EASY ACCESS TO DOUBTFUL TAXONOMIC DECISIONS

[Paleontological Society Theme Session, Toronto, Oct. 28, 1998]

Colin W. Stearn

The conference on Paleontology in the 21st century (convened at Frankfurt) identified databases as a primary concern for paleontologists in the next millennium (Gastaldo, et al, 1998). The General Assembly of Paleo21 unanimously adopted a strong resolution supporting the role of systematics and the need for accessible databases upon which interdisciplinary projects can rely. In the nineties decade the possibility of rapid retrieval of information on paleontological systematics by a world-wide constituency has become a reality. Many specialist paleontologists are compiling taxonomic information into databases that are potentially accessible through the internet. The Paleontological Institute at the University of Kansas has been developing a general retrieval system under the name "Paleobank" for a number of years and through this program much of the information contained in the Treatise of Invertebrate Paleontology will be made available in database form. These databases and their search programs promise to open vast amounts of taxonomic, biostratigraphic, paleoecologic, paleogeographic, diversity, and related information to anyone with a computer, a phone line and a server.

In the past a specialist might spend a lifetime assembling a library of reprints and photocopies of the systematic literature on the group of fossils he or she was specializing in studying. For most fossil groups nearly all of this literature would be in one of five scientific languages: English, French, German, Russian, and Chinese. Very few specialists are able to read all these languages or afford translation of all the species descriptions in large monographs. What a joy it would be if all this literature were available on your computer screen in a single language. But much work is required to complete the entry of a large taxonomic group into a database; not only enough information to identify its species, but information on their ecology, biogeography, and range. In a preliminary version of the Paleobank program provision has been made for recording morphology (several screens), geographic occurrence, environment (marine and non- marine), mode of life (sessile, mobile, etc.), trophic group, lithologic environment (a screen for carbonates and another for clastics), etc.

Databases encompassing the paleontological record of larger taxonomic groups have yielded valuable information on their patterns of radiation, contraction, extinction, and recovery. In large scale studies inconsistencies of taxonomic usage may have little influence on general conclusions. But now paleontologists are anticipating recovering from databases answers to such questions as, "What changes in rhynchonellid diversity took place in shallow water, carbonate-dominated communities in Laurentia during the Emsian?" The answers to such detailed inquiries are subject to differences between taxonomic philosophies.

My own fossil group, Paleozoic stromatoporoids, is a small one but will serve as an example of the problems of placing this taxonomic base, if derived uncritically from the

literature, into computer memory. Recently three colleagues and myself have completed a review of all the Paleozoic genera that includes estimates of the number of species in each (Stearn, et al, 1999). Within the stromatoporoid sponges of Ordovician to Devonian age we found 109 genera that were validly described and placed another 14 in uncertain status. We placed 53 other genera in synonymy; half the number described validly in the literature! In compiling this classification, the authors, each of whom has studied these fossils for a large part of a lifetime, have reduced the generic diversity of the order by one-third. Yet paleontologists (myself included) publish graphs showing trends in diversity in time as if the unassessed literature were a valid sample of the fossil record. Only if the removal of genera by synonymy were random, would the database before and after revision give the same information on biogeography or temporal variations in generic diversity. In fact, since 70% of the genera removed were of Russian provenance, conclusions reached on geographic distribution have been changed systematically. In the revised generic database three times as many genera occur in the Devonian as in the Ordovician. However, about six times the Ordovician number were removed from the Devonian in the revision, considerably decreasing the relative Devonian generic diversity. Diversity both in time and space have been greatly changed.

We estimated that in these genera there were 1,500 species described. These species were assigned to genera without consideration of possible species synonymies. We have not touched the problem of synonymous species which would directly affect estimates of specific diversity.

The number of studies assessing taxonomic consistency and objectivity in the literature is not increasing. The publication of taxonomic reviews such as the Treatise volumes is painfully slow. The number of paleontologists devoting their research efforts to systematics steadily declines and those with expertise are retiring and dying. Meetings of paleontologists, like Paleo21, strongly endorse systematic studies, but how many of the graduate students guided by academics are writing theses on systematics. Many are using other's systematic studies to draw conclusions about the diversity changes caused by impacts or recoveries from mass extinctions, but very few are contributing to the size of the database, or (perhaps more important) to its revision using modern concepts.

Problems of what may be entered in a database can be illustrated by a recent taxonomic review of a single stromatoporoid family (Stearn, 1997). In the end of the last century the genus **Amphipora** was established for a stick-like sponge that is abundant and ubiquitous in Devonian rocks. That single genus was made the basis of a family, the Amphiporidae, in the thirties by Rukhin. Russian and Chinese paleontologists split off 12 new genera from **Amphipora**. About 175 species of these genera have been described. Variation within single skeletons, and within preserved populations of what must have been monospecific gardens of the sponge, indicate that characters that had been considered as diagnostic of 8 of the new genera of amphiporids, are common in the typical species **Amphipora ramosa**. Nearly all of the characters said to be diagnostic of the many species are expressed within a single stem of this species. In this study the number of genera in the family has been reduced from 13 to 5. I have no

idea how many of the 175 species of the family in the literature are synonyms but I suspect that the total species diversity of the group is closer to 20 than 175. I am left wondering whether it is possible from the limited descriptions and inadequate illustrations to resolve the question of variation of the species diversity of **Amphipora** in time, and whether entering this sort of data into a computer memory will only mislead generations to come about the validity of conclusions drawn from such a database.

This example is from the lower invertebrates but recently Nigel Hughes and Conrad Labandeira (1995) have given an example from the trilobites. Taxonomic revision has resulted in a reduction from 30 species to one species in the genus **Dikelocephalus** from the late Cambrian St. Lawrence Formation. Charlotte Jeffery (1998, p. 149) found in studying echinoids that, "In order to avoid the problems of inconsistent taxonomic usages encountered in the literature, a thorough taxonomic revision is a prerequisite for all studies of biological turnover". Among paleontological taxonomists the splitters have lost out to the lumpers and our conception of past animal diversity has shrunk.

This is a major problem of entering the literature of paleontology in a database without critical review. No taxonomic database should be made widely available by giving the task of transferring the literature to disk, to someone who is incapable of making taxonomic judgements. Each database should also be recognized as a subjective product of the compiler's taxonomic bias. Inasmuch as Paleobank is designed to make available the taxonomic experience of the compilers of the Treatise volumes, it will meet this criterion.

There are also legal problems that are common to all databases. In the last 2 years the regulation of databases has become the concern of the Congressional House Judiciary Subcommittee on Courts and Intellectual Property and the United Nations World Intellectual Property Organization (American Geological Institute, 1998). Free access to compilations of scientific data is under attack by those who want some of the costs of making databases available to be borne by those who use them. Legislators are struggling to distinguish and define "fair use" by scientists and what constitutes "piracy" of intellectual property. Scientific organizations so far have managed to postponed decisions on the UN Draft Treaty on Intellectual Property in Respect to Databases and certain similar bills before congressional committees. Although databases of most interest to earth scientists are largely of geophysical nature (climate, weather, earthquakes, etc.), paleontologists should be aware of possible legal impediments to the widespread use of their databases, and that compilers may request user fees to fund what will be the large costs of entering the paleontological literature. So far granting agencies have regarded the costs of producing databases as research, but the day may be approaching when they decide that this work is not uncovering new knowledge but only cataloguing old data.

Colin W. Stearn, Kitchener, 65 Aberdeen Rd., N2M 2Y4.

Acceptance: 10 January 1999

REFERENCES

American Geological Institute, 1998. AGI Scientific community express concern over proposed International Database Treaty. AGI Webpage [http://www.agiweb.org/legis104/wipo and http://www.agiweb.org/legis105/wipoupd].

Gastaldo, R. A., Ashley, G., Lane, H. R., MacLeod, N., O'Neill, B. J., and Wang, C.-Y., 1998. Paleontology in the 21st century: An international workshop 3-9 September 1997. **Palaios**, 13:87-90.

Hughes, N. C. and Labandeira, C. C., 1995. The stability of species in paleontology. **Paleobiology**, 21:401-403.

Jeffery, C. H., 1998. Carrying on regardless: the echinoid **Cyclaster** at the Cretaceous-Tertiary boundary. **Lethaia**, 31:149-157.

Stearn, C. W., 1997. Intraspecific variation, diversity, revised systematics, and type of the Devonian stromatoporoid **Amphipora**. Palaeontology, 40:833-854.

Stearn, C. W., Webby, B. D., Nestor, H., and Stock, C, W., 1999. Revised classification and terminology of Palaeozoic stromatoporoids. **Acta Palaeontologica Polonica,** 1-70.