12

The Future of Systematics

The future of systematics as viewed by numerical taxonomists was the subject of a section in *Principles of Numerical Taxonomy* (Sokal and Sneath, 1963). More detailed discussions of this subject can be found in Ehrlich (1961a), and Sokal (1964a, 1970), discussions that (admittedly) contain a strong propagandistic element in an effort to convince the generally conservative taxonomic establishment of the merits of computer processing for classificatory purposes as well as for information storage and retrieval purposes. Although the views proposed by these authors are far from generally accepted—witness the several replies to the cited prognostications (e.g., Rollins, 1965a,b; Kalkman, 1966; L. A. S. Johnson, 1968) as well as the tone of some recent symposia (Sibley, 1969) or the antiquated outlook of a report by a committee of the U.S. National Academy of Sciences (Handler, 1970)—it would appear that many of the developments forecast by these early "prophets" are well under way and will continue with their own momentum.

It seems appropriate therefore to discuss only those aspects of systematics upon which the development of numerical and computer methods will impinge in one way or another.

The ever increasing application of numerical methods to all kinds of organisms and all types of characters or descriptors must by now be evident even to the casual reader and is documented in detail in the taxonomic lists of Appendix A.

The main impetus for the development of computer methods of handling taxonomic data will come from the great number of taxonomists in various groups experimenting with new types of taxonomic information. Most important among these are the various kinds of chemical characters being obtained by ever more sophisticated and automated technology. These include data on proteins and molecular biology discussed in detail with numerous references in Sections 3.5 and 5.12, but they will also include much information on secondary chemical products found especially in plants (for a review see Alston, 1967 or B. L. Turner, 1969). The need for numerical taxonomy to integrate these many different kinds of characters into the overall body of descriptive taxonomic knowledge is discussed by Sokal and Sneath (1966) and Heywood (1968).

In this connection we should mention various instruments that will automatically yield large data sets about organisms. These include autoanalyzers that automatically obtain the results of numerous chemical analyses, amino-acid sequencers that break down proteins into their constituent parts, and optical scanners that will give morphological descriptors of macroscopic and microscopic structures of various organisms. If an electron microscope is developed that can read molecular structure directly this would open the way to cladistic studies on a large scale from protein sequences by methods discussed in Chapter 6, and even molecular paleontology would then be possible if undegraded proteins could be found in fossil material. Instruments such as scanning electron microscopes coupled to computers (see Heywood, 1969, for just one illustration of the profusion of new information that will result from scanning electron microscopy), or automatic scanning by holography (Gabor, 1965) are already opening up an undreamed-of wealth of new taxonomic information. These developments will interact with those in pattern recognition discussed in Section 11.5. Unless classical taxonomists will categorically rule out evidence obtained by the means of such devices (and there is nothing in the history of systematics or in biological theory to justify such a parochial view), it will be necessary to process this information by computer, regardless of whether the aim of the taxonomist is a phenetic or a cladistic classification.

Machinery of this type and computer processing will become accepted in time as a routine part of taxonomic work. Systematists should therefore have sufficient training in mathematical methods to understand the rationale behind some of the techniques beeing applied, and also training in operating the various devices that yield this information. For a discussion of minimal educational requirements for systematists and ecologists of the coming generation, see Sokal (1970) and Turner (1971). Automatic data gathering machines and computer terminals will be employed not only in description and classification but also in identification. Machines that classify and identify (usually by means of optical scanners coupled to computers) are now available for cytological and pathological research. Their justification for routine identifications of economically and medically important organisms must be investigated. The degree of their employment will largely depend on the economics of the computer field in the years to come. Although heavy capital investment is necessary for computer installations, the overall cost of computing for specific tasks has steadily decreased in recent years by orders of magnitude. The use of electronic data processing (EDP) for the description of organisms, for the preparation of geographic distribution maps, for taxonomic keys, monographs and other taxonomic endeavors has been described by many persons and we need merely cite reviews by Soper and Perring (1967) on map preparation in botany, and by Crovello (1967) on data storage and retrieval. Cutbill (1971) presents articles on various aspects of EDP as it affects biological systematics.

These various developments will clearly affect museum operations, as Sokal (1964a), Sokal and Sneath (1966), Ehrlich (1964), Raven and Holm (1967), and Williams (1967a) have all pointed out. To what degree museum procedures should be modified is an open question that requires considerable study and discussion. A recent report (Steere, 1971) from a committee of natural history museum directors strongly supports computerization of many of the operations in the great systematic biology collections of the United States. A number of authors (Sokal and Sneath, 1966, and Williams, 1967a, among others) have suggested the establishment of international taxonomic centers in which data storage and retrieval will be carried out on a large scale and made available to taxonomic users at remote locations by various means of computerized interfacing.

The realization of these many developments will depend partly on the pressure exerted by the load of material to be processed, partly on technical developments (the speed of which cannot always be accurately predicted), and partly on the training received by the current generation of taxonomists. We cannot know how many of these innovations will be adopted and how rapidly this will take place. We can, however, say that the changes in taxonomic practice are going to be profound and that taxonomists entering the field at this time must acquaint themselves with these techniques and be conversant with them in order to be competent workers in the field. It has become obvious that taxonomy is as dynamic and as changing a field as other aspects of biology and there is little doubt that exciting realms of new endeavor are in store for the taxonomists of the next decade.