Chapter 32

Taxonomic Databases Related to the Flora of Cuba

F. Cejas, M. Esquivel & H. Knüpffer

Tropical countries are in urgent need of accurate estimates concerning the number of species existing within the limits of their territories and the current state of knowledge about certain areas which could prove to be of great interest. Such an inventory could facilitate the study of natural resources as well as taxonomic and ecological investigations, and could also guarantee strong, solid bases for future development and allow for independence in agriculture and pharmaceuticals, excluding the use of alien plant varieties and/or stocks which too often have failed or do not turn out to be satisfactory in tropical climates, e.g., experiments in hybridization in spite of their good results under different conditions.

The great bulk of data that has been accumulated in the past few decades and that is kept in technical literature and specimen collections increases alarmingly day after day and is often lost at the expense of the time wasted in the consultation of such sources. Therefore, in order to process any kind of information, the use of computing methods for speeding up the organization, systematization and analysis of data becomes necessary.

However, tropical countries lack the financial and human resources needed for this sort of task. That is why the aid of financing and scientific agencies that are able to give not only material support but also the technological background for technical needs is required, without losing sight of the special features of each country, since due to the cost the application of foreign computing systems often cannot be put to work. Another reason is the reluctance of botanical institutions to use those systems that could contain ample information but do not always correspond to the material capacities and main objectives of such institutions.

With these facts in mind, Cuba has been developing a project for editing the "Flora of the Republic of Cuba", a work that intends to compile all the information available on the subject up to the present, as well as those changes, additions, etc., that have been made since 1969, the date of publication of the "Supplement of the Flora of Cuba" (Alain 1969). Undoubtedly, such a work will be an obligatory reference for any professional working on the Caribbean area since Cuba has quite a few characteristics of its own, such as its genesis, topography and soil types, all of which make it an important centre of Caribbean speciation, rich in endemics.

As in other countries, several Cuban institutions have begun to create botanical databases, all of them with similar objectives but using different methodologies and software. Specifically, databases concerning Cuban angiosperm endemics as well as wild relatives of timber trees and cultivated plants were brought up to date at the Institute of Ecology and Systematics (IES) of the Cuban Academy of Sciences (Rodríguez *et al.* 1990, cf. also Rodríguez *et al.* 1994). Those databases were created on the basis of herbarium labels of the Herbarium of the Cuban Academy of Sciences (HAC), which alone has about 300,000 specimens, and have already started to serve the purposes of specialists working on phytogeography, plant genetic resources, ecology and chemotaxonomy.

Starting from this point, and keeping in mind the objective of organizing and carrying out all efforts to an optimum level, the building up of a curatorial and taxonomic database for the IES collection of dried specimens of angiosperms was begun. Its structure responds to the experience accumulated in several Cuban institutions, mainly of IES and the National Botanic Garden, as far as the information that should be contained in this database is concerned. This structure could in time be the core

of some other, more complex bases that also include biodiversity and genetic resources and at the same time could interact with others from Latin America and worldwide sources. In order to permit the establishment of patterns that may help to standardize this information and by using the software package foxBASE+ (version 2.0, of 1987), a database management system named HERBARIO was created (Cejas 1990, Cejas & González 1990). This system facilitates the edition and retrieval of data deposited in the database by means of a common model. The data structure is given in Table 32.1.

Table 32.1 Data structure of the system HERBARIO

Field name	Length	Information
FAM	15	Botanical family
GEN	20	Genus
ESP	25	Species epithet
AUT	38	Author of species
INFRASP	37	Category and infraspecific epithet
INFRAUT	38	Author of infraspecific name
NUM	13	Herbarium (abbreviation) and number
EXHER	12	Herbarium of origin
NOCOL	12	Collector's number
TIPO	1	Type of material
DUB	1	Number of duplicates
PAIS	10	Country of collection
LOC	60	Locality
END	1	Cuban endemic (yes/no)
SEC	1	Phytogeographical sector
DIS	2	Phytogeographical district
PRO	2	Province
ANO	9	Year of collection
FEC	8	Month and day of collection
COLEC	15	Collector(s)
HAB	2	Plant habit
FEN	7	Phenological state
CULT	1	Cultivated (yes/no)
VEG	2	Surrounding vegetation
SUEL	15	Soil type
MSNM	9	Altitude (metres above sea level)

The programme was based on the intention of avoiding the material difficulties mentioned above that prevent investigators from using perhaps even more complete but also more sophisticated systems, and was also intended to be used in microcomputers lacking hard disk, and to be operated by users with little knowledge of foxBASE or similar systems. This fact, although reducing the possibilities of our programme, permits its introduction in institutions lacking powerful hardware.

The programme, among others, includes the following functions: registration and editing of records, mass replacements in certain data fields according to the user's specifications, listing of specified fields, printing of herbarium labels, use of working files, collating of files created on diskettes using a hard disk and possibly re-ordering them, distribution of data on diskettes.

The first version of this programme and database has recently been disseminated among the several scientific and pedagogic institutions in Cuba possessing a herbarium, e.g., National Botanic Garden (JBN) which together with the IES have the biggest plant collections in our country. At the same time, and as a demonstration of the interest that this programme has promoted, versions of the system for exclusive use in lichens and fungi collections have been requested and subsequently adapted for that purpose.

The system has recently been adopted as a basis for the National Network of Herbaria. Altogether about 40,000 records have already been registered by HAC and HAJB, the herbarium of the National Botanic Garden; in HAC mainly information related to the endemic species of Cuba.

At present a second version of the programme is in preparation, which will also handle information on the plant uses, synonymy, conservation status, and comments. Furthermore, the database in this second version will occupy less disk space and, by automatic checking procedures, protect the user from entering errors in some of the fields, such as species authors' name, habitat and endemism.

The last objective of this work could be the creation of a database network with similar patterns among Cuban herbaria, so the free exchange of information between institutions and scientists could be facilitated in the future.

The JBN developed a plant register in FoxBase (Camino *et al.* 1990), based on the recommendations of the International Transfer Format for Botanic Garden Plant Records (Anon. 1988).

The taxonomic database of cultivated plants of Cuba (Knüpffer *et al.* 1990) and the national documentation system for plant genetic resources are described in detail in Knüpffer (1992) and Esquivel *et al.* (1994), respectively. The latter paper mentions also other past and present computerization projects in plant genetic resources collections and breeders' collections.

Literature consulted

Alain, Hno., 1969: Flora de Cuba. Suplemento. Caracas, 150 pp.

Anon., 1988: The International Transfer Format for Botanic Garden Plant Records. Version 001.00. Botanic Gardens Conservation Secretariat / Hunt Institute for Botanical Documentation, Carnegie Mellon Univ., Pittsburgh, 64 pp.

Camino, M., A. Leiva & M. Curtiss, 1990: El Registro de Plantas del Jardín Botánico Nacional. In: Resúmenes V Congr. Latinoamericano Botánica, La Habana, 25—30 Junio 1990, p. 342.

Cejas, F., 1990: Proyecto de base de datos y programa para las colecciones de los herbarios cubanos. Rev. Ciencias Biológicas, La Habana, Cuba (in press).

Cejas, F., 1991: Necesidad de la automatización en el control de datos de los herbarios. IV Encuentro de Botánica "Bisse in Memoriam". Camagüey, Cuba.

Cejas, F. & J.L. González, 1990: Propuesta de bases de datos para las colecciones botánicas de los herbarios cubanos. In: Resúmenes V Congr. Latinoamericano Botánica, La Habana, 25—30 Junio 1990, p. 238.

Esquivel, M., A. González, L. López, M. Rodríguez, D. Jiménez & H. Knüpffer, 1994: The National Database of Plant Genetic Resources of Cuba (NDPGRC). In: Hammer, K., M. Esquivel & H. Knüpffer (eds.), "... y tienen faxones y fabas muy diversos de los nuestros ..." — Origin, Evolution and Diversity of Cuban Plant Genetic Resources, pp. 671—691. Inst. Pflanzengenetik u. Kulturpflanzenforsch. Gatersleben.

Esquivel, M., L. López, M. Alvarez & F. Cejas, in press: Bases de datos y sistemas computarizados de procesamiento de información relacionados con la conservación y utilización de los recursos fitogenéticos en Cuba. In: Esquivel, M., L. López & D. Debouck (eds.), Fitorrecursos de

Cuba y su uso: 500 años después. Memorias del seminario científico por el INIFAT, JBN y EXPOCUBA. 3—4 diciembre 1992. Oficina IPGRI para las Américas, Cali, Colombia.

Frost-Olsen, P. & L.B. Holm-Nielsen, 1986: The AAU — Flora of Ecuador Information System. Rep. Bot. Inst. Univ. Aarhus, Denmark 14, 1—39.

Knüpffer, H., 1992: The Database of Cultivated Plants of Cuba. In: Hammer, K., M. Esquivel & H. Knüpffer (eds.), "... y tienen faxones y fabas muy diversos de los nuestros ..." — Origin, Evolution and Diversity of Cuban Plant Genetic Resources, pp. 202—212. Inst. Pflanzengenetik u. Kulturpflanzenforsch. Gatersleben.

Knüpffer, H., M.A. Esquivel & K. Hammer, 1990: DBCPC: a database for the cultivated plants of Cuba. Rev. Jardín Bot. Nacional, La Habana 11, 91—104.

Ordóñez, M. de J. & L. Bojórquez, 1990: Conservation data centre of UNIRBMEX. Ann. Meet. Taxonomic Databases Working Group, 1990. 6 pp.

Pankhurst, R.J., 1986: A package of computer programs for handling taxonomic databases. Computer applications in the Bioscience 2 (1), 33—39.

Rodríguez, M., E. Moreno & A. López, 1990: Especies cubanas relacionadas con el ñame (Dioscorea alata), el boniato (Ipomoea batatas) y la papa (Solanum tuberosum). In: Resúmenes V Congr. Latinoamericano Botánica, La Habana, 25—30 Junio 1990, p. 324.

Rodríguez, M., F. Cejas, A. López & L. Montes, 1994: The wild relatives of cultivated plants in the flora of Cuba. In: Hammer, K., M. Esquivel & H. Knüpffer (eds.), "... y tienen faxones y fabas muy diversos de los nuestros ..." — Origin, Evolution and Diversity of Cuban Plant Genetic Resources, pp. 542—567. Inst. Pflanzengenetik u. Kulturpflanzenforsch. Gatersleben.

Workshop "Application of Databases in Botany", 1990: Final Report. Univ. Concepción, Chile, 5 pp.