: 220-230.
- Pyle RL, Earle JL, Greene BD. 2008. Five new species of the damselfish genus *Chromis* (Perciformes: Labroidae: Pomacentridae) from deep coral reefs in the tropical western Pacific. *Zootaxa* 1671: 3-31.
- Williams PH. 2001. *WORLDMAP, Version 4.20.15 (VIII-26-2001)* (<http://www.nhm.ac.uk/research->

curation/projects/worldmap/). London: published by the author.

Nico M. Franz

Department of Biology, University of Puerto Rico, Mayagüez, USA
<franz@uprm.edu>

Major Transitions in Vertebrate Evolution



Anderson J, Sues H-D (eds.). 2007. Indiana university Press, Bloomington and Indianapolis. ISBN 978-0-253-34926-2

(hardback), US\$49.45.

Vertebrates are the most complex and highly studied organisms, and they have one of the most complete fossil records. Consequently, they have long been an exemplar group for documenting major evolutionary transitions. Robert Lynn "Bob" Carroll has been a leading figure in this research for half a century, and the present festschrift, with contributions from his students and colleagues, is a fitting tribute. There are two chapters on basal vertebrates (Janvier, Wilson et al.), two on amphibians (Anderson, Reisz), two on ontogeny and evolution (Hall and Witten, Larsson), two on mammals (Chiappe and Dyke, Luo), and one each on reptiles and birds (Caldwell, Chiappe and Dyke). Most chapters largely summarise and review recent information, rather than presenting new empirical data, but this makes the volume a much more useful general reference. All are well-written overviews of fairly large areas of research, covering the origin of several different groups. Given the long history of vertebrate palaeontology, it is surprising that a wealth of information for several classic evolutionary transitions has only surfaced in the last few decades. Notable are the feathered

dinosaurs, the fish-tetrapod intermediates, the amphibious basal cetaceans, and the limbed early snakes: all of these (except the fish-tetrapod transition) are all well-covered in the relevant chapters.

How do the current contributions differ from analogous contributions fifty years ago, when Bob Carroll's research began? Almost all recent workers use trees constructed via rigorous phylogenetic analysis, an innovation that spread widely only after the English translation of Hennig's ideas (Hennig 1966). Notably, many the chapters explicitly integrate developmental information into interpreting the fossil record, an innovation that Carroll employed in some of his early works (e.g. Carroll 1970). Conversely, the recent emphasis on testability and rigour in evolutionary biology might have caused a decline of interest in rates and trends of morphological evolution as inferred from the fossil record (e.g. Simpson 1944, 1953): most chapters don't address this topic in detail, or use rather antiquated methods (e.g. using higher taxa as proxies for morphological diversity). But such questions remain of paramount importance, as recognised by other disciplines (e.g. the focus on rates and trends in invertebrate palaeontology, and the amount of research on rate heterogeneity in molecular evolution). They are harder to investigate rigorously in vertebrate fossils given the smaller sample sizes, and the uniqueness of morphological characters, but this challenge deserves to be tackled (e.g. Kemp 2007).

While development has rightly been accorded a central position in this volume, I was surprised that the wealth of phylogenetic information provided by molecular biology remains under-utilised by vertebrate palaeontologists (at least those contributing to this book). Yet this information can be crucial, because without it, palaeontologists will often end up attempting to shoehorn

fossils into an incorrect phylogenetic framework for living taxa. Recent molecular data strongly suggest that urochordates (not cephalochordates) are the closest relatives of vertebrates and share several developmental novelties (e.g. Delsuc et al. 2006); however, one of the agnathan chapters attempts to place several fossil taxa between cephalochordates and vertebrates. Similarly, multiple genes strongly indicate hippopotamids alone (rather than ungulates as whole) are sister to cetaceans: the chapter on whale origins touches on this but does not tackle the issue head-on, interpreting character changes and homologies at the origin of whales using hippopotamids (and their fossil relatives) as the living outgroup.

Incidentally, the formatting of this book appears not terribly efficient, with most pages being 40% blank due to inexplicably wide margins; this extra space could have contained an extra couple of chapters that might have filled some of the gaps above. Admittedly, it is always too easy to suggest what else could have been included in an edited volume: regardless of the above comments, the book is a fine snapshot of current research on vertebrate macroevolution. Most chapters are valuable and highly current reviews that will have a long citation life: unusually for such a work, information from papers within two years of publication (2005 and 2006) is incorporated into the chapters.

References

- Carroll RL. 1970. Quantitative aspects of the amphibian-reptilian transition. *Forma et Functio* 3: 165-178.
- Delsuc F, Brinkmann H, Chourrout D, Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439: 965-968.
- Hennig W. 1966. *Phylogenetic Systematics*. Urbana: University of Illinois