Trends in taxonomy revealed by the published literature

Prominent biologists have stressed the urgent need for systematics in conserving biodiversity (e.g., Wilson 1989, May 1990, Anonymous 1991), and with good reason. According to recent estimates, between 69% and 96% of extant species remain undescribed (Lean et al. 1990, Systematics Agenda 2000 1994). If knowledge of Earth's biota is so incomplete, it follows that even the most basic kind of systematics research, taxonomy—the inventory, description, and classification of organisms—is also far from finished. However, although cladistics and molecular phylogenetic techniques have attracted attention to systematics and awakened new interest among students, reports and surveys on the state of the discipline have all cited problems such as dwindling resources, an aging population of professional systematists, and low recruitment due to job scarcity (Steere et al. 1971, Stuessy and Thompson 1981, Winston 1988, SA2000 1994). In Great Britain alone, 25–30% of evolutionary biology positions are estimated to have disappeared between 1980 and 1990 (Clark 1990); during the same period, the number of theses in systematic biology declined drastically relative to the number in other fields of biology, such as biochemistry (Claridge and Ingrouille 1992). Thus, even if taxonomic data collection is still far from complete, it may be that science is finished with taxonomy—that it is no longer a viable field.

Assessing the state of taxonomy

How can the current state of taxonomy be assessed? One way is by determining the numbers of specimens in collections, the availability of jobs, or the amount of funding for museum research (e.g., Stuessy and Thompson 1981, Clarke 1990). Another possible approach, measuring the numbers of currently active taxonomists, is complicated by the fact that many active taxonomists are not employed in their specialty. They may be officially retired or employed full- or part-time at other jobs (Winston 1988).

In this article, we present a different way to analyze the health of the field. This approach uses the published scientific literature itself, retrievable as citations, keywords, abstracts, and other forms from a variety of electronic databases (e.g., Aquatic Sciences and Fisheries Abstracts, BIOSIS Previews, Dissertation Abstracts Online, Enviroline, Environmental Bibliography, Oceanic Abstracts, Pollution Abstracts, SciSearch, and Zoological Record Online), whose records extend back over the last 20–30 years. Biodiversity-related information retrieved from such databases effectively supplements and complements the information provided by collections databases. We used electronic databases to compile information that could help to answer the following questions about the status of systematics: Is the field growing or declining? Is the basic job of taxonomy finished—that is, is the rate of description of new taxa slowing down, or is it still increasing? Is species description, in particular, still an important part of taxonomy? How much of taxonomists' efforts is really spent on nomenclatural matters? What trends in the field can be identified over the last 28 years?

Is taxonomy in decline?

We used the BIOSIS Previews database on the DIALOG system, which allows online searches of citations and abstracts to the biological literature in approximately 7500 journals and monographs, covering the years 1969 to the present. Our searches ended with 1996 because indexing of the 1997 year was not complete as of our search date (21 October 1997). On the DIALOG system, truncation of search terms is indicated by a question mark, and the symbol (w) indicates that two words must be adjacent. Free text searching was used because it provided the most accurate results, calling up all articles in which the search term appeared in the title, abstract, key words, or the "concept codes" tagged by the database producer. By contrast, we found that limiting our search to the database producer's assigned concept codes severely limited our data collection and missed pertinent citations. BIOSIS differs from other databases because it uses abstracts as printed in journal articles, creating its own abstracts only for book articles that lack them.

Figure 1a shows the growth of the biological literature between 1969 and 1996. During that time, period yearly totals of biological articles indexed increased from 204,326 to 546,496. Most of this growth reflects an increase in the number of biological journals, part of the "information explosion" predicted by librarians in the early 1960s. The growth rate appears to be leveling off over the last few years, but whether this pause is temporary or indicates a real change (e.g., as a result of factors such as negative global economic climate, declines in research funding, the breakup of the scientific apparatus of the former Soviet Union, or even the beginning of a switch to electronic publication) remains to be determined.

Between 1969 and 1988, the number of articles published in taxonomy

by Judith E. Winston and Kristen L. Metzger

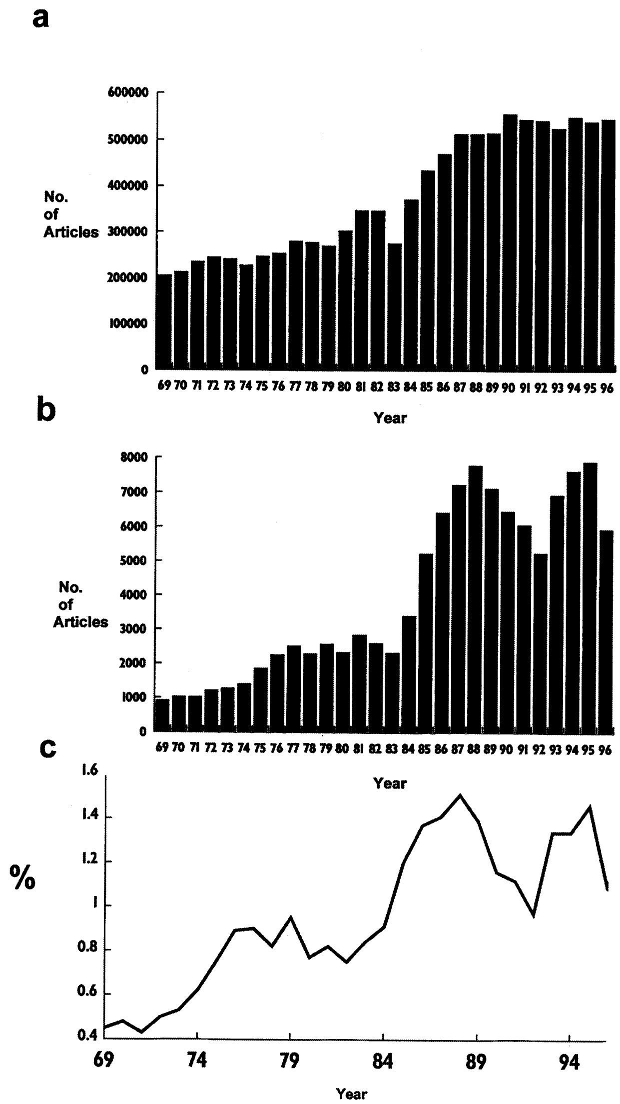

Figure 1. Biological and taxonomic articles, 1969–1996. (a) Growth of the biological literature between 1969 and 1996. (b) Growth of the taxonomic literature (search term: "taxonom?") between 1969 and 1996. (c) Taxonomic literature as a percentage of the total biological literature.

also grew (using the search term "taxonom?"; Figure 1b), as did the percentage of the biological literature represented by articles in taxonomy (Figure 1c). The number of articles with "taxonom" in the title or abstract or among the keywords increased from 910 in 1969 to a peak of 7769 in 1988. Up until 1989 the field was not only growing numerically, but also increasing its share of the total biological literature (from 0.45% to 1.51%). The number of taxonomy articles then declined sharply (to 5230 in 1992), although they peaked again at 7882 in 1994. Taxonomy has lost ground relative to biology as a whole since 1989; its share of the literature declined to 0.98% in 1992, although it rose to 1.46% in 1995. The effects of the decline in numbers of taxonomists (e.g., Holden 1989, Wilson 1989, Feldman and Manning 1991) may only now be becoming apparent due to the time lag between submission and publication of a paper.

Are new species still being described?

The impression that the cataloging stage of taxonomy is far from over is upheld by yearly data for description of new species (Figure 2). The percentage of articles describing new species (search term: "(new(w)species or species(w)nov) and taxonom?") increased from 7.1% in 1969 to 47.4% in 1988, declined to 33.2% in 1992, returned to 43.9% in 1995, but dropped to 33.8% in 1996. Many basic species descriptions are still being published. On average, at least one new species is described in more than one-third (35.7%) of the taxonomic articles published since 1969.

Statistics for the BIOSIS database as a whole also indicate that description of new taxa is a significant part of taxonomists' work. New genera are reported in approximately 8%, and new subspecies in approximately 3%, of articles on taxonomy; higher-level discoveries are rare. New families figure in fewer than one taxonomic article in 100, new orders in fewer than one in a 1000, a new class in one in 4000, and a new phylum or division in only approximately one in 10,000.

Are taxonomists too fussy?

Electronic searches can provide useful data on the nomenclatural aspects of taxonomy. The last 25 years have also seen the prospect of increasing support for revolutionary changes in the present system of biological nomenclature. Proponents of one change, name stabilization, wish to prevent confusing and expensive changes in the names of common or commercially used species by setting up lists of accepted names that would remain immutable (Brandenburg 1991). Professional systematists, on the other hand, defend their scientific freedom to incorporate new knowledge into the discipline, even if this approach means that species names may change over time (e.g., if a species is moved to a new genus; Anderson 1991). In February 1991, an entire Systematics Association symposium was devoted to this topic, without achieving consensus on either the extent of the problem or its solution (Hawksworth 1991).

Name changes fall under the search term "nomenclature" in BIOSIS. As Figure 3 shows, the perception that taxonomists are spending an inordinate amount of time pursuing obscure details of nomenclature appears to be false. The percentage of taxonomy articles indexed under nomenclature (search term: "nomenclature? and taxonom?") showed considerable year-to-year variation but no clear trend. The lowest percentage was 7.2%, in 1988, the highest was 17.7%, in 1993, and the average over the 28-year period was 13.3%.

However, the descriptor "nomenclature" covered a variety of subtopics, which might have obscured any trend; consequently, we also analyzed one subtopic, designation of a neotype, independently. In the present system of biological nomenclature, each scientific name is linked by its author (the person who first publishes a description) to a type specimen. Although the International Zoological and International Botanical Codes of Nomenclature handle types somewhat differently, in both systems a type is essentially a sort of supervoucher specimen. In the course of revisionary work, a taxonomist may find that the type of a particular taxon has been lost or misplaced, or that it was never designated. In such cases, the taxonomist carrying out the revision may pick a new type specimen, or neotype. This procedure is recommended only as a last resort because mistakes can lead to real nomenclatural headaches, for example, if (as has happened), the neotype material later turns out to belong to a completely different taxon than was claimed or is found to be a mixture of taxa. It is reassuring, therefore, that the percentage of articles in which a neotype is designated (search term: "neotype? and taxonom?") is very low (1.4% on average) and has not increased over time. The infrequent use of this procedure indicates that taxonomists are not creating unnecessary work for themselves or others.

Articles dealing with redescription and revision (search term: "(revision or redescription) and taxonom?") increased somewhat over the time period surveyed, from approximately 2% to 8%. Revision, with its goal of increasing the phylogenetic predictiveness of classifications, is often considered to be the most important everyday business of taxonomy, but there is no sign it is being carried to excess—in fact, it might be questioned whether enough revision is being done.

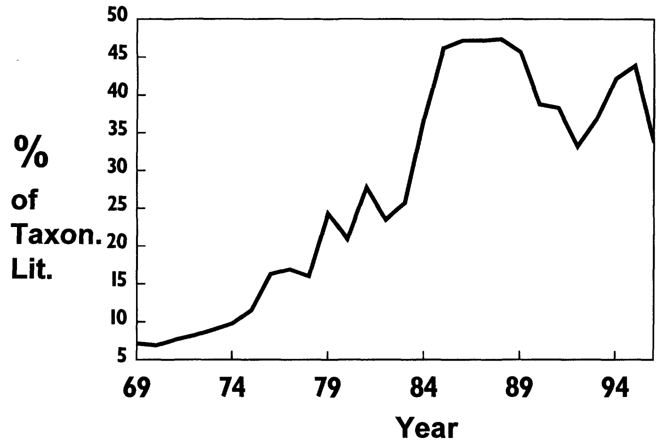

Figure 2. Percentage of taxonomic articles in which at least one new species is described, for each year from 1969 through 1996 (search term: "(new(w)species or species(w)nov) and taxonom?"). Although there was no simple way to search for total number of new species via this database, it was possible to search for articles in which at least one new species was described. On average, at least one new species was described in 35.7% of all taxonomic articles over the last 28 years.

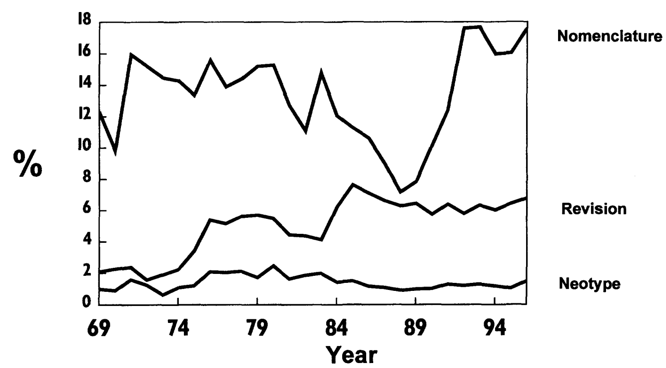

Figure 3. Percentage of taxonomic articles dealing with nomenclature (search term: "nomenclature? and taxonom?"), redescription or revision (search term: "(revision or redescription) and taxonom?"), and the designation of a neotype (search term: "neotype? and taxonom?").

Recent trends

Literature analysis can also effectively document major trends in the field. In 1969, phenetics, especially the quantitative computer-based variety known as numerical taxonomy, was in vogue. Since then, cladistics has come to dominate much of the field (Figure 4a). The number of articles using cladistic methods (search term: "cladistic?") has increased dramatically, as has the percentage of taxonomy articles that mention cladistics (i.e., from 0.7% in 1974, to 3.3% in 1984, to 6.3% in 1996). However, the number of articles based on a phenetic approach (search term: "phenetic?") has not declined but, instead, is still slowly increasing. By examining sample abstracts, we found that most of these articles deal with botanical systematics. For representation of the reticulate evolutionary patterns produced by the genetic systems of some plants, phenetics is apparently still a useful approach (Stace 1989).

The last 28 years have also seen a rapid increase in the use of molecular techniques in taxonomy (Figure 4b). In 1969, protein electrophoresis was the only molecular method available to most taxonomists; it is still the method of choice for certain types of studies and, in addition, is the cheapest and most widely available method (Murphy et. al. 1990). Articles based on electrophoretic methods are still being produced in slowly increasing numbers, but during the 1970s DNA techniques rapidly outdistanced other methods, with sequencing and restriction site analysis now surpassing DNA–DNA hybridization in popularity. In the last few years, articles based on DNA techniques have overtaken those based on protein electrophoresis (search term: "taxonom? and (allozyme? or isozyme? or electrophoresis?)"). Studies using sequencing and restriction site analysis (search term: "taxonom? and (mtDNA or mt(w)DNA or mitochondrial(w)DNA) or sequence? or sequencing?") have pulled ahead of DNA–DNA hybridization (search term: "taxonom? and DNA(w)hybrid?") in popularity. However, articles based on the less expensive electrophoretic methods available to almost any graduate student are still being produced in slowly increasing numbers.

Conclusions

Taxonomy, although acknowledged as a necessity in assessing Earth's biodiversity, is often considered a dying field. A search of the BIOSIS database for the past 28 years shows that, in fact, the field was still growing through 1988, but now it is declining or, at the very least, ceasing to grow. Taxonomic articles describing new species also increased through 1988; their production may have slackened, but they still make up more than one-third of all work published over the time period. Trends in the taxonomic literature show that taxonomists have slightly increased the production rate of their revisionary work, but not of their articles on nomenclature. There has been a large increase in publications based on cladistic methods (without, however, a concomitant decline in

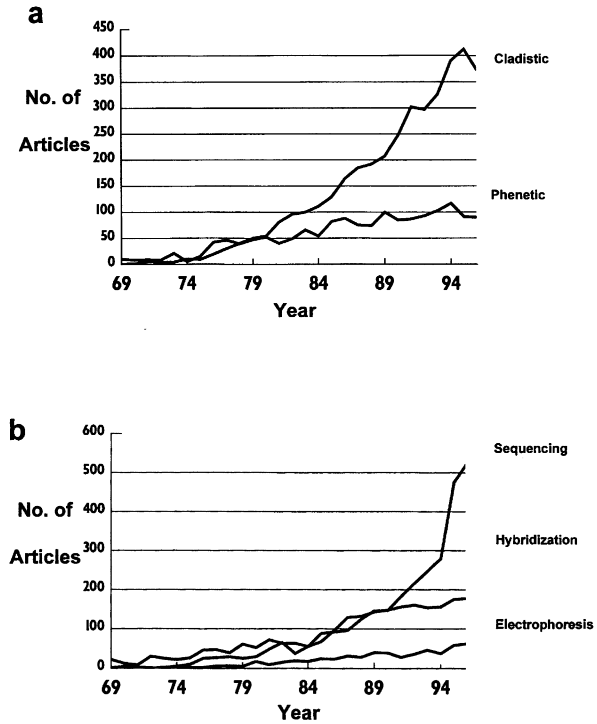

Figure 4. Trends in taxonomic methodology over the last 28 years. (**a**) Number of taxonomic articles using cladistic (search term: "cladistic?") and phenetic (search term: "phenetic?") methods. (**b**) Number of articles using different molecular phylogenetic methods by year (search terms: "taxonom? and (allozyme? or isozyme? or electrophoresis?)" or "taxonom? and (mtDNA or mt(w)DNA or mitochondrial(w)DNA or sequence? or sequencing)" or "taxonom? and DNA(w)hybrid?)").

those using phenetic methods) and in articles based on molecular methods, with sequencing now the most popular technique. This analysis thus indicates that the "crisis" in taxonomy—the basic scientific underpinning of all the efforts of those interested in biodiversity and its preservation—is not a figment of the imaginations of taxonomists but can be objectively documented in the published literature.

Systematists, accustomed to delving into 200 years of literature, have tended to overlook the potential of the three decades of taxonomic literature available online to contribute to taxonomic study, but we believe that these data can be of great practical use in focusing future taxonomic research on the most important or neglected groups and geographic regions. Electronic searches of databases such as BIOSIS or Zoological Record Online can provide a fast, relatively inexpensive assessment of other aspects of taxonomy. For example, a 1991 search using Zoological Record Online, a database from which records of the numbers of taxonomic surveys can be easily compiled by geographic area, showed that although marine habitats cover 70% of the earth's surface, only 13% of published taxonomic surveys were marine oriented. This information would certainly be good ammunition in a request for support for funding of a marine taxonomy project.

Although we looked at trends in taxonomy, similar electronic searches would be useful in other areas. Literature searches cannot replace surveys and inventories, but much information useful in designing and focusing them is already publicly available and can be obtained at a modest cost. Although retrieving synonymic information for a particular taxon usually requires searching the taxonomic literature back much further than the last 20 or 30 years, being able to retrieve the citations for that period in minutes can cut weeks of preparation time from monographic endeavors and may well uncover relevant ecological or other biological citations that would otherwise have been missed. Enlisting the aid of an information specialist from the beginning of a project could avoid the waste of considerable time and money in recompiling or re-documenting such material and would jumpstart any biodiversity program.

References cited

[Anonymous]. 1991. Funding unsexy science. Science 251: 377.

Anderson WR. 1991. Shall we change the code? Concerns of a working taxonomist. Pages 95–103 in Hawksworth DL, ed. Improving the Stability of Names: Needs and Options. Taunus (Germany): Koeltz Scientific Books.

Brandenburg WA. 1991. The need for stabilized plant names in agriculture and horticulture. Pages 23–31 in Improving the Stability of Names: Needs and Options, Hawksworth DL, ed. Taunus (Germany): Koeltz Scientific Books.

Claridge M, Ingrouille FM. 1992. Systematic biology and higher education in the U.K. Pages 39–48 in An Appraisal of Taxonomy in the 1990s. London: The Linnaean Society and the Systematics Association.

Clarke B. 1990. Systematics and evolutionary biology in Britain. Association of Systematic Collections Newsletter 19: 58–59.

Feldman RM, Manning RB. 1991. Crisis in systematic biology in the "Age of Biodiversity."Journal of Paleontology 66: 157–158.

Hawksworth DL, ed. 1991. Improving the Stability of Names: Needs and Options. Taunus (Germany): Koeltz Scientific Books.

Holden C. 1989. Entomologists wane as insects wax. Science 246: 754–756.

Lean G, Hinrichsen D, Markham A. 1990. Atlas of the Environment. New York: Prentice-Hall.

May R. 1990. Taxonomy as destiny. Nature 347: 129–130.

Miller S. 1994. Biological collections databases available on Internet. Association of Systematics Collections Newsletter 22: 57.

Murphy RW, Sites JW, Buth DG, Haufter CH. 1990. Proteins I: Isozyme electrophoresis. Pages 45–126 in Hillis DM, Moritz C, eds. Molecular Systematics. Sunderland (MA): Sinauer.

Stace C. 1989. Plant Taxonomy and Biosystematics. 2nd ed. London: Edward Arnold.

Steere WC, et. al. 1971. The Systematic Biology Collections of the United States: An Essential Resource, Part I. The Great Collections: Their Nature, Importance, Condition, and Future. The Bronx (NY): The New York Botanical Gardens.

Stuessy TF, Thompson KS, eds. 1981. Trends, Priorities, and Needs in Systematic Biology. Lawrence (KS): Association of Systematics Collections.

[SA2000] Systematics Agenda 2000. 1994. Systematics Agenda 2000: Charting the biosphere. New York: SA 2000. Technical Report. Available from: New York Botanical Gardens, The Bronx, NY 10458.

Wilson EO. 1989. The coming pluralization of biology and the stewardship of systematics. BioScience 39: 242–245.

Winston JE. 1988. The systematists' perspective. Memoirs of the California Academy of Sciences 13: 1–6.

Judith E. Winston (e-mail: jwinston@leo.vsla.edu), formerly curator and chair of the Department of Invertebrates at the American Museum of Natural History in New York, is director of research at the Virginia Museum of Natural History, Martinsville, VA 24112. Her research interests include bryozoan systematics and marine biodiversity. Kristen L. Metzger (e-mail: metzger@hboi.edu) is librarian in the research library at the Harbor Branch Oceanographic Institution, Fort Pierce, FL 44946. Her interests include marine science librarianship and information technology.
