

**Ministerio de Ciencia, Tecnología y Medio Ambiente**

**Instituto de Ecología y Sistemática**

**Departamento de Plantas Vasculares**

***BASES DE DATOS Y REQUERIMIENTOS PARA EL MANEJO  
DE COLECCIONES BOTÁNICAS CUBANAS***

**Tesis en Opción al Grado de Master en Ecología y Sistemática Aplicada  
Mención en Curatoría Vegetal**

**Autor: Francisco Cejas Rodríguez**

**Tutor: Dr. Antonio López Almirall**

**Ciudad de la Habana**

**1996**

## **AGRADECIMIENTOS**

Mis sinceros agradecimientos al Doctor Antonio López Almirall, pionero en la aplicación de la computación en nuestra Institución, en quien he tenido siempre un guía y apoyo constante para el desarrollo de este trabajo.

A los integrantes de la Subdirección de Botánica, que con sus señalamientos durante el desarrollo y manejo del programa aquí expuesto se convirtieron en autores del mismo, principalmente la responsable del Herbario HAC Lic. Isora Baró, la Lic. Miriam Prede y el Lic. Armando Payo.

Al Lic. Pedro Herrera, contrapartida humana de los diccionarios botánicos que se van generando, que lo convierte en consulta obligatoria y nuestro asesor principal.

A la Doctora Nancy Ricardo por los consejos brindados para la ejecución del trabajo y la paciencia mostrada en la dirección del proceso de maestría.

A todos los que de una u otra forma han permitido la culminación de este trabajo, gracias.

## **SÍNTESIS**

Con la utilización de avanzados métodos científicos y la infraestructura requerida, nuestras colecciones biológicas permiten la generación e interpretación de información sobre diversidad biológica.

El tiempo y trabajo manual que requiere la compilación y mantenimiento de un banco de datos computarizado es elevado; gasto justificado por lo conveniente que resultan los catálogos automatizados para el manejo de la información de herbarios.

Con el propósito de encauzar los esfuerzos de las instituciones cubanas que comenzaban a crear bancos de datos de información botánica, se emprendió la conformación de un sistema relacional de bases de datos curatorial y taxonómica, que respondiese a la experiencia acumulada por dichas instituciones.

Sobre esta estructura y posibles relaciones se establecen los requerimientos de un programa cuya intención principal fuese:

- (a) permitir a personal no especializado introducir y/o utilizar información
- (b) utilizar ficheros de referencia, rigurosamente verificados, como diccionarios de ayuda que minimicen el margen de error al momento de introducir datos
- (c) controlar la calidad de la entrada de información y velar por la actualidad taxonómica.

Gracias a este programa, durante la captura de datos, sólo requieren del teclado directo por el operador aquellos campos portadores de información irrepetible en cada ejemplar, como los curatoriales y de colecta.

El sistema relacional de bases de datos generado y el programa que lo sustenta, fue adoptado por la Red Nacional de Herbarios.

## ÍNDICE

1. Introducción .....	2
2. Antecedentes .....	7
3. Materiales y Métodos .....	11
4. Resultados y Discusión .....	14
4.1. Elección del sistema de gestión de datos .....	14
4.2. Estructura de las bases de datos .....	15
4.3. Requerimientos del programa que relacione las bases de datos .....	24
4.3.1. Captura y edición de datos .....	28
4.3.2. Posibilidades de ayuda al personal .....	30
4.3.3. Interrelación automática entre la base de trabajo y los ficheros de referencia .....	33
4.3.4. Actualización y revisión de la información taxonómica .....	34
4.4. Potencialidad de utilización de los ficheros de referencia .....	36
Conclusiones .....	40
Recomendaciones .....	43
Referencias	
Anexos	

# ***INTRODUCCIÓN***

## **1. INTRODUCCIÓN**

Las colecciones biológicas, verdaderos bancos de datos e información relativos a los vegetales y animales, son en la actualidad valiosas acumulaciones de material biológico, enriquecidas durante generaciones, que permanecen como entes pasivos en las investigaciones científicas complejas, generalmente utilizadas por:

- Taxónomos que revisan algunos grupos de especies.
- Consultantes que desean identificar algún ejemplar por comparación con los materiales existentes.
- Especialistas de otras disciplinas que necesitan listas de especies u otras informaciones, como colectores, localidades, etc., para realizar sus trabajos.

Esta situación obedece a que la masa de datos acumulada es tan grande, solamente la colección botánica o herbario de la Academia de Ciencias de Cuba contiene alrededor de 350 mil ejemplares, que no es posible concebir su manejo por medios manuales. Los procedimientos computarizados, al mismo tiempo que constituyen la única vía para procesar esta ingente cantidad de información, ofrecen actualmente la posibilidad de alcanzar logros inimaginables sin ellos (Peláez y Ordóñez, 1991).

Se hace necesario, por tanto, la captura de toda esta información espacial y temporal a través de una red de bases de datos computarizada, que establezca un amplio sumario de la información taxonómica básica, atributos biológicos y potencial para la utilización de todas las especies vegetales; y que sea accesible a diferentes usuarios: el biólogo, el ingeniero forestal, el químico, el maestro, el médico, el técnico o cualquier persona interesada, quienes

pueden obtener una serie de datos de mucho valor con fines de aplicación práctica o industrial en el desarrollo de sus trabajos.

El proyecto de constitución de la Red Nacional de Herbarios plantea la necesidad de agilizar el intercambio de la información referente a las colecciones botánicas, y aboga por la creación de una Red de Bases de Datos Botánicas, la cual coordine el trabajo que realizan las diferentes Instituciones que desarrollan bases de datos automatizadas.

Partiendo de estas premisas se concibió la siguiente hipótesis de trabajo:

**Los sistemas automatizados agilizan y facilitan la captura de datos con el menor rango de error posible; si estos principios son efectivos en bancos de datos de herbario, podremos demostrar la factibilidad de aplicar métodos computarizados en la captura y manejo de la información botánica disponible, su ejecución por personal no especializado en estas técnicas y su aprovechamiento, no sólo para trabajos curatoriales, sino también para botánicos en general.**

Para corroborar esta hipótesis nos planteamos los siguientes objetivos:

- 1.- Conformar un sistema de bases de datos relacionales cuya estructura responda, en cuanto a la información que debe contener, a la experiencia y necesidades de las instituciones cubanas y a los patrones que se tratan de fijar a nivel internacional, previendo su enlace con bases de datos extranjeras.
- 2.- Establecer los requerimientos necesarios para el manejo automatizado de información en colecciones botánicas, que respondan a las capacidades materiales, necesidades e intereses de nuestras instituciones.

3.- Explotar las posibilidades que nos brindan los ficheros de referencia que se originan (taxonómicos, de sinónimos y nombres comunes, entre otros); para crear programas accesorios que puedan ser aprovechados en temáticas botánicas no relacionadas directamente con el trabajo de colecciones.

Para el manejo de las colecciones cubanas, se constató la escasa utilización de programas de computación foráneos que, aunque contienen una información científica amplia, no responden a las capacidades materiales y/o financieras de las instituciones interesadas, y a sus necesidades e intereses más urgentes, lo que imposibilita su empleo.

La **novedad científica** del trabajo es que por primera vez, se implementa un sistema automatizado propio, especializado en captura de información de herbarios, que brinda ayuda en particularidades de nuestro país y provisto de sistemas automáticos para actualizar la información taxonómica, lo que permite a personal no especializado asumir la tarea de captura y manejo de la información en los herbarios cubanos.

La **importancia** de este programa es que homogeneiza la información botánica a nivel nacional y facilita la creación de una Red Nacional de Bases de Datos de Herbarios entre las colecciones botánicas, permitiendo a los investigadores, instituciones gubernamentales y al público en general, la ubicación e identificación de los recursos botánicos con potencial para ser aprovechados en el campo de la farmacología, medicina, ingeniería y ecología entre otros, con el objeto de desarrollar diferentes alternativas económicas, tanto industriales como no tradicionales.



En líneas generales, será accesible la información botánica que se encuentra actualmente dispersa, y se conocerá la ubicación geográfica de los recursos botánicos con que cuenta el país para su administración.

# ***ANTECEDENTES***

## **2. ANTECEDENTES**

Diversas instituciones foráneas utilizan métodos automatizados para el manejo de sus colecciones botánicas, lo que incluye el desarrollo de diferentes software o programas, como le llamaremos de aquí en adelante, que para las instituciones cubanas resultan prohibitivos por los esquemas incorporados, sus precios o requerimientos de trabajo. Se conoce entre otros, del programa implementado por el Herbario Nacional de México, que contiene campos o ayudas específicos para este país; el TROPICOS y el ALICE del Missouri Botanical Garden y el Royal Botanical Garden respectivamente, de altos precios, el BG-BASE del Holden Arboretum, que como mínimo ocupa 40 Mbytes en el disco duro de la computadora, o el SYNTAX de la Universidad del Ruhr, que corre sobre Oracle, sistema de gestión de bases poco conocido en nuestro país.

En Cuba, a mediados de los años 80, comienzan los primeros intentos de capturar por medios automatizados la información que sobre antófitos endémicos cubanos aparece en las etiquetas del herbario de la Academia de Ciencias de Cuba (HAC).

Tres equipos de investigadores, dirigidos por el Dr. López Almirall, asumieron esta tarea para los sectores Occidental, Central y Oriental (*sensu* Samek, 1973) del Archipiélago Cubano.

Una vez preestablecidos los campos de las bases de datos, que en general abarcaban información curatorial, taxonómica y de localización; y utilizando solamente las facilidades del sistema de gestión de bases de datos, se comenzó y culminó en dos años el proceso de captura de la información, que como resultado tuvo la creación de las primeras bases de datos

biológicas automatizadas del Instituto de Ecología y Sistemática (CITMA), soporte de numerosos trabajos de distribución y taxonómicos, entre otros, servicios de localización de recursos fitogenéticos, confección de listados florísticos, etc..

Casi paralelamente, otras instituciones nacionales comenzaron a crear proyectos de bases de datos curatoriales (Fleitas et al., 1985; Palazuelos et al., 1985).

Existiendo un consenso general sobre la utilidad de los modelos relacionales de datos para bases curatoriales, se decide continuar la automatización del resto de los especímenes del herbario HAC y extender esta experiencia al resto de las colecciones botánicas, para lo cual se realizan modificaciones y adiciones al formato de la base de datos original, apoyados en la experiencia acumulada y en estrategias planteadas internacionalmente (Allkin, 1991).

Los mayores esfuerzos en las Ciencias Naturales se dirigen en la actualidad a la implementación de patrones, que eviten redundancias y garanticen que la información brindada por cualquier institución científica pueda ser aprovechada por las otras, sin necesidad de gastar esfuerzos en adaptarlas a sus sistemas de trabajo, como indican los objetivos del "Grupo de Trabajo en Bases de Datos Taxonómicas" (TDWG) y de la "Organización Internacional de Información sobre Plantas" (IOPI). Estos son respectivamente el grupo pionero que establece los estándares voluntarios de forma y contenido de datos sobre plantas, y el encargado de preparar bases de datos con información sobre plantas para uso de la comunidad científica.

Sin la existencia de estos patrones y bajo ciertas circunstancias, el uso de las computadoras puede tener algunas desventajas, pues la existencia de grupos que trabajan aisladamente en datos parcialmente congruentes pueden motivar redundancias en la información y diversidad de patrones (Berendsohn, 1992).

La imposibilidad de utilizar en nuestras colecciones botánicas programas automatizados implementados en instituciones extranjeras, hizo que se establecieran los requerimientos para el desarrollo de un programa propio, que permitiera la corrección de los datos taxonómicos que se introducen, mediante la utilización de diversas fuentes de descriptores; ya fueran ficheros de referencia, menús o ventanas de ayuda.

# ***MATERIALES Y MÉTODOS***

### **3. MATERIALES Y MÉTODOS**

Para la conformación de nuestras bases de datos nos basamos fundamentalmente en las experiencias del Instituto de Ecología y Sistemática (CITMA) y el Jardín Botánico Nacional, (MES). Paralelamente se analizaron formatos de bases de datos desarrolladas en otras instituciones, siempre abiertos a posibles modificaciones o inclusión de nuevos campos en nuestra estructura inicial, adecuándolas a las necesidades propias.

Para la estructuración de las bases de datos se siguió la siguiente estrategia:

- Análisis del grado de desarrollo de las bases de datos relacionadas a la botánica en cada una de las instituciones cubanas y foráneas revisadas.
- Análisis del flujo y relaciones de la información presentes en las etiquetas de un herbario (diseño e interpretación de un sistema de manejo de un herbario típico)
- Análisis de algunos sistemas computarizados que administran información botánica, enfocando este análisis hacia la factibilidad de aplicación en nuestro país.

Se eligió el sistema de gestión de datos foxBASE+ ver. 2.00 para el manejo de estas bases de datos, y tratando de evitar las desventajas que puede traer aparejada el uso de computadoras, se trazaron los parámetros a seguir por un programa que:

- Sea lo mas amigable posible, de forma tal que personal no especializado en bases de datos pueda utilizarlo sin mayores dificultades.
- Cuento con la ayuda mas diversa, basada en descriptores contenidos en ficheros de referencia, menús y ventanas; que permita la captura automática de datos, agilice su entrada y reduzca el porcentaje de error humano.

- Garantice la corrección y actualización de los datos sistemáticos introducidos.

Para la creación de nuestras fuentes de descriptores, se ajustaron e incorporaron a ficheros de referencia patrones ya establecidos en diversos campos de la botánica. Dichos descriptores deben ser prácticamente inalterables en el tiempo, con vistas a garantizar la vitalidad de nuestras bases de datos y programa.

Así, se revisaron bases de datos o literatura sobre temas nomenclaturales, de colecciones botánicas, nombres comunes, especies sinantrópicas, tipos nomenclaturales, localización fitogeográfica y formaciones vegetales, entre otros. Donde no se disponía de información, como en fenología o tipos de hábito, se crearon patrones según la experiencia de las instituciones botánicas cubanas.

En los casos pertinentes, se tendió a utilizar un lenguaje corriente en los descriptores generados por los conceptos que estos trabajos contienen, para permitir el manejo del programa por personal no especializado.



## ***RESULTADOS Y DISCUSIÓN***

## **4. RESULTADOS Y DISCUSIÓN**

Un sistema de bases de datos puede ser definido como un "conjunto de elementos interrelacionados, que sirven a un propósito común; sistema influido por aspectos externos que incluyen reglas ya establecidas, políticas institucionales y una gran variedad de otros factores que deben ser considerados" (Berendsohn, 1992).

El sistema de gestión de las bases de datos, la definición de éstas, y la implementación de un programa que ejecute las relaciones definidas entre ellas, constituyen el componente principal de este sistema, y sobre ellos se basó nuestro trabajo.

### **4.1. Elección del sistema de gestión de datos**

Las razones que hicieron nos decidiéramos por el foxBASE+, como sistema de gestión de bases de datos fueron:

- Se basa en el formato .DBF, utilizado en sistemas de amplia distribución como las distintas versiones del dBASE y foxPRO, lo que garantiza el manejo por un amplio espectro de usuarios de la información que se genere.
- En el IES existe una gran experiencia en el manejo y explotación de las posibilidades de este sistema.
- Siendo uno de nuestros objetivos el apoyo a la creación de la Red Nacional de Bases de Datos Botánicas, no es factible desarrollar un programa en sistemas mas poderosos que no puedan ser utilizados por instituciones que no cuentan con los recursos necesarios para su

manejo, y sí poseen colecciones botánicas de valor; por ejemplo los Institutos Pedagógicos establecidos a lo largo del país, que cuentan con herbarios representativos de flóruas regionales o locales.

- Por otra parte, las bases desarrolladas, que siendo contentivas de los datos son el núcleo de cualquier sistema de información, son fácilmente transferibles a otros sistemas.

#### **4.2. Estructura de las bases de datos**

La confección de ficheros automatizados, contentivos de la información proveniente fundamentalmente de las etiquetas de herbario y de notas de campo, permiten ampliar el espectro de trabajo del personal que labora en el herbario, al acelerar ciertas rutinas del trabajo curatorial como:

- Dar entrada a nuevos materiales
- Catalogar los ejemplares de herbario
- Confeccionar etiquetas
- Crear modelos de intercambio y manejo

Las bases de datos basadas en especies son una rica fuente de información para ecólogos y conservacionistas, ahorran tiempo de trabajo a los curadores cuando existan preguntas en estas materias y, por otra parte, constituyen una base importante para las investigaciones botánicas, ya sea en la elaboración de floras, claves de identificación, análisis fenéticos y cladísticos, mapas de distribución, etc., además de evitar el manejo y por tanto contribuir a la conservación de los materiales contenidos en nuestras colecciones botánicas, algunos de los cuales se remontan a finales del siglo pasado.

Cualquier propuesta de bases de datos debe garantizar la continuación y mantenimiento a largo plazo de los datos, en independencia de la futura evolución del software o del hardware.

Con el fin de ampliar las posibilidades de utilización de nuestras bases de datos propusimos una estructura suficientemente flexible para permitir su uso por los más diversos usuarios dentro de la temática que aborda, pero que al mismo tiempo permanezca prácticamente inalterable en el tiempo, con vistas a garantizar su vitalidad, evitando "sobrecargarla" con campos, quizás factibles de utilizar en un momento o tema determinado, pero que generalmente no son empleados por la mayoría de los usuarios, ya sea por dificultades en el acceso a la información que se solicita o por lo específico de ésta. Todo ello sin dejar de responder a las necesidades mas acuciantes de los usuarios potenciales, y considerando la posible intercomunicación con otras similares de instituciones externas.

Se confeccionó una base de trabajo central, alrededor de la cual se situaron diez bases de datos, las que nombraremos indistintamente diccionario, si contienen descriptores que pueden ser utilizados por cualquier registro o *taxa*, y base accesoria, si contienen información referida a un *taxa* o registro en particular, caso éste que sólo ocurre en las bases de datos de nombres comunes y comentarios.

Las bases de datos diseñadas son:

1.- Base de datos de trabajo:

Hace mayor énfasis en datos taxonómicos (campos 1-4), curatoriales (campos 5-9) y de colecta (campos 10-17), y trata más someramente los datos ecológicos (campos 19-22) y fenológicos (campo 18), para un total de 26 campos alfanuméricos que se relacionan en la Tabla I.

Tabla I. Estructura de la base de datos de trabajo.

No.	CAMPO	TIPO	TAMAÑO	CONTENIDO
1	FAM	C	3	Familia
2	GEN	C	24	Género
3	ESP	C	20	Epíteto específico
4	INFRASP	C	24	Epíteto infraespecífico
5	HER	C	13	Número de herbario
6	EXHER	C	12	Número de exherbario
7	NO_COL	C	12	Número de colector
8	TIPO	C	1	Tipo nomenclatural
9	DUB	C	1	Número de dubletas
10	LOC	C	60	Localidad de colecta
11	PAIS	C	10	País
12	SEC	C	1	Sector fitogeográfico
13	DIS	C	2	Distrito fitogeográfico
14	PRO	C	2	Provincia
15	COLEC	C	15	Colector(es)
16	ANO	C	9	Año de colecta
17	FEC	C	8	Fecha
18	FEN	C	7	Estadio fenológico
19	CULT	C	1	Cultivado
20	VEG	C	2	Vegetación circundante
21	SUEL	C	15	Tipo de suelo
22	MSNM	C	9	Altitud
23	HOJA	C	5	Hoja cartográfica
24	CX	C	2	Coordenadas sobre x
25	CY	C	2	Coordenadas sobre y
26	REGISTRO	C	6	Número de registro

Esta estructura es determinada por la dependencia de la cantidad y calidad de la información que provee la principal fuente con que contamos, la etiqueta de herbario, cuyos datos básicos son los que primero se relacionan.

En esta base de trabajo no se incluyen los campos de autores, hábito y procedencia del *taxon*. El contar con un fichero de referencia de *taxa*, como se verá más adelante, hace inútil la presencia de estos campos, que siempre pueden ser referidos y chequeados en esta última, lo que implica un mayor aprovechamiento de las capacidades de almacenamiento de nuestro hardware.

Sin embargo, la base de trabajo acepta los datos de georreferenciación (campos 23-25), contenidos en otro fichero, para asegurar su independencia al manejarse.

## 2.- Base accesoria de comentarios:

Permite acopiar aquella información extra brindada en la etiqueta (coloración, etc.) que no se considera incluir en la base de datos de trabajo por poco frecuente o sorpresiva (Tabla II), reconociéndose los registros de la base de comentarios y los correspondientes en la base de trabajo a través de la similitud en los campos curatoriales y número del registro.

Tabla II. Estructura de la base accesoria de comentarios.

CAMPO	TIPO	TAMAÑO	CONTENIDO
HER	C	13	Número de herbario
EXHER	C	12	Número de exherbario
NO_COL	C	12	Número de colector
COMENT	C	207	Comentario
REGISTRO	C	6	Número de registro

## 3.- Diccionarios de epítetos botánicos:

Se crearon cuatro diccionarios que contienen respectivamente las familias botánicas, géneros, epítetos específicos e infraespecíficos que permiten liberar al usuario de la escritura directa.

El diccionario de familias porta además, como información accesoria, el acrónimo de éstas (Weber, 1982) y volumen de la obra Flora de Cuba en que aparecen (León, 1946, León y Alain, 1951, 1953, 1957, Alain, 1964, 1974) (Tabla III).

Tabla III. Estructura del diccionario de familias.

CAMPO	TIPO	TAMAÑO	CONTENIDO
BOTAFAM	C	20	Familia
ACRON	C	3	Acrónimo
TOMO	C	1	Tomo de Flora de Cuba

El diccionario de géneros (Tabla IV) brinda el acrónimo de la familia a que pertenecen éstos y, en caso de ser un sinónimo, el nombre del género aceptado actualmente (Anexo 1).

Tabla IV. Estructura del diccionario de géneros.

CAMPO	TIPO	TAMAÑO	CONTENIDO
BOTAGEN	C	24	Género
BUEN_NOM	C	24	Género válido
ACRON	C	3	Acrónimo de la familia

Los diccionarios de epítetos específicos (Tabla V) e infraespecíficos (Tabla VI), son los más propensos a incrementarse y por tanto evitan en mayor medida que los dos antes mencionados la introducción de errores ortográficos.

Tabla V. Estructura del diccionario de epítetos específicos.

CAMPO	TIPO	TAMAÑO	CONTENIDO
BOTAESP	C	24	Epíteto específico

Tabla VI. Estructura del diccionario de epítetos infraespecíficos.

CAMPO	TIPO	TAMAÑO	CONTENIDO
BOTAINF	C	24	Epíteto infraespecífico

## 4.- Base accesoria de nombre comunes:

Esta base, cuyos campos se relacionan en la Tabla VII, resulta útil en trabajos etnobotánicos y de identificación rápida por no especialistas de especies conocidas popularmente, recogiendo en el campo creado al efecto los nombres comunes que se vayan capturando de un *taxa* en específico (Anexo 1).

Tabla VII. Estructura de la base accesoria de nombres comunes.

CAMPO	TIPO	TAMAÑO	CONTENIDO
GEN	C	24	Género
ESP	C	20	Epíteto específico
INFRASP	C	24	Epíteto infraespecífico
COMUN	C	207	Nombres comunes

## 5.- Diccionario de colecciones botánicas:

Contiene los nombres de las instituciones botánicas representadas en nuestros herbarios y las siglas con que se identifican (Tabla VIII).

Tabla VIII. Estructura del diccionario de colecciones botánicas.

CAMPO	TIPO	TAMAÑO	CONTENIDO
EXHER	C	6	Siglas de herbarios
BOTAEXH	C	60	Nombres de herbarios



Este diccionario evita las confusiones que existan respecto a las siglas de las diferentes colecciones botánicas, y que son propensas a ser introducidas erróneamente durante la captura de información

#### 6.- Diccionario de localidades de colecta:

Relaciona las localidades de colecta, la hoja cartográfica y coordenadas del mapa a escala 1:250 000 según la proyección cónica conforme de Lambert (Anexo 1), provincia, sector y distrito fitogeográfico (*sensu* Samek, 1973), (Tabla IX). El aporte automático de los datos de georreferenciación y fitogeográficos, difíciles de añadir manualmente a cada registro creado, permiten búsquedas más rápidas a los interesados en esta información.

Tabla IX. Estructura del diccionario de localidades.

CAMPO	TIPO	TAMAÑO	CONTENIDO
BOTALOC	C	60	Localidades
HOJA	C	5	Hoja cartográfica
CX	C	2	Coordenadas sobre x
CY	C	2	Coordenadas sobre y
SEC	C	1	Sector fitogeográfico
DIS	C	2	Distrito fitogeográfico
PRO	C	2	Provincia

#### 7.- Diccionario de sinónimos e información bibliográfica:

La estrategia que se sigue con este diccionario, permite definir la información que contiene cada registro (Anexo 1), de acuerdo a los campos ocupados conjuntamente con los referidos a la especie (Tabla X).

- (a) Campos de bibliografía y sinónimo ocupados, mas llamado de atención en el campo de basonimia, indica el basónimo de la especie y la bibliografía de donde se conoce el cambio.
- (b) Campos de bibliografía y sinónimo ocupados indica uno de los sinónimos de la especie y la bibliografía de donde se conoce el cambio, creándose tantos registros como sinónimos aparezcan.
- (c) Campo de bibliografía solamente ocupado, indica la bibliografía original de la especie.

Tabla X. Estructura del diccionario de sinonimia/bibliografía.

CAMPO	TIPO	TAMAÑO	CONTENIDO
GEN	C	24	Género
ESP	C	20	Epíteto específico
INFRASP	C	24	Epíteto infraespecífico
LITER	C	64	Referencia bibliográfica
BASON	C	1	Marca de basónimo
SINON	C	69	Sinónimo

8.- Diccionario de *taxa* (binomios):

En este diccionario los campos taxonómicos están acompañados de datos que nunca varían en los binomios; a saber, autor(es), hábito y procedencia (Tabla XI).

Tabla XI. Estructura del diccionario de *taxa*.

CAMPO	TIPO	TAMAÑO	CONTENIDO
FAM	C	3	Familia
GEN	C	24	Género
ESP	C	24	Epíteto específico
INFRASP	C	24	Epíteto infraespecífico
AUT	C	38	Autor de la especie
INFRAUT	C	38	Autor de la infraespecie
END	C	1	Procedencia
HAB	C	6	Hábito

En la base de datos de trabajo, diccionarios de *taxa*, nombres comunes y sinonimia/bibliografía se repiten los campos que identifican a un binomio en específico: género mas epítetos específico e infraespecífico, y a veces el de familia. De utilizarse un sistema de codificación de especies, digamos desde el diccionario de *taxa*, sólo es necesario utilizar el código de referencia en el resto de las bases de datos mencionadas y éstas ocupan menos espacio en los soportes magnéticos.

Sin embargo, para poder codificar los binomios, primero hay que relacionar todos los conocidos, y aún no se cuenta con ningún listado actualizado de las especies que conforman nuestra flora.

Para soslayar esta falta se hicieron intentos de lograr la codificación automática por el propio programa, basados en la posible transformación numérica de los epítetos botánicos, que al no resultar hicieron necesaria la aparente redundancia en la información, dado que uno de los objetivos del sistema de bases de datos y programa acompañante es su adopción por las diferentes instituciones que conforman la Red Nacional de Herbarios, y no debe ocurrir que cada una de las instituciones codifique independientemente sus binomios según los vayan incorporando.

Esta codificación independiente dificulta la fusión de las bases que se crean para conformar un banco de datos central, propuesto por la Red Nacional de Herbarios a establecerse en el IES.

La estructuración de nuestro sistema de bases de datos relacionales presentó similitudes con los patrones mostrados en la 8va Reunión del "Grupo de Trabajo en Bases de Datos

Taxonómicas" (TDWG) y 2da de la "Organización Internacional de Información sobre Plantas" (IOPI) en 1992, lo que confirma la efectividad del análisis realizado.

### **4.3. Requerimientos de un programa que relacione las bases de datos**

Un modelo relacional organiza cada base de datos como una tabla bidimensional, donde las columnas constituyen los campos y las filas, los registros. Con esta estructura se puede manipular interactivamente una gran cantidad de información, lo que incidió se constituyeran en aplicaciones populares para microordenadores (Tsu-der, ?).

Sin embargo, es normal que durante el tiempo de captura y tamizaje de información se detecten errores humanos.

En el caso de la información sobre antófitos endémicos, los errores más comunes eran, a nivel de grupos de trabajo, debido a una ortografía deficiente y por la diversidad de descriptores utilizados. Al tratar de relacionarnos con bases de otras instituciones, surgían además problemas de incompatibilidad de formatos y metodología.

Para evitar esta situación, y una vez determinada la estructura de las bases de datos, es necesario aplicar un programa, basado en las experiencias de las ciencias tanto de computación como biológicas en el del programa, que vele por la homogeneidad, veracidad y actualidad de los datos que se introducen y sea capaz de llevar a cabo, sin la participación consciente del operador, todas las relaciones que se hallan preestablecido para las bases.

Cualquier proyecto para desarrollo de nuevas bases de datos debe demostrar que el software existente ha sido considerado inadecuado. Entre las limitaciones que presentan los programas foráneos revisados, las más usuales son:

- I.- Por su sofisticación o alto costo, no responden a las capacidades materiales y/o financieras de muchas de nuestras instituciones.
- II.- Algunos de estos sistemas corren en programas que no son los de más amplio uso en nuestro país, lo que dificulta la compatibilización con bases de datos ya desarrolladas internamente.
- III.- Es imposible pretender que brinden ayuda en particularidades de nuestro país, como coordenadas geográficas por localidades y tipos de vegetación; portando, por el contrario, esquemas que no se ajustan a nuestra realidad.
- IV.- Hay campos que en la práctica resultan difíciles de llenar o evaluar, lo que acumula espacios vacíos.
- V.- Existen campos cuyo contenido, al carecer de los patrones de ejecución, se presta a diversas interpretaciones según el colector, y al final resultan irre recuperables.

Puede ocurrir también que:

- VI.- No tienen forma de controlar los errores en la captación de datos.
- VII.- Carecen de sistemas automáticos para actualizar la información taxonómica, por lo que periódicamente se hacen necesarias largas sesiones de corrección.

Según la experiencia de trabajo de otras instituciones, estos inconvenientes han traído como consecuencia:

- 1.- Mucha de la información contenida en algunas de las bases de datos más prestigiosas del mundo es irre recuperable, pues resulta errada o atrasada.

2.- Estos sistemas de información sólo se usan para manejar colecciones biológicas en trabajos de organización, pero nunca en investigaciones científicas.

Todo esto determina la implementación de un programa, para cuyo diseño y desarrollo nos basamos en la experiencia de las instituciones cubanas, sus necesidades y capacidades materiales.

Aunque centrado en brindar herramientas para la captura de información, el programa que se propone para el manejo de estas bases de datos alivia al usuario de tareas comunes en un sistema de gestión de datos, como son la creación y puesta en funcionamiento de los ficheros de trabajo, edición de datos, ejecución de listados, búsquedas rápidas o basadas en varias condiciones y recorridos a lo largo de las bases de datos, entre otras; todo mediante un sistema interactivo en el que el operador no precise conocer los comandos del sistema de gestión de datos.

Un ejemplo de lo anterior lo constituyen la generación de listados y el establecimiento de condiciones de trabajo, momento en el que se deben escribir los comandos propios del sistema de gestión y la condición o relación lógica. Esta última requiere el conocimiento por el usuario del manejo de diferentes operadores relacionales, su agrupamiento por prioridades y del tipo de objeto de búsqueda que debe coincidir con el tipo de la clave de búsqueda. Estos pasos resultan los de mayor complicación en la interacción con el sistema de gestión de datos.

Para facilitar la selección de campos y/o condiciones a personas no familiarizadas con el dBASE o foxBASE+, y que desean generar listados o realizar modificaciones en determinados

registros, se propone un menú básico que relaciona los campos contenidos en las bases de datos.

Primero se seleccionan el o los campos que se deseen listar simplemente marcando éstos. Si el listado o modificación debe realizarse bajo determinadas condiciones se ejecutan los siguientes pasos:

- Seleccionar el campo relacionado con el objeto de búsqueda o condición
- Indicar al programa si se trabajará con los registros que cumplan la condición deseada o, por el contrario, con todos los registros excepto los que contienen dicha condición.
- Teclar la condición deseada para el campo de interés. Los campos que disponen de diccionarios acceden a los mismos (Fig. 1).

FAM		PRO
<b>GEN</b>		COLEC
ESP	<b>GEN igual a</b>	ANO
INFRASP	GEN desigual a	FEC
HER		FEN
EXHER		CULT
NO_COL	<b>CAL</b>	VEG
TIPO		SUEL
DUB		MSNM
LOC		HOJA
PAIS		CX
SEC		CY
DIS		
TODOS LOS REGISTROS No DE REGISTRO		

Fig. 1. Menú para determinar condiciones. Como ejemplo se desea trabajar con registros que contienen al género *Calophyllum*.

Se puede especificar mas de una condición. La escritura de los comandos, la ubicación de los operadores relacionales y su ordenamiento por prioridades, los realiza el programa automáticamente según se vayan solicitando diferentes condiciones, de forma que puedan ser comprendidas por el sistema de gestión de datos.

Así, supongamos un caso hipotético en que deseamos trabajar con los registros que incluyan los ejemplares de *Calophyllum utile* y *Calophyllum antillanum*, florecidos y colectados en Holguín. Mediante la ventana interactiva explicada y tecleando solamente las tres primeras letras del género y del epíteto específico, gracias a los diccionarios disponibles, los pasos a seguir serían:

GEN	IGUAL A	CALOPHYLLUM	
ESP	IGUAL A	UTILE	
ESP	IGUAL A	ANTILLANUM	
FEN	IGUAL A	2	(código de fenología)
PRO	IGUAL A	11	(código de provincia)

El programa, sin la intervención o conocimientos del operador, ordena estas condiciones y las traslada al tipo de comandos del sistema de gestión de datos, obteniéndose:

```
locate for gen='CALOPHYLLUM' .and. (esp='UTILE' .or. esp='ANTILLANUM') .and. '2'$fen .and. pro='11'
```

#### 4.3.1. Captura y edición de datos

No es objetivo de este trabajo mostrar cómo se programaron todas las tareas enunciadas al inicio de este acápite, pero consideramos necesario explicar aunque sea someramente el funcionamiento "interno" del módulo referido a la captura y edición de datos; razón de ser del programa diseñado.



A través de una ventana o formato de pantalla se provee al operador de las herramientas y ayuda necesaria para llevar a cabo la entrada de datos al sistema.

A lo largo del tiempo y en la misma medida que el trabajo con el programa hacía necesario diversos ajustes y adiciones, esta ventana no varió o sólo ínfimamente; lo que no es extraño si consideramos que la misma es casi un reflejo fiel de la estructura de las diversas bases de datos implementadas (Fig. 2).

No de registro	Trabajo	Base de
Herbario: (v)		
Exherbario: (d)		
No de colector: (v,d)		
Tipo: (m)		
Familia: (d)		
Género: (d)		
Especie: (d)		
<b>Autor(es): (d)</b>		
Infraespecie: (d)		
<b>Infraautor(es): (d)</b>		
No Duplicatas:		
<b>Endémico: (m,d)</b> País:		
Colector(es): (v)		
Localidad: (d)		
Año:		
Fecha: (v)		
<b>Hábito: (m)</b>		
Fenología: (m)		
Cultivado: (v)		
SALVAR Y CONTINUAR		
	Abandonar	Localización
	<b>Comentario (ba)</b>	
	<b>Sinonimia (d)</b>	
	<b>Nombre Común (ba)</b>	
	Ayuda	
	Pag. Anterior	
	Inicio Fichero	
	Final Fichero	
	Registro No ...	
	Modificar	
	Vegetación Circundante: (m)	
	Suelo:	
	MSNM: (v)	

Fig. 2. Ventana de captura y edición de datos. Se indica existencia de ayuda a través de (d) diccionarios, (ba) bases accesorias, (m) menús y (v) ventanas. En negritas se muestran los campos a guardar en ficheros de referencia y no en la base de trabajo.

#### 4.3.2. Posibilidades de ayuda al personal.

Hasta que el personal responsabilizado con la introducción de datos no domine el uso de la computadora y del programa especializado, la aplicación de cualquier sistema implica un elevado costo de salario y tiempo.

De una etapa en la que todos los datos tenían que ser tecleados directamente por el operador, se ha pasado a otra en que sólo requieren de ello aquellos campos portadores de información irrepetible en cada ejemplar, como los curatoriales y de colecta, gracias a la inclusión paulatina de diversas fuentes de descriptores:

1.- Los ficheros de referencia (diccionarios y bases accesorias):

Las bases de datos descritas en el acápite 4.2., excepto la de trabajo y de comentarios, logran unificar la información susceptible de ser automatizada en los diferentes herbarios y permiten ser utilizados como ficheros de referencia para brindar ayuda accesoria en la introducción de la información. Además, los diccionarios de *taxa* y sinonimia/bibliografía juegan un papel fundamental en la actualización y revisión de los binomios que se introducen.

Todos estos diccionarios no sólo ofrecen los datos que contienen (Fig. 3) sino que son capaces a su vez de incorporar aquella información que no posea, digamos por ejemplo, un nuevo epíteto específico (Anexo 2). Este es el único momento que se precisa de escritura directa por parte del operador, pero el amplio volumen de información que han acumulado nuestros diccionarios desde 1987 a la fecha (Anexo 1), ha incidido en una disminución brusca de incorporación de nuevos binomios y epítetos, mayormente de especies exóticas o variedades cultivadas, lo que agiliza todo el procedimiento de introducción.

Ejemplo de ello es la base de datos que en la actualidad se desarrolla para el herbario de plantas útiles del Dr. Juan Tomás Roig. Cuando la base de trabajo contaba con 340 especies en 670 registros, sólo 25 de éstas fueron incorporadas por el operador al diccionario de *taxa* como nuevos binomios, mientras el resto se capturaba automáticamente.

CALOPHYLLUM ANTILLANUM (Revisada)	
CALOPHYLLUM BRASILIENSE VAR ANTILLANUM (Revisada)	
CALOPHYLLUM CALABA (Revisada)	
CALOPHYLLUM INOPHYLLUM	
CALOPHYLLUM PINETORUM (Revisada)	
CALOPHYLLUM RIVULARE (Revisada)	
<b>CALOPHYLLUM UTILE (Revisada)</b>	
CALOPOGON PULCHELLUS (Revisada)	
CALOPOGON TUBEROSUS (Revisada)	
CALOPOGONIUM COERULEUM (Revisada)	
CALOPOGONIUM MUCUNOIDES (Revisada)	
CALOPOGONIUM PUMILUM (Revisada)	
CALOTROPIS PROCERA (Revisada)	
CALYCOGONIUM ACUNANUM (Revisada)	
CALYCOGONIUM ANGULATUM (Revisada)	
<b>-- NUEVA ESPECIE --</b>	<b>-- MAS --</b>

Fig. 3. Ejemplo de captura a través de diccionario. En el ejemplo, diccionario de *taxa* luego de señalado el género *Calophyllum*, en el que la captura de uno de los binomios trae aparejada la del nombre de sus autores, hábito y procedencia. Con los comandos adicionales se pueden agregar nuevas especies o continuar la búsqueda en el diccionario.

Los ficheros de referencia de *taxa*, epítetos específicos e infraespecíficos, localidades de colecta y sinonimia/bibliografía, están formados enteramente a partir del propio esfuerzo de los usuarios que los han empleado desde sus inicios

Otros pudieron basarse total o parcialmente en bases de datos ya existentes o en descriptores establecidos internacionalmente.

Entre éstos últimos están las bases de familias (Weber, 1982), géneros (Index ??), colecciones botánicas (Holgrem, 1992) y nombres comunes, los últimos tomados, entre otras fuentes, de los volúmenes de la obra Flora de Cuba (León, 1946, León y Alain, 1951, 1953, 1957, Alain, 1964, 1974) y de las bases de datos etnobotánicas desarrolladas en el IES. En el diccionario de *taxa* se utilizó la base de datos de especies sinantrópicas (Ricardo et al, 1995) para validar la procedencia de las especies incluidas.

## 2 - Menús de descriptores:

De acuerdo con la estrategia de utilización de patrones internacionales se confeccionaron los menús de descriptores para aquellos campos que no necesitan de ficheros de referencia amplios y proclives a seguir desarrollándose (Fig. 4).

Herbario:	
Exherbario:	
No de colector:	
Tipo:	<input type="text"/>
Familia:	
Género:	
Especie:	
<b>Autor(es):</b>	
Infraespecie:	
<b>Infrautor(es):</b>	
No Dubletas:	

(I)sotipo
<b>(H)olotipo</b>
(C)otipo
(P)aratipo
(T)opotipo
(S)intipo
(L)ectotipo
(F)ototipo
(X)Ninguno

Fig. 4. Ejemplo de captura a través de un menú de descriptores. En este caso de tipos nomenclaturales.

Dichos menús aparecen en los campos contentivos de información sobre tipos nomenclaturales (Font Quer, 1975), localización fitogeográfica (Samek, 1973) y formaciones vegetales (Capote y Berazaín, 1984), garantizando la homogeneidad de los datos que se introducen. También se emplean en los campos de fenología y tipos de hábito, donde se

crearon patrones que cubrían sus posibilidades, basados en la experiencia de las instituciones botánicas cubanas (Anexo 3).

### 3 - Recuadros o ventanas de ayuda:

Los recuadros aparecen siempre en el ángulo superior derecho (Fig. 5) en los campos que precisan de cierta regularidad en la forma de introducir los datos pero que no cuentan con ningún tipo de referencia, a saber: curatoriales, fechas de colecta, altitudinales y de colector(es).

Herbario: <input type="text"/>	
Exherbario:	
No de colector:	
Tipo:	
Familia:	
Género:	
Especie:	
<b>Autor(es):</b>	
Infraespecie:	

Herbario (sección) y número actual

Ejemplo: HAC(I)13253

Herbario de la Academia de Ciencias

sección de investigación

ejemplar número 13253

Fig. 5. Ejemplo de recuadro de ayuda. En este caso para número de herbario.

A medida que se activan los campos cuya información está contenida en los ficheros de referencia, van ocurriendo los enlaces que se hayan preestablecido entre éstos y la base de trabajo. Caso de que se desee trabajar con el diccionario de sinonimia/bibliografía o las bases accesorias de comentarios y nombres comunes, existe un submenú, también en el ángulo superior derecho, que permite que se realice el enlace.

#### 4.3.3. Interrelación automática entre la base de trabajo y los ficheros de referencia

- Cuando se va a introducir una especie, se visualizan los binomios del género en cuestión disponibles en el diccionario de *taxa*. La captura de uno de estos binomios trae aparejada la del nombre de sus autores, hábito y procedencia. Actualmente se está efectuando un tamizaje de toda la información contenida en este diccionario (Anexo 1), por lo que al operador se le informa si los datos correspondientes al binomio que está capturando han sido revisados o no.
- Caso de que el binomio que interese no aparezca en el Diccionario de *taxa*, se procede a su conformación (familia, género, epíteto específico e infraespecífico), utilizando los diccionarios de epítetos a los que se accede individualmente mediante el teclado de las tres primeras letras del nombre que nos interese, y añadiendo los autores, hábito y procedencia.
- Las bases de comentarios, nombres comunes y sinonimia/bibliografía se activan cuando se desea introducir información de este tipo, mostrando los datos que contengan referidos al *taxa* o registro de que se trate, con lo que se evitan redundancias.
- La base de localidades brinda las conocidas hasta el momento y aporta los datos de georreferenciación.

#### **4.3.4. Actualización y revisión de la información taxonómica**

Los errores humanos y del programa, una vez en el sistema, son difíciles de encontrar y arreglar.

El desarrollo de un fichero de referencia sobre sinónimos, propicia la creación de un módulo de trabajo que vela por la actualización de la información taxonómica que se introduce:

- una vez editado o completado un nuevo registro, se chequea, el binomio recién capturado para ver si aparece dentro del diccionario mencionado. De ser así muestra la fuente en la que se conoce el cambio (Fig. 6).
- En caso afirmativo se puede proceder a su actualización (Fig. 7)..

Calophyllum brasiliense CAMB. var antillanum (BRITT.) STANDL.

En el diccionario de sinónimos aparece:  
 CALOPHYLLUM BRASILIENSE VAR ANTILLANUM STANDL.  
 En: J. Arnold Arboretum, 43:397. 1962  
 ¿Coincide su autor(es) con el de la especie introducida? (S/N): S

Fig. 6. Pantalla de aviso sobre posible sinónimo

CALOPHYLLUM BRASILIENSE VAR ANTILLANUM STANDL.  
 se considera sinónimo de:  
 Calophyllum antillanum  
 En: J. Arnold Arboretum, 43:397. 1962  
 ¿Utilizo esta última combinación? (S/N):

Localidad: Lomas que circundan Punta Gaviota, Siboney  
 Provincia: Santiago de Cuba

Fig. 7. Pantalla de aviso sobre intención de actualización de un posible sinónimo. La localidad de colecta puede ayudarnos a dilucidar algunos casos de sinónimos *pro parte*

- Posteriormente se controlan en el diccionario de *taxa*, los autores, hábito y procedencia de la especie en cuestión (Fig. 8). De esta forma se llama la atención sobre cualquier modificación que exista y se introducen correcciones de ser pertinente o, en caso de que el binomio no aparezca en la base, se adiciona a la misma.

CALOPHYLLUM BRASILIENSE VAR ANTILLANUM (CLU)

Tomo: 3
Anteriormente como autor(es) de esta infraespecie aparecía: (BRITT.) STANDL.
¿Sustituyo por STANL.? (S/N): N

Fig.8. Pantalla de llamada de atención sobre modificación en nombre de autor, introducido erróneamente como STANL.. Similar pantalla aparece para los datos de hábito y procedencia.

Al quedar verificada la información, ocurren las modificaciones y/o adiciones necesarias en las bases encargadas de guardarla, ya sea en la de trabajo o en los ficheros de referencia.

#### **4.4. Potencialidad de utilización de los ficheros de referencia**

Los listados generados por un banco de datos botánico pueden ser asimilados por otros sistemas, como es el caso del programa FITOMAP (Caballero et al, 1993), desarrollado en el IES para la realización de esquemas de distribución de plantas cubanas.

También, ante el volumen de datos contenidos en los ficheros de referencias creados, existe la posibilidad de usar los mismos individualmente en trabajos diferentes de aquel para el que han sido creados.

Por ejemplo, la confección de un listado actualizado de nuestra flora, a partir del diccionario de *taxa*, ya está en vías de dejar de ser una idea acariciada por aquellos deseosos de contar con un compendio de las especies descritas y modificaciones ocurridas desde la publicación de la Flora de Cuba (León, 1946, León y Alain, 1951, 1953, 1957, Alain, 1964, 1974) sobre todo



después de la realización del “Estudio Nacional de la Biodiversidad de la República de Cuba” (1995).

Otro caso más hipotético pudiera ser, utilizando la base de datos de nombres comunes, la confección de claves rápidas, si no de identificación al menos de acercamiento rápido a la misma, a partir de la regularidad que muchas veces aparece en la aplicación de nombres comunes basados en caracteres morfológicos o propiedades visibles de las plantas, como es el caso de "Chicharrón" que se encontró atribuido a un grupo de 18 plantas de hojas coriáceas con bordes dentados (*Antirhea aristata*, *Casearia ilicifolia* y *Drypetes mucronata*, entre otras).

Por otra parte, los ficheros de referencia pueden también incorporarse al desarrollo de diversos utilitarios, como ya se ha comprobado con los diccionarios taxonómicos.

Uno de estos utilitarios permite la elaboración de listados florísticos, susceptibles de ser incorporados a bases botánicas que lo requieran. Otro, basado en algunos módulos del programa de colecciones botánicas, posibilita la revisión de los campos taxonómicos de cualquier base de datos botánica, siempre que la estructura de los campos contentivos de esta información concuerden con la nuestra.

Este último aspecto ha motivado que diversos grupos de trabajo dentro del IES, que laboran en campos tan disímiles como Fitoquímica (Base de Datos de Alcaloides y Saponinas), Etnobotánica (Base de Datos de Plantas Útiles), Endemismo (Base de Datos de Antófitos Endémicos) y Palinología (procesamiento de imágenes en plantas melíferas) entre otros, que

están desarrollando bases botánicas y conocen de las potencialidades del programa y/o ficheros de referencia desarrollados por el HAC, sigan la estructura desarrollada por éste último, al menos en lo que a campos taxonómicos se refiere.

# ***CONCLUSIONES***

## **5. CONCLUSIONES**

- 1.- Se diseñó y puso en funcionamiento una estructura de bases de datos flexible para asegurar su uso por las instituciones botánicas cubanas y garantizar su vitalidad; integrada por una base de trabajo central y diez ficheros de referencia: diccionarios de epítetos botánicos, colecciones botánicas, localidades de colecta, sinónimos/bibliografía y *taxa*, más dos bases accesorias de comentarios y nombres comunes
2. Se establecieron los requerimientos para la implementación de un programa de manejo de las bases de datos creadas.
- 3.- Se diseñaron módulos que alivian al usuario de tareas comunes en el trabajo con bases de datos, como son la creación y puesta en funcionamiento de éstas, captura y edición de datos, ejecución de listados, búsquedas rápidas o basadas en varias condiciones y recorridos a lo largo de las bases de datos, entre otras; mediante un sistema interactivo en el que el operador no precisa conocer los comandos del sistema de gestión de datos.
- 4.- A través de distintas fuentes de descriptores, formadas a partir de patrones ya establecidos en diversos campos de la botánica, incluidos aquellos particulares de nuestro país (coordenadas geográficas por localidades, tipos de vegetación, etc.) se brinda un volumen de ayuda que agiliza la captura de datos y minimiza el margen de error.
- 5.- Se estableció un control de la calidad y actualidad de la información taxonómica mediante los diccionarios de *taxa* y sinonimia creados.

- 6.- Los ficheros de referencia desarrollados están siendo empleados tanto independientemente como en la implementación de diversos programas, por ejemplo de revisión taxonómica, elaboración de listados, etc.; de utilidad en otros campos de la botánica.
  
- 7.- Un programa diseñado según el formato propuesto, al que se nombró HERBARIO, fue adoptado por la Red de Bases de Datos Botánicas, que agrupa las Instituciones que automaticen sus colecciones botánicas. Entre los propósitos de esta Red está que el IES mantenga un banco central que se nutre de las bases de datos de los demás herbarios y a su vez brinda a éstos toda la información que se le solicite, además de asesorar en el manejo del programa y mantenimiento de las bases de datos.

## ***RECOMENDACIONES***

## 6. RECOMENDACIONES

1.- Específicamente para el programa propuesto, pueden pensarse algunas modificaciones, como:

- cuando el diccionario de *taxa* sea lo suficientemente extenso, estimamos posible eliminar los de epítetos específicos e infraespecíficos, y realizar la búsqueda de estos epítetos en la primera base
- pasar a la georreferenciación mediante coordenadas geográficas, para lo cual ya existe un trabajo previo de captura de éstas realizado en el herbario HAC
- continuar los ensayos para lograr una codificación de las especies, siempre que la misma la realice el software automáticamente y de forma tal que coincidan las codificaciones que se asignen en instituciones diferentes

2.- Las bases botánicas en desarrollo deben seguir los patrones aquí establecidos, al menos en los campos taxonómicos, para que así puedan beneficiarse de los programas que se desarrollen paralelamente basados en los ficheros auxiliares.

3.- A pesar de las particularidades que las diferencian de las plantas superiores, las colecciones de hongos y helechos pueden tener una estructura similar a las aquí mostradas en campos sistemáticos, curatoriales y ecológicos, aunque varían en aspectos vegetativos y fenológicos, de incluirse éstos. Esta comunidad de campos posibilitaría la utilización en las diferentes colecciones de los programas que se desarrollen.

4.- En las colecciones zoológicas, pudieran aprovecharse las facilidades de revisión y actualización de *taxa*, y las estrategias de ayuda desarrolladas, aunque prácticamente los puntos de interrelación serían en los campos ecológicos, de hábitat, datos de colecta y distribución; pues las diferencias en la formas de registro y manejo de ambos tipos de colecciones, impiden la existencia de un patrón común de campos curatoriales para las mismas. Las variaciones que existen en cuanto a las reglas de nomenclatura que se aplican a plantas y animales, también inciden en la conformación de estrategias diferentes en el establecimiento de los campos sistemáticos. Igualmente, en los campos morfológicos, además de las diferencias naturales entre plantas y animales, también existen caracteres excluyentes entre las propias clases y órdenes zoológicos.

5.- Cualquier propuesta para un proyecto de bases de datos debe demostrar un compromiso y factibilidad de la continuación y mantenimiento a largo plazo de los datos, en independencia de la futura evolución del software o del hardware.

6.- Cualquier proyecto para desarrollo de nuevas bases de datos debe demostrar:

- a) que el software existente ha sido considerado inadecuado
- b) la competencia técnica y experiencia del equipo de investigación para el apropiado diseño y desarrollo del programa

7.- Si se ha considerado necesario el desarrollo de un nuevo programa, entonces:

- a) Los fondos de financiamiento del diseño, implementación y prueba del programa deben separarse y pedirse de antemano a la captura, manejo y diseminación de los datos. Eso evitará la pérdida de recursos botánicos en proyectos que no puedan llevar a cabo los



objetivos para el desarrollo del programa establecido. Una vez que el programa esté implementado, entonces los fondos pueden liberarse para otros recursos necesarios para construir la base de datos.

- b) para la fase de diseño del programa debe asignarse suficiente tiempo; el 80% de los errores del sistema resultan de un diseño inadecuado.
- c) Los recursos deben ser asignados para incluir las experiencias de las ciencias tanto de computación como botánicas/zoológicas en el proceso de diseño del programa.

## ***REFERENCIAS***

## **BIBLIOGRAFÍA CITADA**

1. Alain, Hno. (1964): **Flora de Cuba. Tomo V.** Assoc. Est. Ccias. Biológicas. 362 pp.
2. Alain, Hno. (1974): **Flora de Cuba. Suplemento.** Inst.Cub.Libro. 150 pp.
3. Allkin, Robert (1991): **Informe final. Taller: Aplicación de bases de datos en Botánica.** Concepción, Chile. 5 pp.
4. Berendsohn, W.G. (1992): **CASE: Techniques in Botanical Database Design.** 8va Reunión de la "Taxonomic Databases Working Group" y 2da Reunión Anual de "International Organization for Plant Information", México.
5. Caballero, E.; H. Ferrás y F. Cejas (1993): Programa FITOMAP para la mapificación de especies botánicas cubanas. **Resúmenes IV Simposio de Botánica,** La Habana.
6. Capote, R.P. y R. Berazaín (1984): Clasificación de las formaciones vegetales de Cuba. **Rev. Jard. Bot. Nac.,** 5(2)1-49.
7. Colectivo de autores (1995): Estudio Nacional de la Biodiversidad de la República de Cuba
8. Fleitas, I.E.; E. Orozco e I. Baró (1985): Propuesta para la generación de las bases de datos del herbario del Instituto de Botánica de la Academia de Ciencias de Cuba. **Resúmenes 1er Congreso de Botánica,** Tomo II. La Habana. 660-664.
9. Font Quer, P (1975): **Diccionario de Botánica.** VI Ed. Labor S.A. España. 1046-1047.
10. Holgrem, P. (1992): **Index Herbariorum.** New York Bot. Garden.
11. León, Hno. (1946): Flora de Cuba. Tomo I. **Contrib. Ocas. Museo Hist. Nat. La Salle,** Habana. 441 pp
12. León, Hno. y Hno. Alain (1951): Flora de Cuba. Tomo II. **Contrib. Ocas. Museo Hist. Nat. La Salle,** Habana. 456 pp.

13. León, Hno. y Hno. Alain (1953): Flora de Cuba. Tomo III. **Contrib. Ocas. Museo Hist. Nat. La Salle**, Habana. 502 pp.
14. León, Hno. y Hno. Alain (1957): Flora de Cuba. Tomo IV. **Contrib. Ocas. Museo Hist. Nat. La Salle**, Habana. 556 pp.
15. Palazuelos, R. ; E. Orozco; T. Tosco y R. Nieto (1985): Procesamiento automatizado del listado de plantas existentes en el Jardín Botánico de Cienfuegos. **Resúmenes 1er Congreso de Botánica**, Tomo II. La Habana. 651-659.
16. Peláez, A. y M.J. Ordoñez (1991): Perspectivas actuales para la conformación de bancos de información biológica. **Simposio sobre bases de datos botánicas. X Congreso Mexicano de Botánica.**
17. Ricardo, N. E., E. Pouyú y P. Rojas (1995): The sinantropic flora of Cuba. **Fontqueria**, 42:367-429.
18. Samek, V. (1973): Regiones fitogeográficas de Cuba. **Ser. For. Acad. Ccias. Cuba**, 15:63 pp.
19. Tsu-der, G. (??): **dBASE III. Guía del programador**. 423 pp.
20. Weber, W.A. (1982): Mnemonic three letter acronyms for the families of vascular plants: a device for more effective herbarium curation. **Taxon**, 31(1):74-88

## **BIBLIOGRAFIA CONSULTADA**

21. Allkin, Robert (1990): **ALICE: A Bio-diversity Database System**. ALICE software partnership. 3 pp.
22. Beach, J.H., S. Pramanik y J.H. Beaman (1990): Hierarchic structures for taxonomic database systems. **Resúmenes Taller Inteligencia Artificial y Biología Sistemática**, Napa Valley, California. 14 pp.
23. Bisby, F.A. (1994): **Plant Names in Botanical Databases**. TDWG, 30 pp.
24. Cook, F. Y L.H. Hastings (1992): **Economic Botany Data Collection Standards**. TDWG, 43 pp.
25. Esquivel, M.A (1992): **Estudio de los recursos fitogenéticos y su documentación mediante un sistema de información computadorizado: el género *Arachis* L. en Cuba**. Tesis en opción al grado de Doctor en Ciencias Agrícolas. Ciudad de la Habana.
26. Esquivel, M.A.; D. Jiménez; A. González y H. Knupffer (1990): **Sistema de Documentación Integrado de Recursos Fitogenéticos, Versión 1.0**. Manual del Usuario. INIFAT, Ciudad Habana, Cuba.
27. Guánchez, F. (1993): Red de Información Botánica de Venezuela. **Resúmenes III Jornada Científico Técnica COMPACC'94**.
28. Hollis, S. Y R. K. Brummitt (1992): **World Geographical Scheme for recording plant distributions**. TDWG, 104 pp.
29. IOPI (1994): **Global Plant Checklist Project Plant** (K. Wilson, Eds).

30. Leon, C., D. Mackinder y, P. Rooney y H. Synge (1989): **Plant occurrence and status scheme**. World Conservation Monitoring Centre. 14 pp.
31. Pankhurst, R. J. (1992): **TDWG descriptors subgroup**. TDWG, 10 pp.
32. Rosete, S. Y F. Cejas (1994): Propuesta de base de datos y programa para el manejo de información sobre plantas útiles. **Resúmenes Jornada de Medicina Popular "Juan Tomás Roig"**, Pinar del Río.
33. Taxonomic Databases Working Group (1992): **Proposed New Constitution**. 5 pp.

## **BIBLIOGRAFIA DEL AUTOR REFERIDA AL TEMA**

34. Cejas, F. (1990): Proyecto de base de datos y programa para las colecciones de los herbarios cubanos. **Resúmenes V Congreso Latinoamericano de Botánica. Ciudad Habana, Cuba.**
35. ----- (1991): Necesidad de la automatización en el control de datos de los herbarios. **Resúmenes IV Encuentro de Botánica "Bisse in Memoriam".** Camagüey.
36. ----- (1991): Taxonomic database related with the Flora of Cuba. En: **Origin, evolution and diversity of Cuban plant genetic resources.** 23-27.
37. ----- (1992): Programa y base de datos y programa para las colecciones de los herbarios cubanos. **Ciencias Biológicas** (24):147-151.
38. ----- (1992): Programa HERBARIO para automatizar las colecciones de plantas cubanas. **Resúmenes VII Forum de Piezas de Repuesto.** La Habana.
39. ----- (1992): Automatización en bases de datos botánicas cubanas. Actualidad y perspectivas. **Resúmenes 8va Reunión de la "Taxonomic Databases Working Group" y 2da Reunión Anual de "International Organization for Plant Information",** México.
40. ----- (1992): Bases de datos y sistemas computarizados relacionados con la conservación y utilización de los recursos fitogenéticos en Cuba. **Evento "Fitorecursos de Cuba y Sus Usos: 500 Años Despues",** La Habana.
41. ----- (1992): Programa HERBARIO. Metodología y posibilidades. **Resúmenes II Jornada Científico-Técnica, CEDISAC,** La Habana.
42. ----- (1993): Programa HERBARIO, versión 1.02. **Resúmenes IV Simposio de Botánica,** La Habana.



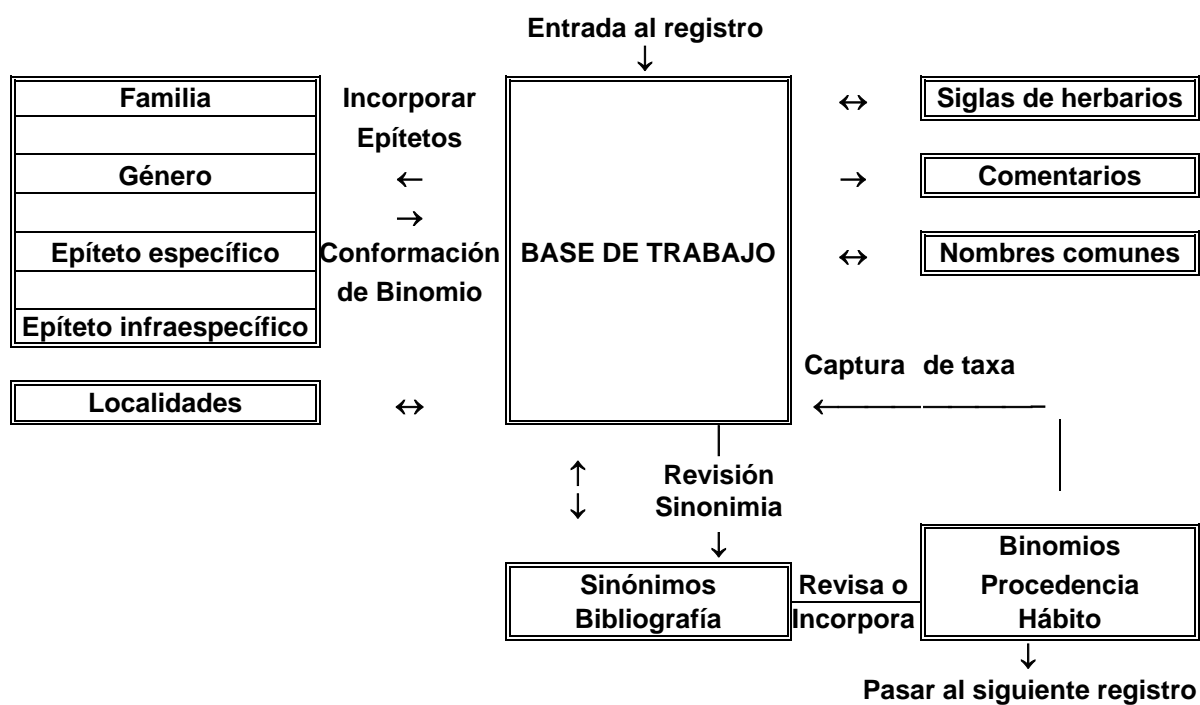


***ANEXOS***

Anexo 1. Cantidad de registros contenidos en los ficheros de referencia hasta agosto de 1996.

◆ Epítetos Familia	776
◆ Epítetos Géneros	16117
- Señalados como Sinónimos	2168
◆ Epítetos Específicos	4372
◆ Epítetos Infraespecíficos	904
◆ Taxa con Nombres Comunes	2229
- Nombres Comunes	3277
◆ Siglas de Colecciones Botánicas	31
◆ Localidades de Colecta	2557
- Georreferenciadas Lambert	1664
- Georreferenciadas Geográficas	1142
- Ubicación Fitogeográfica Samek	2299
◆ Sinónimos / Bibliografía	4740
- Basónimos	1064
- Sinónimos	2641
- Bibliografía Original	1035
◆ Taxa (Binomios)	9479
- Chequeadas	7587

Anexo 2. Utilización de los ficheros de referencia como ayuda en la introducción y/o revisión de la información. El sentido de las flechas indica captura (hacia la base de trabajo) o incorporación (hacia los ficheros de referencias) de datos.



Anexo 3. Códigos establecidos para los menús de descriptores.

<p><b>A- Tipo nomenclatural:</b>  H- Holotipo  I - Isotipo  C- Cotipo  P- Paratipo  T- Topotipo  S- Sintipo  L- Lectotipo  F- Fototipo  X- Ninguno</p>	<p><b>D- Distrito Fitogeográfico:</b>  01 Península de Guanahacabibes  02 Arenas Blancas Guane-Sabanalamar  03 Sabana Centro-Meridional  04 Alturas de Pizarras  05 Mogotes  06 Sierra del Rosario  07 Cajalbana  08 Colinas Bahía Honda-Cabañas  09 Anafe  10 Sur Isla de la Juventud  11 Arenas Blancas Indios-Siguanea  12 Centro Isla de la Juventud  13 Costa Norte Habana-Matanzas  14 Colinas Habana-Limonar  15 Llanura Centro-Occidental  16 Motembo  17 Zapata  18 Cayería Sur  19 Escambray  20 Costa Trinidad  21 Serpentinias Santa Clara  22 Serpentinias Camagüey  23 Serpentinias Holguín  24 Cuba Centro-Oriental  25 Costa y Cayos Septentrional  26 Costa Media Luna-Cabo Cruz-Baconao  27 Sierra Maestra  28 Cordillera del Turquino  29 Gran Piedra  30 Valle Central Oriental  31 Sierra de Nipe  32 Sierra de Cristal &lt;Pinares de Mícaro&gt;  33 Serpentinias Moa-Toa-Baracoa  34 Costa Norte Bahía de Nipe-Bahía de Cebollas  35 Baracoa  36 Costa meridional Maisí-Guantánamo  37 Sierra de Imías  38 Colinas Septentrionales del Valle Central  39 Centro de las Montañas Septentrionales &lt;Santa Catalina&gt;  ? Desconocido</p>
<p><b>B- Procedencia:</b>  N- No Endémica  S- Endémica  I - Introducida  X- Exótica  D- Origen Desconocido  ? - Duda</p>	
<p><b>C- Vegetación:</b>  01 Pluvisilva  02 Monte nublado  03 Bosque siempreverde  04 Bosque semideciduo  05 Bosque de ciénaga  06 Bosque galería  07 Manglar  08 Pinar  09 Manigua costera  10 Cuabal  11 Charrascal  12 Monte fresco  13 Semidesértica  14 Salinas  15 Herbazal ciénaga  16 Mogote  17 Costa rocosa  18 Costa arenosa  19 Secundaria  20 Sabana o pastizal  21 Comunidad acuática  22 Vegetación ruderal</p>	

<b>F- Provincia:</b> 01 Pinar del Río 02 La Habana 03 Ciudad de la Habana 04 Matanzas 05 Villa Clara 06 Cienfuegos 07 Sancti Spiritus 08 Ciego de Avila 09 Camagüey 10 Las Tunas 11 Holguín 12 Gramma 13 Santiago de Cuba 14 Guantánamo 15 Isla de la Juventud	<b>G- Sector Fitogeográfico:</b> 1 Occidental 2 Central 3 Oriental ? Desconocido  <b>H- Hábito:</b> 1 Herbáceo 2 Arbusto 3 Árbol <incluidas palmas altas> 4 Trepadora 5 Liana <mayor consistencia> 6 Epífita 7 Parásita 8 Suculenta 9 Roseta 10 Rastrera
---	--