NOTE

Statement on the WFCC Center for Storage, Retrieval, and Classification of Data on Microorganisms

V. B. D. SKERMAN

Department of Microbiology, University of Queensland, St. Lucia, Brisbane, Australia

As a sequel to a request from the Japanese Federation of Culture Collections to the United Nations Organization, a group of specialists met in Paris in July, 1966, under the auspices of the United Nations Educational, Scientific, and Cultural Organization (UNESCO) to consider a variety of problems relating to culture collections. Such items as the establishment of a world federation of culture collections, the training of culture collection personnel, and the establishment of collections in developing countries were discussed. It soon became apparent that very little could be done at the international level until the location, the list of the personnel, and the list of the contents of the various collections of the world were obtained. Therefore, it was recommended that a survey of culture collections be carried out. At its meeting on July 13, 1966, the then IAMS Section on Culture Collections agreed to undertake such a survey with the aim of preparing a World Directory of Collections of Cultures of Microorganisms.

Financial support for the project has been supplied by UNESCO, the World Health Organization (WHO), and the Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia. The National Research Council of Canada and the University of Queensland, Brisbane, Australia, have both contributed significantly by way of staff and material assistance.

The finance provided by UNESCO to this Department over the period 1969 to 1971 amounted to U.S.$6,000; the contribution from WHO was U.S.$1,500, and that from the CSIRO was A$5,000. The original intention was that this money be used for the production of the directory using computer techniques. However, the very slow manner in which information came to hand from various collections suggested to us that it would be just as economical to produce the directory by hand while using the available finances to produce programs for the storage and retrieval of strain data published on microorganisms, which data would also contain the information required for the generation of future catalogues. Most of the WHO grant was used to pay the expenses of Michah Krichevsky, of the National Institute of Dental Research at Bethesda, Maryland, to come to Australia for a short period to discuss with us the design of programs intended to make use of the coding procedures for bacteria which had been published by Rogosa, Krichevsky, and Colwell, after extended consultation with myself and others, in the International Journal of Systematic Bacteriology 21, 1971. The programming was undertaken under our direction and at our expense by a firm of commercial computer programmers on a Univac 1108 computer of UCCompunet Pty. Ltd., Sydney, Australia. Since the start of these programs, a further A$6,000 has been provided from research funds of the University of Queensland towards the incorporation of a method for classification of data, making use of the primary steps in the retrieval program for the initial comparison of strains. The program is a complete linkage one designed by Skerman and DeWitt for the handling of very large masses of data in the most economic manner possible. At present the acceptance of data is almost unlimited, but limitations are placed on classification by the size of the existing computer which permits us to analyze data for up to 2,600 strains with 3,000 characters per strain (or other combinations of these, e.g., 18,000 strains with approximately 500 characters per strain or 25,000 strains with approximately 300 characters per strain). This is ample for the immediately foreseeable future, and
requests are being sent to publishers of data to supply the center with their information.

When publishing the methods for coding of characteristics of bacteria, M. Krichevsky agreed to make available the forms upon which people can submit the data coded by their methods. These forms are also available from this office. However, it is our view that, because of the novelty of the coding method, it is more desirable that research workers submit directly to the data center their normal laboratory information, thereby allowing us to do the coding on a uniform basis in the center.

Although the programs have been written in Fortran V for use on a Univac 1108, it is likely that they will be converted to the University's own PDP 10 computer which, in the next 3 years, will be extended in capacity to cope with the needs of the center at least to, and most likely beyond, the capability of the Univac 1108. It is difficult at this stage to calculate the cost of conversion. This Department has also installed an Olivetti computer terminal. Once programs have been translated to the PDP 10, we will be in a position to operate the center on a "on-line" basis on the PDP 10 in addition to the Batch system at present used with the Univac 1108. Present investigations suggest that the Olivetti terminal used in conjunction with the PDP 10 will provide a versatile means of updating and printing future editions of the World Directory.

At an early stage in the development of this center, I expressed the opinion that we should be aiming towards the development of branches of the center elsewhere. At present M. Krichevsky is actively developing a center at the National Institute of Dental Research in the United States. Care has been taken in our development to insure that a format for storage of data is acceptable in the two centers for exchange of information. Communications have already been established with the Japanese Federation of Culture Collections with a view to establishing a center in Japan.

The center was formally established by the World Federation for Culture Collections (WFCC) at the Xth International Congress for Microbiology in Mexico City in August 1970 and located for the time being in the Department of Microbiology at the University of Queensland. The center is now functionally ready for the acceptance of data.

Provisions for other forms of microbes. All data coded previously have related solely to the bacteria, although the basis upon which the coding operates is one which can be used for the coding of viruses and fungi (and other biological forms, with some modification). A data-coding committee has been set up within the WFCC, with M. Krichevsky as chairman and myself as secretary. The membership of this committee includes Adrian Gibbs representing virologists. Negotiations are in progress with A. Gibbs relating to coding procedures for the viruses. Some difficulty has been encountered at the international level with the establishment of an international committee on the nomenclature of fungi. Consequently, the attempt to make progress with the coding of the characters of fungi has been delayed. An offer of assistance in the development of a code for the fungi has been received from London. Any mycologists interested and willing to assist would be welcome. When coding facilities are available for viruses and fungi, these facilities will be added to the present activities of the center.

Services of the center. With the acquisition of data, the present programs will permit services of the following nature: (i) identification of an unknown organism; (ii) the supply of data on any particular taxon, within the limits of our store; (iii) acceptance of data from research workers, matching it against data in the center, and supplying the output to the research worker either unclassified or classified; (iv) report of the location of cultures in various culture collections. This function is a very important one as the World Directory of Collections of Cultures of Microorganisms cannot be expected to be an up-to-date document. New information regarding the accession of cultures to, and deletion of cultures from, collections, will be used to maintain an up-to-date statement on the contents of collections within the center, so that up-to-date information can be issued at any particular time on any specific request. At intervals it will be possible to generate a new edition of the catalogue. Some additional programs may be needed for the latter purpose. In this connection facilities are also available here for production of camera-ready copy in different fonts using the Olivetti terminal and an ASCII-TTS translation program in conjunction with a PDP 8 computer; (v) provision of information on the geographical distribution of organisms, host ranges, or the occurrence of organisms with specific properties; (vi) provision of a list of contents of any collection; (vii) provision of a list of all collections maintaining particular organisms; (viii) supplying culture collections with their own collection data in punch-card form if they wish to indulge in some local operations using the data. Fuller details of the
programs in use in the center will be published separately.

The principal needs of the center are continuing financial support from international sources and the supply of strain data from research laboratories. Any assistance would be welcome.

The World Directory of Collections of Cultures of Microorganisms was generated in camera-ready form in this Department, with instructions in English, French, German, Spanish, Russian, and Japanese, and published by J. Wiley & Sons, New York, in 1972. Copies (220) of the Directory were distributed gratis to those collections which supplied us with information on the holdings of cultures in their collections. Royalty payments have been assigned to the WFCC. The intention, approved by the Executive Board of the WFCC, is that money accruing in this form will be invested and only the interest from the investment will be used for expenditure purposes.