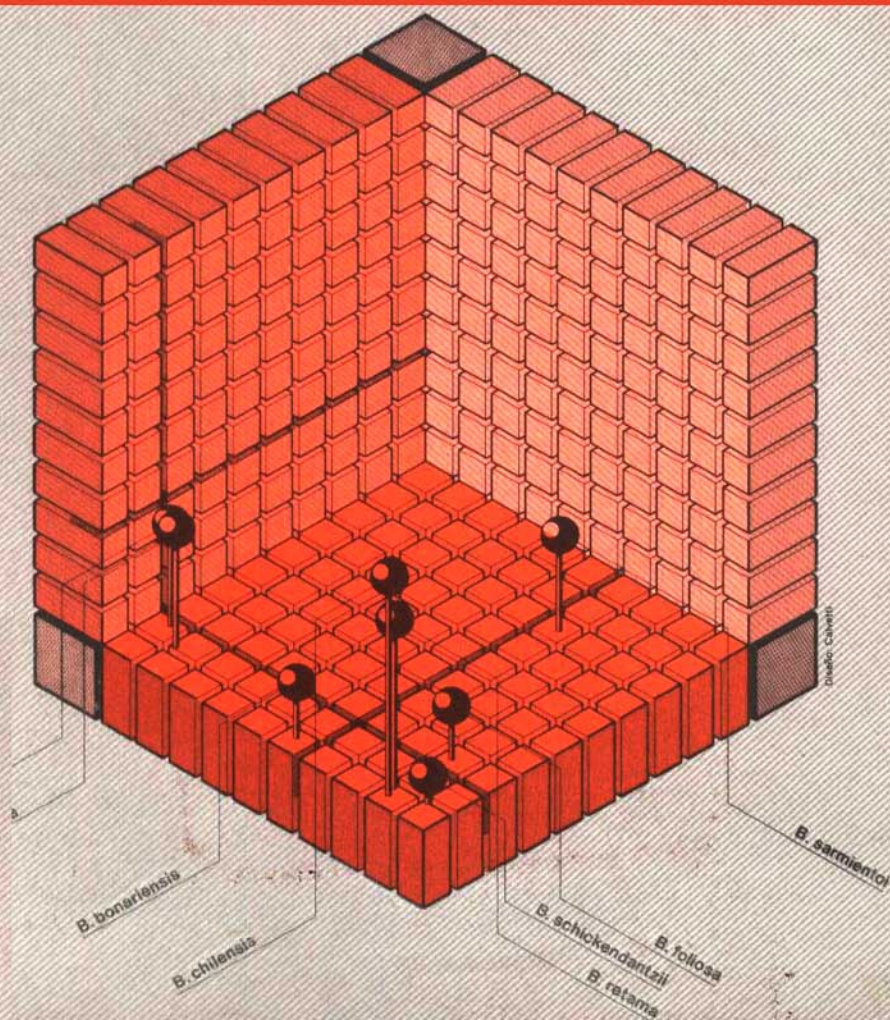


# INTRODUCCION A LA TEORIA Y PRACTICA DE LA TAXONOMIA NUMERICA

Programa Regional de Desarrollo Científico y Tecnológico  
Departamento de Asuntos Científicos  
Secretaría General de la  
Organización de los Estados Americanos





# **INTRODUCCION A LA TEORIA Y PRACTICA DE LA TAXONOMIA NUMERICA**

por

**Jorge Víctor Crisci**

y

**Maria Fernanda López Armengol**  
Cátedra de Introducción a la Taxonomía  
Facultad de Ciencias Naturales  
Universidad Nacional de La Plata  
La Plata, ARGENTINA

**Secretaría General de la  
Organización de los Estados Americanos  
Programa Regional de Desarrollo Científico y Tecnológico  
Washington, D.C. - 1983**

©Copyright 1983 by  
The General Secretariat of the  
Organization of American States  
Washington, D.C.

Derechos Reservados, 1983  
Secretaría General de la  
Organización de los Estados Americanos  
Washington, D.C.

Esta monografía ha sido preparada para su publicación en el Departamento de Asuntos Científicos y Tecnológicos de la Secretaría General de la Organización de los Estados Americanos.

Editora: Eva V. Chesneau

Asesor Técnico: Dr. Otto T. Solbrig  
The Gray Herbarium of Harvard University  
Cambridge, Ma., Estados Unidos



# A los lectores

El programa de monografías científicas es una faceta de la vasta labor de la Organización de los Estados Americanos, a cargo del Departamento de Asuntos Científicos y Tecnológicos de la Secretaría General de dicha Organización, a cuyo financiamiento contribuye en forma importante el Programa Regional de Desarrollo Científico y Tecnológico.

Concebido por los Jefes de Estado Americanos en su Reunión celebrada en Punta del Este, Uruguay, en 1967, y cristalizado en las deliberaciones y mandatos de la Quinta Reunión del Consejo Interamericano Cultural, llevada a cabo en Maracay, Venezuela, en 1968, el Programa Regional de Desarrollo Científico y Tecnológico es la expresión de las aspiraciones preconizadas por los Jefes de Estado Americanos en el sentido de poner la ciencia y la tecnología al servicio de los pueblos latinoamericanos.

Demostrando gran visión, dichos dignatarios reconocieron que la ciencia y la tecnología están transformando la estructura económica y social de muchas naciones y que, en esta hora, por ser instrumento indispensable de progreso en América Latina, necesitan un impulso sin precedentes.

III

El Programa Regional de Desarrollo Científico y Tecnológico es un complemento de los esfuerzos nacionales de los países latinoamericanos y se orienta hacia la adopción de medidas que permitan el fomento de la investigación, la enseñanza y la difusión de la ciencia y la tecnología; la formación y perfeccionamiento de personal científico; el intercambio de informaciones, y la transferencia y adaptación a los países latinoamericanos del conocimiento y las tecnologías generadas en otras regiones.

En el cumplimiento de estas premisas fundamentales, el programa de monografías representa una contribución directa a la enseñanza de las ciencias en niveles educativos que abarcan importantísimos sectores de la población y, al mismo tiempo, propugna la difusión del saber científico.

La colección de monografías científicas consta de cuatro series, en español y portugués, sobre temas de física, química, biología y matemática. Desde sus comienzos, estas obras se destinaron a profesores y alumnos de ciencias de los primeros años de la universidad; de éstos se tiene testimonio de su buena acogida.

Este prefacio brinda al Programa Regional de Desarrollo Científico y Tecnológico de la Secretaría General de la Organización de los Estados Americanos la ocasión de agradecer al doctor Jorge Víctor Crisci y a la licenciada María Fernanda López Armengol, autores de esta monografía, y a quienes tengan el interés y buena voluntad de contribuir a su divulgación.



# INDICE

	Página
A los Lectores.....	iii
CAPITULO 1. INTRODUCCION.....	1
CAPITULO 2. LA TAXONOMIA NUMERICA COMO TEORIA CLASIFICATORIA.....	3
Definición de Términos.....	3
Naturaleza de la Clasificación.....	4
Objetivos de la Clasificación Biológica.....	5
Fundamentos de la Clasificación Biológica.....	6
El Feneticismo y el Enfoque Empírico y Operativo de la Clasificación Biológica.....	10
Relaciones Taxonómicas.....	11
CAPITULO 3. APLICACION DE LAS TECNICAS DE LA TAXONOMIA NUMERICA.....	15
Pasos Elementales de las Técnicas Numéricas.....	15
Elección de las OTU.....	17
Evidencia Taxonómica: Caracteres.....	18
Homología.....	18
Tipos de Caracteres.....	19
Elección de los Caracteres.....	20
Número de Caracteres Utilizados.....	20
El Problema de la Importancia de Cada Carácter.....	20
CAPITULO 4. LOS CARACTERES COMO DATOS CIENTIFICOS.....	23
Tipos de Datos.....	23
Datos Doble-estado y su Codificación.....	24
Datos Multiestados y su Codificación.....	25
El Problema de la Variación Intra-OTU.....	29
Un Ejemplo: Estudio de Taxonomía Numérica en el Género <i>Bulnesia</i> (Zygophyllaceae).....	30
Matriz Básica de Datos.....	34
CAPITULO 5. ESTIMACION DEL PARECIDO TAXONOMICO: COEFICIENTES.....	39
Coeficientes de Distancia.....	39
Coeficientes de Correlación.....	41
Coeficientes de Asociación.....	42
Escala.....	45
Elección del Coeficiente de Similitud.....	46
Matriz de Similitud.....	46
CAPITULO 6. LA BUSQUEDA DE LA ESTRUCTURA TAXONOMICA: ANALISIS DE AGRUPAMIENTOS.....	51
Análisis de Agrupamientos.....	51
Representación Gráfica de las Técnicas de Análisis de Agrupamientos.....	54
Construcción de Fenogramas.....	54
Medición de la Distorsión.....	64
Los Caracteres como OTU: Técnica R.....	67
Interpretación de los Fenogramas.....	67

CAPITULO 7. LA BUSQUEDA DE LA ESTRUCTURA TAXONOMICA:	
MÉTODOS DE ORDENACION.....	71
Análisis de Componentes Principales.....	71
Representación Gráfica del Análisis de	
Componentes Principales.....	74
Interpretación de los Resultados del Análisis	
de Componentes Principales.....	80
Relación entre las Técnicas y Criterios	
para Seleccionarlas.....	81
CAPITULO 8. LA COMPARACION DE CLASIFICACIONES:	
CONGRUENCIA TAXONOMICA.....	83
Tipos de Experimentos.....	84
Evaluación de la Congruencia Taxonómica.....	86
Algunos Resultados Experimentales.....	87
Explicación de los Resultados.....	88
Connotaciones Sistemáticas.....	91
CAPITULO 9. LA RECONSTRUCCION DE LA FILOGENIA	
POR METODOS NUMERICOS.....	93
Componentes de la Filogenia.....	93
Pasos Elementales en la Reconstrucción de la	
Filogenia y Determinación de la Polaridad.....	94
Construcción del Arbol Mediante la Técnica de Wagner.....	96
Un Ejemplo Hipotético.....	98
Retículos.....	103
CAPITULO 10. TAXONOMIA NUMERICA Y COMPUTADORAS.....	107
Una Rápida Mirada a las Computadoras.....	107
Comunicación del Hombre con la Computadora:	
Programas y Lenguaje.....	108
Programas de Computadoras para Aplicar la	
Taxonomía Numérica.....	109
CAPITULO 11. VALOR Y LIMITACIONES DE LA TAXONOMIA NUMERICA..	111
Una Evaluación del Feneticismo.....	111
Una Evaluación de las Técnicas Numéricas.....	112
CAPITULO 12. EPILOGO: UNA BUSQUEDA INCESANTE.....	115
AGRADECIMIENTOS.....	117
BIBLIOGRAFIA.....	119

## INTRODUCCION

"Aunque esto es locura, hay cierto método en lo que dice".

W. Shakespeare. Hamlet

La ciencia exhibe una acusada y predominante contemporaneidad. Una proporción muy amplia de los acontecimientos científicos de cualquier clase correspondientes a todas las épocas está ocurriendo ahora ante nuestros ojos (Price, 1973). Ejemplo patente de este peculiar fenómeno es la sistemática biológica, disciplina con más de dos milenios de antigüedad, pero que en los últimos cincuenta años ha experimentado numerosos cambios, reorganizaciones y avances notables (Anónimo, 1974).

Esta fecunda contemporaneidad de la sistemática biológica se debe a causas múltiples e interactuantes, pero cuatro de ellas merecen destacarse:

1. el reexamen crítico de las bases filosóficas de la sistemática, incluyendo conceptos, procedimientos y reglas;

2. la tendencia entre los sistemáticos a abandonar el "descriptivismo radical", que imperó durante milenios en la disciplina. A la pregunta ¿cómo?, se ha agregado ahora ¿por qué?;

3. la utilización en sistemática de técnicas moleculares precisas;

4. el creciente progreso tecnológico que ha puesto al alcance del sistemático las computadoras digitales de enorme capacidad de memoria.

Dentro de este dinámico contexto histórico surge la taxonomía numérica en la década de 1950 y se establece como un método válido para clasificar a los seres vivos. Fue el resultado de la preocupación entre los sistemáticos por la práctica intuitiva y poco precisa de su disciplina.

La taxonomía numérica ha sido definida como la evaluación numérica de la afinidad o similitud entre unidades taxonómicas y el agrupamiento de estas unidades en taxones, basándose en el estado de sus caracteres (Sokal y Sneath, 1963).

Para ilustrar el enorme desarrollo de la taxonomía numérica mencionaremos el hecho de que entre los años 1957 y 1961 se publicaron 60 trabajos en los que se utilizó este método; solamente en 1975 el número de trabajos sobre el tema superó los 200.

Muchos de los procedimientos de la taxonomía numérica y sus justificaciones teóricas han sido objeto de críticas por parte de algunos taxónomos. Dejando de lado estas controversias se puede decir, sin temor a equivocarse, que la taxonomía numérica ha sido un gran estímulo en la revisión de los principios taxonómicos y de los propósitos de la clasificación.

El enfoque planteado por la taxonomía numérica comprende dos aspectos: uno filosófico, basado en la teoría clasificatoria denominada "feneticismo", y el otro, el de las técnicas numéricas, que son el camino operativo para aplicar dicha teoría.

Esta monografía constituye una introducción a los aspectos teóricos del feneticismo y a la aplicación de las técnicas numéricas. En ella se tendrán en cuenta dos preguntas que, a menudo, se hacen muchos taxónomos: ¿Qué es la taxonomía numérica? y ¿qué puede hacer la taxonomía numérica por mí?

Las técnicas numéricas no se presentan en forma exhaustiva, sino ilustrativa y sin proceder a un riguroso examen matemático, ya que éste trasciende el propósito de introducción al tema de esta monografía.

Este trabajo no es en absoluto original, pues intenta resumir nociones ya consagradas. La importancia relativa atribuida a diferentes actividades, la selección de los ejemplos que se presentarán y el capítulo sobre valor y limitaciones de la taxonomía numérica reflejan, es cierto, tendencias personales.

Esta contribución está dirigida, en principio, a los biólogos y estudiantes de biología. Pero entre los posibles lectores se ha considerado a investigadores de otros campos debido a la difusión de la taxonomía numérica fuera de la disciplina en la cual tuvo su origen.

Con esta monografía se ha querido subsanar la falta de trabajos sobre el tema en idioma castellano y al redactarla nos ha guiado la convicción de que un nuevo enfoque puede iluminar con un nuevo día antiguos problemas.



## LA TAXONOMIA NUMERICA COMO TEORIA CLASIFICATORIA

La clasificación es el agrupamiento de objetos en clases sobre la base de atributos que poseen en común y/o sus relaciones. El término relaciones abarca, por ejemplo, el caso de objetos que se encuentran en un mismo tiempo y/o lugar. Una relación entre el ente A y el ente B puede considerarse como un atributo del par formado por A y B en ese orden. Por lo tanto, "atributos" podría incluir a "relaciones".

Ante la diversidad que lo rodea, el hombre recurre a la clasificación como medio para evitar la confusión e, instintiva o conscientemente, clasifica su mundo circundante. El lenguaje mismo no sería posible sin una clasificación implícita, ya que cada objeto individual sería considerado como tal sin llegar a una generalización; así, por ejemplo, cada mesa debería tener nombre propio, ya que el término "mesa" implica un concepto colectivo inaplicable sin una clasificación previa.

El origen de la ciencia de la clasificación se remonta hasta los antiguos griegos. No obstante, el proceso de clasificar, el reconocimiento de similitudes y el agrupamiento de organismos u objetos en función de esas similitudes, comienza con el hombre primitivo.

En la diversidad que el hombre enfrenta está el mundo viviente y es en éste donde ha tenido lugar un monumental proceso clasificatorio sin paralelo en otras disciplinas. No es entonces casual que este proceso clasificatorio haya sido el contexto donde nació la taxonomía numérica (Crisci, 1978).

En consecuencia, una introducción a la taxonomía numérica, como teoría clasificatoria, debe incluir una visión de la disciplina que le da origen y nutre su mensaje. En otras palabras, para comprender lo que es la taxonomía numérica hay que considerarla dentro del desenvolvimiento de la clasificación biológica y del vertiginoso renacimiento y transformación experimentados en los últimos 30 años por los aspectos teóricos de esta disciplina.

Por tal motivo, esta presentación de la taxonomía numérica como teoría clasificatoria irá acompañada de una concisa y limitada introducción a los aspectos teóricos de la clasificación biológica.

## DEFINICION DE TERMINOS

Ningún vocablo está exento de ambigüedad, de ahí que sea necesario definir con la mayor precisión posible algunos términos importantes de la clasificación biológica.

Ante todo, es preciso distinguir entre clasificación y determinación. Este último término se define como: la ubicación de un objeto no identificado en la clase o grupo al que corresponde conforme a una clasificación construida previamente. Por ejemplo, quien ha coleccionado una "margarita" y, luego, con la ayuda de la bibliografía adecuada la ubica en la división *Angiospermae*, clase *Dicotyledonae*, familia *Compositae*, está determinando, es decir, ubicando una entidad en una clasificación previamente establecida. Algunos matemáticos y filósofos incluyen dentro de clasificación a la determinación; pero en esta mono-

grafía seguiremos la difundida costumbre entre los biólogos de distinguir con precisión ambos términos.

"Taxonomía" es el estudio teórico de la clasificación, incluyendo sus bases, principios, procedimientos y reglas. El motivo de estudio de las clasificaciones son los objetos y organismos a clasificar. El motivo de estudio de la taxonomía son las clasificaciones. La taxonomía es la disciplina que trata de explicar cómo se clasifica y cómo se determina.

El vocablo "sistemática" se emplea para definir el estudio científico de las clases y la diversidad de los organismos y de sus interrelaciones; comprende la clasificación, la taxonomía y la determinación.

El término taxon (plural: taxa o taxones) se aplica a un grupo de organismos considerado como unidad de cualquier rango en un sistema clasificatorio. La familia Acanthaceae, el género *Nassauvia*, la especie *Nassauvia ameghinoi* son ejemplos de taxones. Un taxon es un grupo de organismos; sus componentes tienen existencia real.

Conviene distinguir claramente entre taxon y categoría. Este último vocablo designa un rango o nivel en una clasificación jerárquica; por ejemplo, la categoría género. Como es sabido, en la clasificación biológica se utilizan por convención las categorías de la jerarquía linneana: especie, género, familia, orden, clase, división (en zoología filum), reino, etc.

Grupo monotético es un grupo que ha sido constituido por rígidas y sucesivas divisiones lógicas y en el cual todos sus miembros comparten un conjunto de atributos que son condición necesaria y suficiente para convertirse en integrante del grupo.

Grupo politético es un grupo donde sus miembros comparten un gran número de atributos, pero ninguno de esos atributos es compartido por todos los miembros con la excepción del atributo de pertenecer a ese grupo (véase la siguiente sección de este capítulo). Cada miembro posee algunos atributos de ese conjunto. Por lo tanto, ningún atributo es condición necesaria o suficiente para convertir a un objeto en integrante del grupo.

En general, los taxones son grupos politéticos ya que es muy difícil encontrar características "esenciales".

En sistemática, la idea de grupo monotético se utiliza con frecuencia en la construcción de claves con fines de determinación.

#### NATURALEZA DE LA CLASIFICACION

¿Cuáles son los requisitos lógicos que debe cumplir un agrupamiento de objetos en clases para que sea considerado una clasificación?

Deben cumplirse tres axiomas (Williams y Dale, 1965):

a) En cada clase de más de un objeto (no unitarias) debe existir, para cada objeto de la clase, otro distinto que comparta con él como mínimo un atributo.

b) El ser miembro de una clase no es en sí mismo un atributo.

c) Cada objeto de cualquier clase debe diferir, al menos, en un atributo de cada objeto de cualquier otra clase.

Según el primer axioma, un objeto no puede ser clasificado si no se sabe nada de él.

El segundo axioma tiene dos consecuencias. Primero, excluye cualquier formación de clases, definida solamente por la posesión de un número establecido de objetos; por ejemplo, la división de un conjunto de objetos en clase de a cinco. En segundo lugar, permite a cualquiera de las clases recibir nuevos objetos.

El tercer axioma no sólo dispone que los objetos idénticos no pueden ser distribuidos en clases diferentes, sino que también establece un criterio para las clases unitarias.

Estos axiomas son suficientes para estructurar una clasificación, pero no necesariamente una que sea útil, pues la clasificación de un conjunto de objetos (sean organismos o no) no es única, ya que las posibilidades de clasificar un mismo conjunto son numerosísimas. Por ejemplo, si diez alumnos constituyen el conjunto a clasificar, aplicando el procedimiento más simple que es la división en dos subconjuntos, podemos dividirlos en los que tienen ojos claros y en los que tienen ojos oscuros, podemos dividirlos en los que aprobaron Introducción a la Botánica y los que no lo lograron; podemos también dividirlos en hombres y mujeres, como asimismo construir numerosas clasificaciones más. Cada una de estas clasificaciones será válida y hasta óptima para un propósito en particular e irrelevante para otros. De esto se infiere que dado un conjunto de objetos, las exigencias lógicas de una clasificación son satisfechas por una inmensa variedad de soluciones.

El problema no es clasificar, lo que aparentemente no presenta serias dificultades, sino encontrar razones para elegir una solución entre las muchas que se nos presentan.

5

¿Qué elemento de juicio se utiliza para elegir una clasificación entre varias posibles de un mismo conjunto de objetos? La solución a elegir depende del propósito para el cual va a ser utilizada la clasificación. Por lo tanto, carece de sentido preguntar si una clasificación es "mejor" que otra, si no se especifica el objetivo de la misma.

Cuando el conjunto a clasificar son los organismos vivos, cabe preguntar: ¿Cuál es el interés o propósito especial que puede tener el botánico o zoólogo para preferir una clasificación de los seres vivos entre las varias posibles?

#### OBJETIVOS DE LA CLASIFICACION BIOLOGICA

Se han formulado diversas opiniones respecto a los objetivos de la clasificación biológica (por ejemplo, Warburton, 1967; Williams, 1967; Mayr, 1969; Inglis, 1970; Heywood, 1975).

Conforme a una opinión, todavía sostenida por algunos sistemáticos, se atribuye a la clasificación biológica una función de inventario. Esta suerte de "filatelia" es, sin duda, un objetivo muy pobre que dejaría fuera del quehacer de la ciencia a esta disciplina y la colocaría en el simple papel de protociencia descriptiva.

Si, por el contrario, colocamos a la clasificación biológica dentro del ámbito de la ciencia, su propósito primordial sería ampliar el conocimiento acerca de los organismos y la comprensión más profunda de sus propiedades, semejanzas, diferencias e interrelaciones. De esto resulta que la clasificación biológica es una ciencia teórica, con una gran dosis de descripción, pero no una ciencia puramente descriptiva.

Podríamos, entonces, decir que el objetivo de la clasificación biológica es el conocimiento, no de tal o cual organismo en particular, sino de las leyes generales que los rigen y de las relaciones causales existentes entre ellos. Por tanto, una clasificación biológica será "mejor" que otra en la medida que sugiera más leyes científicas y contribuya mejor a la formulación de hipótesis explicativas. Una buena clasificación biológica será fecunda como principio organizador de nuestro conocimiento.

La anterior definición de los objetivos de la clasificación biológica puede "traducirse" a un lenguaje pragmático: la mejor clasificación será la más *estable*, la más *robusta*, la más *predictiva*. Entendiéndose por *estable* que no se modifica en forma drástica por la incorporación de nueva información; por *robusta* que no es alterada radicalmente por la incorporación de nuevas entidades y por *predictiva* que una propiedad conocida para la mayoría de las entidades de un grupo está garantizada con una gran probabilidad de que existe en aquellas entidades de ese grupo todavía no examinadas con miras a buscar esa propiedad.

Un interesante ejemplo de clasificación con objetivos científicos, fuera de la biología, es la clasificación de los elementos químicos en la tabla de Mendeleiev. Quienes tengan algunos conocimientos de química, saben que esta tabla ha demostrado en el curso del tiempo coherencia teórica, capacidad explicativa y predictiva, estabilidad y robustez, es decir, todos los requisitos que debe satisfacer una clasificación con objetivos que se circunscriban en el ámbito de la ciencia.

6

Este ejemplo podría suscitar una controversia ya que, para algunos biólogos, los elementos químicos y los seres vivos no son comparables, pues los elementos carecen de historia. Sin ánimo de iniciar una controversia, consideramos que, a pesar de diferencias sustanciales en el grado de complejidad entre seres vivos y elementos, el ejemplo de la tabla de Mendeleiev es totalmente válido como clasificación científica y como tal comparable con cualquier otra.

Una vez determinado el objetivo, podemos preguntarnos: ¿Cuál es el camino operativo para construir una clasificación biológica que responda a las expectativas y definiciones antes mencionadas? o ¿cuáles son los fundamentos de la clasificación biológica?

#### FUNDAMENTOS DE LA CLASIFICACION BIOLOGICA

Los fundamentos de la clasificación biológica son objeto de una de las más intensas controversias en biología, ya que hay varias corrientes de pensamiento acerca de los fundamentos teóricos a aplicar en dicha clasificación. La abundante bibliografía al respecto da fe de esta estimulante controversia (por ejemplo: Heywood y McNeill, 1964; Mayr, 1965, 1974, 1981; Rollins, 1965; Colless, 1967; Gilmartin, 1967; Williams, 1967; Johnson, 1968; Michener, 1970; Hull, 1970; Ruse, 1973; Schlee, 1975; Sokal, 1975; Duncan y Baum, 1981).

Algunos sistemáticos sostienen que estas disputas sobre los fundamentos de la clasificación biológica son vanas, según ellos, la clasificación no es algo que se deba pensar demasiado, sino sólo acometerla. Huelga casi señalar, que estos mismos sistemáticos, al construir sus clasificaciones, adhieren implícita y hasta puede decirse involuntariamente a alguna de las teorías de la clasificación que pasaremos a examinar.

A pesar de la variedad de opiniones podemos considerar que existen, básicamente, cuatro doctrinas sobre la clasificación: esencialismo, cladismo, evolucionismo y feneticismo (Hull, 1970).

El objetivo de este capítulo es presentar una sola de esas doctrinas --el feneticismo-- y no analizar críticamente o evaluar cada una de ellas. Sin embargo, haremos una breve demarcación de cada una antes de abordar de lleno el feneticismo.

**Esencialismo.** Fue la teoría dominante por muchos siglos, está basada en la lógica aristotélica y responde a los puntos de vista sostenidos por Platón y muchos de sus discípulos.

A pesar de haber sido descalificada por biólogos y filósofos (por ejemplo, Popper, 1950; Simpson, 1961; Hull, 1965; Mayr, 1969), varios sistemáticos sostienen y/o practican algunos de los principios esencialistas (por ejemplo, Thompson, 1952, 1962; Borgmeier, 1957; Blackwelder y Boyden, 1952).

El esencialismo sostiene que es tarea de la ciencia descubrir la "verdadera naturaleza" de los objetos, es decir, su realidad oculta o esencial.

La práctica del esencialismo en la clasificación biológica comprende los siguientes principios:

a) Las esencias ("verdadera naturaleza o forma") de los organismos existen y pueden ser descubiertas y discriminadas con la ayuda de la intuición intelectual.

b) La tarea de la clasificación biológica es descubrir y discriminar las esencias.

c) Las esencias tienen un nombre, pueden ser descritas con palabras y a esa descripción se le denomina "definición".

d) Los seres vivos reflejan una serie básica de tipos y formas inmutables.

e) Todos los organismos miembros de un mismo agrupamiento clasificatorio (taxon) reflejan la misma naturaleza esencial y corresponden al mismo tipo básico.

f) La variación dentro de un mismo taxon es desechable a los fines clasificatorios, por ser el producto de una desviación de los arquetipos básicos.

Por postular la existencia de "tipos básicos", esta escuela se ha denominado también "tipología".

De la lectura de los principios se infiere que para el esencialismo la clasificación no se construye, sino que se descubre.

Los numerosos siglos de esencialismo dejaron su resabio en la práctica de la clasificación biológica y no son pocos los sistemáticos que siguen intuitivamente algunos principios esencialistas (por ejemplo, definición aristotélica de los taxones). La misma nomenclatura taxonómica no escapa a la influencia esencialista.

**Cladismo.** El término "cladismo" proviene del griego y su raíz etimológica significa "rama", denominación que recibe esta escuela por la

importancia otorgada al árbol genealógico en la construcción de la clasificación y a las ramificaciones de dicho árbol. Algunos de los sostenedores de esta escuela la denominan "filogenética" (Wiley, 1981). En esta monografía seguiremos la tradicional costumbre de denominar "cladista" a esta teoría para evitar confusiones con la escuela evolucionista que también reclama para sí el nombre de "filogenética".

Según esta teoría, la clasificación biológica debe basarse en la filogenia ("historia evolutiva") de los organismos, de ahí que se le ha dado también el nombre de "enfoque genealógico".

Los criterios para aplicar esta teoría son los siguientes:

a) Cada grupo o taxon a formar debe ser monofilético (Rensch Hennig, 1968), es decir, todos sus miembros comparten entre sí un antecesor común más reciente que con cualquier otro miembro de algún otro grupo de igual rango.

b) Se eligen caracteres en los cuales pueda determinarse el estado primitivo (estado presente en el más reciente antecesor común del grupo). Esta determinación de las tendencias evolutivas de los caracteres se basa en estudios del grupo y de los grupos afines (véase el capítulo 9).

c) Se establece la secuencia de ramificaciones del árbol genealógico y la posición relativa en el tiempo de esas ramificaciones. Este paso se concreta con la ayuda de los caracteres elegidos y de fósiles cuando los hay.

8

d) Los grupos que se forman se basan en la posesión de estados evolucionados en común, el monofiletismo y la cercanía temporal a un antecesor común. El recurrir sólo a los estados evolucionados de los caracteres tiene por objeto eliminar la distorsión producida por las diferentes velocidades de evolución de distintos caracteres a lo largo de diferentes líneas filéticas.

e) La categoría taxonómica conforme a la jerarquía linneana asignada a cada grupo está en relación directa con su desprendimiento en el tiempo de otra línea evolutiva (ramificaciones). Por ejemplo, un taxon desprendido en el Precámbrico constituye una división, entre Cámbrico y Devónico una clase y así sucesivamente. Cuanto más antiguo, más elevado el rango dentro de la jerarquía. Actualmente, muchos seguidores del cladismo descartan este criterio.

La idea básica es que la clasificación debe expresar las relaciones filogenéticas, teniendo en cuenta las ramificaciones del árbol genealógico que es necesario reconstruir previamente a la clasificación.

Esta teoría se halla expuesta en las obras de Hennig, 1968; Brundin, 1968; Schlee, 1969; Janvier y colaboradores, 1980, y Wiley, 1981.

**Evolucionismo.** Constituye un enfoque ecléctico que combina varios criterios con información genealógica. Esta escuela no pretende que la clasificación exprese la filogenia, sólo que sea consecuente con ella.

Esta teoría se basa en los siguientes criterios:

a) Filogenia.

b) Cantidad de modificaciones evolutivas desde el antecesor común, con el fin de determinar el grado de diversificación y el tamaño de la



discontinuidad entre los posibles grupos a formar. Este criterio contempla no sólo el monofilietismo, sino también el grado en que se han alejado, en similitud o parecido, en el transcurso del tiempo, los posibles grupos a formar. Por ejemplo, dos grupos pueden tener un antecesor común cercano en el tiempo, pero difieren de tal manera que la discontinuidad entre ellos (en cuanto a similitud) es muy pronunciada.

c) Homogeneidad interna (en cuanto a similitud) en cada grupo formado (taxon).

d) Asociación del taxon con un determinado hábitat.

e) El rango adjudicado al taxon debe ser, en lo posible, adecuado a su tamaño. Por ejemplo, no es conveniente tener una familia de un solo género, salvo que la discontinuidad sea lo suficientemente apreciable como para admitirlo.

f) Debe existir una cierta equivalencia entre las categorías adjudicadas a grupos afines. Por ejemplo, el concepto de la categoría "género" debe conservar su uniformidad en una misma familia.

La importancia concedida a cada uno de estos criterios depende de cada autor.

La diferencia entre el cladismo y el evolucionismo radica principalmente en que el primero intenta expresar en la clasificación las ramificaciones del árbol evolutivo, en tanto que el segundo sostiene que la clasificación debe ser consecuente con esas ramificaciones y representar a su vez otros factores de la filogenia, tales como el grado de diversificación y divergencia en cuanto a similitud.

Esta teoría ha sido descrita por Simpson, 1961; Mayr, 1969 y Bock, 1973.

**Feneticismo.** El feneticismo sostiene los siguientes principios:

a) Las clasificaciones deben efectuarse con un gran número de caracteres, que deben ser tomados de todas las partes del cuerpo de los organismos y de todo su ciclo vital.

b) Todos los caracteres utilizados tienen la misma significación e importancia en la formación de los grupos.

c) La similitud total (o global) entre dos entidades es la suma de la similitud en cada uno de los caracteres utilizados en la clasificación.

d) Los grupos de taxones a formar se reconocen por una correlación de caracteres diferentes.

e) La clasificación es una ciencia empírica, en la cual la experiencia sensible desempeña el papel preponderante y, por lo tanto, está libre de inferencias genealógicas.

f) Las clasificaciones deben basarse exclusivamente en la similitud fenética. Se entiende por "fenético" cualquier tipo de carácter utilizable en la clasificación, incluyendo los morfológicos, fisiológicos, ecológicos, etológicos, moleculares, anatómicos, citológicos y otros.

g) El número de taxones establecido en cualquier rango es arbitrario, aunque siempre debe ser coherente con los resultados obtenidos.

Una vez establecido debe continuarse aplicando el mismo criterio de delimitación en todo el grupo estudiado.

Para el feneticismo es imposible llevar a cabo clasificaciones que expresen la filogenia o sean consecuentes con ella, por el desconocimiento de detalles suficientes acerca de la historia evolutiva de la mayoría de los organismos. En muchos grupos de organismos, debido a la falta de fósiles o de otro tipo de información no se conoce la genealogía respectiva o ésta es en alto grado especulativa.

Es muy importante destacar que el feneticismo no cuestiona la Teoría de la Evolución ni la existencia de una genealogía de los organismos; simplemente excluye del proceso clasificatorio la información filogenética. Por otra parte, considera válido el estudio de la filogenia una vez efectuada la clasificación del grupo.

Esta teoría clasificatoria tampoco descarta la posibilidad de que clasificaciones basadas en sus principios reflejen relaciones genealógicas.

Algunos sistemáticos practican una suerte de "feneticismo atemperado" no numérico, de pocos caracteres casi siempre morfológicos.

Aún no se ha resuelto la controversia histórica sobre el origen de esta teoría de la clasificación. Según algunos autores (por ejemplo, Sneath y Sokal, 1973), fue Adanson, botánico del siglo XVIII, quien primero esbozó lo que hoy es el fundamento doctrinal del feneticismo. Otros (por ejemplo, Guédès, 1967) sostienen que las ideas de Adanson fueron mal interpretadas, que de ninguna forma tenían un "tinte" feneticista y que el origen de esta teoría es reciente. Un análisis más profundo, sin embargo, revela que dicha teoría tiene sus raíces en el empirismo y el operacionismo, escuelas epistemológicas que tratan del origen del conocimiento.

## **EL FENETICISMO Y EL ENFOQUE EMPIRICO Y OPERATIVO DE LA CLASIFICACION BIOLOGICA**

El empirismo es una doctrina epistemológica muy antigua, pero su verdadero fundador es John Locke (1632-1704). Locke y sus continuadores sostienen que el conocimiento sólo deriva de la experiencia sensible. El empirismo opone a la tesis del racionalismo (según la cual el pensamiento, la razón, es la verdadera fuente del conocimiento) la antítesis de que la única fuente del conocimiento humano es la experiencia (Hessen, 1938; Romero, 1973). Para el empirismo, cualquiera que sea el origen de una hipótesis científica la experiencia debe comprobarla. En resumen, todos nuestros conceptos, incluso los más generales y abstractos, proceden de la experiencia.

El feneticismo ha llevado esta idea a la clasificación biológica, al desterrar de ella toda especulación filogenética y poner de relieve un proceso clasificatorio empírico, es decir, basado en las decisiones taxonómicas que provienen de la experiencia sensible. El feneticismo señala como ejemplo de concepto taxonómico no empírico el concepto biológico de especie, donde es imposible encontrar un correlato o criterio observable que permita aplicarlo (Sokal y Crovello, 1970; Crisci, 1981).

El gran realce que da el feneticismo al enfoque empírico de la clasificación biológica ha llevado a algunos autores a una posición derivada del empirismo: el denominado "operacionismo" (Sokal y Camín, 1965).

El operacionismo fue fundado en 1927 por el físico P.W. Bridgman en su ya famoso tratado *The Logic of Modern Physics*. En el operacionismo un concepto sólo es válido si pueden describirse las operaciones que conducen a establecerlo. Casi siempre, al determinar estas operaciones se entra en el ámbito de las mediciones.

Un símbolo adquiere interpretación operacional y operativa si se establece una correspondencia entre él y los resultados de una operación real o posible, arbitrada para medirlo (Hull, 1968; Bunge, 1969).

El enfoque operativo de la clasificación biológica exige que cualquier hipótesis clasificatoria sea comprobada por la observación, la experimentación y la medición y que se establezcan criterios para definir categorías taxonómicas y operaciones clasificatorias.

En trabajos taxonómicos es frecuente encontrar frases como la siguiente: "A está más relacionado con B que con C".

Para el operacionismo esta frase no es válida, a menos que vaya acompañada de:

- una clara definición de lo que se entiende por "más relacionado";
- un criterio explícito por el cual puede observarse ese "más relacionado";
- un criterio explícito por el cual ese "más relacionado" puede medirse, y
- una enumeración de los pasos a seguir para completar las operaciones de observación y medición.

11

Lo esencial es explicitar, en la forma más clara posible, todas las operaciones taxonómicas, dándoles sentido empírico y operacional. En otras palabras, se evitan las especulaciones sin base y los conceptos ambiguos.

El feneticismo no rechaza de plano todos los conceptos no operacionales (Sneath y Sokal, 1973), lo que es imposible de hacer, pero deja en claro que los conceptos operacionales deben ser predominantes en el proceso clasificatorio.

Este realce del enfoque empírico en la clasificación biológica ha sido propuesto no sólo por los feneticistas sino también por otros biólogos (por ejemplo, Ehrlich y Holm, 1962; Davis y Heywood, 1963) y forma parte de una tendencia general para reducir la ambigüedad y el papel de la intuición en la sistemática (Crisci, 1978). Por eso, se ha tratado de precisar al máximo los conceptos claves de la clasificación biológica entre los cuales figura el de relación taxonómica, al que se ha prestado considerable atención y al cual dedicaremos nuestra próxima sección.

## RELACIONES TAXONOMICAS

El concepto de relaciones taxonómicas ha sido objeto de diferentes interpretaciones a lo largo de la historia de la clasificación biológica; esta diversidad ha creado confusión. Cualquier intento de aclarar esta confusión exige ante todo un análisis de las relaciones posibles entre los seres vivos, a fin de poder decidir cuál o cuáles de ellas pueden considerarse como parte de las "relaciones taxonómicas".

Las relaciones posibles entre dos organismos o entre dos taxones son las siguientes:

a) **Relaciones fenéticas o de similitud.** Se basan en el parecido entre los organismos o, en otras palabras, en las propiedades observadas en ellos, pero sin considerar el proceso genealógico por el cual aparecieron esas propiedades. Estas relaciones se expresan como proporción de las similitudes y diferencias existentes entre los organismos (Sneath, 1978).

b) **Relaciones de parentesco** (también llamadas genealógicas o filogenéticas). Son aquellas que indican el grado por el cual dos o más organismos están relacionados a un antecesor común. Estas relaciones se expresan fundamentalmente por el grado de cercanía relativa con el antecesor común y el grado de cambio evolutivo que ha ocurrido desde el mismo.

c) **Relaciones cronísticas** (denominadas también temporales). Indican el grado de cercanía en el tiempo de dos o más organismos.

d) **Relaciones espaciales o geográficas.** Denotan el grado de situación espacial relativa entre dos o más organismos.

Esta lista no agota las relaciones posibles, pero incluye las que pueden considerarse de mayor trascendencia.

Sin duda, la teoría de clasificación a adoptar es importante en el momento de definir relaciones taxonómicas. Por ejemplo, para un cladista o un evolucionista las relaciones de parentesco son las valiosas; para un feneticista, relaciones taxonómicas es sinónimo de relaciones fenéticas. Todos ellos utilizan las relaciones espaciales como accesorias: los feneticistas las incluyen en las relaciones fenéticas y los cladistas y los evolucionistas las consideran como elementos de juicio en la reconstrucción filogenética.

Podemos decir que en el ámbito de la paleontología las relaciones cronísticas adquieren vital importancia, sea cual fuere la teoría aplicada.

El centro de gravedad del problema "relaciones taxonómicas" está en la respuesta a la pregunta: ¿Son congruentes las relaciones fenéticas y las de parentesco? La respuesta puede ser afirmativa o negativa, según cada taxon, ya que la fenética puede expresar las relaciones de parentesco o, por el contrario, puede darse el caso de que como consecuencia de fenómenos evolutivos --el paralelismo, la convergencia y las reversiones-- el parentesco y la fenética expresen diferente información (Sokal, 1966). Un ejemplo hipotético de este último caso se presenta en la figura 1, donde las relaciones fenéticas son más estrechas entre C y D que entre D y E, a pesar de que genealógicamente D y E están más íntimamente ligadas que C y D. Esta posibilidad de incongruencia entre fenética y filogenética lleva a concluir que, si bien es posible que las relaciones fenéticas sean equivalentes a las relaciones genealógicas, esto dista mucho de ser una verdad absoluta. Por lo tanto, las relaciones taxonómicas son consideradas en un enfoque estrictamente feneticista como relaciones de similitud o parecido fenético.

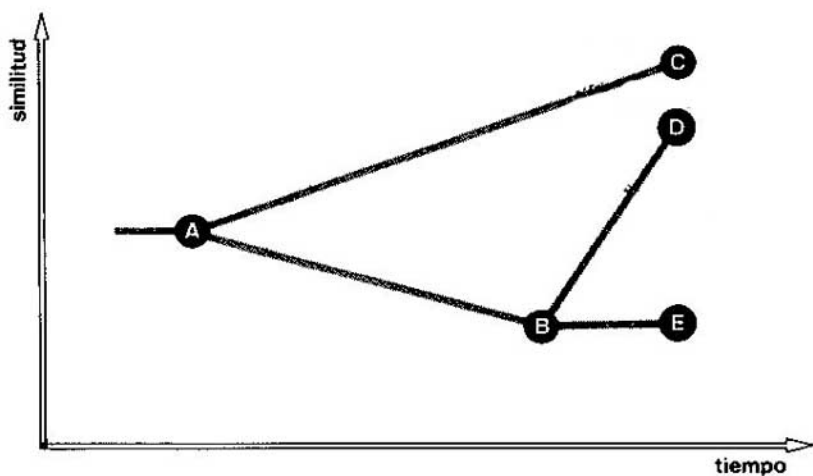


Fig. 1. Representación gráfica de las relaciones entre cinco taxones: A, B, C, D y E. Las relaciones están expresadas en tres factores: tiempo, similitud y parentesco.





## APLICACION DE LAS TECNICAS DE LA TAXONOMIA NUMERICA

Unas pocas décadas atrás la matemática desempeñaba un modesto papel en la biología. Hoy día prácticamente toda disciplina biológica tiene a su alcance técnicas matemáticas para la construcción de sus teorías. Este lenguaje matemático ha llegado a la clasificación biológica gracias a lo que se denomina técnicas numéricas.

Se entiende por técnicas numéricas la rama de la taxonomía numérica que, mediante operaciones matemáticas, calcula la afinidad entre unidades taxonómicas a base del estado de sus caracteres.

El uso de las técnicas no implica necesariamente una aceptación total de los principios del feneticismo, como podrá apreciarse a lo largo de los próximos capítulos y en especial en el capítulo 9.

La asociación de conceptos sistemáticos con variables numéricas ha dado como resultado una inmensa cantidad y variedad de técnicas numéricas. A pesar de esta diversidad, es posible hallar en casi todas ellas una serie de pasos comunes. En este capítulo presentaremos esos pasos elementales y analizaremos en detalle dos de ellos: la elección de las unidades de estudio y la elección de los caracteres.

15

### PASOS ELEMENTALES DE LAS TECNICAS NUMERICAS

Los pasos elementales comunes a casi todas las técnicas numéricas son los siguientes:

1. Elección de las unidades. Se eligen los organismos a estudiar y se definen las unidades a clasificar denominadas "Unidades Taxonómicas Operativas" (OTU).

2. Elección de los caracteres. Se eligen los caracteres que describan a las OTU y se registra el estado de los caracteres presentes en ellas.

3. Construcción de una matriz básica de datos. Con la información obtenida en los pasos anteriores se construye una matriz básica de datos (MBD) de OTU por estados de los caracteres.

4. Obtención de un coeficiente de similitud para cada par posible de OTU. A base de la MBD y utilizando un coeficiente adecuado a los datos que contiene, se calcula la similitud para cada par posible de las unidades taxonómicas.

5. Construcción de una matriz de similitud. Con los valores de similitud calculados en el paso anterior se construye una matriz de similitud OTU por OTU.

6. Conformación de grupos. A base de la matriz de similitud del paso anterior y mediante la aplicación de distintas técnicas (por ejemplo, análisis de agrupamientos) se obtiene la estructura taxonómica del grupo en estudio.

7. Generalizaciones. Se formulan las generalizaciones acerca de los taxones, tales como: elección de caracteres discriminitorios, relación entre los organismos, inferencias acerca de los taxones, etc.

La figura 2 constituye un diagrama de flujo de los pasos elementales de las técnicas numéricas.

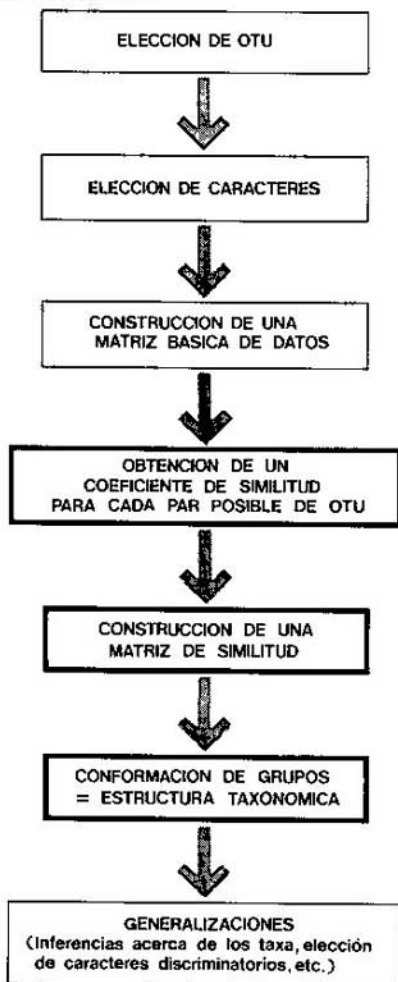


Fig. 2. Diagrama de flujo de los pasos elementales para la aplicación de las técnicas de taxonomía numérica. Los pasos recuadrados con líneas de mayor grosor se realizan, por lo general, con la ayuda de computadoras.

Conviene tener en cuenta los siguientes tres puntos:

- que las generalizaciones acerca de los taxones no pueden formularse antes de que los mismos hayan sido reconocidos al construirse la estructura taxonómica;

- que los taxones no pueden ser reconocidos antes de calcular el valor de similitud entre los organismos, y
- que esa similitud no puede calcularse antes de que los organismos y sus caracteres hayan sido estudiados.

En consecuencia, el orden de los pasos no puede ser alterado sin destruir la racionalidad del proceso clasificatorio.

Dada la gran cantidad de datos que estos estudios requieren, suele ser necesario utilizar computadoras para concretar algunos de los pasos (cálculo de coeficiente de similitud y construcción de la estructura taxonómica).

## ELECCION DE LAS OTU

El primer paso en cualquier proceso clasificatorio consiste en elegir las unidades a clasificar, es decir las OTU. Esta sigla proviene de la denominación en inglés: "Operational Taxonomic Unit" (Sneath y Sokal, 1973). En idioma español se han propuesto otras siglas para estas unidades, por ejemplo UTO, pero dada la enorme difusión de la sigla OTU es preferible no utilizar otro nombre en la ya abarrotada terminología taxonómica.

El taxónomo puede tener ante sí una gran variedad de entidades; individuos, poblaciones, especies, géneros, etc. La existencia objetiva de algunas de estas entidades es motivo de controversia (Mayr, 1969; Sokal y Crovello, 1970; Crisci, 1974, 1981; Raven, 1976; Levin, 1979). Aquí utilizaremos libremente los términos que se corresponden con ellas sin entrar a discutir el aspecto filosófico acerca de su realidad.

La elección de la unidad de estudio entre esas entidades dependerá, en gran medida, de la estrategia y de los objetivos del trabajo taxonómico. Por ejemplo, si se trata de revisar el género *Leucontheria* las unidades de estudio serán las especies. Si la finalidad del estudio es la variación geográfica en *Leucontheria sineticoides*, las unidades serán las poblaciones de esa especie. Claro está que cualquiera que sea la entidad unitaria, siempre está compuesta por individuos y, por lo tanto, éstos deberían ser la unidad universal. Ante la imposibilidad de examinar todos los individuos que componen un taxon, las muestras compuestas por individuos o entidades conceptuales (por ejemplo, una especie) son válidas como OTU. La única regla general en la elección de las OTU es que cada una de ellas debe ser internamente lo más homogénea posible. Excepto en el caso del individuo como unidad, siempre existirá cierta variación interna imposible de eludir (véase el capítulo 4: El problema de la variación intra-OTU).

Los problemas especiales que se plantean con ejemplares de diferentes estadios del ciclo de vida o con material fragmentario exigen, en todos los casos, un cuidadoso análisis preliminar y el conocimiento de la biología del grupo que es objeto de estudio. De no ser así, podría darse la situación de colocar en unidades diferentes a una zoospora y el plasmodio que se originó de ella.

Respecto al número de OTU se aconseja utilizar tantas como se necesiten. En algunos casos, el límite estará dado por el número de taxones en estudio; en otros, por ejemplo cuando la unidad es la población, el límite dependerá del número máximo de OTU que permita utilizar el programa de computadora, generalmente unos pocos cientos.

Una vez elegidas las OTU, el paso siguiente es la búsqueda de la evidencia taxonómica que permita clasificar esas unidades.

## EVIDENCIA TAXONOMICA: CARACTERES

Todo proceso clasificatorio se basa en las diferencias existentes entre los objetos a clasificar. Esa variación es la fuente de la evidencia taxonómica, también llamada caracteres.

El carácter puede definirse como cualquier propiedad que varía en las OTU en estudio. Los posibles valores que ese carácter pueda presentar se los considera sus estados (Sneath y Sokal, 1973).

Por ejemplo, si la especie A con hojas aserradas se distingue de la especie B que posee hojas enteras, el carácter es "margen de la hoja"; "aserrado" y "entero" son los estados de ese carácter.

Cuando se comparan dos organismos en un proceso de clasificación es importante que las estructuras comparadas se correspondan: por ejemplo, que la longitud del pétalo de la OTU A sea comparada con la longitud del pétalo de la OTU B. El reconocimiento de que las dos estructuras son pétalos y de que la variable a medir en las dos OTU es longitud se denomina "determinación de homología", paso necesario y de enorme importancia en cualquier proceso clasificatorio.

## HOMOLOGIA

Por regla general, puede establecerse lo siguiente: todos los estados de un mismo carácter deben ser homólogos. Para aplicar esta regla habrá que saber: ¿Qué es homología? y ¿cuál es el criterio para reconocerla?

La noción de homología es anterior a la biología evolutiva (De Beer, 1971; Bertalanffy, 1975; Cain, 1976), pero a partir de 1859, con la aparición de la teoría darwiniana, este vocablo adquirió un sentido biológico más preciso. Tal es así que las definiciones vigentes se expresan en términos evolutivos. Citaremos algunas definiciones de homología:

a) Dos o más estados de un mismo carácter de dos o más organismos son homólogos si se originan filogenéticamente en el mismo estado del más reciente antecesor común de los organismos comparados (Bock, 1978).

b) Estados homólogos son aquéllos que pueden ser considerados como sucesivos estados de transformación de un mismo estado inicial. Se considera transformación al proceso histórico real de la evolución (Hennig, 1968).

c) Homólogos son los estados de los caracteres de dos o más organismos cuyo origen puede determinarse en el mismo estado del carácter del antecesor común de esos organismos (Mayr, 1969).

d) Dos estados son homólogos si se originan filogenéticamente uno del otro, o de una base hereditaria común (Bertalanffy, 1975).

El problema más serio no es definir que es homología, sino reconocerla. El criterio más frecuente para ello es el de la similitud, que puede ser de dos tipos: homóloga y no homóloga u homoplásica. La similitud homóloga es debida al antecesor común; la homoplásica puede ser producto de paralelismo, convergencia, analogía, mimetismo y el azar (Simpson, 1961). La similitud homóloga es reconocible (pero no siempre) por la correspondencia en la composición y estructura de los estados que intervienen y su parecido profundo, no superficial.

Causalmente, dos estados son parecidos porque son homólogos. Metodológicamente, se puede seguir el camino inverso para determinar la homología: dos estados son homólogos porque son parecidos.

Ello explica porque el feneticismo propuso la llamada homología operativa (Sneath y Sokal, 1973): dos estados son homólogos cuando se corresponden en su composición y en su estructura. Por correspondencia en la composición se entiende la similitud cualitativa, desde el punto de vista biológico y/o químico de sus constituyentes. Por correspondencia estructural se entiende la similitud, en cuanto a orden de sus partes u orden espacio-temporales, o en la estructura de sus fenómenos bioquímicos, o en el orden secuencial de las sustancias o estructuras organizadas.

Como no siempre la homología es fácilmente distinguible en función de la similitud, todavía queda mucho margen para la polémica sobre este tema.

Dejando de lado esta discusión, que excede el ámbito de esta monografía, optamos por el criterio operativo al mismo tiempo que reconocemos sus limitaciones.

#### TIPOS DE CARACTERES

Es imposible contar con una lista exhaustiva y única de los tipos de caracteres de cada grupo de organismos. Sólo los especialistas están en condiciones de definirlos y describirlos.

Una posible clasificación general de los tipos de caracteres podría ser la siguiente:

##### 1. Morfológicos

- a) externos
- b) internos (anatomía)
- c) embriológicos
- d) palinológicos
- e) citológicos
- f) ultraestructurales

##### 2. Fisiológicos

##### 3. Químicos

##### 4. Etológicos

##### 5. Ecológicos

- a) hábitat
- b) parásitos
- c) alimentos
- d) variaciones estacionales

##### 6. Geográficos

- a) distribución
- b) relación entre poblaciones (simpatria, alopatría)

##### 7. Genéticos

Respecto a una lista más detallada de tipos de caracteres en zoología, consúltese a Mayr, 1969; en botánica, a Davis y Heywood, 1963 y Skerman, 1967, y en microbiología, a Lockhart y Liston, 1970 y Sneath, 1978.

## ELECCION DE LOS CARACTERES

En un trabajo de taxonomía numérica es aconsejable elegir todo tipo de carácter y de todas las partes del ciclo vital. Solamente deben excluirse los siguientes caracteres:

a) **Caracteres sin sentido biológico.** Ejemplo de este tipo de carácter es el número dado por el colector al ejemplar de herbario. Lo que se persigue es evitar la utilización de aquellos caracteres que no tengan una relación importante con lo que se observa en los organismos propiamente dichos.

b) **Caracteres correlacionados lógicamente.** Ejemplos de este tipo de carácter son:

Carácter 1: diámetro del tallo.

Carácter 2: radio del tallo.

Se debe excluir como redundante toda propiedad que sea consecuencia lógica de otra propiedad ya utilizada. Hay que evitar la utilización por dos o más vías diferentes de la misma información. Sólo uno de los caracteres correlacionados es aceptado como válido.

c) **Caracteres invariables en las OTU en estudio.** Ejemplo de este tipo de carácter es considerar la "presencia de vértebras" como carácter en un estudio de un grupo de vertebrados.

20

Es preciso excluir los caracteres que no varían en las OTU en estudio, ya que carecen de poder discriminatorio y, por lo tanto, de valor taxonómico. En realidad, por definición no son caracteres, aunque puedan serlo para otros grupos e incluso para un grupo de jerarquía superior que contenga al grupo objeto de estudio.

## NUMERO DE CARACTERES UTILIZADOS

En un estudio sobre taxonomía numérica y a diferencia de los estudios clásicos, deben utilizarse numerosos caracteres. Pero ¿cuál es la cifra de caracteres que satisface este requisito? No existe una respuesta absoluta a este interrogante. En una época se estimó que este número no debía ser inferior a 60 (Sokal y Sneath, 1963), pero esta recomendación carece de bases teóricas o empíricas. El número de caracteres posibles es casi ilimitado y, por lo tanto, desde el punto de vista estadístico no puede sostenerse que el conjunto de todos los caracteres posibles forman una población al azar ("Random population") en la cual puede indicar el tamaño de una muestra representativa. Sneath (1978) señaló que en la práctica los caracteres se comportan como variables "cuasialeatorias" y que, según su experiencia, 50 era el número mínimo a utilizar. Esto es todavía una afirmación intuitiva.

## EL PROBLEMA DE LA IMPORTANCIA DE CADA CARACTER

¿Existe algún carácter que sea más importante que los otros? En caso afirmativo, ¿cómo reconocerlo? La respuesta del feneticismo a estas preguntas es motivo de controversia. Para el feneticismo no existen *a priori* caracteres más importantes que otros y, si así fuera, es imposible reconocerlos.

Sin embargo, no se exige que los complejos de caracteres, tales como los de las flores u hojas, tengan la misma significación o peso. Cada uno tiene un valor diferencial en proporción a su complejidad o



contenido de información que está expresado por el número de caracteres que lo componen (Sneath y Sokal, 1973).

Para entender la posición del feneticismo se debe distinguir significación o peso *a priori* de significación o peso *a posteriori* (Cain, 1959; Heywood, 1968; Mayr, 1969).

La significación *a priori* consiste en otorgar mayor valor taxonómico a uno o a más caracteres antes de confeccionar la clasificación. Esos caracteres se reconocen usando algún criterio establecido. Por ejemplo, porque son estables genéticamente, porque se comportan como caracteres diagnósticos en otros grupos, porque exhiben escasa plasticidad fenotípica, o bien por sospechar que son indicadores de relaciones filogenéticas (Mayr, 1969).

La significación *a posteriori* es resultado de la clasificación y consiste en reconocer el o los caracteres que mejor discriminan o diagnostican los grupos formados. El carácter "presencia de capítulo" es muy importante para reconocer la familia de las Compuestas, pero esa importancia sólo se reconoce una vez construida una clasificación que contenga al taxon Compuestas, no antes.

Generalmente, por la falta de evidencia al aplicar los criterios establecidos para reconocer la importancia *a priori*, ésta debe ser desechada, no así la importancia *a posteriori* que resulta del proceso clasificatorio y que es clave para el proceso de determinar.

El feneticismo reivindica la significación *a posteriori* y rechaza la significación *a priori*, dándole a todos los caracteres el mismo peso o importancia al comenzar una clasificación.



## LOS CARACTERES COMO DATOS CIENTÍFICOS

Los caracteres taxonómicos forman parte del universo denominado "datos científicos" y responden a las exigencias de éste. El científico observa hechos y los registra en datos. Los hechos suceden o subsisten, son eventos y/o estados. Los datos son representaciones simbólicas de los eventos y/o estados y se obtienen por la observación (Kneller, 1978).

Una observación científica debe ser *sistemática, detallada y variada*. Es *sistemática*, pues debe ser controlada por una hipótesis o por una idea precisa del fenómeno estudiado. Es *detaillada* por el uso de instrumentos poderosos y/o por concentrarse en una propiedad particular del fenómeno estudiado. Es *variada*, ya que el fenómeno es captado bajo diferentes condiciones o en forma experimental cuando se añade a la observación el control de ciertos factores.

Los datos obtenidos por la observación deben ser *objetivos y precisos*. *Objetivos* (a pesar de que no existe dato que no esté coloreado por nuestros preconceptos) en el sentido de que cualquier otro científico, capacitado para la observación y que lleve a cabo las mismas operaciones, logre reconocer los mismos hechos que fueron registrados y, por lo tanto, obtenga los mismos datos. Con este fin, los datos son expresados en un lenguaje de validez universal, más que en función de sensaciones únicas del observador. Los datos son *precisos* cuando describen los hechos y los diferencian, en el mayor grado posible, de hechos similares.

23

Los datos más objetivos y precisos son los expresados en forma cuantitativa. La taxonomía numérica exige que todos los datos sean expresados en forma cuantitativa, de modo que sean computables; es decir, que con ellos se puedan realizar operaciones de cálculos mediante números.

No todos los datos miden relaciones cuantitativas en sentido estricto, de ahí que algunos deban ser sometidos a una actividad lógica, la codificación, para ser transformados en datos cuantitativos.

## TIPOS DE DATOS

Existen distintas categorías de datos y se han propuesto numerosas clasificaciones para expresar esa variabilidad (Walker, 1968; Bunge, 1969; Cohen y Nagel, 1971; Sneath y Sokal, 1973; Clifford y Stephenson, 1975).

La clasificación propuesta en esta monografía puede observarse en el Cuadro I. A continuación, y basándonos en esa clasificación, examinaremos los distintos tipos de datos y su posible codificación.

Cuadro I. Tipos de Datos y Algunos Ejemplos de Cada Uno de Ellos

TIPOS DE DATOS		EJEMPLO	
		CARACTER	ESTADOS
DOBLE-ESTADO	Presencia / ausencia	Bandas de color en la conchilla	Presencia Ausencia
	Estados excluyentes	Posición de la inflorescencia	Terminal Axilar

MULTIESTADO	Cualitativos	Sin secuencia lógica	Margen de la hoja	Aserrado
				Lobulado
				Entero
				Gleba
	Cuantitativos	Con secuencia lógica	Pubescencia de la hoja	Poco abundante
				Muy abundante
		Continuos	Longitud del abdomen	—
		Discontinuos	Número de flores de la inflorescencia	—

#### DATOS DOBLE-ESTADO Y SU CODIFICACION

Son los llamados datos binarios o predicados dicotómicos, es decir aquéllos que tienen sólo dos estados. Pueden indicar presencia/ausencia y estados excluyentes.

a) **Datos doble-estado, presencia/ausencia.** Son también denominados datos "todo" o "nada" y representan características que están o no presentes; por ejemplo:

Carácter	Estados
Rayas en el abdomen	Presencia de rayas
	Ausencia de rayas

24

**Codificación de los datos doble-estado presencia/ausencia.** Se expresan numéricamente como 1 (presencia) y 0 (ausencia). El número aquí desempeña una función de rótulo o marca de identificación que facilita el tratamiento cuantitativo, pero bien podrían expresarse como + (presencia) y - (ausencia), o cualquier otra forma convencional.

Al codificar el anterior ejemplo de carácter doble-estado presencia/ausencia resulta:

Carácter	Estados	Codificación
Rayas en el abdomen	Presencia de rayas	1
	Ausencia de rayas	0

b) **Datos doble-estado, estados excluyentes.** Representan caracteres cualitativos que tienen dos estados solamente; por ejemplo:

Carácter	Estados
Tipo de fruto	dehiscente
	indehiscente

Codificación de los datos doble-estado, estados excluyentes. En este caso, puede expresarse como 1 y 0, otorgándole 0 a cualquiera de los estados, o utilizando como expresión numérica el 1 y el 2 (esto permite distinguir este tipo de datos de los anteriores). El número aquí es también utilizado en forma convencional y a manera de rótulo. Al codificar el ejemplo anterior de carácter doble-estado, estados excluyentes, resulta:

Carácter	Estados	Codificación	
Tipo de fruto	dehiscente	Alternativa 1	Alternativa 2
		1	2
	indehiscente	0	1

#### DATOS MULTIESTADOS Y SU CODIFICACION

Son aquellos datos que poseen tres o más estados y pueden ser de dos tipos: cualitativos y cuantitativos.

##### Datos Multiestados Cualitativos

Expresan cualidades no mensurables y se dividen en cualitativos sin secuencia lógica y cualitativos con secuencia lógica.

a) Datos multiestados cualitativos sin secuencia lógica o desordenados. Son datos cualitativos que no pueden ser ordenados en una secuencia de grados del atributo. Por ejemplo, el carácter "tipo de superficie pilosa de la hoja" con los siguientes estados: escabrosa, estrigosa, hispida, hirsuta, serícea, estrellada.

Codificación de los datos multiestados cualitativos sin secuencia lógica o desordenados. Es el tipo de dato más difícil de codificar, ya que por no presentar una secuencia o grados es imposible representarlo bien con números. En el ejemplo dado podemos atribuir 1-2-3-4-5 y 6 a los estados nombrados en el orden dado. Los números cumplen aquí funciones convencionales y de rótulo. Pero ¿por qué otorgar 6 al estado estrellado y no 2? Es decir, carecemos de fundamentos para establecer un orden. En el caso de los datos doble-estado este inconveniente no existe, ya que son posibles dos secuencias, equivalentes entre sí.

El problema podría resolverse transformando a cada uno de los estados en datos presencia-ausencia; por ejemplo:

Carácter	Estados	Cod.
Superficie de la hoja escabrosa	Presente	1
	Ausente	0

Carácter	Estados	Cod.
Superficie de la hoja estrigosa	Presente	1
	Ausente	0

Carácter	Estados	Cod.
Superficie de la hoja hispida	Presente	1
	Ausente	0

Carácter	Estados	Cod.
Superficie de la hoja hirsuta	Presente	1
	Ausente	0

Carácter	Estados	Cod.
Superficie de hoja sericea	Presente	1
	Ausente	0

Carácter	Estados	Cod.
Superficie de la hoja estrellada	Presente	1
	Ausente	0

Esta solución tiene el inconveniente de que, bajo determinadas circunstancias (por ejemplo, cuando se utiliza el coeficiente de asociación "Simple Matching", véase el capítulo 5: Coeficientes de asociación), confiere "peso" al carácter original (en este caso "tipo de superficie de la hoja") que, al transformarse en varios caracteres independientes, aumenta su valor taxonómico en perjuicio de caracteres no codificados de esta forma.

26

b) **Datos multiestados cualitativos con secuencia lógica.** Es el caso de caracteres cualitativos que pueden ser ordenados en una secuencia de magnitud de la cualidad estudiada; por ejemplo:

Carácter	Estados
Presencia de pelos	rara
	común
	abundante

**Codificación de los datos multiestados cualitativos con secuencia lógica.** Si se codifica el ejemplo anterior, observamos que existen distintas posibilidades, de las cuales damos una a continuación:

Carácter	Estados	Codificación
Presencia de pelos	rara	1
	común	2
	abundante	3

Aquí los números desempeñan una función ordinal, ya que indican la posición de una cualidad en una secuencia de grados (otra codificación

válida consistiría en otorgar los números 1, 1,5 y 2 a los estados "rara", "común" y "abundante").

Los valores ordinales correspondientes a cada estado del carácter no pueden someterse a operaciones aritméticas como la adición. Por eso, no es posible afirmar que el estado "abundante" (3) tiene tres veces más pelos que el estado "rara" (1).

A este tipo de carácter se aplica, en algunos casos, el llamado código aditivo (Sneath y Sokal, 1973), terminología no muy feliz ya que, como hemos señalado antes, los números ordinales otorgados no pueden ser objeto de adición. En el código aditivo cada estado es transformado en un carácter presencia-ausencia, pero con la siguiente diferencia: cada vez que el carácter "abundante" está presente se considera que los otros dos estados también lo están, o sea la presencia de "abundante" conlleva la presencia de "común" y de "rara". Cuando el carácter presente es "común", también indica la presencia de "rara". Este sistema tiene el inconveniente de conferir peso al carácter original.

A continuación ofrecemos un ejemplo de la codificación aditiva del carácter "presencia de pelos" en las OTU A, B y C; la OTU A presenta el estado "rara"; la B "común" y la C "abundante":

OTU \ Estados	Rara	Común	Abundante
OTU A	1	0	0
OTU B	1	1	0
OTU C	1	1	1

27

#### Datos Multiestados Cuantitativos

Son los datos cardinales, llamados también magnitudes o simplemente cantidades. Miden relaciones cuantitativas en sentido estricto.

El campo de variabilidad de los datos es un conjunto de números. En este caso los números indican relaciones cuantitativas entre cualidades y pueden ser, por lo tanto, sometidos a operaciones matemáticas (aunque a veces con restricciones, por ejemplo, pueden sumarse los números de estomas de dos hojas, pero no las densidades de estomas de esas dos hojas). En general, estos datos, no necesitan codificación y pueden ser de dos categorías: discontinuos y continuos.

a) **Datos multiestados cuantitativos discontinuos.** Representan cualidades que son expresables sólo por números enteros, sin fracciones, y, por lo tanto, presentan variabilidad discontinua; por ejemplo, el carácter "número de pétalos", donde las expresiones serán siempre números enteros.

b) **Datos multiestados cuantitativos continuos.** Expresan dimensiones continuas, es decir, cualidades cuya variabilidad se distribuye en una escala continua. La expresión de estos datos puede ser un número entero o fraccionario. Ejemplos de este tipo de datos son todas las expresiones de tamaño: "longitud de la hoja", "altura de la planta", "ancho del pétalo", etc.

Finalmente, los datos cuantitativos pueden ser transformados en datos "doble-estado" si los cálculos así lo exigen, como veremos más adelante al tratar los coeficientes de similitud.

Un ejemplo sería el carácter "número de pétalos" con variabilidad de 2 a 10, que puede ser subdividido en intervalos que se transformarán en caracteres binarios:

Carácter	Estados	Codificación
De 2 a 4 pétalos	Presente	1
	Ausente	0

Carácter	Estados	Codificación
De 5 a 7 pétalos	Presente	1
	Ausente	0

Carácter	Estados	Codificación
De 8 a 10 pétalos	Presente	1
	Ausente	0

28

Los datos continuos, como, por ejemplo, el carácter "longitud de la hoja", con variabilidad entre 2 y 10 cm, pueden ser transformados en:

Carácter	Estados	Codificación
Longitud entre 2 y 3,99 cm	Presente	1
	Ausente	0

Carácter	Estados	Codificación
Longitud entre 4 y 5,99 cm	Presente	1
	Ausente	0

Carácter	Estados	Codificación
Longitud entre 6 y 7,99 cm	Presente	1
	Ausente	0



Carácter	Estados	Codificación
longitud entre 8 y 10 cm	Presente	1
	Ausente	0

Esta transformación es a veces necesaria, por ejemplo, para aplicar los coeficientes de similitud del tipo de asociación. Pero tiene el inconveniente del peso que otorga al carácter (en este caso, "longitud de la hoja") y de la relativa arbitrariedad al determinar los intervalos.

#### EL PROBLEMA DE LA VARIACION INTRA-OTU

Cuando las OTU son entidades de nivel superior al de organismo (población, especie, género, etc.) es posible que esos agregados de organismos presenten caracteres que varían no sólo entre las OTU, sino también dentro de ellas. ¿Cuál es el tratamiento que debe recibir esa variación intra-OTU?

El problema admite cuatro soluciones posibles: tres son aplicables a datos multiestado y una a datos doble-estado.

Una primera solución consiste en considerar que la variación intra-OTU puede representarse mediante una medida de posición estadística, por ejemplo la media o la moda.

La media es utilizada para datos multiestados cuantitativos continuos. Sin embargo, no es recomendable para datos multiestados cualitativos con secuencia lógica y cuantitativos discontinuos, pues el valor obtenido podría no corresponderse con un valor real (por ejemplo, una media de 3,5 para el carácter número de pétalos).

La media y la moda no reflejan la dispersión de los valores, por lo que algunos autores (Sneath y Sokal, 1973) complementan esta solución agregando como nuevo "carácter" una medida de dispersión, por ejemplo, la desviación estándar. El carácter original se transforma, entonces, en dos caracteres para cada OTU: la media (o moda) y la desviación estándar.

Una segunda solución al problema de la variación intra-OTU consiste en elegir al azar un organismo de los que componen la OTU y considerar que los estados presentes en ese organismo son los estados de la OTU. Con esta solución (denominada método del ejemplar) se corre el riesgo de elegir un organismo que presente estados atípicos para la OTU en cuestión.

La tercera solución consiste en utilizar coeficientes de similitud que, en los cálculos, tengan en cuenta la variación intra-OTU. Un ejemplo de este tipo de coeficiente es la distancia de Crovello (véase el capítulo 5: Coeficientes de distancia).

Los datos doble-estado (y los multiestados cualitativos sin secuencia lógica, transformados a doble-estado, véase en este capítulo: Datos multiestados y su codificación) exigen una solución diferente a las anteriores. Esta solución se encuentra mediante la codificación, expresando la variación intra-OTU como un estado independiente. En un estudio numérico de la tribu Nassauvinae de la familia de las Compuestas (Crisci, 1974), donde cada OTU representaba un género, se utilizó

el carácter "pubescencia en el receptáculo" que originalmente tiene dos estados posibles: presente (1) y ausente (0). Pero se encontró que algunos géneros incluían, especies con pubescencia en el receptáculo y especies sin pubescencia. La codificación que tuvo en cuenta esa variación fue la siguiente:

Carácter	Estados	Codificación
Pubescencia en el receptáculo	ausente	0
	ausente y presente	1
	presente	2

**UN EJEMPLO: ESTUDIO DE TAXONOMIA NUMERICA EN EL GENERO *Bulnesia* (ZYGOPHYLLACEAE)**

En adelante y en todas las secciones de los tres próximos capítulos y con el objeto de ilustrar las explicaciones teóricas, analizaremos un ejemplo tomado del estudio realizado en el género *Bulnesia* (Crisci y colaboradores, 1979). La estrategia para realizarlo está representada en el diagrama de flujo de la figura 3.

Las técnicas no previstas en dicho estudio serán ilustradas con ejemplos hipotéticos o provenientes de otras fuentes.

La ocho especies del género *Bulnesia* (*B. bonariensis*, *B. foliosa*, *B. chilensis*, *B. retama*, *B. sarmentosa*, *B. schickendantzii*, *B. arborea* y *B. carrago*) constituyeron las ocho OTU.

Los datos consistieron en 43 caracteres registrados para las ocho OTU. El conjunto de caracteres comprendía 19 cualitativos, 19 cuantitativos continuos y cinco cuantitativos discontinuos.

Los valores que se volcaron a la matriz básica de datos correspondían al valor medio en el caso de los caracteres cuantitativos continuos, la moda en relación con el caso de los cuantitativos discontinuos (con excepción del carácter "número de óvulos", en el que se registró la media) y los cualitativos fueron codificados.

En algunos de los caracteres cualitativos, tal como hábito que tienen dos estados --árbol y arbusto-- es posible hallar especies cuyos individuos son todos árboles, especies cuyos individuos son todos arbustos y especies que presentan individuos árboles e individuos arbustos.

Es claro que la mejor codificación para este carácter es la siguiente:

Carácter	Estados	Codificación
Hábito	arbustos	0
	arbustos y árboles	1
	árboles	2

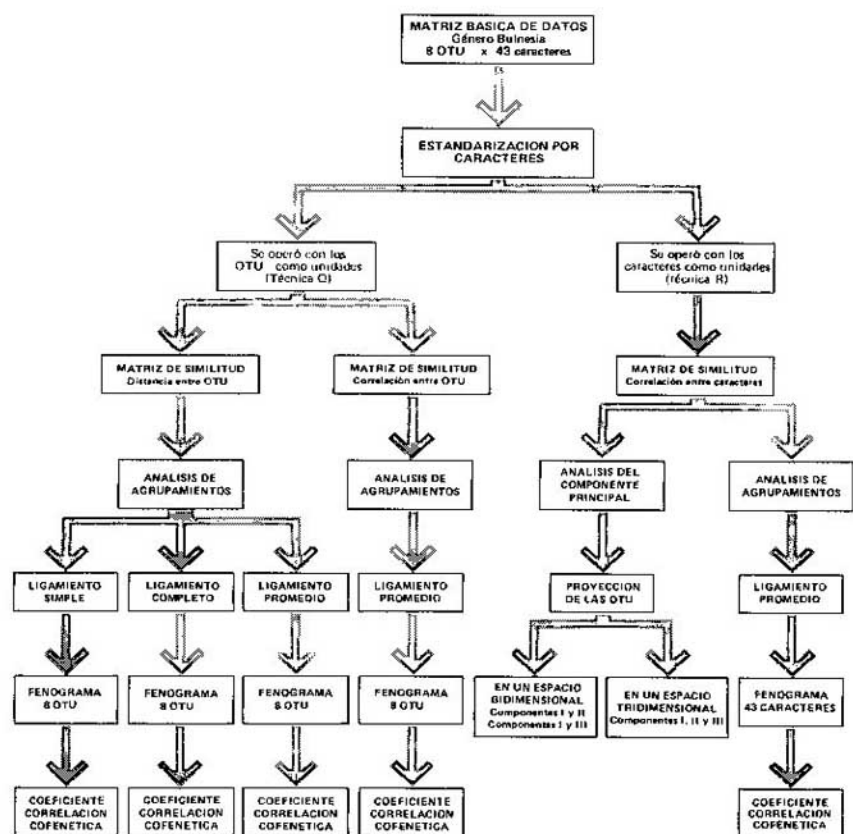


Fig. 3. Diagrama de flujo de la estrategia seguida en el estudio de taxonomía numérica realizado en el género de plantas *Bulnesia* (Zygophyllaceae), que será utilizado como ejemplo en esta monografía.

Los caracteres fueron registrados respecto a cada OTU de la siguiente forma:

Carácter	Estados	Codificación
1. Hábito	arbustos	0
	arbustos y árboles	1
	árboles	2
2. Longitud del internodio (cm)	-	-

3. Diámetro del internodio (cm)	-	-
4. Longitud de la hoja (cm)	-	-
5. Ancho de la hoja (cm)	-	-
6. Longitud del pecíolo (cm)	-	-
7. Número de folíolos	-	-
8. Presencia de peciólulos	folíolos no sésiles	0
	folíolos sésiles o no	1
	folíolos sésiles	2
9. Disposición de los folíolos en el raquis	folíolos alternos	0
	folíolos subopuestos	1
	folíolos opuestos	2
10. Pubescencia	ausente	0
	ausente y presente	1
	presente	2
11. Longitud del folíolo (mm)	-	-
12. Ancho del folíolo (mm)	-	-
13. Número de nervaduras primarias del folíolo	-	-
14. Posición de los folíolos terminales	paralelos	0
	paralelos y divergentes	1
	divergentes	2
15. Presencia de mucrón en folíolos	folíolos no mucronados	0
	folíolos mucronados	1
16. Tipo de inflorescencia	flores solitarias	1
	inflorescencia en dicasio	2
17. Longitud del pedúnculo (mm)	-	-
18. Longitud del sépalo (mm)	-	-
19. Ancho del sépalo (mm)	-	-

20. Color de los pétalos	blanco	1
	amarillo	2
21. Longitud del pétalo (mm)	-	-
22. Ancho del pétalo (mm)	-	-
23. Número de nervaduras del pétalo	-	-
24. Tipo de estambres	heterogéneos	0
	heterogéneos y homogéneos	1
	homogéneos	2
25. Modificación de los estambres	no modificados	0
	modificados y no modificados	1
	modificados	2
26. Presencia de gran escama junto al estambre	ausente	0
	ausente y presente	1
	presente	2
27. Presencia de pelos en la base del filamento estaminal	sin pelos	0
	con o sin pelos	1
	con pelos	2
28. Presencia de una escama suplementaria junto al estambre	ausente	0
	presente	1
29. Agrupación de los estambres	no agrupados	0
	agrupados o no	1
	agrupados	2
30. Longitud del filamento (mm)	-	-
31. Longitud de la antera (mm)	-	-
32. Longitud de la escama (mm)	-	-
33. Presencia de ápice laciniado en la escama estaminal	sin ápice laciniado	0
	con o sin ápice laciniado	1
	con ápice laciniado	2

34. Número de carpelos	en n° de 3	0
	en n° de 3 y 5	1
	en n° de 5	2
35. Curvatura del estilo	estilo no curvado	0
	curvado o no	1
	curvado	2
36. Número de óvulos por carpelo	-	-
37. Pubescencia del fruto	glabro	0
	pubescente	1
38. Longitud del fruto (mm)	-	-
39. Ancho del fruto (mm)	-	-
40. Desarrollo del carpóforo	reducido	1
	bien desarrollado	2
41. Longitud del carpóforo (mm)	-	-
42. Forma de la semilla	semicircular o semi elíptica	1
	oblongo-reniforme	2
43. Longitud de la semilla (mm)	-	-

#### MATRIZ BÁSICA DE DATOS

Los datos obtenidos se presentan en forma de cuadro o tabla denominada Matriz Básica de Datos (MBD). Esta es una matriz  $n \times t$  (Fig. 4) donde las  $n$  columnas representan los caracteres y las  $t$  filas representan las OTU. La alternativa OTU = columnas, caracteres = filas, también es válida. Cada casillero de la matriz  $X_{ij}$  representa el valor del carácter  $j$  en la OTU  $i$ .

Es necesario asimismo tener un símbolo para datos no obtenibles, hecho frecuente en estudios taxonómicos, generalmente debido a la deficiencia de las colecciones. El símbolo NC (no comparable) es el más utilizado. Los "no comparables" son descartados durante el procesamiento de los datos.

La matriz básica de datos puede ser estudiada desde dos puntos de vista (Cattell, 1952): el de la asociación de caracteres, llamada técnica R, y el otro es la práctica inversa, es decir, la asociación de las OTU, llamada técnica Q. En taxonomía numérica se usan ambas técnicas, aunque con mayor frecuencia la técnica Q.

M.B.D.		Caracteres				
		1	2	3	• • •	n
OTU	1	$X_{11}$	$X_{21}$	$X_{31}$	• • •	$X_{n1}$
	2	$X_{12}$	$X_{22}$	$X_{32}$	• • •	$X_{n2}$
	3	$X_{13}$	$X_{23}$	$X_{33}$	• • •	$X_{n3}$
	•	•	•	•	• • •	•
	•	•	•	•	• • •	•
	•	•	•	•	• • •	•
	t	$X_{1t}$	$X_{2t}$	$X_{3t}$	• • •	$X_{nt}$

Fig. 4. Matriz básica de datos, donde  $X_{nt}$  corresponde al valor del carácter n para la OTU t.

La matriz básica de datos de *Bulnesia* es la siguiente:

Caracteres OTU	1	2	3	4	5	6	7	8	9
<i>B. arborea</i>	2	34,9	2,1	84,9	56,6	7,7	13	2	0
<i>B. carmelo</i>	2	35,7	1,6	96,6	70,8	9	7	2	0
<i>B. chilensis</i>	0	23,5	2,6	13,5	8,9	1,8	8	2	1
<i>B. bonariensis</i>	0	20,4	1,3	25,9	17,8	3,4	14	1	0
<i>B. retama</i>	1	40,4	2	12,5	11,3	3,1	5	1	1
<i>B. foliosa</i>	0	18,8	1,3	28,4	25	5,8	4	1	1
<i>B. schickendantzii</i>	0	10,4	1,5	19,7	11,6	2,7	10	0	0
<i>B. sammienti</i>	2	21,6	1,4	21,4	27,1	5,1	2	2	2

Caracteres OTU	10	11	12	13	14	15	16	17	18
<i>B. arborea</i>	2	29,6	8,6	6	0	0	2	17,2	7,1
<i>B. carmelo</i>	2	39,7	16,4	5,8	1	0	2	18,1	6,4
<i>B. chilensis</i>	0	5,2	2,4	NC	2	0	1	9	7,1
<i>B. bonariensis</i>	2	8,9	2	1	2	0	1	11,5	6,8
<i>B. retama</i>	2	6,6	2,6	2	1	1	1	10	7,4

<i>B. foliosa</i>	2	13,6	7,8	3	2	0	1	13,2	5,3
<i>B. schickendantzii</i>	2	5,7	1,9	1	2	0	1	10,4	5,9
<i>B. sammentoi</i>	1	16,8	12	5	2	0	1	3,9	2,9

Caracteres	19	20	21	22	23	24	25	26	27
OTU									
<i>B. arborea</i>	3,4	2	22,4	17,3	16	0	1	2	0
<i>B. carrapo</i>	5,8	2	24,3	18,5	12	0	2	2	0
<i>B. chilensis</i>	4,2	2	9,1	5,7	8	0	0	0	2
<i>B. bonariensis</i>	3,9	2	17,8	10,2	10	0	1	1	0
<i>B. retama</i>	4,4	2	7,7	4,6	7	1	0	0	1
<i>B. foliosa</i>	3	2	8,5	2,7	6	2	0	0	0
<i>B. schickendantzii</i>	3,1	2	9,3	4,3	5	1	0	0	0
<i>B. sammentoi</i>	2,3	1	11,5	7	6	2	0	0	0

36

Caracteres	28	29	30	31	32	33	34	35
OTU								
<i>B. arborea</i>	0	2	10,2	1,5	5,2	2	2	1
<i>B. carrapo</i>	0	2	9,9	1,4	5,3	2	2	2
<i>B. chilensis</i>	0	2	7	2	4,4	1	1	NC
<i>B. bonariensis</i>	0	1	10,6	1,6	4,4	0	2	2
<i>B. retama</i>	0	1	7,1	2,2	3,1	1	2	1
<i>B. foliosa</i>	0	0	6,2	1,6	3,9	1	2	1
<i>B. schickendantzii</i>	1	0	7,4	1,7	4,4	2	2	1
<i>B. sammentoi</i>	0	0	4,1	1,1	2,9	2	0	0

Caracteres	36	37	38	39	40	41	42	43
OTU								
<i>B. arborea</i>	2	0	45,7	40,8	2	7,5	1	13,4
<i>B. carrapo</i>	2	0	56,2	51,8	2	5,3	1	12,1
<i>B. chilensis</i>	7	0	13,4	12,3	2	0,6	2	2,7



<i>B. bonariensis</i>	1	0	36,9	32,6	2	4,8	1	10,8
<i>B. retama</i>	8	0	22,5	18,8	1	0,8	2	11
<i>B. foliosa</i>	4	1	16,3	13,3	1	0,7	2	4,9
<i>B. schickendantzi</i>	4	1	11,8	12,9	1	0,4	2	5,3
<i>B. sarmientoi</i>	2	0	51,8	47,7	2	5,2	1	13,5



## ESTIMACION DEL PARECIDO TAXONOMICO: COEFICIENTES

¿Puede expresarse en forma cuantitativa el parecido entre dos OTU?

El parecido o similitud es cuantificable aplicando un coeficiente de similitud. Con el uso de estos coeficientes en operaciones matemáticas pueden calcularse las similitudes (o su complemento; las diferencias) respecto a cada par posible de OTU de una matriz básica de datos.

Debido al gran auge de la taxonomía numérica se han formulado numerosos coeficientes de similitud. Por la imposibilidad de examinar todos ellos, en este capítulo únicamente nos referimos a aquéllos que han alcanzado mayor difusión. Los coeficientes presentados serán divididos en tres grandes grupos: de distancia, de correlación y de asociación.

## COEFICIENTES DE DISTANCIA

Se aplican sobre matrices básicas que presentan datos doble-estado o multiestados, o en las que poseen ambos tipos de datos (datos mixtos).

Una representación geométrica de los coeficientes de distancia podría ser como sigue: imaginemos cuatro OTU --A, B, C y D-- y tres caracteres para clasificarlos --1, 2 y 3--. Supongamos ahora la existencia de un "espacio fenético" de tres dimensiones, una para cada carácter; por lo tanto, los representaremos como tres ejes de coordenadas (Fig. 5). Cada OTU puede ubicarse en el espacio formado por esos tres ejes, de acuerdo con el valor que tenga en cada uno de los caracteres. La distancia entre las OTU es una posible cuantificación de las diferencias. A mayor distancia, menor similitud.

Con más de tres caracteres estaríamos ante un espacio multidimensional. Pero, no por eso dejaría de tener validez la computación de las distancias en ese espacio, ya que, a pesar de que es imposible representar gráficamente más de tres ejes, está demostrado algebraicamente que la geometría del espacio tridimensional es aplicable a espacios euclidianos de más de tres dimensiones, al menos en los aspectos que interesan aquí.

Veamos algunos coeficientes capaces de medir distancias entre OTU.

El coeficiente de distancia más comúnmente usado es el llamado "Mean Character Difference" (MCD) que fue propuesto como medida taxonómica por Cain y Harrison (1958) y se expresa como la sumatoria del valor absoluto de la diferencia entre cada estado de los caracteres de las OTU, dividido por el número de caracteres. El "Mean Character Difference" está expresado en la fórmula:

$$MCD = \frac{1}{n} \sum_{j=1}^n |x_{ij} - x_{jk}| \quad [11]$$

donde:  $x_{ij}$  = valor del carácter  $i$  en la OTU  $j$ ,  
 $x_{jk}$  = valor del carácter  $i$  en la OTU  $k$ , y  
 $n$  = número de caracteres.

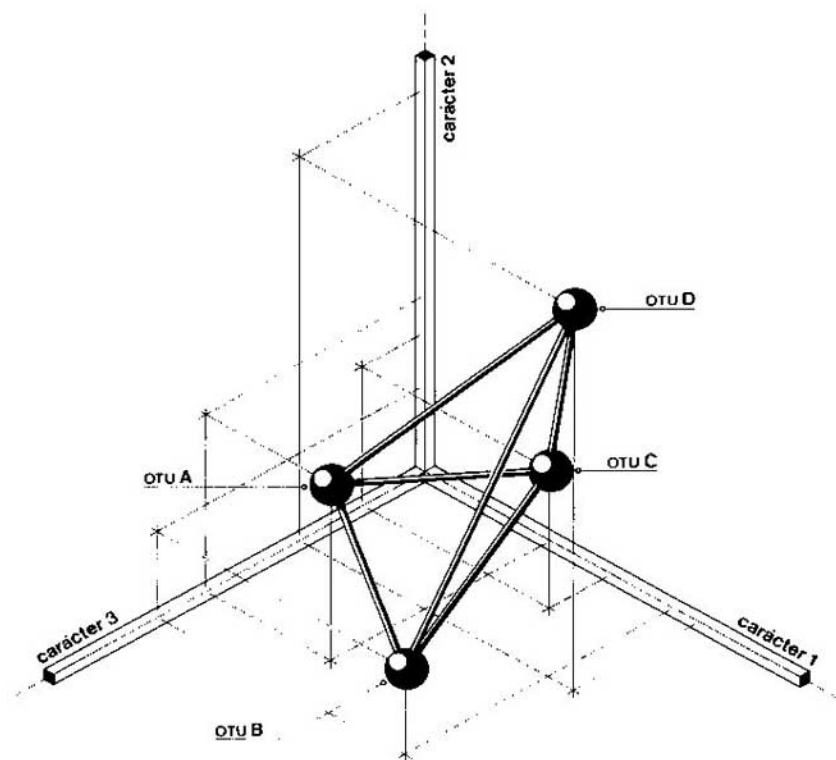


Fig. 5. Representación de cuatro OTU --A, B, C y D-- en un espacio tridimensional determinado por los caracteres 1, 2 y 3. Una medida de la similitud entre las OTU está dada por las distancias entre ellas, en este caso representadas por las barras que las unen.

Este coeficiente recibe el nombre de "Manhattan Distance" cuando no se divide por el número de caracteres.

Otro coeficiente de distancia muy usado es el denominado "Taxonomic Distance" (TD), propuesto por Sokal (1961), y expresado en la fórmula:

$$TD = \left[ \sum_{i=1}^n (x_{ij} - x_{ik})^2 \right]^{1/2} \quad [2]$$

Debido a que el número de caracteres influye en la estimación de la distancia mediante el "Taxonomic Distance", es común que en los cálculos se utilice una distancia promedio (MTD); esto se logra con la aplicación de la siguiente fórmula:

$$MTD = \sqrt{\frac{(TD)^2}{n}} \quad [3]$$

Un coeficiente de distancia que tiene en cuenta la variación intra-OTU (véase el capítulo 4: El problema de la variación intra-OTU) es el propuesto por Crovello (Sneath y Sokal, 1973) y expresado por:

$$CD = \sum_{i=1}^n [(\bar{x}_{ij} - \bar{x}_{ik})^2 + (s_{ij} - s_{ik})^2]^{\frac{1}{2}} \quad [4]$$

donde:  $\bar{x}_{ij}$  = media del carácter i para la OTU j;

$\bar{x}_{ik}$  = media del carácter i para la OTU k;

$s_{ij}$  = desviación estándar del carácter i para la OTU j;

$s_{ik}$  = desviación estándar del carácter i para la OTU k.

Los valores obtenidos a partir de la aplicación de los coeficientes de distancia varían de cero a infinito; cero es la máxima similitud.

Todos los coeficientes de distancia se basan en el postulado de que los estados de los caracteres de los organismos pueden ser usados para obtener una disposición de puntos representando a las OTU en un adecuado espacio euclideo. Williams y Dale (1965) han sugerido el abandono de este postulado y el uso de modelos que incluyan espacios no euclideos. Sin embargo, hasta ahora no se ha propuesto ninguna medida taxonómica no euclidea.

#### COEFICIENTES DE CORRELACION

En la figura 6 es posible cuantificar la similitud midiendo la separación angular formada por las dos líneas que parten del origen de las coordenadas y pasan por las OTU j y k. Los coeficientes de correlación son funciones de esos ángulos. El más empleado de estos coefi-

41

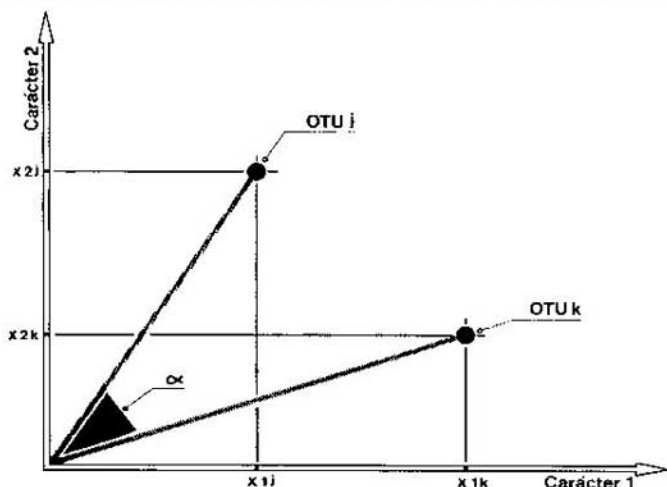


Fig. 6. Representación de dos OTU --j y k-- en un espacio bidimensional determinado por los caracteres 1 y 2. Una medida de similitud entre las OTU está dada por la separación angular determinada por las dos líneas que parten del origen de las coordenadas y pasan por las OTU.

cientes es el coeficiente de Pearson, también llamado coeficiente de correlación del momento-producto. Fue introducido en la taxonomía numé-

rica por Michener y Sokal (1957). Este coeficiente de correlación está expresado en la fórmula:

$$r = \frac{\sum_{i=1}^n (\bar{x}_{ij} - \bar{x}_j)(\bar{x}_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^n (\bar{x}_{ij} - \bar{x}_j)^2 \sum_{i=1}^n (\bar{x}_{ik} - \bar{x}_k)^2}} \quad (5)$$

donde:  $\bar{x}_j$  = media para todos los valores de los estados de la OTU j;

$\bar{x}_k$  = media para todos los valores de los estados de la OTU k.

Los valores de este coeficiente estadístico oscilan entre 1 y -1, siendo 1 el valor de la máxima similitud. El valor -1 difícilmente se alcanza, ya que es imposible una correlación negativa tan alta entre OTU.

El uso de este coeficiente estadístico en taxonomía numérica ha sido objeto de críticas (Eades, 1965; Minkoff, 1965), ya que los caracteres (las columnas de la matriz básica de datos) no son independientes entre sí. Por lo tanto, no se basa en los mismos postulados estadísticos ni son aplicables a él las pruebas de significación. Otro factor crítico es que en algunos casos, la codificación de caracteres ejerce considerable influencia en los resultados. Asimismo, no permite reconocer diferencias de tamaño cuando las OTU son iguales en cuanto a forma. A pesar de estas dificultades es un coeficiente útil, cuando los datos son adecuados a sus propiedades.

#### COEFICIENTES DE ASOCIACION

Estos coeficientes miden las coincidencias y diferencias en los estados de caracteres entre dos OTU. Esta medición exige datos del tipo doble-estado, aunque bajo determinadas condiciones algunos de estos coeficientes (por ejemplo el "simple matching coefficient") pueden aplicarse sobre datos multiestados cualitativos sin secuencia lógica (Sneath y Sokal, 1973).

Se han propuesto innumerables coeficientes de asociación (Cole, 1949, 1957; Dagnelie, 1960; Goodman y Kruskal, 1954, 1959, 1963; Cheethan y Hazel, 1969; Sneath y Sokal, 1973; Clifford y Stephenson, 1975). Antes de presentarlos procede establecer cierta simbología.

Si se comparan dos OTU, j y k, para un carácter doble-estado (presencia/ausencia) se tienen cuatro posibilidades:

1. que ambas OTU tengan presente el carácter comparado (1,1);
2. que ambas OTU tengan ausente el carácter comparado (0,0);
3. que la primera de ellas tenga el carácter presente y la segunda ausente (1,0).
4. que la primera de ellas tenga el carácter ausente y la segunda presente (0,1).

Si se trasladan esas posibilidades a una matriz de 2 x 2, se obtendrá una representación como la indicada en la figura 7.

Donde  $\underline{a}$  (1,1) es el número de caracteres en los cuales el mismo estado es compartido por ambas OTU,  $\underline{b}$  (1,0) es el número de caracteres en los cuales un estado es poseído por la primera OTU, pero no por la segunda,  $\underline{c}$  (0,1) es el número de caracteres en los cuales un estado está ausente en la primera OTU y presente en la segunda y  $\underline{d}$  (0,0) es el nú-

mero de caracteres en los cuales el mismo estado está ausente de ambas OTU.

		OTU j	
		1	0
OTU k	1	1, $\underline{a}$	1, $\underline{b}$
	0	0, $\underline{c}$	0, $\underline{d}$

Fig. 7. Matriz de 2 x 2 que muestra las combinaciones posibles que pueden obtenerse al comparar dos OTU --j y k-- para un mismo carácter del tipo doble-estado.

El resultado de la suma de  $\underline{a}$ ,  $\underline{b}$ ,  $\underline{c}$  y  $\underline{d}$  es el número total de caracteres utilizado ( $=n$ ).

A continuación, presentaremos distintos coeficientes de asociación conjuntamente con el cálculo realizado a base de la siguiente matriz básica de datos hipotética, constituida por las OTU A y B y 10 caracteres del tipo doble-estado, presencia/ausencia.

	1	2	3	4	5	6	7	8	9	10
OTU A	0	1	1	0	1	0	1	0	1	0
OTU B	1	1	0	0	1	1	0	0	1	1

43

El coeficiente de asociación más sencillo es el llamado "Simple Matching Coefficient" (SMC) y está expresado en la fórmula:

$$SMC = \frac{a + d}{a + b + c + d} \quad [6]$$

El valor de este coeficiente oscila entre 1 y 0, siendo 1 la máxima similitud y 0 la mínima.

De su aplicación en las OTU A y B de la matriz básica de datos anterior resulta:

$$SMC(A,B) = \frac{3 + 2}{3 + 2 + 3 + 2} = \frac{5}{10} = 0,5$$

Hay otros coeficientes que no consideran a  $\underline{d}$  (0,0) como elemento en favor de la similitud; por ejemplo, el Coeficiente de Asociación de Jaccard (CAJ). Los valores de similitud obtenidos a partir de la aplicación de este coeficiente varían entre 0 (mínima similitud) y 1 (máxima similitud). Este coeficiente se expresa:

$$CAJ = \frac{a}{a + b + c} \quad [7]$$

Al aplicarse a las OTU A y B de la matriz básica de datos hipotética resulta:

$$CAJ(A,B) = \frac{3}{3+2+3} = \frac{3}{8} = 0,37$$

Otros coeficientes acentúan distintos aspectos de la similitud. Por ejemplo, el coeficiente de Rogers y Tanimoto (RT) da mayor importancia a las diferencias. Los valores de este coeficiente también varían entre 0 y 1, equivalentes a la mínima y máxima similitud, respectivamente. Este coeficiente se expresa:

$$RT = \frac{a+d}{a + (2b) + (2c) + d} \quad [8]$$

Al aplicarse a las OTU A y B de la matriz básica de datos hipotética resulta:

$$RT(A,B) = \frac{3+2}{3 + (2.2) + (2.3) + 2} = \frac{5}{3+4+6+2} = \frac{5}{15} = 0,33$$

Procede mencionar igualmente el coeficiente denominado Dice o de Sørensen (SD) que confiere mayor peso a las coincidencias. Los valores de este coeficiente también varían entre 0 y 1, que equivalen a los valores de mínima y máxima similitud, respectivamente. Este coeficiente se expresa como sigue:

$$SD = \frac{2a}{2a + b + c} \quad [9]$$

Al aplicarse a las OTU A y B de la matriz básica de datos hipotética resulta:

$$SD(A,B) = \frac{2.3}{(2.3) + 2 + 3} = \frac{6}{6+2+3} = \frac{6}{11} = 0,54$$

El coeficiente de Sokal y Sneath (SS) concede mayor significación a las coincidencias y ausencias conjuntas. Los valores de similitud oscilan entre 1 y 0 que equivalen a la máxima y mínima similitud, respectivamente. Este coeficiente se expresa:

$$SS = \frac{2(a+d)}{2(a+d) + b + c} \quad [10]$$

Al aplicarse a las OTU A y B de la matriz básica de datos hipotética resulta:

$$SS(A,B) = \frac{2(3+2)}{2(3+2) + 2 + 3} = \frac{2.5}{10+2+3} = \frac{10}{15} = 0,66$$

El coeficiente de Hamann (H) destaca las diferencias. Los valores de similitud oscilan entre 1 y -1, que equivalen a la máxima y mínima similitud, respectivamente. Este coeficiente se expresa como sigue:

$$H = \frac{(a+d) - (b+c)}{a + b + c + d} \quad [11]$$

Al aplicarse a las OTU A y B de la matriz básica de datos hipotética resulta:



$$H = \frac{(3+2) - (2+3)}{3 + 2 + 3 + 2} = \frac{0}{10} = 0$$

La objeción práctica que puede hacerse a los coeficientes de asociación es que sólo pueden ser utilizados para datos doble-estado. Por lo tanto los datos cuantitativos continuos deben transformarse a datos doble-estado, esto implica una relativa arbitrariedad al determinar los intervalos (Jardine y Sibson, 1971).

Existen también los coeficientes denominados probabilísticos, donde a la coincidencia entre las OTU respecto a estados de caracteres poco frecuentes se asigna más valor o "peso" que a la coincidencia en estados frecuentes (Sneath y Sokal, 1973). Este tipo de coeficiente es útil en trabajos ecológicos, donde la frecuencia de determinada especie puede ser tratada con buen fundamento de manera diferente a las otras.

## ESCALA

En la aplicación de los coeficientes a datos multiestados cuantitativos continuos surge una dificultad porque pueden coexistir diferentes escalas de medida en relación con una misma OTU. Por ejemplo, el carácter "diámetro del grano de polen" expresado en micrones y el carácter "altura de la planta" expresado en metros. Si tratásemos de igualar las escalas y expresáramos la altura de la planta en micrones, habría que colocar en la matriz básica de datos una cifra enorme que distorsionaría los cálculos. Otro ejemplo sería el caso de la coexistencia de un carácter que mida longitud en centímetros y otro que mida peso en gramos. En estos casos, hay que transformar la matriz básica de datos para que de allí en adelante exprese todos los valores en una sola escala.

45

La técnica de transformación más utilizada es la denominada *estandarización*, que consiste en expresar los valores de la matriz básica de datos en unidades de desviación estándar. El cálculo mencionado se expresa:

$$X_{ij} \text{ estandarizado} = \frac{X_{ij} - \bar{X}_i}{S_i} \quad [12]$$

en que  $X_{ij}$  es el valor del carácter  $i$  para la OTU- $j$ ;  $\bar{X}_i$  la media del carácter  $i$  y  $S_i$  la desviación estándar del carácter  $i$ .

Para dar un ejemplo, si tenemos dos estados de un carácter: 2 y 4, la media es 3 y la desviación estándar 0,7, los estados se transforman de la siguiente manera:

Valor Inicial	Cálculo	Valor Estandarizado
2	$2-3/0,7$	-1,428
4	$4-3/0,7$	1,428

Debido a la estandarización de la matriz básica de datos la media de un carácter debe expresarse como cero y la varianza como la unidad. La mayor parte de los valores se encuentran entre -3 y 3.

Esta técnica de transformación tiene algunos inconvenientes; por ejemplo, el hecho de que iguala la variación: un carácter con escaso

rango de variabilidad una vez transformada la matriz tiene la misma variabilidad (de -3 a 3) que un carácter con amplio rango.

En los estudios acerca de la influencia de la transformación de la matriz básica de datos sobre los resultados (Rohlf, 1962, Rohlf y Sokal, 1965; Sokal y Michener, 1967, y Moss, 1968), se demuestra que, a pesar de sus inconvenientes, es aconsejable la estandarización de los datos multiestados cuantitativos continuos.

#### ELECCION DEL COEFICIENTE DE SIMILITUD

La elección del coeficiente de similitud que se va a utilizar está supeditada al tipo de datos que contiene la matriz básica de datos. Por ejemplo, los coeficientes de asociación sólo se aplican a datos doble-estado y el coeficiente de correlación no da buenos resultados con esos datos.

En el Cuadro II se resume la información sobre los coeficientes, incluyendo nombre, fórmula, tipo de dato al que se aplican, valor máximo y mínimo de similitud.

Se ha demostrado empíricamente (Sneath y Sokal, 1973) que en aquellos estudios en los cuales predominan los caracteres doble-estado conviene transformar, mediante codificación, los datos multiestados restantes en datos doble-estado y utilizar coeficientes de asociación.

En aquellos estudios en los cuales predominan los caracteres multiestados es aconsejable la estandarización y la utilización de coeficientes de distancia y correlación.

Dos comentarios finales son pertinentes. El primero respecto a la validez de comparar valores de similitud obtenidos de diferentes estudios: el valor de similitud entre dos OTU sólo es comparable a valores obtenidos en ese mismo estudio (que surjan de la misma matriz básica de datos) y no es comparable con ningún otro valor que no pertenezca a ese estudio, es decir, a otra matriz básica de datos. Por ejemplo, el valor de un coeficiente de Jaccard de 0,52 entre dos especies de Compuestas no es comparable con un valor 0,80 entre dos especies de Mamíferos y, por lo tanto, no puede afirmarse que esas dos especies de Compuestas son menos similares entre sí que las dos especies de Mamíferos.

El segundo comentario tiene que ver con la pregunta: ¿Puede el valor obtenido con un coeficiente de similitud someterse a pruebas de significación estadística? Hasta el momento la respuesta es ambigua. En muy pocos estudios estadísticos se ha profundizado sobre este tema. Tal vez como medida de precaución convenga considerar a ese valor sólo como número descriptivo que, cuantitativamente, expresa una realidad taxonómica.

#### MATRIZ DE SIMILITUD

Los resultados obtenidos de la aplicación de cualquiera de los coeficientes de similitud para los pares posibles de OTU ordenados en forma tabular constituyen la matriz de similitud (Fig. 8).

Las OTU ocupan tanto las filas como las columnas, siguiendo el mismo orden en ambas; de esta manera se logra comparar cada OTU consigo misma y con las restantes OTU. Así estructurada la matriz, cada valor de la diagonal principal ( $S_{11}$ ,  $S_{22}$ ,  $S_{33}$ , ...,  $S_{tt}$ ) representa a cada OTU comparada consigo misma; este valor corresponde al de la máxi-

Cuadro II. Tipos de Coeficientes y Ejemplos: Nombres, Fórmulas, Tipos de Datos Necesarios y Valores Máximos y Mínimos de Similitud

TIPO DE COEFICIENTE	NOMBRE	FÓRMULA	TIPO DE DATOS SOBRE LOS QUE SE APLICA	MAXIMA SIMILITUD	MINIMA SIMILITUD
DISTANCIA	Mean character difference	$\frac{1}{n} \sum_{i=1}^n  X_{ij} - X_{ik} $	• Doble - estado	0	$\infty$
	Manhattan distance	$\sum_{i=1}^n  X_{ij} - X_{ik} $	• Multiestados		
	Taxonomic distance	$\left[ \sum_{i=1}^n (X_{ij} - X_{ik})^2 \right]^{\frac{1}{2}}$	• Mixtos		
CORRELACION	Momento-producto de Pearson	$\frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k)}{\left[ \sum_{j=1}^n (X_{ij} - \bar{X}_j)^2 \sum_{k=1}^n (X_{ik} - \bar{X}_k)^2 \right]^{\frac{1}{2}}}$	• Multiestados cuantitativos	1	-1
			• Mixtos (Con predominancia de multiestados cuantitativos)		
ASOCIACION	Simple matching	$\frac{a+d}{a+b+c+d}$	• Doble - estado	1	0
	Jaccard	$\frac{a}{a+b+c}$			
	Rogers y Tanimoto	$\frac{a+d}{a+(2b)+(2c)+d}$			
	Dice	$\frac{2a}{2a+b+c}$			
	Sokal y Sneath	$\frac{2(a+d)}{2(a+d)+b+c}$			
	Hamann	$\frac{(a+d) - (b+c)}{a+b+c+d}$			
				1	-1

M.S.		OTU							
		1	2	3	4	5	• • •	t	
OTU	1	$S_{11}$							
	2	$S_{12}$	$S_{22}$						
	3	$S_{13}$	$S_{23}$	$S_{33}$					
	4	$S_{14}$	$S_{24}$	$S_{34}$	$S_{44}$				
	5	$S_{15}$	$S_{25}$	$S_{35}$	$S_{45}$	$S_{55}$			
	•	•	•	•	•	•	•		
	•	•	•	•	•	•	• •		
	•	•	•	•	•	•	• • •		
	t	$S_{1t}$	$S_{2t}$	$S_{3t}$	$S_{4t}$	$S_{5t}$	• • •	$S_{tt}$	

Fig. 8. Matriz de similitud entre OTU, donde  $S_{12}$  corresponde al valor de similitud entre la OTU 1 y la OTU 2.

48

ma similitud; 1 en el caso de los coeficientes de asociación y correlación y 0 respecto a los coeficientes de distancia.

La similitud entre la OTU 1 y la OTU 2 ( $S_{12}$ ) es la misma que entre la OTU 2 y la OTU 1 ( $S_{21}$ ), por eso la parte superior derecha de la matriz es la imagen en espejo de la parte inferior izquierda. En las publicaciones es convencional usar solamente el triángulo inferior izquierdo.

La figura 9 representa las matrices de similitud obtenidas de la aplicación de los coeficientes de correlación y distancia ("Taxonomic Distance") a partir de la matriz básica de datos, previamente estandarizada, de nuestro ejemplo en el género *Bulnesia*.

Matriz de Similitud (distancia)								
	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.	0,68	0						
B.chil.	1,65	1,80	0					
B.bon.	1,19	1,31	1,22	0				
B.ret.	1,58	1,72	1,01	1,21	0			
B.fol.	1,59	1,66	1,18	1,12	0,99	0		
B.schic.	1,70	1,84	1,23	1,13	1,13	0,73	0	
B.sar.	1,67	1,74	1,57	1,52	1,53	1,27	1,55	0

Fig. 9. a) Matriz de similitud entre ocho OTU del género *Bulnesia*, obtenida a partir de un coeficiente de distancia.

Matriz de Similitud (correlación)

b)

	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	1							
B.car.	0,71	1						
B.chil.	0,39	-0,49	1					
B.bon.	0,16	0,17	-0,16	1				
B.ret.	-0,47	-0,56	0,35	-0,24	1			
B.fol.	-0,59	-0,44	-0,07	-0,30	0,16	1		
B.schic.	-0,49	-0,53	0,02	-0,02	0,11	0,53	1	
B.sar.	-0,06	0,01	-0,25	-0,41	-0,24	-0,03	-0,34	1

Fig. 9. b) matriz de similitud entre ocho OTU del género *Bulnesia* obtenida a partir de un coeficiente de correlación.



## LA BUSQUEDA DE LA ESTRUCTURA TAXONOMICA: ANALISIS DE AGRUPAMIENTOS

La matriz de similitud es insuficiente para expresar relaciones entre la totalidad de las OTU, pues sólo expone similitudes entre pares de dichas unidades.

Se dispone de una gran variedad de técnicas de análisis de matrices de similitud (Sneath y Sokal, 1973), cuyo objeto es sintetizar la información de la matriz de similitud a fin de permitir el reconocimiento de las relaciones entre la totalidad de las OTU.

En éste y en el capítulo próximo presentaremos dos de las técnicas más utilizadas: el análisis de agrupamientos ("cluster analysis") y el método de ordenación ("ordination").

## ANALISIS DE AGRUPAMIENTOS

El análisis de agrupamientos comprende técnicas que, siguiendo reglas más o menos arbitrarias, forman grupos de OTU que se asocian por su grado de similitud.

Esta definición es poco precisa y ello se debe a dos factores: primero, el escaso acuerdo entre los investigadores acerca de cómo reconocer los límites entre grupos, y segundo, la enorme variedad de técnicas propuestas.

Las numerosas técnicas de análisis de agrupamientos han sido estudiadas por Ball, 1965; Williams y Dale, 1965; Wishart, 1969 a; Spence y Taylor, 1970; Cormack, 1971 y Hartigan, 1975, entre otros.

Dichas técnicas se clasifican a base de los siguientes dilemas:

- a) técnicas que forman grupos exclusivos *versus* técnicas que forman grupos no exclusivos;
- b) técnicas que forman grupos jerárquicos *versus* técnicas que forman grupos no jerárquicos;
- c) técnicas divisivas *versus* técnicas aglomerativas, y
- d) técnicas secuenciales *versus* técnicas simultáneas.

**a) Técnicas exclusivas:** Originan grupos donde las OTU son exclusivas del grupo del cual forman parte y no pueden pertenecer a otro grupo que se halle en un mismo rango o nivel.

**Técnicas no exclusivas:** Originan grupos donde las OTU pueden pertenecer a más de un grupo en un mismo nivel o rango.

**b) Técnicas jerárquicas:** Originan conjuntos que presentan rangos, en los cuales las OTU o grupos de OTU subsidiarios forman parte de un grupo mayor o inclusivo.

**Técnicas no jerárquicas:** Originan conjuntos que no exhiben rangos.

c) **Técnicas aglomerativas:** Son las que partiendo de  $n$  OTU separadas, las agrupa en sucesivos conjuntos (siempre en número menor que  $n$ ) para llegar finalmente a un solo conjunto que contiene a las  $n$  unidades.

**Técnicas divisivas:** Son aquéllas que partiendo de un conjunto que contiene a las  $n$  OTU las divide en subconjuntos.

d) **Técnicas secuenciales:** Cada grupo es formado uno por vez hasta que se agota el conjunto total.

**Técnicas simultáneas:** Los grupos son formados simultáneamente, en una sola operación.

De las alternativas presentadas en los dilemas, las más utilizadas son las exclusivas, jerárquicas, aglomerativas y secuenciales, las cuales se combinan caracterizando a las técnicas de agrupamientos que describiremos. Dentro de las mismas hemos elegido, por ser las más sencillas, las del llamado "grupo par" ("pair group"), en las cuales solamente puede ser admitida una OTU o un grupo de OTU por nivel. Esto significa que los grupos formados en cualquier etapa de los agrupamientos contienen sólo dos miembros.

A continuación describiremos la técnica operativa, con sus variantes (un ejemplo de su aplicación se presentará más adelante):

1. Se examina la matriz de similitud para localizar el mayor valor de similitud existente en ella, descartando, lógicamente, la diagonal principal. Se identifica así a las dos OTU que formarán el denominado núcleo del primer grupo. Núcleo es todo conjunto formado por dos OTU y grupo es todo conjunto formado por más de dos OTU. En algunos casos, puede haber más de un valor máximo de similitud, es decir otro par o pares de OTU presentan igual valor que el anterior; en ese caso, se construyen a ese nivel dos o más núcleos separados.

2. Se busca en la matriz de similitud el próximo valor de mayor similitud. En las primeras etapas del proceso de agrupamiento, el hallazgo de este nuevo valor puede llevar a:

- la formación de nuevos núcleos;
- la incorporación de una OTU a un núcleo ya existente para formar un grupo, y
- la fusión de los núcleos existentes.

3. Se repite la segunda etapa del proceso hasta que todos los núcleos y grupos estén unidos y en ellos se incluya la totalidad de las OTU.

El primer paso es común a todas las técnicas, el segundo (incorporación de nuevas OTU a núcleos y grupos existentes) puede realizarse por tres caminos diferentes denominados:

- a) Ligamiento simple ("simple linkage").
- b) Ligamiento completo ("complete linkage").
- c) Ligamiento promedio ("average linkage").

a) **Ligamiento simple.** Las OTU se incorporan a grupos o núcleos ya formados tomando en cuenta que el valor de similitud entre la OTU



candidato a incorporarse y el grupo o núcleo es igual a la similitud entre el candidato y la OTU integrante del grupo o núcleo más parecido a ella; en otras palabras, el de mayor valor de similitud. Si el candidato a incorporarse es un grupo o núcleo en sí mismo, el valor de similitud será igual a la máxima similitud hallada entre dos OTU provenientes una de cada grupo o núcleo.

b) **Ligamiento completo.** En este caso, se considera que el valor de similitud entre la OTU candidato a incorporarse y el grupo o núcleo es igual a la similitud entre el candidato y la OTU integrante del grupo o núcleo menos parecido a él, en otras palabras el de menor valor de similitud. Si el candidato a incorporarse es un grupo o núcleo en sí mismo, el valor de similitud será igual a la mínima similitud hallada entre dos OTU provenientes una de cada grupo o núcleo.

c) **Ligamiento promedio.** En este caso se considera que el valor de similitud entre la OTU candidato a incorporarse y el grupo o núcleo es igual a una similitud promedio resultante de los valores de similitud entre el candidato y cada uno de los integrantes del grupo o núcleo. Como existen varios tipos de medias, es posible contar con más de una técnica de ligamiento promedio. La más utilizada es la media aritmética no ponderada (UPGMA, "unweighted pair-group method using arithmetic averages"). Si el candidato a incorporarse es un grupo o núcleo en sí mismo, el valor de similitud será un promedio de los valores de similitud entre los pares posibles de OTU provenientes uno de cada grupo o núcleo.

El reconocimiento del valor de similitud al que se formarán nuevos núcleos o grupos, o se incorporarán nuevas OTU, o al que se fusionarán los núcleos o grupos existentes, se ve facilitado por la obtención de *matrices derivadas*.

53

Estas matrices tienen su origen en la matriz de similitud original y se obtienen con anterioridad a cada una de las etapas del proceso de agrupamiento. El cálculo de los valores que constituirán estas matrices varía según la técnica de agrupamiento utilizada (ligamiento simple, completo o promedio).

La primera etapa del proceso de agrupamiento consiste en reconocer en la matriz de similitud el par de OTU con el valor de mayor similitud. La primera matriz derivada diferirá de la matriz de similitud original en que el núcleo formado será considerado como unidad con respecto a las restantes OTU.

Si comparamos la matriz de similitud de la cual se parte con la matriz derivada obtenida, ésta disminuye en tamaño en una columna y en una fila.

Es en la matriz derivada donde debe buscarse el próximo valor de mayor similitud. Sabremos aquí cuál de las posibilidades mencionadas se presenta.

Se vuelve a obtener una nueva matriz derivada, diferente a la anterior, en la cual los nuevos núcleos o grupos formados se consideran como unidades con respecto a las OTU no integrantes de núcleos o grupos. Se repite así el procedimiento hasta lograr que todos los núcleos y grupos constituyan un grupo que contenga a todas las OTU.

Es común a todas las técnicas de agrupamientos que la primera matriz derivada se obtenga de la matriz de similitud. En el ligamiento simple, en el ligamiento completo y en algunos tipos de ligamiento

promedio (los llamados "ponderados") las matrices derivadas son calculadas a partir de los valores de la matriz derivada inmediatamente anterior. En el ligamiento promedio aritmético "no ponderado" (UPGMA) al calcular las matrices derivadas se vuelve siempre a la matriz de similitud original.

#### REPRESENTACION GRAFICA DE LAS TECNICAS DE ANALISIS DE AGRUPAMIENTOS

La estructura taxonómica obtenida de la matriz de similitud con las técnicas de análisis de agrupamientos puede representarse gráficamente de varias formas, pero la más utilizada es el fenograma (Fig. 10).

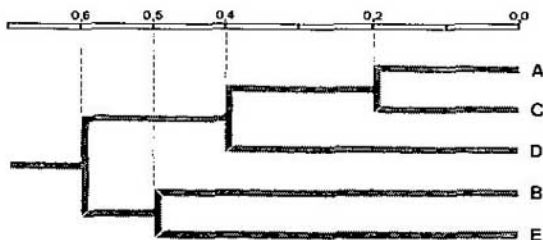


Fig. 10. Fenograma de cinco OTU: A, B, C, D y E.

El fenograma es un diagrama arborescente que muestra la relación en grado de similitud entre dos OTU o grupos de OTU.

54

Los valores de similitud se expresan en una escala que suele encontrarse en su extremo superior.

Las OTU se colocan en el extremo derecho y dan origen cada una a un eje horizontal. Los ejes horizontales se unirán mediante ejes verticales que expresan, en relación con la escala, el valor de similitud existente entre las OTU, o conjuntos de OTU.

Podríamos decir que un grupo cualquiera se constituye por las paralelas a la escala que nacen de un mismo eje vertical y las ramificaciones que contiene.

Respecto a la terminología, Sokal y Camin (1965), Mayr (1965) y Wiley (1981) distinguen entre fenogramas, que representan relaciones fenéticas, y cladogramas, donde las ramificaciones están basadas en las conexiones filogenéticas entre las OTU. El término dendrograma incluye a los fenogramas y a los cladogramas.

El fenograma de la figura 10 contiene cinco OTU (A, B, C, D y E). Las OTU A y C se unen a un nivel 0,2 de similitud; de allí el eje A-C se une con la OTU D a un nivel de 0,4. El núcleo B-E se forma a un nivel de 0,5 y todas las OTU quedan reunidas a un nivel 0,6.

El orden de las OTU en el lado derecho es de escasa importancia, ya que los ejes pueden rotar en cualquier grado (180° en el papel) sin alterar los patrones de relación entre las OTU. La figura 11 muestra otras tres formas de representar correctamente el fenograma de la figura 10.

#### CONSTRUCCION DE FENOGRAMAS

Tomaremos la matriz de similitud de distancia del género *Bulnesia* (Fig. 9a) como ejemplo ilustrativo y realizaremos todos los pasos del

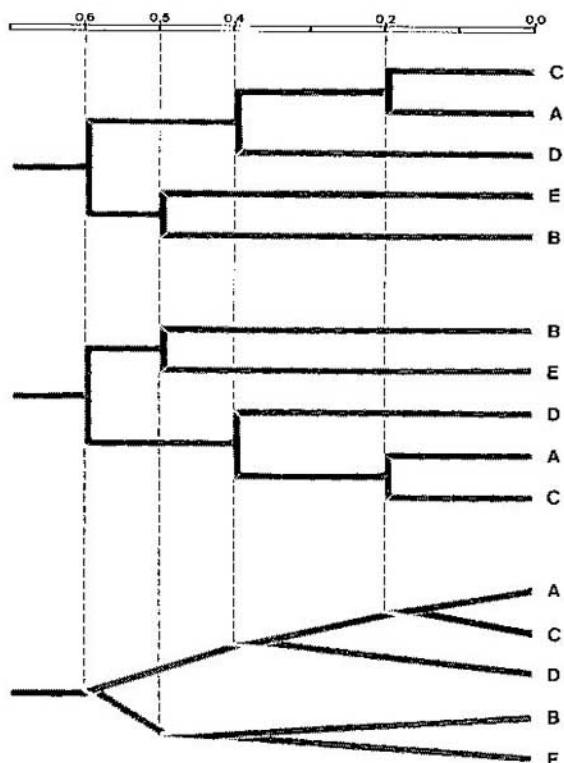


Fig. 11. Tres formas válidas de representar un mismo fenograma.

análisis de agrupamientos a partir de las tres técnicas presentadas: ligamiento simple, ligamiento completo y ligamiento promedio (UPGMA).

**Ligamiento simple** (Figs. 12 y 13). El valor de mayor similitud encontrado en la matriz de similitud es el que poseen B.arb. (*B. arborea*) y B.car. (*B. carapo*) que forman un núcleo, unido a un valor de similitud de 0,68.

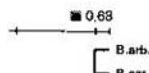
La primera matriz derivada considera a B.arb. y B.car. como un núcleo con respecto a las restantes OTU. Como la técnica utilizada es la del ligamiento simple los valores que se extraen de la matriz de similitud derivan de la elección del valor de similitud numéricamente menor (el más parecido) entre el par B.arb.-B.car. y las restantes OTU.

En esta primera matriz derivada se observa que el próximo valor de mayor similitud es el que poseen B.fol. (*B. foliosa*) y B.schic. (*B. schickendantzii*) que se unen formando un nuevo núcleo a un valor de similitud de 0,73.

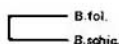
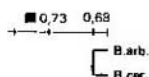
La segunda matriz derivada considera no sólo a B.arb. y a B.car. como un núcleo, sino también a B.fol. y B.schic. y se extraen los menores valores de distancia existentes entre:

- a) el nuevo núcleo y el anterior, y
- b) el nuevo núcleo y las restantes OTU.

M.S.	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.	0,68	0						
B.chil.	1,65	1,80	0					
B.bon.	1,19	1,31	1,22	0				
B.ret.	1,58	1,72	1,01	1,21	0			
B.fol.	1,59	1,86	1,18	1,12	0,99	0		
B.schic.	1,70	1,84	1,23	1,13	1,13	0,73	0	
B.sar.	1,67	1,74	1,57	1,52	1,53	1,27	1,55	0



I	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.								
B.chil.	1,65	0						
B.bon.	1,19	1,22	0					
B.ret.	1,58	1,01	1,21	0				
B.fol.	1,59	1,18	1,12	0,99	0			
B.schic.	1,70	1,23	1,13	1,13	0,73	0		
B.sar.	1,67	1,57	1,52	1,53	1,27	1,55	0	



II	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.								
B.chil.	1,65	0						
B.bon.	1,19	1,22	0					
B.ret.	1,58	1,01	1,21	0				
B.fol.	1,59	1,18	1,12	0,99	0			
B.schic.								
B.sar.	1,67	1,57	1,52	1,53	1,27	1,55	0	

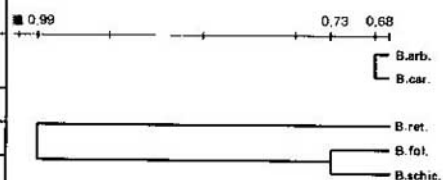
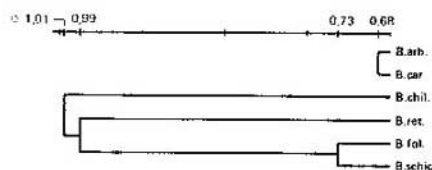


Fig. 12. Tres primeros pasos en la construcción de un fenograma mediante la técnica del ligamiento simple. Cada paso comprende la matriz derivada y el agrupamiento correspondiente (continúa en la Fig. 13).

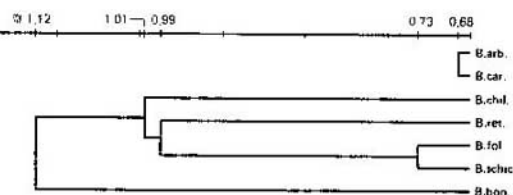
III

B.arb.	B.car.	B.chil.	B.bon.	B.fol.	B.schic.	B.sar.
B.arb.	0					
B.car.						
B.chil.	1,65	0				
B.bon.	1,19	1,22	0			
B.fol.	1,58	1,01	1,12	0		
B.schic.						
B.sar.	1,67	1,57	1,52	1,27	0	



IV

B.arb.	B.car.	B.bon.	B.chil.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0						
B.car.							
B.bon.	1,19	0					
B.chil.	1,58	1,12	0				
B.ret.							
B.fol.							
B.schic.							
B.sar.	1,67	1,52	1,27	0			



V

B.arb.	B.car.	B.chil.	B.ret.	B.fol.	B.schic.	B.bon.	B.sar.
B.arb.	0						
B.car.							
B.chil.	1,19	0					
B.ret.							
B.fol.							
B.schic.							
B.bon.							
B.sar.	1,67	1,27	0				



VI

B.arb.	B.car.	B.chil.	B.ret.	B.fol.	B.schic.	B.bon.	B.sar.
B.arb.	0						
B.car.							
B.chil.							
B.ret.							
B.fol.							
B.schic.							
B.bon.							
B.sar.	1,27	0					

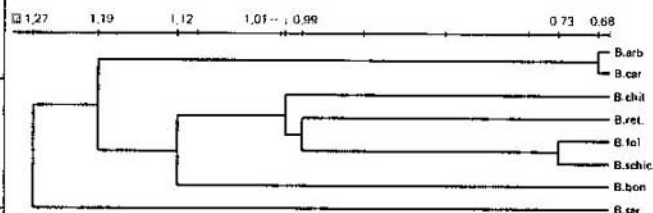


Fig. 13. Pasos finales en la construcción de un fenograma mediante la técnica del ligamiento simple.

En esta matriz se observa que el valor de mayor similitud es el que posee B.ret. (*B. retama*) y el núcleo B.fol.-B.schic. Esto significa que B.ret. se une a dicho núcleo originando un grupo a un valor de similitud de 0,99.

La tercera matriz derivada contiene los valores de similitud del núcleo B.arb.-B.car. y del grupo B.ret.-B.fol.-B.schic. entre sí y de ambos con las restantes OTU. Como se observa, el valor de mayor similitud de esta matriz es de 1,01 y une a B.chil. (*B. chilensis*) al grupo B.ret.-B.fol.-B.schic.

La cuarta matriz derivada contiene los valores de similitud del núcleo B.arb.-B.car. y del grupo B.chil.-B.ret.-B.fol.-B.schic. entre sí y de ambos con las restantes OTU. Se observa que el nuevo valor de mayor similitud es de 1,12 y es a este nivel que se une B.bon. (*B. bonariensis*) al grupo B.chil.-B.ret.-B.fol.-B.schic.

La quinta matriz derivada considera el valor de similitud entre el núcleo B.arb.-B.car. con respecto al grupo B.chil.-B.ret.-B.fol.-B.schic.-B.bon. y el de ambos con respecto a B.sar. (*B. szymientoi*). El resultado de este paso es la unión del núcleo B.arb.-B.car. al grupo B.chil.-B.ret.-B.fol.-B.schic.-B.bon. a un nivel de 1,19.

La sexta y última matriz derivada nos da el valor al que se une B.sar. a las restantes OTU, o sea a 1,27.

**Ligamiento completo** (Figs. 14 y 15). El valor de mayor similitud hallado en la matriz de similitud es el que poseen B.arb. y B.car. que forman un núcleo a un valor de similitud de 0,68.

La primera matriz derivada considera a B.arb. y a B.car. como un núcleo con respecto a las restantes OTU. Como la técnica utilizada es la del ligamiento completo, los valores obtenidos de la matriz de similitud provienen de la elección del valor de similitud numéricamente mayor (el menos similar) entre el par B.arb.-B.car. y las restantes OTU.

En esta primera matriz derivada se observa que el próximo valor de mayor similitud es el que poseen B.fol. y B.schic. que se unen formando un nuevo núcleo a un valor de similitud de 0,73.

La segunda matriz derivada considera no sólo a B.arb. y a B.car. como un núcleo, sino también a B.fol. y a B.schic. y se extraen los mayores valores de distancia existentes entre:

- a) el nuevo núcleo y el anterior, y
- b) el nuevo núcleo y las restantes OTU.

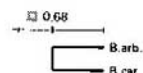
En esta matriz se observa que el valor de mayor similitud es el que posee B.ret. y B.chil. dando origen a un tercer núcleo, a un valor de similitud de 1,01.

La tercera matriz derivada considera el mayor valor de similitud existente entre:

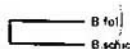
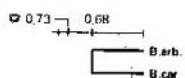
- a) cada núcleo comparado con los restantes núcleos, y
- b) cada núcleo comparado con las restantes OTU.

El valor de mayor similitud es el que posee el par B.fol.-B.schic. y B.bon. Esto significa que B.bon. se une a dicho núcleo originando un grupo a un valor de similitud de 1,13.

M.S.	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.	0,68	0						
B.chil.	1,65	1,80	0					
B.bon.	1,19	1,31	1,22	0				
B.ret.	1,58	1,72	1,01	1,21	0			
B.fol.	1,59	1,66	1,18	1,12	0,99	0		
B.schic.	1,70	1,84	1,23	1,13	1,13	0,73	0	
B.sar.	1,67	1,74	1,57	1,52	1,53	1,27	1,55	0



I	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.	0							
B.chil.	1,80	0						
B.bon.	1,31	1,22	0					
B.ret.	1,72	1,01	1,21	0				
B.fol.	1,66	1,18	1,12	0,99	0			
B.schic.	1,84	1,23	1,13	1,13	0,73	0		
B.sar.	1,74	1,57	1,52	1,53	1,27	1,55	0	



II	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.	0							
B.chil.	1,80	0						
B.bon.	1,31	1,22	0					
B.ret.	1,72	1,01	1,21	0				
B.fol.	1,84	1,23	1,13	1,13	0			
B.schic.	1,74	1,57	1,52	1,53	1,55	0		

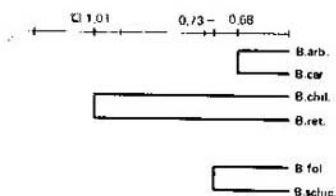
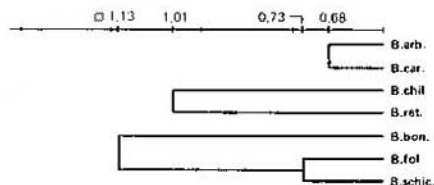
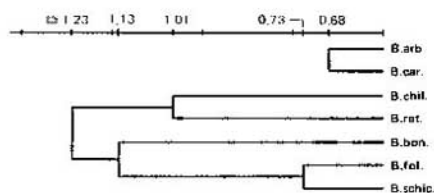


Fig. 14. Tres primeros pasos en la construcción de un fenograma mediante la técnica del ligamiento completo. Cada paso comprende la matriz derivada y el agrupamiento correspondiente (continúa en la Fig. 15).

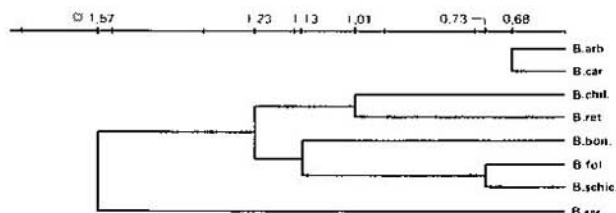
III	B.arb B.car.	B.chil B.ret.	B.bon.	B.fol B.schic.	B.sar.
B.arb B.car.	0				
B.chil B.ret.	1,80	0			
B.bon.	1,31	1,22	0		
B.fol B.schic.	1,84	1,23	113	0	
B.sar	1,74	1,57	1,52	1,55	0



IV	B.arb B.car.	B.chil B.ret.	B.bon B.fol B.schic.	B.sar
B.arb B.car.	0			
B.chil B.ret.	1,80	0		
B.bon B.fol B.schic.	1,84	1,23	0	
B.sar	1,74	1,57	1,55	0



V		B.chil. B.ret B.bon. B.fol B.sar	
B.arb B.car	0		
B.chil. B.ret B.bon. B.fol. B.schie.	1,84	0	
B.sar.	1,74	1,57	0



VI		B.chil B.ret. B.bon B.fol. B.schie. B.sar
B.arb. B.car	0	
B.chil B.ret. B.bon B.fol B.schie B.sar	1,84	0

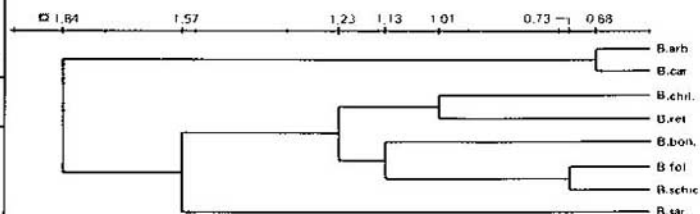


Fig. 15. Pasos finales en la construcción de un fenograma mediante la técnica del ligamiento completo.



La cuarta matriz derivada contiene los valores de similitud de los núcleos B.arb.-B.car.-B.chil.-B.ret. y del grupo B.hon.-B.fol.-B.schic. entre sí y con las restantes OTU. El valor de mayor similitud es de 1,23 entre el núcleo B.chil.-B.ret. y el grupo B.hon.-B.fol.-B.schic. y ambos se fusionan en un grupo.

La quinta matriz derivada considera el valor de similitud entre el núcleo B.arb.-B.car. con respecto al grupo B.chil.-B.ret.-B.hon.-B.fol.-B.schic. y de ambos con B.sar. El valor de mayor similitud es de 1,57 y relaciona a B.sar. con B.chil.-B.ret.-B.hon.-B.fol.-B.schic.

La sexta y última matriz derivada nos da el valor 1,84 al que se une el núcleo B.arb.-B.car. con las restantes OTU.

**Ligamiento promedio** (Figs. 16 y 17). El mayor valor de similitud encontrado en la matriz de similitud es el que poseen B.arb. y B.car. que forman un núcleo a un valor de similitud de 0,68.

La primera matriz derivada considera a B.arb. y a B.car. como un núcleo con respecto a las restantes OTU.

La técnica utilizada es la del ligamiento promedio no ponderado (UPGMA). Los valores que se vuelcan a la matriz derivada provienen de la media aritmética extraída de los valores de similitud del núcleo B.arb.-B.car. con respecto a las demás OTU.

En esta primera matriz derivada se observa que el próximo valor de mayor similitud es el que poseen B.fol. y B.schic. que se unen formando un nuevo núcleo, a un valor de similitud de 0,73.

61

La segunda matriz derivada considera no sólo a B.arb. y a B.car. como un núcleo sino también a B.fol. y a B.schic. Se extraen las medias entre:

- a) el nuevo núcleo y el anterior, y
- b) el nuevo núcleo y las restantes OTU.

La similitud entre el núcleo ya formado B.arb.-B.car. con respecto al nuevo núcleo B.fol.-B.schic. se calcula de la siguiente manera:

B.arb.-B.fol.	1,59
B.arb.-B.schic.	1,70
B.car.-B.fol.	1,66
B.car.-B.schic.	<u>1,84</u>
	6,79

$$\text{promedio} = \frac{6,79}{4} = 1,69$$

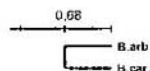
Este cálculo se ha obtenido con el método no ponderado.

Si recurrimos a un método que atribuya peso al candidato ingresante, procederíamos de la siguiente forma:

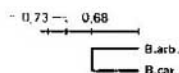
(B.arb.-B.car.) B.fol.	= 1,62
(B.arb.-B.car.) B.schic.	= <u>1,77</u>
	3,39

$$\text{promedio pesado} = \frac{3,39}{2} = 1,69.$$

M.S.	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.	0,68	0						
B.chil.	1,65	1,80	0					
B.bon.	1,19	1,31	1,22	0				
B.ret.	1,58	1,72	1,01	1,21	0			
B.fol.	1,59	1,66	1,18	1,12	0,99	0		
B.schic.	1,70	1,84	1,23	1,13	1,13	0,73	0	
B.sar.	1,67	1,74	1,57	1,52	1,53	1,27	1,55	0



I	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.		0						
B.chil.	1,72	0						
B.bon.	1,25	1,22	0					
B.ret.	1,65	1,01	1,21	0				
B.fol.	1,62	1,18	1,12	0,99	0			
B.schic.	1,77	1,23	1,13	1,13	0,73	0		
B.sar.	1,70	1,57	1,52	1,53	1,27	1,55	0	



II	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.		0						
B.chil.	1,72	0						
B.bon.	1,25	1,22	0					
B.ret.	1,65	1,01	1,21	0				
B.fol.	1,69	1,20	1,12	1,06	0			
B.schic.						0		
B.sar.	1,70	1,57	1,52	1,53	1,41	0		

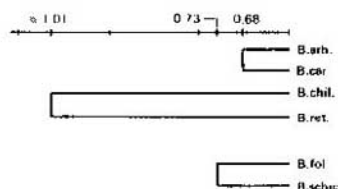
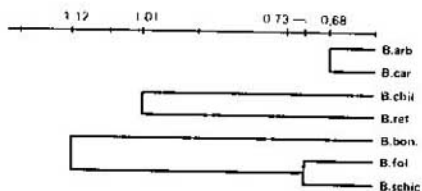
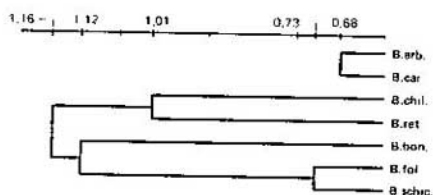


Fig. 16. Tres primeros pasos en la construcción de un fenograma mediante la técnica del ligamiento promedio. Cada paso comprende la matriz derivada y el agrupamiento correspondiente (continúa en la Fig. 17)

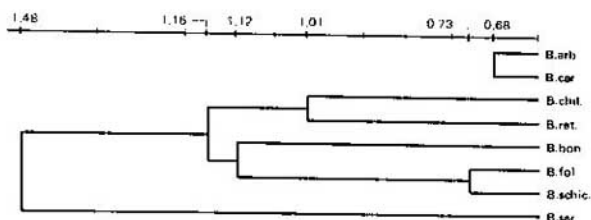
III	B.arb. B.car.	B.chil. B.ret.	B.bon.	B.fol. B.schic.	B.sar.
B.arb. B.car.	0				
B.chil. B.ret.	1,68	0			
B.bon.	1,25	1,21	0		
B.fol. B.schic.	1,89	1,13	1,12	0	
B.sar.	1,70	1,55	1,52	1,41	0



IV	B.arb. B.car.	B.chil. B.ret.	B.bon. B.fol.	B.schic. B.sar.
B.arb. B.car.	0			
B.chil. B.ret.	1,68	0		
B.bon. B.fol.	1,54	1,16	0	
B.schic. B.sar.	1,70	1,55	1,44	0



V	B.arb. B.car.	B.chil. B.ret.	B.bon. B.fol.	B.schic. B.sar.
B.arb. B.car.	0			
B.chil. B.ret.	1,60	0		
B.bon. B.fol.	1,70	1,48	0	
B.schic. B.sar.	1,70	1,48	0	



VI	B.arb. B.car.	B.chil. B.ret.	B.bon. B.fol.	B.schic. B.sar.
B.arb. B.car.	0			
B.chil. B.ret.	1,62	0		
B.bon. B.fol.	1,62	0		
B.schic. B.sar.	1,62	0		

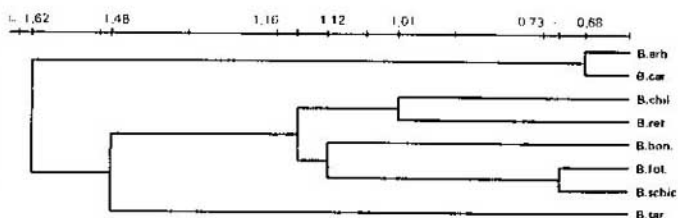


Fig. 17. Pasos finales en la construcción de un fenograma mediante la técnica del ligamiento promedio.

Como puede apreciarse, con el método no ponderado se retorna en cada paso a la matriz de similitud original; en cambio, con el método ponderado se utiliza la matriz derivada inmediatamente anterior.

Retornemos a la segunda matriz derivada, donde se observa que el valor de mayor similitud es el que posee B.ret. y B.chil., dando origen a un tercer núcleo a un valor de similitud de 1,01.

La tercera matriz derivada considera los promedios de similitud existentes entre:

- a) cada núcleo comparado con los restantes núcleos, y
- b) cada núcleo comparado con las restantes OTU.

El valor de mayor similitud es el que posee el par B.fol.-B.schic. y B.bon. Esto significa que B.bon. se une a dicho núcleo originando un grupo a un valor de similitud de 1,12.

La cuarta matriz derivada contiene los valores de similitud de los núcleos B.arb.-B.car., B.chil.-B.ret. y del grupo B.bon.-B.fol.-B.schic. entre sí y las restantes OTU.

Como se observa, el valor de mayor similitud es el de 1,16 entre el núcleo B.chil.-B.ret. y el grupo B.bon.-B.fol.-B.schic.; ambos se fusionan en un gran grupo.

La quinta matriz derivada considera el valor de similitud entre el núcleo B.arb.-B.car. con respecto al grupo B.chil.-B.ret.-B.bon.-B.fol.-B.schic. y de ellos con B.sar. Siendo 1,49 el valor de mayor similitud. B.sar. pasa a formar parte del grupo constituido por B.chil.-B.ret.-B.bon.-B.fol.-B.schic.

La sexta y última matriz derivada nos da el valor 1,62, al que se une el núcleo B.arb.-B.car. y el grupo constituido en el paso anterior.

La técnica del ligamiento promedio (UPGMA) fue aplicada también sobre la matriz de similitud de correlación (Fig. 9b) y se obtuvo el fenograma que se observa en la figura 18.

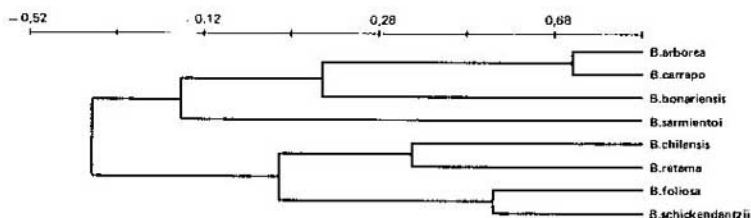


Fig. 18. Fenograma de ocho OTU del género *Bulnesia* construido por la técnica del ligamiento promedio (UPGMA) sobre la base de la matriz de similitud del coeficiente de correlación.

#### MEDICION DE LA DISTORSION

Si se examinan con atención las técnicas de construcción de un fenograma, se comprenderá que es imposible que el fenograma sea un reflejo exacto de la matriz de similitud. Algunas de las relaciones de simi-

litud serán, necesariamente, distorsionadas al realizar una representación bidimensional. Se han propuesto varias técnicas para medir el grado en que el fenograma representa los valores de la matriz de similitud. La técnica más conocida es la del coeficiente de correlación cofenética, establecido por Sokal y Rohlf (1962). Consiste en construir una nueva matriz de similitud a partir de los valores del fenograma y se la denomina "matriz cofenética".

Dispondremos, entonces, de la matriz de similitud que dio origen al fenograma y de la matriz cofenética que representa el fenograma. Se computa el coeficiente de correlación del momento-producto entre ambas matrices. Una alta correlación entre ellas es señal de escasa distorsión. Generalmente, los valores oscilan entre 0,6 y 0,95. Se ha demostrado empíricamente (Sneath y Sokal, 1973) que los valores superiores a 0,8 indican una buena representación de la matriz de similitud por parte del fenograma y que la técnica del ligamiento promedio es la que origina menor distorsión.

El coeficiente de correlación cofenética es una medida de la distorsión interna de la técnica, pero no una evaluación de la clasificación que se construirá a partir del fenograma (Parris, 1969; Rohlf, 1970).

Continuando con el ejemplo de *Bulnesia*, tomamos la matriz de similitud obtenida a partir del coeficiente de distancia y el fenograma resultante de esa matriz empleando la técnica del ligamiento promedio. El valor de similitud entre *B. arborea* y *B. retana* en la matriz de distancia es de 1,58 y en el fenograma esa relación está a nivel de 1,62. Por lo tanto, en la matriz cofenética la relación entre *B.arb.*-*B.ret.* será de 1,62. Si volcamos en la matriz cofenética todos los valores de similitud entre pares de OTU tal como están representados en el fenograma, podremos calcular la distorsión aplicando, entre ambas matrices, el coeficiente de correlación cofenética. La figura 19 muestra las dos matrices y el fenograma, junto al valor del coeficiente de correlación cofenética (CCC) obtenido en este caso (0,90).

65

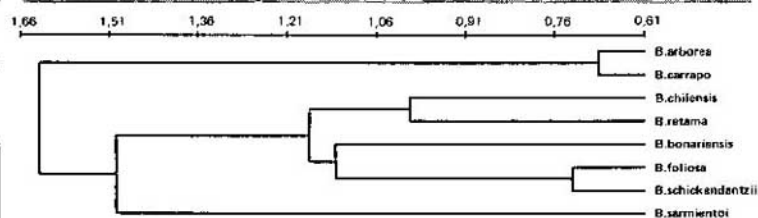
Los valores del coeficiente de correlación cofenética para los restantes fenogramas publicados en esta monografía son los siguientes:

- Para el fenograma de OTU de la figura 13, obtenido a partir de la matriz de similitud de distancia y empleando la técnica de ligamiento simple: 0,82.
- Para el fenograma de OTU de la figura 15, obtenido a partir de la matriz de similitud de distancia y utilizando la técnica del ligamiento completo: 0,89.
- Para el fenograma de OTU de la figura 17, obtenido a partir de la matriz de similitud de distancia y empleando la técnica del ligamiento promedio: 0,90.
- Para el fenograma de OTU de la figura 18, obtenido a partir de la matriz de similitud de correlación y usando la técnica del ligamiento promedio: 0,89.
- Para el fenograma de caracteres de la figura 20, obtenido a partir de la matriz de similitud de correlación y utilizando la técnica del ligamiento promedio: 0,82.

### Matriz de Similitud / distancia

	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.	0,88	0						
B.chil.	1,65	1,80	0					
B.bon.	1,19	1,31	1,22	0				
B.ret.	1,58	1,72	1,01	1,21	0			
B.fol.	1,59	1,66	1,18	1,12	0,99	0		
B.schic.	1,70	1,84	1,23	1,13	1,13	0,73	0	
B.sar.	1,67	1,74	1,57	1,52	1,53	1,27	1,55	0

### Fenograma / Ligamiento Promedio



### Matriz Cofenética

	B.arb.	B.car.	B.chil.	B.bon.	B.ret.	B.fol.	B.schic.	B.sar.
B.arb.	0							
B.car.	0,68	0						
B.chil.	1,62	1,62	0					
B.bon.	1,62	1,62	1,16	0				
B.ret.	1,62	1,62	1,01	1,16	0			
B.fol.	1,62	1,62	1,16	1,12	1,16	0		
B.schic.	1,62	1,62	1,16	1,12	1,16	0,73	0	
B.sar.	1,62	1,62	1,49	1,49	1,49	1,49	1,49	0

Fig. 19. Coeficiente de correlación cofenética: matriz de similitud original (distancia) entre ocho OTU del género *Eulnesia*, fenograma obtenido a partir de ella y matriz cofenética obtenida a partir del fenograma; CCC es el valor del coeficiente de correlación cofenética.

### LOS CARACTERES COMO OTU: TÉCNICA R

Como señalamos en el capítulo 4, es posible examinar la matriz básica de datos asociando los caracteres en vez de las OTU. En otras palabras, considerar a los caracteres como OTU y a las OTU como caracteres. La técnica R, como ha dado en llamarse a este examen de la matriz básica de datos, se aplica con el propósito de establecer cuáles caracteres forman grupos y, por lo tanto, están altamente correlacionados y cuáles caracteres son relativamente independientes entre sí.

Este tipo de análisis puede generar ideas e hipótesis acerca del origen, valor selectivo y patrones de variación de los caracteres. Incluso, se logra determinar posibles complejos adaptativos a partir de los grupos formados.

Si tomamos el ejemplo de *Bulnesia* y retornamos a la matriz básica de datos original (capítulo 4), podemos aplicar un coeficiente de correlación entre cada par posible de caracteres (columnas en este caso).

Habremos obtenido, entonces, una matriz de similitud de 43 x 43 (demasiado extensa para reproducirla aquí) mediante la aplicación del coeficiente de correlación del momento-producto.

Sobre la matriz de similitud resultante aplicamos la técnica del análisis de agrupamientos para construir fenogramas por el ligamiento promedio.

En la figura 20 se aprecia el fenograma que agrupa a los caracteres.

#### INTERPRETACION DE LOS FENOGRAMAS

La interpretación de un fenograma es una operación sencilla. Visualmente, se reconocen primero, los grandes grupos, es decir, los que se han originado a bajos niveles de similitud. Luego, se analizan dichos grupos separándolos en subgrupos, conjuntos y subconjuntos hasta llegar a los núcleos que representan la máxima similitud hallada en los organismos en estudio.

En el fenograma de la figura 17 (técnica Q, coeficiente de distancia, ligamiento promedio), se reconoce un amplio grupo constituido por *B. retama*, *B. chilensis*, *B. bonariensis*, *B. foliosa*, *B. schickendantzii* y *B. sarmentoi* y un núcleo relativamente aislado del grupo anterior y formado por *B. arborea* y *B. carrapa*.

Dentro del amplio grupo se observa una OTU aislada de las demás *B. sarmentoi* y dos subgrupos: el primero constituido por *B. chilensis* y *B. retama* y el segundo, por *B. bonariensis*, *B. foliosa* y *B. schickendantzii*. Dentro de este último subconjunto *B. foliosa* y *B. schickendantzii* se encuentran estrechamente asociados.

Examinemos ahora un fenograma algo más complicado, como es el de la figura 20 (técnica R, coeficiente de correlación, ligamiento promedio). Este fenograma considera a los caracteres como OTU y a diferencia del fenograma anterior no intenta establecer relaciones de similitud entre OTU, sino determinar el grado de asociación entre caracteres; para deducir a partir de esa asociación: el origen, el valor selectivo y los patrones de variación de los caracteres.

A un bajo nivel de similitud se originan dos grandes grupos. El primero está compuesto por los caracteres 1, 38, 39, 41, 43, 4, 16, 5, 11, 6, 21, 26, 22, 25, 23, 8, 13, 12, 40, 2, 19, 29, 33, 7, 17, 32, 30, 35, 18, 20, 34, 10 y el segundo, por los caracteres número 3, 27, 31, 36, 42, 15, 9, 24, 14, 28 y 37.

En el primer gran grupo encontraremos dos subgrupos: uno constituido por los caracteres 1, 38, 39, 41, 43, 4, 16, 5, 11, 6, 21, 26, 22, 25, 23, 8, 13, 12, 40, 2, 19, 29 y 33 y el otro por los caracteres 7, 17, 32, 30, 35, 18, 20, 34 y 10. El primer subgrupo, a su vez, se divide en tres conjuntos y una OTU aislada:

primer conjunto: 1, 38, 39, 41, 43

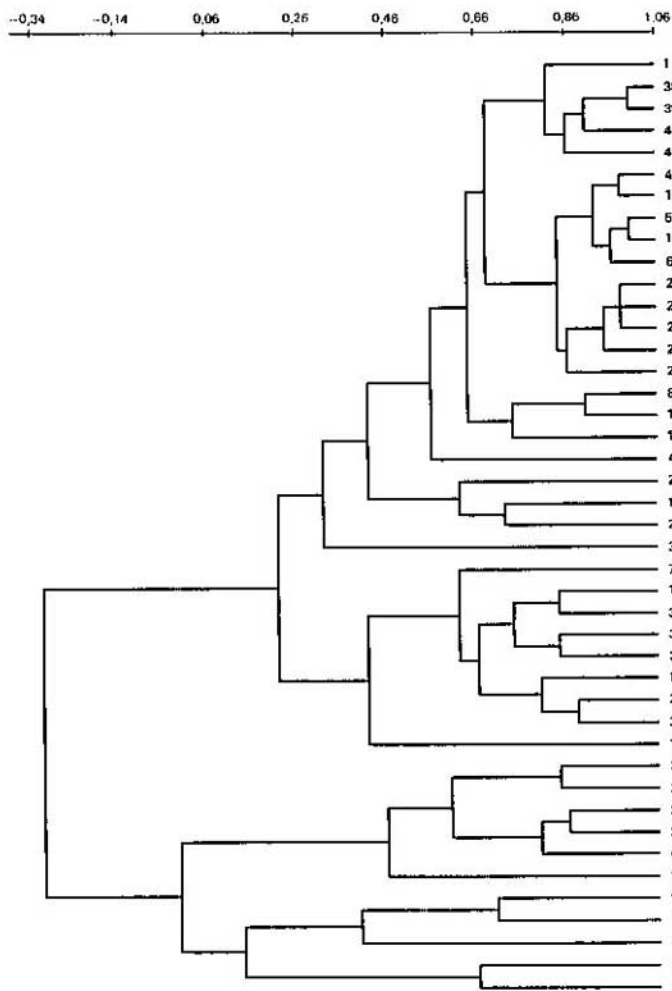


Fig. 20. Técnica R, fenograma de los 43 caracteres del estudio sobre el género *Bulnesia*: obtenido por el método del ligamiento promedio a base de la matriz de similitud del coeficiente de correlación entre caracteres, 43 x 43.

segundo conjunto: 4, 16, 5, 11, 6, 21, 26, 22, 25, 23

tercer conjunto: 8, 13, 12, 40, 2, 19, 29

OTU aislada: 33

El primer conjunto se divide en un subconjunto formado por los caracteres 38 (longitud del fruto), 39 (ancho del fruto), 41 (longitud del carpóforo), 43 (longitud de la semilla) con un núcleo (38, longitud del fruto y 39, ancho del fruto) y un carácter aislado, el número 1, (hábito). Se continúa así el análisis de los otros grupos.



Esta interpretación es una descripción del fenograma y no debe confundirse con el paso final (formulación de las inferencias), en el que entran en juego el juicio del taxónomo y los conocimientos que éste posee sobre el grupo. En esta etapa final habrá que explicar el porqué de los agrupamientos. Es decir, formular hipótesis sobre las causas biológicas y metodológicas a que se deben.



## LA BUSQUEDA DE LA ESTRUCTURA TAXONOMICA: METODOS DE ORDENACION

Para encontrar el patrón de relaciones entre la totalidad de las OTU se utilizan también los métodos de ordenación.

Podemos imaginar un espacio multidimensional donde ubicar a las OTU y donde cada dimensión represente un carácter, es decir, tantas dimensiones como caracteres. Los métodos de ordenación reducen, sin gran pérdida de información, el número de dimensiones y de esa manera facilitan la representación de las OTU y sus relaciones en función de los caracteres empleados.

A diferencia del análisis de agrupamientos, estos métodos no trazan límites en el espacio que separen a grupos, tarea que corresponde al investigador al interpretar los resultados. Las relaciones entre las OTU están reflejadas en la posición en que se disponen en ese espacio. Cuanto más cerca se encuentran entre sí dos OTU más estrechamente relacionadas están.

Se han propuesto numerosas técnicas para la ordenación (Gower, 1966; Whittaker, 1973; Sneath y Sokal, 1973; Benzecri, 1980), entre las cuales procede mencionar: el análisis de componentes principales, el análisis de coordenadas principales, el análisis factorial múltiple y el análisis factorial de correspondencias.

71

Estas técnicas pueden clasificarse según la manera de proceder. Algunas siguen el camino de las técnicas Q, que parten de una matriz de similitud entre las OTU, por ejemplo, el análisis de coordenadas principales. Otras se ciñen a las técnicas R, que parten de una matriz de similitud entre caracteres; ejemplo de esta última técnica es el llamado análisis de componentes principales. Sin embargo, a pesar de que el punto de partida es una matriz de similitud entre caracteres, la representación gráfica final se refiere a las relaciones entre las OTU.

A continuación, presentaremos la técnica del análisis de componentes principales por ser una de las más difundidas.

### ANALISIS DE COMPONENTES PRINCIPALES

La técnica del análisis de componentes principales tiene su origen en los trabajos realizados a principios de este siglo por Karl Pearson. Pero fue Hotelling (1933) quien consolidó su uso para representar, según un modelo lineal, un conjunto numeroso de caracteres mediante un número reducido de variables hipotéticas, llamadas componentes principales. Estos componentes principales no están correlacionados entre sí y, por lo tanto, se interpretan independientemente unos de otros. El número de dichos componentes depende del número de caracteres originales y el número máximo posible es igual o menor al número de estos últimos.

Cada componente contiene una parte de la variabilidad total de los caracteres. El primer componente es el que contiene la mayor variabilidad. De la variabilidad restante, el segundo componente es el que incluye más información. El tercer componente posee la mayor variabi-

lidad no contenida en los componentes anteriores. Así se continúa hasta que toda la variabilidad ha sido distribuida diferencialmente entre los componentes.

Cada componente contiene información de todos los caracteres pero en diferentes proporciones.

Examinaremos en términos generales los principios básicos de esta técnica. Para un examen más detallado y matemáticamente más riguroso, consúltese a: Seal, 1964; Harman, 1967; Morrison, 1967; Blackith y Reyment, 1971; Clifford y Stephenson, 1975.

Los fundamentos subyacentes del análisis de componentes principales pueden visualizarse con un modelo geométrico constituido por dos caracteres, es decir, dos dimensiones.

En la figura 21a se han ubicado ocho OTU en el espacio bidimensional originado por dos caracteres (1 y 2). De la figura queda claro que los dos caracteres están correlacionados.

La dispersión de las OTU en relación con cualquiera de los dos ejes depende de la escala utilizada. Por lo tanto, teniendo en cuenta las operaciones matemáticas que han de realizarse, es necesario que ambos caracteres estén expresados en la misma unidad de medida. Cuando esto no suceda es conveniente estandarizar los caracteres.

Una consecuencia importante de la estandarización es que las OTU están referidas ahora a un nuevo par de ejes ortogonales que se cortan en un punto que corresponde al valor medio de cada carácter. Cada OTU tendrá un nuevo par de coordenadas definidas en función de unidades de desviación estándar. La figura 21b grafica la distribución de las OTU en función del par de ejes estandarizados. Para hallar la máxima separación entre las OTU en una sola dimensión, es necesario proyectarlas sobre una línea. Puede pensarse que esto se logrará proyectando las entidades sobre la abscisa o sobre la ordenada, pero a simple vista se advierte que ninguna de estas líneas es la adecuada. La línea apropiada debe inclinarse entre los dos ejes. Geométricamente, la disposición espacial de las OTU en relación con dos caracteres correlacionados es el de una nube elíptica. La línea buscada coincide pues con el eje mayor de esa elipse y ese es el eje que se deseaba encontrar o primer componente principal. Si queremos determinar ahora la máxima separación entre las OTU en una segunda dimensión, la línea buscada será ortogonal a la anterior y coincidente con el eje menor de la elipse y recibirá el nombre de segundo componente principal (Fig. 21c).

Si tuviésemos tres caracteres, la elipse de la figura 21c se transformaría en un elipsoide y el tercer componente principal estaría representado por el tercer eje del elipsoide.

Si el estudio incluye más de tres caracteres, se necesitarán dimensiones adicionales cuya representación geométrica no puede ser visualizada, pero podrá aplicarse el tratamiento matemático.

Los ejes de la elipse o componentes principales se reconocen, pues las distancias cuadráticas de las OTU con respecto a ellos es mínima. Matemáticamente, este reconocimiento se basa en el cálculo de la matriz de correlación entre los caracteres. También puede basarse sobre una matriz de varianza-covarianza cuando todos los caracteres estén originalmente expresados en la misma unidad de medida y, por lo tanto, sean no estandarizados.

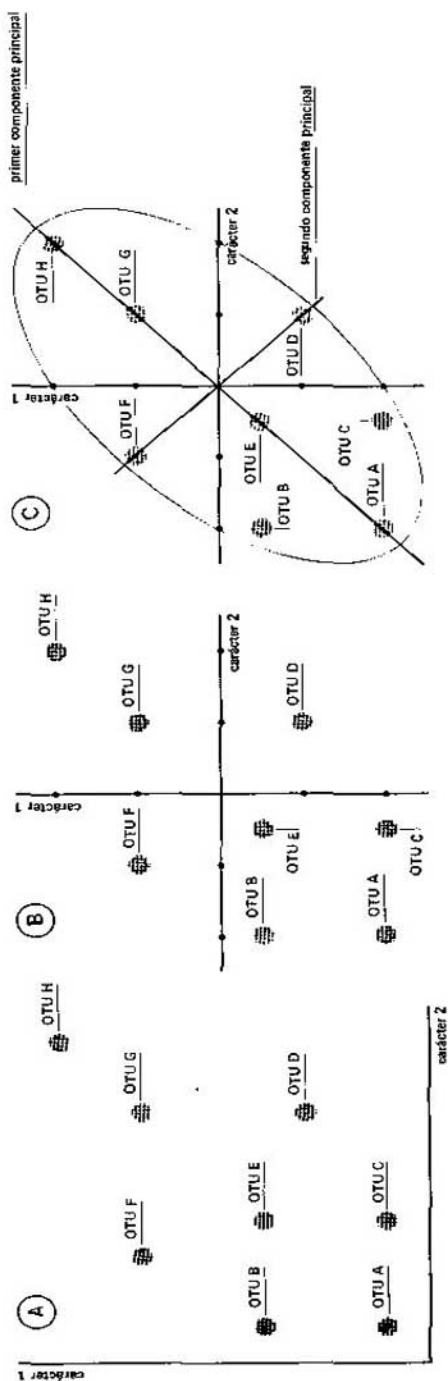


Fig. 21. a) Ocho OTU --A, B, C, D, E, F, G y H-- ubicadas en el espacio bidimensional determinado por los caracteres 1 y 2; b) el espacio está determinado ahora por los caracteres 1 y 2 estandarizados; c) la distribución espacial de las OTU posee una forma elipsoidal; la máxima separación entre ellas está determinada por el eje mayor de la elipse (primer componente principal). La máxima separación entre las OTU en una segunda dimensión coincide con el eje menor de la elipse (segundo componente principal).

A partir de las interrelaciones exhibidas por la matriz de correlación de caracteres y mediante transformaciones matemáticas se construye un nuevo conjunto de variables denominadas componentes principales. Estos componentes son ortogonales entre sí y por lo tanto no están correlacionados.

La contribución individual de un carácter a un componente principal está expresada por el coeficiente de regresión del componente con respecto a ese carácter. Todos los caracteres contribuyen a todos los componentes, pero de manera diferencial, es decir, el carácter 1 puede ser un importante aporte para el primer componente, pero pobre para el segundo.

El cuadrado de la contribución de un carácter para un componente representa la varianza de ese carácter para el citado componente. La sumatoria de las varianzas de todos los caracteres para un determinado componente principal recibe el nombre de eigen-valor o raíz latente.

Los eigen-valores son diferentes para cada componente. El componente con mayor eigen-valor será el primer componente principal, el que le sigue será el segundo y así sucesivamente.

La suma de los eigen-valores constituye la varianza total de las OTU para los caracteres utilizados; por consiguiente, puede establecerse el porcentaje de variación (porcentaje de traza) contenido en cada componente principal según su aporte a esa suma. Cada componente principal es independiente de los otros, es decir, contiene una porción de la variación no expresada en ningún otro componente principal.

#### 74 REPRESENTACION GRAFICA DEL ANALISIS DE COMPONENTES PRINCIPALES

Los resultados de esta técnica se grafican sobre ejes ortogonales que representan los componentes principales que delimitan un espacio bi o tridimensional, según se utilicen dos o tres ejes por vez. Obviamente, es imposible la representación gráfica simultánea de más de tres ejes.

Las OTU se sitúan dentro del espacio delimitado por los componentes según los valores de sus coordenadas con respecto a éstos.

Generalmente, para estas representaciones se eligen los tres primeros componentes, ya que son los que contienen la mayor parte de la variabilidad. No es aconsejable desdeñar *a priori* los otros componentes, pues a pesar de su escaso contenido de variabilidad pueden brindar información importante acerca de la estructura taxonómica del grupo en estudio.

Las combinaciones más comunes para la elaboración de los gráficos bidimensionales son: primer componente con el segundo; primero con el tercero y segundo con el tercero. A veces, se combinan el primero con el cuarto y el segundo con el cuarto. La posición de uno u otro componente en la abscisa o en la ordenada es indistinta.

Los gráficos tridimensionales se basan en los tres primeros componentes.

Las representaciones gráficas deben ir acompañadas de tablas en las que figure la siguiente información acerca de los componentes principales: eigen-valores, porcentaje de traza, acumulación del porcentaje a medida que se extraen componentes y valor de la contribución de cada carácter a cada componente.

La suma de porcentajes contenidos en los ejes graficados da una idea de la cantidad de variación expresada en el dibujo. Por ejemplo, si el primer componente contiene 50% de la variación total y el segundo componente 20%, el gráfico bidimensional de estos componentes expresaría un 70% de la variación total de las OTU.

Generalmente, más del 90% de la variación de las OTU está contenida en los primeros componentes.

Al aplicar esta técnica a nuestro ejemplo de *Bulnesia*, se tomó la matriz básica de datos estandarizada y se calculó el coeficiente de correlación entre caracteres. De la matriz de correlación resultante, de 43 x 43, se extrajeron los componentes principales y los eigen-valores. Los siete primeros componentes extraídos acumularon el 100% de la variación total de las OTU.

La siguiente tabla muestra el eigen-valor de cada uno de los componentes, el porcentaje de la variación total (porcentaje de traza) expresado por ese eigen-valor y la acumulación de los porcentajes a medida que extraemos los componentes hasta llegar al 100%, valor que se alcanza en el séptimo componente.

Componente Principal	Eigen-Valor	Porcentaje de Traza	Acumulación de Porcentaje
Primer componente	19,51	45,39	45,39
Segundo componente	9,34	21,73	67,12
Tercer componente	5,72	13,31	80,43
Cuarto componente	3,39	7,89	88,32
Quinto componente	2,60	6,05	94,37
Sexto componente	1,65	3,85	98,22
Séptimo componente	0,97	2,26	100

75

Los tres primeros componentes expresan más del 80% de la variación y son elegidos para las representaciones gráficas.

La figura 22a corresponde al gráfico bidimensional de los componentes primero y segundo. Esta representación expresa un 45,39% en el primer componente y un 21,73% en el segundo componente, o sea un total de 67,12% de variación expresada en el gráfico. Las OTU están dispuestas en el espacio bidimensional con arreglo a los valores de sus coordenadas en relación con estos dos primeros componentes.

La figura 22b constituye el gráfico bidimensional que representa la combinación del primer y tercer componente. Se aprecia un 58,70% de variación (45,39% del primer componente y 13,31% del tercero).

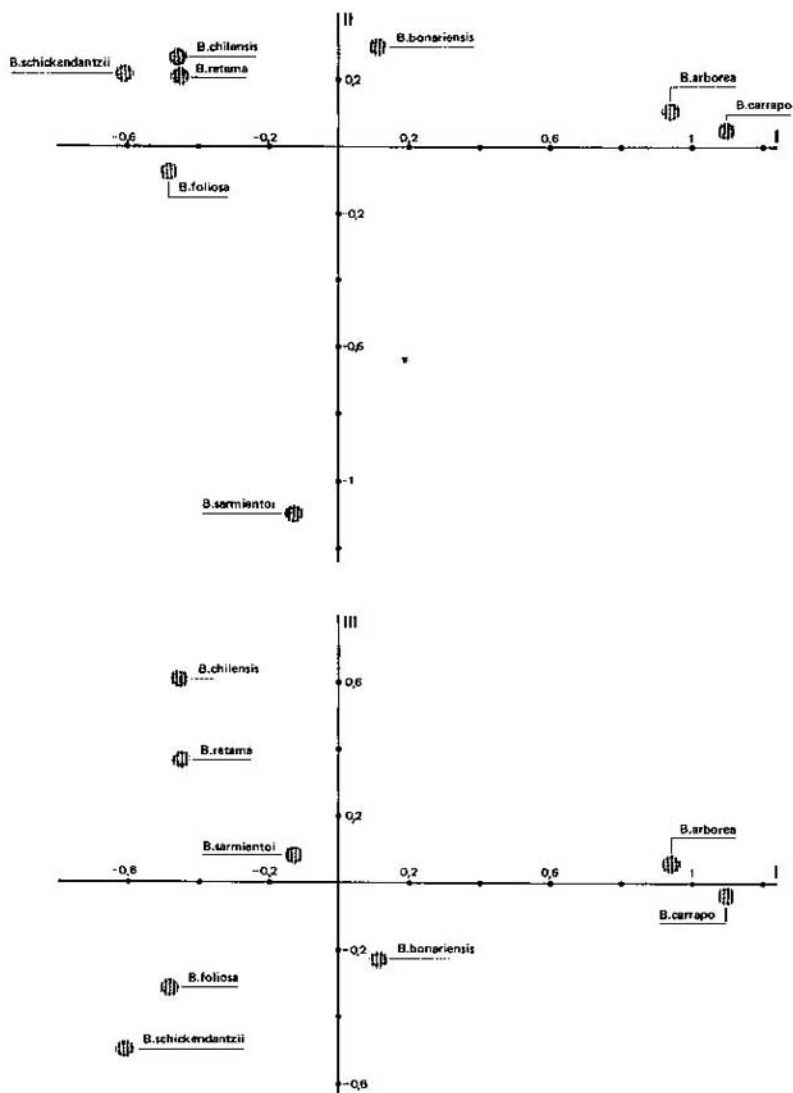


Fig. 22. Análisis de los componentes principales en el género *Bulnesia*. Gráficos bidimensionales con la proyección de las OTU en los tres primeros componentes (I, II y III) tomados de a dos por vez; a) componentes I y II, esta representación expresa 67,12% de la variación total (I = 45,59%, II = 21,73%); b) componentes I y III, esta representación expresa un 58,70% de la variación total (III = 13,31%).

La figura 23 corresponde al gráfico tridimensional de los componentes primero, segundo y tercero. El porcentaje de variación expresado en el gráfico es del 80,43%.



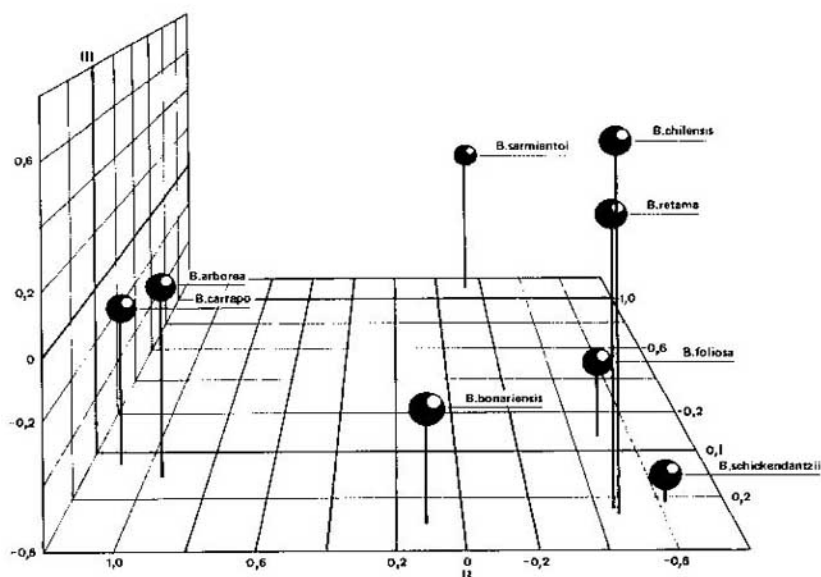


Fig. 23. Análisis de los componentes principales en el género *Bulnesia*. Gráfico tridimensional con la proyección de las OTU en los tres primeros componentes (I, II y III). Esta representación expresa el 80,43% de la variación total.

77

La siguiente tabla nos muestra el aporte de cada uno de los 43 caracteres a los tres primeros componentes principales. Cuanto más alto es el valor de esa contribución (sin importar el signo), mayor es el aporte del carácter al componente.

Contribución de los caracteres a los tres primeros componentes:

Carácter	1er. Componente	2do. Componente	3er. Componente
1. Hábito	0,720	-0,519	0,223
2. Longitud del internodio	0,554	0,137	0,578
3. Diámetro del internodio	0,019	0,349	0,833
4. Longitud de la hoja	0,949	0,058	-0,114

5. Ancho de la hoja	0,937	-0,129	-0,101
6. Longitud del pecíolo	0,848	-0,243	-0,199
7. Número de folíolos	0,358	0,640	-0,219
8. Presencia de peciólulos	0,590	-0,376	0,639
9. Disposición de los folíolos en el raquis	-0,477	-0,732	0,424
10. Pubescencia	0,303	0,277	-0,701
11. Longitud del folíolo	0,922	-0,188	-0,071
12. Ancho del folíolo	0,682	-0,558	-0,052
13. Número de nervaduras primarias del folíolo	0,769	-0,486	0,324
14. Posición de los folíolos terminales	-0,711	-0,200	-0,229
15. Presencia de mucrón en folíolos	-0,274	0,213	0,404
16. Tipo de inflorescencia	0,937	0,099	0,015
17. Longitud del pedúnculo	0,713	0,532	-0,220
18. Longitud del sépalo	0,180	0,931	0,283

19. Ancho del sépalo	0,473	0,546	0,303
20. Color de los pétalos	0,082	0,959	-0,086
21. Longitud del pétalo	0,975	0,084	-0,124
22. Ancho del pétalo	0,987	0,070	0,031
23. Número de nervaduras del pétalo	0,893	0,253	0,155
24. Tipo de estambres	-0,555	-0,659	-0,276
25. Modificación de los estambres	0,913	0,215	-0,133
26. Presencia de gran escama junto al estambre	0,973	0,203	-0,085
27. Presencia de pelos en la base del filamento estaminal	-0,399	0,323	0,821
28. Presencia de una escama suplementaria junto al estambre	-0,369	0,186	-0,565
29. Agrupación de los estambres	0,649	0,458	0,582
30. Longitud del filamento	0,668	0,672	-0,159
31. Longitud de la antera	-0,476	0,728	0,424

32. Longitud de la escama	0,677	0,528	-0,189
33. Presencia de ápice lacinia- do en la esca- ma estaminal	0,337	-0,432	-0,094
34. Número de carpelos	0,209	0,806	-0,401
35. Curvatura del estilo	0,446	0,745	-0,221
36. Número de óvulos por carpelo	-0,614	0,333	0,580
37. Pubescencia del fruto	-0,505	0,094	-0,702
38. Longitud del fruto	0,829	-0,483	0,021
39. Ancho del fruto	0,823	-0,494	-0,002
40. Desarrollo del carpóforo	0,639	-0,230	0,352
41. Longitud del carpóforo	0,864	-0,314	-0,036
42. Forma de la semilla	-0,802	0,375	0,101
43. Longitud de la semilla	0,674	-0,405	0,024

#### INTERPRETACION DE LOS RESULTADOS DEL ANALISIS DE COMPONENTES PRINCIPALES

Esta técnica permite determinar dos aspectos del grupo en estudio: las relaciones entre las OTU y el valor discriminatorio de los caracteres con respecto a las relaciones establecidas en el primer punto.

Las relaciones entre OTU se establecen por su proximidad en el espacio delimitado por los componentes: cuanto más próximos se encuentran, más relacionados están.

El análisis de la figura 22a muestra estrecha relación entre *B. arborea* - *B. carrapo*, al igual que el par *B. chilensis* - *B. retama*. *B. schickendantzii* está cercano a este último par. Las demás OTU están ligeramente aisladas con excepción de *B. sarmentoi* que se encuentra significativamente alejada de las demás.

En la figura 22b algunas relaciones se mantienen con respecto a la figura 22a, pero se observa que aparecen los siguientes cambios: una mayor relación entre *B. schickendantzii* y *B. foliosa* y un acercamiento de *B. sarmentoi* a las OTU restantes.

En la figura 23 se han integrado en un solo dibujo las relaciones de las figuras 22a y b.

Si analizamos el segundo aspecto del análisis de componentes principales, referido a los caracteres, es importante observar cuáles de ellos son eficaces en la discriminación entre las OTU. Con este fin, se seleccionaron de la tabla anterior los cinco caracteres que tienen el más alto valor de contribución, independientemente del signo que posean. El signo denota el tipo de correlación existente entre los caracteres.

El primer componente es un buen discriminante entre los grupos conformados por *B. arborea* - *B. carrapo* y *B. chilensis* - *B. retama* - *B. foliosa* - *B. schickendantzii* y los caracteres más importantes son:

22. ancho del pétalo;
21. longitud del pétalo;
26. presencia de gran escama junto al estambre;
4. longitud de la hoja, y
5. ancho de la hoja

El segundo componente es un buen discriminante entre *B. sarmentoi* y las restantes OTU. Los caracteres más importantes son:

20. color de los pétalos;
18. longitud del sépalo;
34. número de carpelos;
35. curvatura del estilo, y
9. disposición de los folíolos en el raquis.

El tercer componente es un buen discriminante entre *B. chilensis*, *B. retama* y *B. schickendantzii* y los caracteres más importantes son:

3. diámetro del internodio;
27. presencia de pelos en la base del filamento estaminal;
37. pubescencia del fruto;
10. pubescencia, y
8. presencia de peciólulos.

#### RELACION ENTRE LAS TÉCNICAS Y CRITERIOS PARA SELECCIONARLAS

Las técnicas, ya sean de análisis de agrupamientos o de ordenación, permiten obtener diferentes resultados gráficos y, en muchos casos, difieren en las relaciones que sugieren. Por ejemplo, Rohlf (1968) ha demostrado que el análisis de componentes principales refleja con mayor fidelidad relaciones entre grupos formados a bajos niveles de similitud.

En cambio, es menos fiel en reflejar relaciones muy estrechas. El caso inverso se da con las técnicas del análisis de agrupamientos donde las relaciones estrechas son distorsionadas en menor proporción que las conexiones de baja similitud del fenograma.

Esto nos lleva a concluir que en el procesamiento de los datos es aconsejable el uso de más de una técnica. En lo posible una de análisis de agrupamientos y una de ordenación, con el fin de minimizar los efectos metodológicos. Las conclusiones finales deben surgir de una complementación de técnicas diferentes.

Es tan amplia la gama de técnicas propuestas que el investigador tendrá siempre que afrontar el problema de seleccionar cuáles utilizar entre muchas posibles. Con frecuencia, la elección depende de criterios pragmáticos que pueden resumirse en los cuatro siguientes:

1. Programas de computadora al alcance del investigador. Muchos centros de cómputo no poseen programas relacionados con procesos de taxonomía numérica. Por esto, el usuario debe ante todo localizar un centro donde se encuentren esos programas.

2. Tipos de datos. En los capítulos anteriores hemos señalado que los distintos tipos de datos condicionan el uso de técnicas determinadas. Por ejemplo, si todos los caracteres de la matriz básica de datos son de tipo doble-estado, no puede utilizarse el análisis de componentes principales.

3. Costo del uso de la computadora. Este factor limita el número de técnicas a aplicar y descarta el uso de otras en determinadas circunstancias. Por ejemplo, la ordenación de una gran cantidad de OTU y caracteres resulta prohibitiva por su costo.

4. Coherencia entre el objetivo de la investigación y el resultado de las técnicas a aplicar. Es importante conocer la capacidad de la técnica para conseguir el propósito deseado. Por ejemplo, para identificar los caracteres discriminitorios en determinado grupo, la técnica aconsejable es el análisis de componentes principales.

Estos cuatro criterios son, obviamente, insuficientes. Pero por el momento es imposible determinar en forma universal cuáles son las técnicas más adecuadas.

La diferencia de resultados que se obtienen con diversas técnicas, causa principal del problema planteado en esta sección, es parte de una cuestión más amplia e importante: la congruencia taxonómica. A este tema dedicaremos el capítulo siguiente.

## LA COMPARACION DE CLASIFICACIONES: CONGRUENCIA TAXONOMICA

No es raro el caso de un grupo de organismos para el que se haya propuesto más de una clasificación, ya sea porque se toman en cuenta diferentes caracteres o porque, si bien se consideran los mismos caracteres, se utilizan diferentes técnicas clasificatorias, o por ambos motivos. Esta conflictiva situación, que no es sólo propia de la taxonomía numérica, ha obligado a investigar las causas de las similitudes y diferencias entre clasificaciones de un mismo conjunto de organismos, lo que da lugar al concepto de congruencia taxonómica.

Congruencia taxonómica es el grado de correspondencia entre diferentes clasificaciones de un mismo conjunto de organismos. Dos clasificaciones de un mismo conjunto de organismos son perfectamente congruentes si postulan los mismos agrupamientos (Farris, 1971; Micevich, 1978).

La taxonomía numérica ha sido un factor importante en la formulación del concepto de congruencia taxonómica, pues los resultados obtenidos por técnicas precisas y claramente explícitas permiten mejores comparaciones.

Un ejemplo relacionado con el género *Bulnesia*, en el cual se efectuó un estudio multidisciplinario (Hunziker, 1980), incluye el análisis de:

- a) la morfología (cuyos resultados sirvieron de ejemplo a los capítulos anteriores);
- b) las proteínas seminales mediante electroforesis (Comas y colaboradores, 1979), y
- c) sus compuestos fenólicos mediante cromatografía (Poggio, 1980).

Los datos obtenidos por los tres procedimientos fueron tratados en forma independiente utilizando técnicas numéricas. Así se compararon fenogramas obtenidos con cada procedimiento. La figura 24 muestra tres fenogramas, obtenidos todos ellos a partir de un coeficiente de distancia y la técnica del ligamiento promedio.

El fenograma basado en las proteínas muestra dos OTU pertenecientes a la especie *B. arborea*, uno de las OTU representa a poblaciones de Venezuela y la otra a poblaciones de Colombia. Por falta de material en este fenograma, no se encuentra representada *B. chilensis*.

El problema de la congruencia taxonómica es un espejo de cuestiones claves: respecto a los caracteres, cuestiones como selección, significación y factores biológicos representados por ellos; respecto a la metodología, predictividad y estabilidad de las clasificaciones establecidas con diferentes métodos.

En este capítulo presentaremos los experimentos taxonómicos posibles sobre congruencia taxonómica, una lista de coeficientes y otros medios de evaluación de la congruencia taxonómica y algunos resultados

experimentales sobre el tema. Trataremos también de delinear un marco teórico apropiado para los resultados obtenidos.

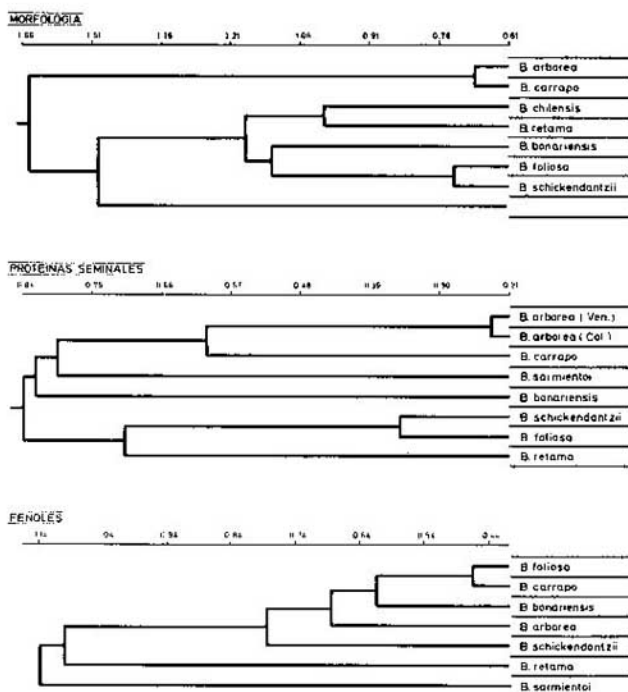


Fig. 24. Congruencia taxonómica en el género *Bulnesia*. Fenogramas (distancia, UPGMA) obtenidos a partir de tres tipos diferentes de caracteres: morfológicos, proteínas seminales y fenoles.

#### TIPOS DE EXPERIMENTOS

El problema de la congruencia debe ser encarado de manera experimental. Se toma un grupo, se provocan deliberadamente variaciones clasificatorias y se observan e interpretan los resultados. Los dos tipos básicos de experimentos taxonómicos pueden realizarse:

- variando la metodología, y
- variando la fuente de caracteres.

El primer tipo de experimento consiste en estudiar la congruencia entre clasificaciones basadas en los mismos caracteres, pero estructuradas mediante la aplicación de diferentes técnicas clasificatorias (Fig. 25). Un simple experimento consistiría en clasificar a las Pteridófitas conforme a caracteres químicos utilizando una técnica clasificatoria en particular. Luego, y con los mismos caracteres, se utilizaría una técnica diferente a la anterior y se observaría, finalmente, el grado de correspondencia entre esas dos clasificaciones.



# POSIBLE EXPERIMENTO SOBRE CONGRUENCIA TAXONOMICA 1

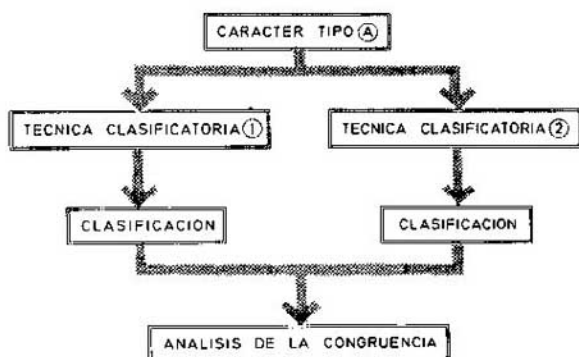


Fig. 25. Tipo de experimento sobre congruencia taxonómica 1, donde varía la técnica clasificatoria utilizada.

En el segundo tipo de experimento (Fig. 26) puede estudiarse la congruencia entre clasificaciones basadas en diferentes órganos o regiones del cuerpo (por ejemplo, hoja *versus* flor), diferentes manifestaciones en organismos dimórficos (macho *versus* hembra), diferentes etapas del ciclo vital (larvas *versus* adultos), diferentes tipos de caracteres (etológicos *versus* ecológicos), entre otras posibilidades. Tomando el ejemplo precedente podría clasificarse las Pteridófitas según sus caracteres químicos; luego con el mismo método se clasificarían conforme a sus caracteres morfológicos y, finalmente, se observaría el grado de correspondencia entre esas dos clasificaciones.

85

# POSIBLE EXPERIMENTO SOBRE CONGRUENCIA TAXONOMICA 2

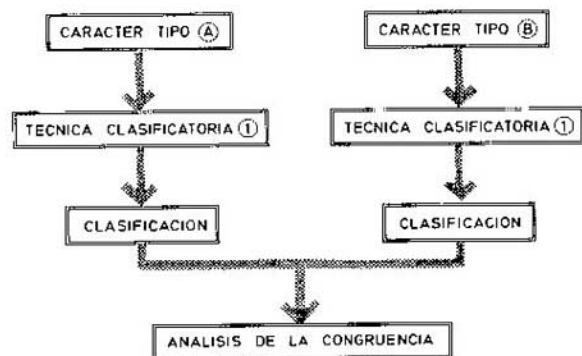


Fig. 26. Tipo de experimento sobre congruencia taxonómica 2, donde varía la fuente de caracteres.

Existe un tercer tipo de experimento (Fig. 27), que combina ambos tipos básicos, y que consiste en estudiar la congruencia entre dos o más clasificaciones de diferentes caracteres utilizando una determinada

técnica. En seguida, a los mismos datos se aplica otra técnica y se observa si la congruencia obtenida con la primera técnica es igual o diferente a la obtenida con la segunda. Por ejemplo, clasificar a la Pteridófitas con arreglo a sus caracteres químicos y después clasificarlas utilizando la misma técnica, pero según sus caracteres morfológicos. Luego, el experimento se repite con los mismos caracteres, pero cambiando la técnica, para finalmente comparar la congruencia obtenida con cada una de las técnicas.

Con el primer tipo de experimento (variando el método) se procura determinar cuáles son las técnicas clasificatorias que permiten obtener resultados más semejantes entre sí.

#### POSIBLE EXPERIMENTO SOBRE CONGRUENCIA TAXONÓMICA 1

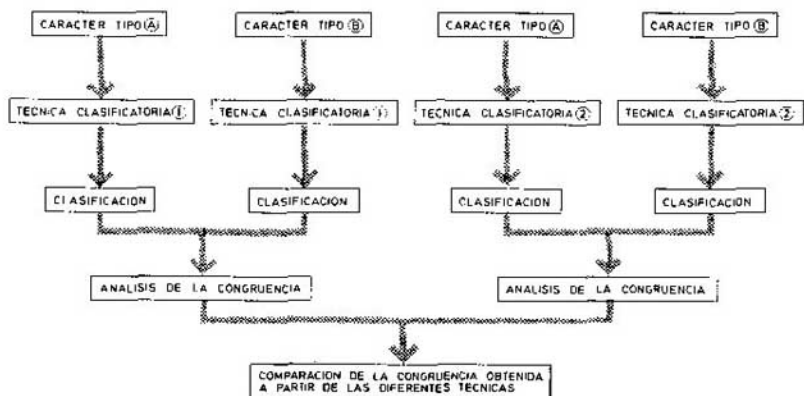


Fig. 27. Tipo de experimento sobre congruencia taxonómica 3, donde varía la fuente de caracteres y la técnica clasificatoria.

Lo que se persigue con el segundo tipo (variando la fuente de caracteres) es encontrar respuesta a preguntas como ésta: ¿Cuál es el grado de interdependencia causal entre caracteres morfológicos y caracteres químicos?

Con el tercer tipo (variando los caracteres y el método) se trata de determinar si es posible que la morfología y la química logren resultados congruentes utilizando una determinada técnica clasificatoria, e incongruentes con otra.

#### **EVALUACION DE LA CONGRUENCIA TAXONÓMICA**

Se ha recurrido a numerosos medios para evaluar la congruencia taxonómica. Los más difundidos configuran cuatro grandes grupos, tres de los cuales se aplican a clasificaciones numéricas y uno a clasificaciones no numéricas (véase la Fig. 28).

En dos de los grupos presentados la evaluación se realiza con coeficientes y mediciones precisas; en los otros dos, mediante la comparación visual de los resultados.

## TIPOS DE EVALUACIÓN DE LA CONGRUENCIA

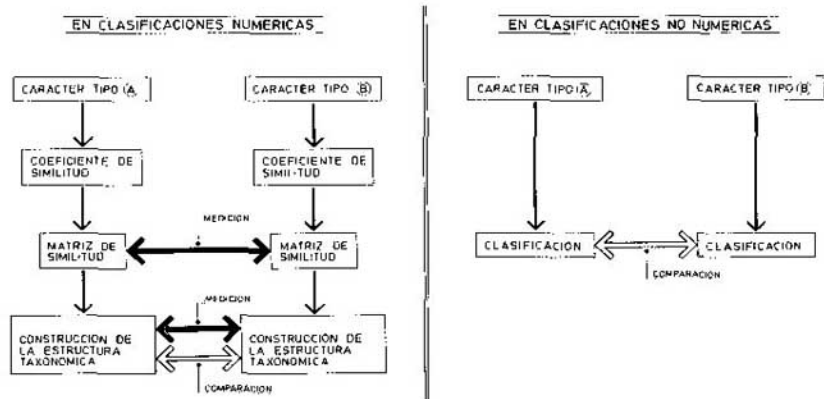


Fig. 28. Tipos de evaluación de la congruencia.

A continuación se describen los tres grupos relativos a clasificaciones numéricas.

1. Se mide la coincidencia entre las matrices de similitud.

Pueden aplicarse dos coeficientes:

- a) coeficiente de correlación (Sneath y Sokal, 1973), y
- b) coeficiente de discordancia (Edmonds, 1978).

2. Se mide la coincidencia entre las estructuras taxonómicas derivadas de las matrices de similitud. Pueden aplicarse dos coeficientes:

- a) coeficiente de distorsión (Farris, 1973), y
- b) índice de consenso (existen varias versiones de este índice, Adams, 1972; Micevich, 1978).

3. Se comparan visualmente, sin ningún tipo de medición, las estructuras taxonómicas obtenidas de las matrices de similitud.

Los taxónomos no han llegado a un acuerdo respecto al tipo de evaluación más adecuado. Por lo tanto, y con el objeto de ilustrar el tema, nos referiremos al coeficiente de correlación, el más utilizado y el más sencillo de los propuestos.

### ALGUNOS RESULTADOS EXPERIMENTALES

En los Cuadros III y IV se consignan los resultados de trabajos en los que se utilizó el coeficiente de correlación. La información se presenta en el orden siguiente: grupo estudiado, caracteres comparados, número de caracteres en cada conjunto, coeficientes de similitud usados y el valor de la correlación entre las matrices de similitud.

obtenidas. La información ha sido tomada de Sneath y Sokal (1973); Hsiao (1975); Roback y Moss (1978). La mayor parte de los experimentos son del tipo 3, es decir, varían los caracteres y el método.

La congruencia está expresada por el valor de la correlación entre las matrices de similitud (última columna de los Cuadros III y IV). El valor máximo posible es 1.

Rohlf (1965) ha efectuado pruebas de muestreo al azar para establecer la significación de las matrices de correlación de algunos de los estudios presentados en los Cuadros III y IV. Los resultados obtenidos por Rohlf demuestran que los valores de correlación entre subconjuntos de caracteres de diferentes tipos, o diversas partes del cuerpo, o distintos estadios vitales son significativamente menores que los valores que se obtienen si los subconjuntos se forman al azar. Por ejemplo, si se comparan 20 caracteres morfológicos con 30 caracteres químicos y luego de los 50 caracteres totales se forman al azar dos subconjuntos uno de 20 y otro de 30 la congruencia, según Rohlf, será mayor entre los subconjuntos tomados al azar. Esto significa que cada tipo de carácter expresa relaciones, en parte, distintas de los otros tipos y, por lo tanto, parte de la incongruencia observada es el resultado de la diferencia entre los tipos de caracteres.

#### EXPLICACION DE LOS RESULTADOS

Respecto al primer tipo de experimento, en el capítulo 7: Relación entre técnicas y criterios para seleccionarias, se ha señalado que diferentes técnicas poseen diferentes propiedades y que, a veces, permiten obtener diferentes resultados; por esto insistimos en la complementación de las técnicas. Nos referiremos ahora a los otros dos tipos de experimentos.

Los resultados indican que nunca se obtiene una completa congruencia, pero tampoco una completa incongruencia. Los valores varían según el grupo estudiado, los caracteres analizados, el método aplicado y la forma en que se evalúa la congruencia.

En teoría, debería existir una perfecta congruencia y existen dos argumentos principales al respecto.

1. Sokal y Sneath (1963) formularon la hipótesis de la no especificidad de los caracteres, conforme a la cual no existen distintas clases de genes que afecten exclusivamente a una clase de caracteres o que afecten a regiones especiales del organismo, como la cabeza, el esqueleto o las hojas. En otras palabras, un genotipo no puede ser subdividido en clases de genes, de tal manera que cada una de éstas influya exclusivamente en una sola clase de caracteres fenotípicos. En consecuencia, no cabría esperar, por ejemplo, que un conjunto de genes controle caracteres morfológicos y otro conjunto de genes, independiente del anterior, controle caracteres químicos. De ahí que una amplia muestra del genotipo podría conseguirse de un gran número de caracteres de una parte restringida del fenotipo.

Esta hipótesis se basa en la idea de que cada carácter es en realidad el producto final de muchos genes y que entre éstos existen numerosas interrelaciones derivadas de fenómenos como el ligamiento, la pleiotropía y la epistasia.

Si así fuera debería haber congruencia entre las clasificaciones obtenidas a partir de diferentes tipos de caracteres.

**Cuadro III. Congruencia Taxonómica Medida con el Coeficiente de Correlación.** (La información fue tomada de Sneath y Sokal, 1973; Hsiao, 1975; Roback y Moss, 1978. Continúa en el Cuadro IV).

1 CONGRUENCIA ENTRE CLASIFICACIONES MEDIDA CON EL COEFICIENTE DE CORRELACION				
ORGANISMOS	COMPARACION		COEFICIENTE DE SIMILITUD UTILIZADO	CORRELACION ENTRE LAS MATRICES DE SIMILITUD
	CARACTERES COMPARADOS	Nº DE CARACTERES EN CADA CONJUNTO		
*ESPECIES DEL COMPLEJO HOMOTIS (LARIAS)	CABEZA vs RESTO DEL CUERPO	60 62	CORRELACION DISTANCIA	0.81 0.73
*ESPECIES DEL COMPLEJO HOMOTIS (LARIAS)	MACHOS vs HEMBRAS	53 69	CORRELACION DISTANCIA	0.79 0.25
*GENEROS DE PARLONDAE (MARIPOSAS)	CABEZA vs TORAX	39 105	CORRELACION DISTANCIA	0.98 0.31
	CABEZA vs ABDOMEN	38 27	CORRELACION DISTANCIA	0.76 0.35
	TORAX vs ABDOMEN	105 27	CORRELACION DISTANCIA	0.27 0.69
	EXTERNOS vs INTERNOS	100 96	CORRELACION DISTANCIA	0.26 0.70
	EXTERNOS vs MUSCULATURA	100 75	CORRELACION DISTANCIA	0.76 0.18
*ESPECIES DE DOS GENEROS DE CROSOPODIPLOE (LEUCOPHAGA Y PARALUCOPHAGA)	CABEZA vs TORAX	39 51	DISTANCIA	0.45
	CABEZA vs ABDOMEN	39 62	DISTANCIA	0.25
	TORAX vs ABDOMEN	51 62	DISTANCIA	0.30
*ESPECIES DE AGARTIS DE LOS GENEROS (DERMANYSSES Y LIPONYSSOIDES)	OSTEOSOMA vs RESTO DEL CUERPO	15 120	CORRELACION DISTANCIA	0.75 0.82
	BIOSOMA DORSAL vs RESTO DEL CUERPO	45 90	CORRELACION DISTANCIA	0.35 0.80
	BIOSOMA VENTRAL vs RESTO DEL CUERPO	40 95	CORRELACION DISTANCIA	0.32 0.85
	PIATAS vs RESTO DEL CUERPO	35 100	CORRELACION DISTANCIA	0.52 0.71
*ESPECIES DE MOSQUITOS DE LOS GENEROS (ARDES Y PSOROPHORA)	PIATAS vs RESTO DEL CUERPO	67 91	CORRELACION DISTANCIA	0.42 0.78
*ESPECIES DE MOSQUITOS DEL GENERO (AEDES)	ADULTOS vs LARVAS	77 71	CORRELACION DISTANCIA	0.79 0.55



**Cuadro IV. Congruencia Taxonómica Medida con el Coeficiente de Correlación.** (La información fue tomada de Sneath y Sokal, 1973; Hsiao, 1975; Roback y Moss, 1978).

2 CONGRUENCIA ENTRE CLASIFICACIONES MEDIDA CON EL COEFICIENTE DE CORRELACION				
ORGANISMOS	COMPARACION		COEFICIENTE DE SIMILITUD UTILIZADO	CORRELACION ENTRE LAS MATRICES DE SIMILITUD
	CARACTERES COMPARADOS	Nº DE CARACTERES EN CADA CONJUNTO		
*ESPECIES DE INFIEROS DE LAS TRIBUS (MACROPODIPLOE Y ANATOPHYDRI)	ADULTOS vs PUPAS	16 22	DISTANCIA	0.59
	ADULTOS vs LARVAS	16 16	DISTANCIA	0.68
	LARVAS vs PUPAS	16 22	DISTANCIA	0.57
*GENEROS DE PAJAROS GALINACEOS	MUSCULOS DE LAS PATAS vs MUSCULOS DE LAS PATAS	82 93	CORRELACION	0.67
	MUSCULOS DE LAS PATAS vs MUSCULOS DEL PE	77 56	SIMPLE MATCHING	0.72
	MUSCULOS EXTENSIONES vs MUSCULOS FLEXIONES	52 81	JOINT MATCHING	0.74
*GENEROS Y ESPECIES DE LARI (PAJAROS)	EXTERNOS vs ESQUELETO	72 51	CORRELACION DISTANCIA	0.51 0.45
*ESPECIES DE ANGIOSPERMAS GENERO RARICOSTEMMA	FLORES vs VEGETATIVOS	61 72	CORRELACION DISTANCIA	0.73 0.47
*ESPECIES DE ANGIOSPERMAS GENERO SALIX	VEGETATIVOS vs SEXUALES	72 125	0 SIMILITUD	0.29
	SEXUALES vs SEXUALES FEMORINOS	60 65	0 SIMILITUD	0.83
*ESPECIES DE ANGIOSPERMAS GENERO PLATANUS	MORFOLOGIA vs MORFOLOGIA REPRODUCTIVA	23 19	CORRELACION	0.41
	QUIMICA HOJA vs QUIMICA FRUTOS (FENCILES)	51 29	CORRELACION	0.26
	MORFOLOGIA HOJA vs QUIMICA HOJA	23 51	CORRELACION	0.41
	MORFOLOGIA REPRODUCTIVA vs QUIMICA HOJA	19 51	CORRELACION	0.28
	MORFOLOGIA HOJA vs QUIMICA FRUTOS	23 29	CORRELACION	0.37
	MORFOLOGIA REPRODUCTIVA vs QUIMICA FRUTOS	19 29	CORRELACION	0.32
*CELAS DE BACTERIOS DEL GENERO CHROMOBACTERIUM	MORFOLOGICAS vs FISIOLOGICAS Y QUIMICAS	29 76	ADCORD	0.61

2. Según Hennig (1968), dos conjuntos de caracteres de un mismo taxon deben dar lugar a clasificaciones idénticas, ya que todos los caracteres de un organismo comparten su cladogénesis (genealogía). Pero, esta perfecta congruencia sólo se alcanza mediante la aplicación de técnicas acordes con la filosofía clasificatoria del cladismo que, por definición, basa la clasificación en el cladograma de los organismos. Para un determinado taxon, sólo puede existir un cladograma, sin importar cuáles tipos de caracteres contribuyen a su detección, ya que el proceso cladogenético de cualquier taxon es único. Similares razonamientos han sido usados por Mickevich y Johnson (1976). Sin embargo, al aplicar este argumento (Mickevich 1978, 1980) tampoco se ha obtenido una congruencia perfecta.

Los dos argumentos parecen estar bien fundamentados, aunque en cierto modo se oponen. El hecho principal es que nunca se ha conseguido una congruencia perfecta: ni con muchos caracteres, ni aplicando una técnica cladista. Nos preguntamos entonces: ¿Por qué no se da una perfecta congruencia? ¿cuáles son las causas que impiden obtener un 100% de congruencia?

Existen dos tipos de causas: causas biológicas y causas metodológicas.

Las primeras afectan principalmente a los enfoques feneticista y evolucionista de la sistemática y están dadas por el grado de interdependencia causal entre los conjuntos de caracteres utilizados. Los siguientes fenómenos interactuantes reducen esa interdependencia y, por lo tanto, la incongruencia aumenta:

a) la diferente actividad de los genes en diferentes células de un mismo organismo. Por ejemplo, una célula epidérmica posee el mismo genotipo que una célula hepática, pero la manifestación de ese genotipo es diferente;

b) la plasticidad fenotípica, es decir, la posibilidad de que un mismo genotipo se exprese de distinta forma en ambientes diferentes;

c) las mutaciones somáticas, a las cuales se debe el que dentro de un mismo individuo existan genotipos ligeramente diferentes;

d) las diferentes presiones de selección (diferentes en naturaleza, intensidad y dirección) que actúan sobre diferentes estadios o partes de un mismo organismo, especialmente cuando esos estadios o partes no son sincrónicos, y

e) la evolución en mosaico. Está bien demostrado que la velocidad de evolución de los diferentes tipos de caracteres puede ser distinta (Mayr, 1968; Michener, 1977). A veces la velocidad con que evolucionan diferentes órganos es manifestamente distinta. Algunos pueden haber avanzado mucho en tanto que otros pueden estancarse.

Teóricamente, una técnica acorde con el cladismo no debería ser afectada por las causas biológicas de incongruencia, pero no existe un solo caso de perfecta congruencia para dos clasificaciones construidas por una técnica cladista. Esto se debe a que un cladograma construido a partir de un conjunto de caracteres es sólo una estimación de la cladogénesis real y como tal está expuesto a errores producidos por lo que llamaremos causas metodológicas.

Las causas metodológicas son todos aquellos pasos técnicos que pueden producir distorsiones ya sea en la estimación de la cladogénesis

real (y en este caso afecta principalmente al cladismo) o en la estimación de la interdependencia entre caracteres (y en este caso afecta principalmente al evolucionismo y al feneticismo). Estos pasos técnicos incluyen entre otros muchos: la determinación de homoloqías, la cantidad y calidad de los caracteres, la codificación, el tipo de evaluación de la congruencia, etc.

En relación con estas causas metodológicas ha surgido una controversia sin solución prevista. Por ejemplo, Micevich (1978, 1980) ha tratado de demostrar que las técnicas numéricas del cladismo son las que provocan menos distorsiones metodológicas, posición ésta no compartida por Rohlf y Sokal (1980) y Colless (1980).

## CONNOTACIONES SISTEMÁTICAS

El núcleo de estas connotaciones se puede plantear en dos interrogantes que intentaremos responder.

El primero es el siguiente: Si nunca hallamos una perfecta congruencia ¿existe un tipo de carácter que, como indicador de relaciones taxonómicas, sea más importante que los otros?

Antes de poder establecer el valor taxonómico de los distintos tipos de caracteres, hay que ser muy claros acerca de qué se entiende por relaciones taxonómicas, ya que como vimos en el capítulo 2, existe más de una relación posible entre dos taxones. De las relaciones posibles dos son las más importantes desde un punto de vista taxonómico: las relaciones de parentesco y las relaciones fenéticas.

Las relaciones de parentesco pueden subdividirse, a su vez, en relaciones patristicas, que son las similitudes debidas al antecesor común (similitudes homólogas) y relaciones cladísticas que expresan la cercanía relativa con el antecesor común.

Las distintas escuelas clasificatorias dan diferente valor a estas relaciones. Para un evolucionista, relaciones taxonómicas son sinónimo de relaciones de parentesco en su totalidad; para un cladista son sinónimo de relaciones cladísticas, y para un feneticista son sinónimo de relaciones fenéticas.

Todos los caracteres de un mismo taxon comparten las relaciones cladísticas, pero no necesariamente las relaciones patristicas y fenéticas. Por lo tanto, en un enfoque puramente cladista todos los tipos de caracteres (químicos, morfológicos, ecológicos, etc.) tienen el mismo valor. Por el contrario, en un enfoque evolucionista o fenético estamos expuestos a la posibilidad de incongruencia debida a las relaciones patristicas y fenéticas. Por consecuencia, en un enfoque fenético o evolucionista, un tipo de carácter es importante si se halla conectado causalmente con muchos otros y por lo tanto comparte con muchos otros tipos de caracteres sus relaciones patristicas y fenéticas. Por el momento, no es posible establecer para los distintos grupos qué tipo de carácter satisface estos requisitos en mayor medida que los demás.

De cualquier modo está claro que:

- Un enfoque cladista producirá resultados congruentes en la misma medida que la técnica utilizada sea capaz de estimar las reales relaciones cladísticas.

- Un enfoque evolucionista o fenético está expuesto a incongruencia debido a fenómenos biológicos, como, por ejemplo, la evolución en mo-

saico. Por lo tanto, el análisis de nuevos caracteres producirá, generalmente, cambios en los resultados ya obtenidos.

- Los estudios biosistemáticos, que analizan al organismo a lo largo de todo su ciclo vital y consideran todos sus aspectos, son los más aconsejables.

El segundo interrogante que se plantea es el siguiente: ¿Es la congruencia un buen criterio para determinar las bondades de una técnica clasificatoria?

La congruencia es un buen criterio para un enfoque cladista, pero no lo es para el enfoque evolucionista o feneticista, ya que la incongruencia puede ser el resultado de fenómenos biológicos que estas dos escuelas clasificatorias deben tener en cuenta en la construcción de sus clasificaciones.



## LA RECONSTRUCCION DE LA FILOGENIA POR METODOS NUMERICOS

La filogenia es el estudio de la historia evolutiva de un grupo o taxon.

Los resultados de los estudios filogenéticos se expresan gráficamente por líneas continuas divergentes de variable longitud, las cuales reflejan una hipótesis con respecto a esa historia. Estos gráficos denominados árboles evolutivos fueron en un principio el resultado de la asociación entre las evidencias halladas y la intuición del investigador. Actualmente, los árboles que se elaboran son el resultado del análisis de la evidencia realizado conforme a un conjunto de criterios y reglas claramente establecidos.

Gran difusión han alcanzado los enfoques numéricos que intentan expresar cuantitativamente esos criterios y reglas (Funk y Stuessy, 1978). Entre estos enfoques hallamos la técnica del árbol de Wagner, que no sólo es utilizada por los feneticistas para reconstruir la filogenia, sino que es usada con fines clasificatorios por los cladistas, quienes la consideran compatible con los resultados del análisis cladístico no cuantitativo (Mickey, 1978; Wiley, 1981).

En este capítulo trataremos en general de las técnicas numéricas de reconstrucción filogenética y, en particular, de la técnica del árbol de Wagner.

93

## COMPONENTES DE LA FILOGENIA

La filogenia tiene tres parámetros estimables (Kluge, 1976): polaridad (dirección de cambio), patrística (cantidad relativa de cambio) y cladística (cercanía relativa a un antecesor común).

Para un árbol evolutivo en particular que simula la filogenia de un grupo, la polaridad está representada por las tendencias evolutivas; la patrística, por la similitud entre los integrantes del árbol, y la cladística, por el origen, el número y la secuencia de las ramificaciones del árbol. En otras palabras, la polaridad está dada por la secuencia que va de primitivo a evolucionado; la patrística, por el grado de divergencia alcanzado, y la cladística, por la genealogía.

La similitud medida por la patrística es la que tiene su origen en el antecesor común; por lo tanto, es una similitud homóloga y no incluye a la similitud homoplásica. La similitud patrística más la similitud homoplásica constituyen la similitud fenética.

En los enfoques numéricos, la polaridad se determina a base de criterios establecidos *a priori* y la patrística y la cladística mediante técnicas cuantitativas.

Puede considerarse que la filogenia posee un cuarto parámetro: el tiempo (cronística), que precisa el momento en que sucedieron los fenómenos evolutivos de un grupo en particular.

## PASOS ELEMENTALES EN LA RECONSTRUCCION DE LA FILOGENIA Y DETERMINACION DE LA POLARIDAD

En la reconstrucción de la filogenia por métodos numéricos se distinguimos los pasos elementales siguientes:

1. elección de las OTU (denominados también "evolutionary unit", Estabrook, 1977);
2. elección de los caracteres;
3. confección de la matriz básica de datos (OTU x caracteres);
4. determinación de la polaridad (estimación del estado primitivo de los caracteres), y
5. construcción del árbol evolutivo (cladística y patrística).

Los tres primeros pasos deben ceñirse a las indicaciones dadas en capítulos precedentes acerca de la aplicación de técnicas numéricas, con las siguientes diferencias:

a) cada OTU debe, necesariamente, ser monofilética, es decir, todos los miembros deben poseer un ancestro común (monofiletismo, *sensu* Ross, 1974);

b) el grupo en estudio como totalidad debe ser monofilético;

c) no es necesario que el número de caracteres sea elevado;

d) la similitud medida será la similitud homóloga, y

e) el reconocimiento de la polaridad del grupo en estudio es condición indispensable.

Para reconocer la polaridad es preciso determinar el estado primitivo (ancestral, plesiomorfo) de cada carácter utilizado y, en consecuencia, reconocer el estado o los estados derivados (avanzados, apomorfos, evolucionados).

El estado primitivo de un carácter es aquél que se halla o se infiere que se hallaba en el más reciente antecesor común del grupo, cuya historia evolutiva se está determinando (Crisci y Stuessy, 1980). Este concepto de primitivismo es relativo: un estado de un carácter puede ser primitivo en un taxon en particular, pero no necesariamente en otro. Por lo tanto, el uso del término primitivo y, por ende, el de derivado, carece de sentido si no se refiere a un grupo o taxon en particular.

En un determinado taxon, un carácter presenta un estado primitivo y el estado o los estados restantes son, en mayor o menor medida, derivados en proporción a su alejamiento del antecesor.

Varios autores han intentado resumir la situación actual del conocimiento sobre los criterios para determinar estados primitivos (Crisci y Stuessy, 1980; De Jong, 1980; Eldredge y Cracraft, 1980; Stevens, 1980; Watrous y Wheeler, 1981).

En esta sección nos referiremos a los nueve criterios presentados en el trabajo de Crisci y Stuessy, 1980, clasificados en dos grupos: los del primer nivel y los del segundo nivel.

**Criterios de primer nivel:** Son aquéllos de los cuales no se tiene conocimiento previo acerca de condiciones primitivas dentro del grupo estudiado o de los grupos más afines.

1. Evidencia fósil. Es probable que el estado de carácter hallado en el más antiguo representante fósil conocido, dentro de dicho grupo, sea el primitivo.

2. El estado más común es el primitivo. Existen dos posibilidades: "común dentro del grupo" y "común en los grupos afines al estudiado".

a) Común dentro del grupo estudiado. Es probable que el estado de un carácter sea el primitivo si es el más frecuente dentro del grupo estudiado.

b) Común en los grupos afines al estudiado. Si el estado de un carácter se encuentra en el taxon sometido a estudio y es frecuente en el grupo (o grupos) más afín a él, ese estado es probablemente el primitivo.

3. Ontogenia. Ante dos o más estados de un determinado carácter, podría considerarse que el primitivo es aquél que aparece en etapas más tempranas del desarrollo ontogenético de los integrantes del grupo en estudio.

4. Anormalidades menores de la organogénesis (no monstruosidades). Algunos casos teratológicos menores, debidos a factores genéticos, podrían ser considerados como reversiones al estado primitivo del carácter.

5. Organos vestigiales. Si un carácter referido a un órgano o estructura presenta los siguientes tres estados:

- presente y funcional,
- presente y no funcional, y
- ausente.

El estado "presente y funcional" es probablemente el primitivo.

6. Coocurrencia de estados primitivos. Si dos o más estados de dos o más caracteres se encuentran frecuentemente en los mismos taxones, esos estados son probablemente primitivos, ya que los estados primitivos tienden a ocurrir juntos más a menudo que los estados derivados.

**Criterios de segundo nivel:** Son aquéllos que dependen del conocimiento previo de una condición primitiva dentro del grupo estudiado o grupos afines. Este conocimiento se adquiere aplicando los criterios del primer nivel.

7. Asociación. Si conocemos el grupo ancestral que dio origen al taxon en estudio, el estado del carácter más similar al del grupo ancestral es probablemente el primitivo.

8. Correlación. El estado de un carácter es probablemente el primitivo si su presencia está siempre correlacionada con la presencia del estado primitivo de otro carácter, cuya polaridad fue previamente determinada por la aplicación de otro criterio.

9. Tendencias de grupos. Dentro de grandes grupos algunas tendencias evolutivas se repiten en forma paralela. Si se determina el estado primitivo en una de ellas, podría considerarse que esa polaridad se repite en las otras.

Existen controversias respecto al grado de validez de cada criterio. Por ejemplo, para los sostenedores del cladismo (Wiley, 1981) "el estado más común en grupos afines es el primitivo", constituye el criterio que menos excepciones presenta.

Importa señalar que ninguno de los criterios es absoluto y universal, todos en mayor o menor medida tienen excepciones y por esto el investigador debe proceder con cautela al aplicarlos (Stuessy y Crisci, 1983).

#### CONSTRUCCION DEL ARBOL MEDIANTE LA TECNICA DE WAGNER

Numerosos coeficientes de similitud y técnicas de agrupamientos han sido propuestos entre los enfoques numéricos de reconstrucción filogenética.

Las técnicas pueden clasificarse de acuerdo con el concepto de compatibilidad de caracteres o el de simplicidad (Funk y Stuessy, 1978).

Las técnicas basadas en el postulado de la compatibilidad de caracteres encuentran el mayor conjunto de caracteres compatibles entre sí en su polaridad y las relaciones filogenéticas se estiman a base de ese conjunto (Wilson, 1965; Le Quesne, 1969; Estabrook, 1972; Meachan, 1981).

Las técnicas que utilizan el concepto de simplicidad se basan en el siguiente postulado: los cambios evolutivos se realizan de la forma más económica posible, desde el punto de vista de los sucesos genéticos requeridos. La evolución orgánica se efectúa con la máxima eficacia, y en todos los cambios de un estado de un carácter a otro recorre el camino más directo entre ellos ("parsimony", Camin y Sokal, 1965). En otras palabras, la evolución sigue siempre el camino más corto.

La exigencia de una simplicidad imperativa en el estudio de la filogenia es claramente incompatible con el conocimiento actual sobre los procesos evolutivos. Por lo tanto, la simplicidad en filogenia es una prescripción metodológica y no una ley de la evolución (Crisci, 1982).

Se han propuesto varias técnicas para obtener a partir de un grupo en particular, el árbol filogenético con menos cambios evolutivos. Todavía no existe acuerdo respecto a cuál de ellas permite aplicar mejor el postulado de simplicidad (Estabrook, 1978).

La técnica del árbol de Wagner (Kluge y Farris, 1969) es la más difundida (Crisci, 1980) no sólo entre los métodos que utilizan el postulado de simplicidad, sino también entre todas las técnicas de reconstrucción filogenética por métodos cuantitativos. Esta técnica para la simulación de árboles evolutivos está basada en un concepto formulado por W.H. Wagner (1961) que ha sido computarizado y transformado en un procedimiento numérico por Kluge y Farris (1969).

A continuación, describiremos los pasos de la técnica del árbol de Wagner.

Una vez elegidas las OTU y los caracteres y determinada la polaridad, se calcula una matriz de similitud (OTU x OTU) utilizando el coe-

ficiente "Manhattan Distance", cuyo cálculo para las OTU A y B está representado en la fórmula:

$$D(A,B) = \sum_{i=1}^n |X(A,i) - X(B,i)| \quad (13)$$

donde:

D = distancia;  
 X(A,i) = es el estado del carácter i para la OTU A;  
 X(B,i) = es el estado del carácter i para la OTU B, y  
 n = número de caracteres.

Esta técnica construye, a partir de un antecesor señalado por el investigador (basándose en la polaridad), un árbol que conecta a todas las OTU en el menor número de cambios evolutivos posibles, es decir, que su longitud expresada en los valores del coeficiente de similitud utilizado será mínima. Se entiende por longitud la sumatoria de todos los intervalos del árbol. El resultado es un gráfico bidimensional que representa: la polaridad (dada por el antecesor), la historia cladística y las relaciones patrísticas entre las OTU.

El antecesor puede ser una OTU real que posea una gran cantidad de estados primitivos, o también es posible construir un vector que represente una OTU hipotética con estados primitivos en todos sus caracteres.

Con el fin de acortar la longitud del árbol, la técnica va construyendo taxones intermedios hipotéticos entre las OTU; estos intermedios denominados unidades taxonómicas hipotéticas ("Hypothetical Taxonomic Units", HTU, Farris, 1970), representan taxones artificiales y son los puntos de ramificación del árbol. Las HTU contendrán en cada carácter la mediana de los estados presentes en las tres OTU (u otras HTU) a las cuales está conectada.

97

El árbol contiene inicialmente un solo elemento --el antecesor-- que puede ser hipotético, en cuyo caso es tratado como OTU y no como HTU, ya que los estados de sus caracteres son precisados por el investigador y no calculados por la técnica.

Una vez elegido el antecesor, se prosigue en esta técnica con los pasos siguientes:

a) hallar en la matriz de similitud la OTU que está a menor distancia (máxima similitud) del antecesor y conectarla a él por un intervalo, cuyo valor es igual a la distancia entre ambos;

b) hallar entre las OTU restantes la que esté a menor distancia del antecesor;

c) hallar el intervalo del árbol con el cual la OTU hallada en b) esté a menor distancia. La distancia entre una OTU, por ejemplo C y un intervalo, por ejemplo A-B, se determina aplicando la fórmula (Farris, 1970):

$$D[C, \text{intervalo (A-B)}] = [D(C,B) + D(C,A) - D(B,A)] 1/2 \quad (14)$$

d) conectar la OTU hallada en b) con el intervalo determinado en c) por un segmento de valor igual a la distancia entre ellos. Construir un intermedio hipotético (HTU) y colocarlo en el árbol en el lugar de la conexión entre la OTU y el intervalo;

e) si queda alguna OTU sin ubicar, retornar a b); en caso contrario finalizar aquí.

Esta técnica permite la reversión de los estados de los caracteres, lo que significa que los estados ancestrales no son necesariamente conservadores, o sea que un estado primitivo puede reaparecer con posterioridad a un estado derivado.

En la figura 29 se presenta un ejemplo de los tres primeros pasos de construcción de un árbol con su antecesor y la colocación de las OTU A, B y C.

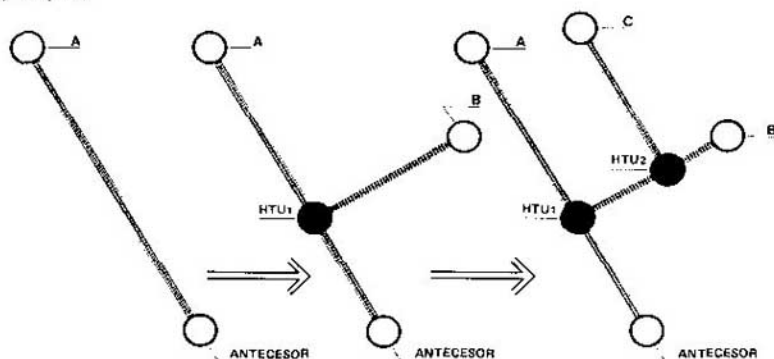


Fig. 29. Técnica del árbol de Wagner; los tres primeros pasos en la construcción del árbol, con su antecesor y la colocación de las OTU A, B y C. Los intermedios hipotéticos (HTU<sub>1</sub> y HTU<sub>2</sub>) son colocados con el fin de acortar la longitud del árbol.

Una vez construido el árbol, es posible medir el grado de paralelismo, convergencia y/o reversiones que presenta mediante el índice de consistencia ( $C = \text{"index of consistency"}$ , Kluge y Farris, 1969), el cual se determina con la fórmula:

$$C = \frac{R}{L} \quad [15]$$

donde:

$L$  = longitud total del árbol, y

$R$  = sumatoria de los rangos (rango es la diferencia entre el estado numéricamente mayor y el estado numéricamente menor de un carácter).

El valor de este índice oscila entre 0 y 1. Se aproxima a 1 cuando no se producen sucesos evolutivos "extras" y tiende a cero a medida que aumenta el paralelismo, la convergencia y/o las reversiones.

#### UN EJEMPLO HIPOTÉTICO

Hemos elegido un ejemplo hipotético en el que la matriz básica de datos está constituida por ocho OTU, para las que se registraron 18 caracteres.

La codificación presenta para cada uno de los caracteres tres estados: 0; 0,5 y 1. La matriz básica de datos obtenida es la siguiente:

Caracteres																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
OTU A	0,5	0	1	0	0,5	0	0,5	1	0,5	0,5	1	0,5	0,5	1	0	0,5	0,5	1
OTU B	1	0,5	1	0	0,5	0	0,5	0,5	1	1	0,5	1	0,5	0	0,5	0,5	0,5	1
OTU C	0	1	1	0	0	1	0	0	0	0,5	0	0,5	0,5	0,5	0	0,5	0	1
OTU D	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0,5	0
OTU E	0	0	0,5	0	0	0	0,5	0	0	0	0	0	1	0	0	0	1	0,5
OTU F	0	0	0	0,5	0,5	0,5	0,5	0	0	0,5	0	0,5	0	0	0	0,5	0	1
OTU G	0	0,5	1	1	1	0,5	0,5	0	0	0,5	0,5	0,5	0,5	0	1	1	0,5	1
OTU I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Para cada carácter, el valor 0 corresponde al estado primitivo, el 0,5 a una condición derivada intermedia y el 1 al estado derivado final.

Existe una OTU (OTU I), que será considerada como el antecesor, que exhibe estados primitivos en todos sus caracteres y a partir de la cual comenzaremos a construir el árbol.

De la aplicación del coeficiente "Manhattan Distance" entre cada par posible de OTU, se obtiene la siguiente matriz de similitud de distancia:

	A	B	C	D	E	F	G	I
A	0							
B	5	0						
C	7	8	0					
D	8	9	7	0				
E	8	9	7	3	0			
F	7	8	5	6	6	0		
G	7,5	6,5	6,5	8,5	8,5	5,5	0	
I	9,5	10,5	6,5	2,5	3,5	4,5	10	0

Describiremos a continuación el análisis de la matriz de similitud y los pasos para construir el árbol.

a) La OTU con menor distancia del antecesor es la OTU D y se conecta a éste por el intervalo cuyo valor es igual a la distancia entre ambos: 2,5 (Fig. 30a).

b) Entre las OTU restantes la de menor distancia del antecesor (3,5) es E.

c) Se conecta la OTU E al intervalo I-D a una distancia que se determina aplicando la fórmula:

$$\text{Distancia E, int. (I-D)} = \frac{D(E,I) + D(E,D) - D(I,D)}{2}$$

$$D = \frac{3,5 + 3 - 2,5}{2} = 2$$

d) Se construye la HTU<sub>1</sub> que conectará la OTU E al intervalo I-D. La HTU<sub>1</sub> contendrá en cada carácter la mediana de los estados presentes en las tres OTU a las cuales estará conectada.

La siguiente matriz confirma que los estados de los caracteres de la HTU<sub>1</sub> son el resultado de las operaciones antes descritas.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
OTU I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU D	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0,5	0
OTU E	0	0	0,5	0	0	0	0,5	0	0	0	0	0	1	0	0	0	1	0,5
HTU <sub>1</sub>	0	0	0,5	0	0	0	0,5	0	0	0	0	0	0	0	0	0	0,5	0

Si se calcula la distancia de acuerdo con la matriz precedente, la HTU<sub>1</sub> deberá estar a 1,5 de la OTU I y a 1 de la OTU D (Fig. 30b).

e) Entre las OTU restantes, la OTU F es la que está a menor distancia del antecesor (4,5). Hay tres intervalos a los cuales puede incorporarse la citada OTU:  $\text{int}(I\text{-}HTU_1)$ ,  $\text{int}(HTU_1\text{-}D)$  e  $\text{int}(HTU_1\text{-}E)$ . Es preciso calcular la distancia entre la OTU ingresante y cada uno de los intervalos anteriormente nombrados. Se conecta la ingresante al intervalo con el que se encuentra a menor distancia.

$$DF, \text{int}(I\text{-}HTU_1) = \frac{4,5 + 5 - 1,5}{2} = 4$$

$$DF, \text{int}(HTU_1\text{-}D) = \frac{5 + 6 - 1}{2} = 5$$

$$DF, \text{int}(HTU_1\text{-}E) = \frac{5 + 6 - 2}{2} = 4,5$$

Por lo tanto, la OTU F debe conectarse al intervalo(I-HTU<sub>1</sub>) a una distancia de 4. Se construye la HTU<sub>2</sub>, que conectará la OTU F a dicho intervalo. Los estados de los caracteres de la HTU<sub>2</sub> son el resultado de la determinación de la mediana de los estados de los caracteres de la OTU I, la OTU F y la HTU<sub>1</sub>.

La siguiente matriz muestra el resultado de estos cálculos:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
OTU I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU F	0	0	0	0,5	0,5	0,5	0,5	0	0	0,5	0	0,5	0	0	0	0,5	0	1
HTU <sub>1</sub>	0	0	0,5	0	0	0	0,5	0	0	0	0	0	0	0	0	0	0,5	0
HTU <sub>2</sub>	0	0	0	0	0	0	0,5	0	0	0	0	0	0	0	0	0	0	0

Si calculamos la distancia, de acuerdo con la matriz precedente, la HTU<sub>2</sub> deberá estar a 0,5 de la OTU I y a 1 de la HTU<sub>1</sub> (Fig. 30c).

f) Entre las OTU restantes, la OTU C, es la que se encuentra a menor distancia del antecesor (6,5). Existen cinco intervalos a los cuales puede incorporarse la mencionada OTU:  $\text{int}(I\text{-HTU}_2)$ ;  $\text{int}(\text{HTU}_2\text{-F})$ ;  $\text{int}(\text{HTU}_2\text{-HTU}_1)$ ;  $\text{int}(\text{HTU}_1\text{-D})$ ;  $\text{int}(\text{HTU}_1\text{-E})$ . Hay que calcular las distancias entre la OTU ingresante y cada uno de los intervalos anteriormente nombrados. Se conecta la ingresante al intervalo con el que se encuentra a menor distancia.

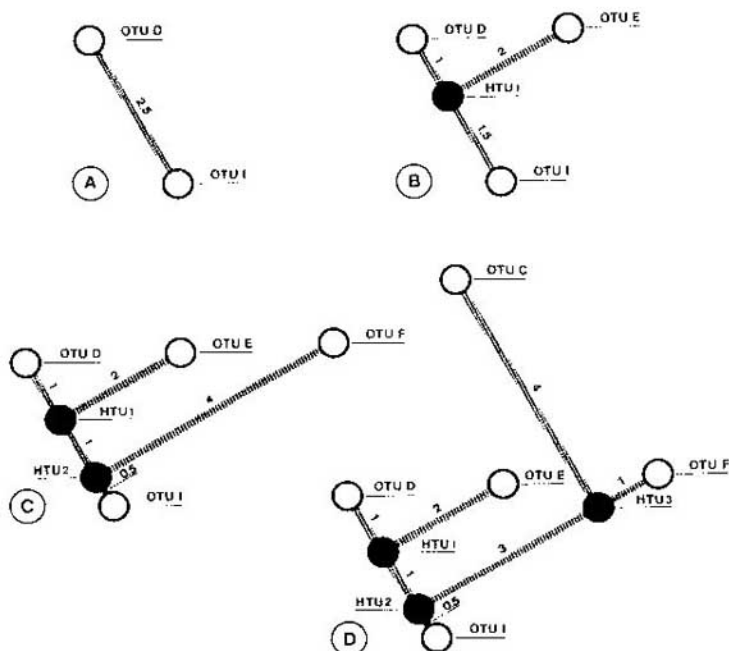


Fig. 30. Técnica del árbol de Wagner; construcción del árbol del ejemplo hipotético con ocho OTU, detallado para sus cuatro primeros pasos.

El mecanismo de cálculo se repite en los pasos sucesivos, por lo cual no continuaremos detallando las operaciones matemáticas.

El intervalo al que debe ingresar la OTU C es el  $\text{int}(\text{HTU}_2\text{-F})$  y a una distancia de 4. La HTU<sub>3</sub> que conectará a C con el  $\text{int}(\text{HTU}_2\text{-F})$  estará a una distancia de 3 de la HTU<sub>2</sub> y a una distancia de 1 de la OTU F (Fig. 30d).

El lector, a modo de ejercicio, puede verificar que siguiendo los procedimientos precedentes, las restantes OTU se irán incorporando en la siguiente secuencia: OTU A, OTU G y, finalmente, OTU B, con la construcción de: HTU<sub>4</sub>, HTU<sub>5</sub> y HTU<sub>6</sub>.

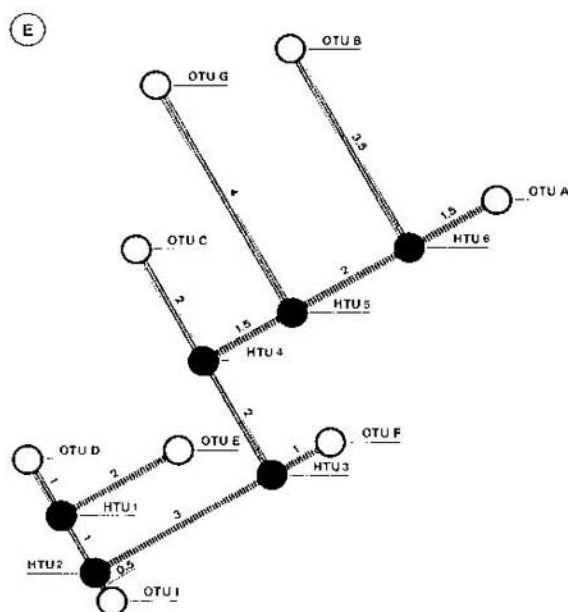
La figura 31 ilustra la representación final del árbol obtenido para el ejemplo hipotético. La longitud (L) del árbol es de 25. La di-

ferencia entre el valor numérico máximo y mínimo de cada uno de los caracteres es 1; por lo tanto, la sumatoria de los rangos (R) es 18. Si el árbol no tuviese reversiones, paralelismos y convergencias debería tener 18 de longitud, por consiguiente, el índice de consistencia sería 1.

En el ejemplo hipotético el número de cambios del árbol (longitud) es de 25, por lo que el índice de consistencia es:

$$C = \frac{18}{25} = 0,72$$

Este valor demuestra que fue escaso el número de sucesos evolutivos "extras".



103

Fig. 31. Técnica del árbol de Wagner; resultado final del ejemplo hipotético con ocho OTU.

## RETICULOS

Los retículos ("networks") son técnicas que permiten diseñar diagramas de las ramificaciones sin precisar antecesor. Al igual que el árbol de Wagner, se basan en el postulado de simplicidad, es decir, que la suma de la totalidad de las distancias es mínima.

En algunos retículos ("Wagner networks") se construyen HTU con el objeto de minimizar aún más esa sumatoria. La diferencia fundamental entre el árbol de Wagner y los retículos es que estos últimos carecen de polaridad.

Los retículos pueden aplicarse en trabajos taxonómicos como una forma de análisis de agrupamientos. En reconstrucción filogenética son muy utilizados en el análisis preliminar de los datos.

El retículo de Prim (Prim, 1957) es uno de los más utilizados y fue ideado con el fin de determinar la línea más corta que conecta un conjunto dado de terminales telefónicas. Ha sido aplicado a problemas biológicos, sistemáticos y evolutivos (Edwards y Cavalli-Sforza, 1964; Solbrig, 1970, 1971; Crisci, 1974; Petriella y Crisci, 1978).

Esta técnica requiere una matriz de distancia ("Manhattan Distance") y, básicamente, consiste en los siguientes pasos:

a) se inspecciona la matriz de similitud en busca del par de OTU que posea el mayor valor de similitud;

b) se unen ambas OTU mediante un segmento cuyo valor es igual a la distancia entre ellas;

c) se inspecciona la matriz de similitud en busca de la OTU que posea el mayor valor de similitud en relación con cualquiera de los integrantes del retículo;

d) la OTU hallada se une al retículo, conectándola con la OTU con la cual poseía el mayor valor de similitud. El intervalo será igual a la distancia entre ellas.

e) se repite el procedimiento a partir de c) hasta que todas las OTU se hayan incorporado al retículo.

104

De la aplicación de la técnica del retículo de Prim a la matriz de similitud (distancia), descrita en este capítulo (Un ejemplo hipotético), resultan los siguientes pasos (Fig. 32):

1. el par más similar entre sí está compuesto por las OTU I y D que se unen a la distancia de 2,5 (Fig. 32a);

2. de las restantes OTU, E es la que presenta el mayor valor de similitud con respecto a alguno de los integrantes del retículo, en este caso la OTU D. Se incorpora la OTU E al retículo, conectándola con la OTU D a una distancia de 3 (Fig. 32b);

3. la próxima ingresante es la OTU F que se encuentra a una distancia de 4,5 con respecto a la OTU I (Fig. 32c);

4. la siguiente ingresante es la OTU C que está a una distancia de 5 con respecto a la OTU F (Fig. 32d);

5. la siguiente ingresante es la OTU G que está a una distancia de 5,5 con respecto a la OTU F (Fig. 32e);

6. la siguiente ingresante es la OTU B que se encuentra a una distancia de 6,5 de G (Fig. 32f), y

7. la última ingresante es la OTU A que se encuentra a una distancia de 5 de la OTU B (Fig. 32g).

La longitud total del retículo es de 32.

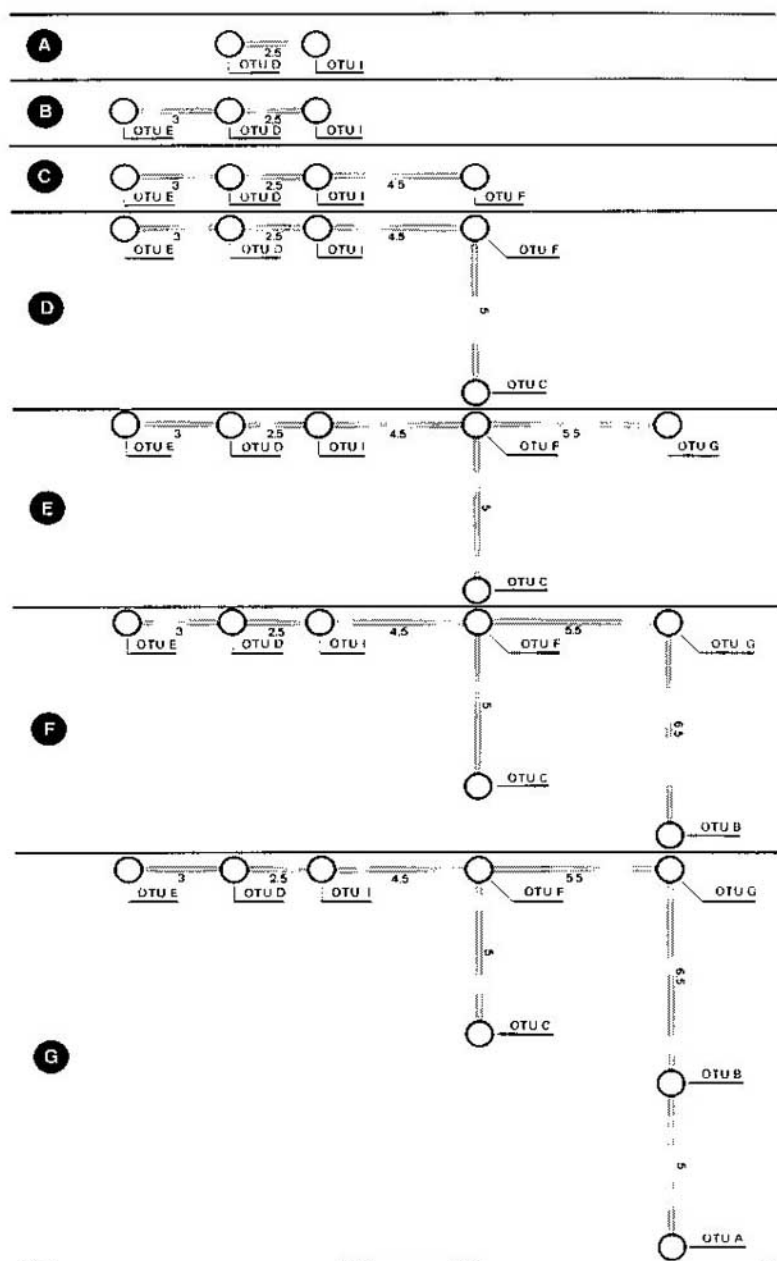


Fig. 32. Técnica del retículo de Prim; pasos en la construcción del retículo del ejemplo hipotético con ocho OTU.



En el desenvolvimiento de la taxonomía numérica las computadoras han desempeñado dos papeles complementarios: han sido protagonistas e instrumentos. Protagonistas, por originar una tendencia al cálculo, a la medición precisa, a una clara descripción de cada uno de los pasos de la clasificación y al acercamiento a la matemática. Instrumento, al permitir el manejo de una gran cantidad de información, la resolución de cálculos complejos de manera confiable y la realización de todos los procesos a una gran velocidad (Sokal, 1974).

Dada la estrecha relación entre taxonomía numérica y computadoras, es necesario que quien intente aplicar este método posea, por lo menos, conocimientos básicos sobre ellas.

El constante avance tecnológico en este campo y en especial el alcanzado por las poderosas computadoras personales o microcomputadoras limita cualquier presentación sobre este tema. Sin embargo, en este capítulo intentaremos presentar las nociones elementales sobre computadoras en función de cuatro interrogantes: ¿Qué es una computadora? ¿Cuáles son sus componentes? ¿cómo nos comunicamos con ella? ¿cuáles son los programas más comunes en taxonomía numérica?

#### UNA RAPIDA MIRADA A LAS COMPUTADORAS

La computadora es una máquina con capacidad para efectuar cálculos, como asimismo para almacenar datos y programas.

Programa es el conjunto de instrucciones que el investigador introduce en la computadora y que determina el ordenamiento de las operaciones (cómputos y decisiones lógicas) que ésta debe realizar (Farina, 1972).

Las computadoras son básicamente de dos tipos: digitales y analógicas; ambas se distinguen principalmente por la manera en que realizan los cálculos.

Las computadoras digitales se denominan así debido a que las magnitudes que en ellas se almacenan varían en forma discreta, por dígitos; es decir, el cálculo se realiza contando. En el caso de las computadoras analógicas las magnitudes varían de una forma continua; emplean circuitos y magnitudes capaces de simular por analogía los más variados fenómenos físicos, es decir, el cálculo se realiza midiendo.

Por ser las computadoras digitales las más comunes, se hará referencia a ellas en el resto del capítulo.

La computadora digital consta de cinco elementos (Michelow, 1980): unidad de entrada, memoria central, unidad aritmética y lógica, unidad de control y unidad de salida. La relación entre estos cinco componentes puede verse en la figura 33.

1. Unidades de entrada. Permiten el ingreso de la información (datos originales y programas) en la memoria central. Estas unidades de

entrada son de distintos tipos: máquina de escribir, lectora de tarjetas perforadas, lectora de cintas magnéticas, lectora de discos magnéticos.

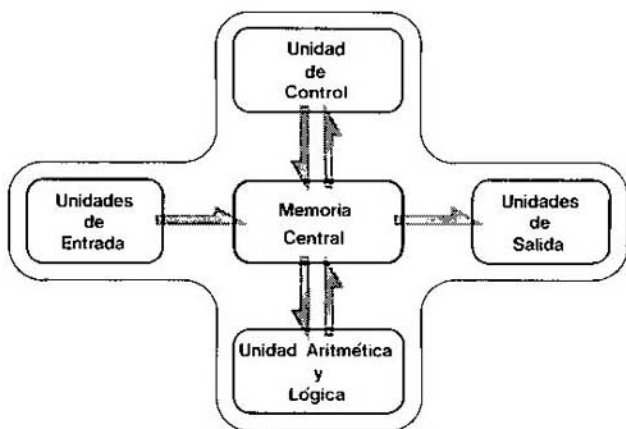


Fig. 33. Componentes de una computadora digital.

2. Unidad de memoria. Almacena la información recibida y los resultados de las operaciones intermedias y finales.

3. Unidad aritmética y lógica. Realiza operaciones de suma, resta, multiplicación, división y comparación de dígitos. Asimismo toma decisiones lógicas basándose en las instrucciones y los resultados de las operaciones intermedias.

4. Unidad de control. Toma las instrucciones de la memoria central y las ejecuta en forma ordenada y secuencial. Hace uso, cuando es necesario, de la unidad aritmética y lógica. A su vez, dirige el almacenamiento en la memoria central.

5. Unidades de salida. Permiten la salida de los resultados una vez finalizado el procesamiento de los datos. Hay varios tipos de unidades de salida: máquina de escribir, pantalla con teclado, impresora, grabadora de cintas magnéticas y grabadora de discos magnéticos. Físicamente, algunas de las unidades de salida son al mismo tiempo unidades de entrada.

#### COMUNICACION DEL HOMBRE CON LA COMPUTADORA; PROGRAMAS Y LENGUAJE

Los órdenes que desencadenan la ejecución de las operaciones se denominan "instrucciones" y los conjuntos de éstas constituyen los programas. Los programas deben ser escritos en un lenguaje comprensible para la computadora e incluir con detalle todos los pasos necesarios para el procesamiento de los datos.

Cada modelo de computadora otorga un código a cada tipo de instrucciones. El programa escrito en este *lenguaje máquina* es directamente comprensible para la máquina computadora, pero presenta dificultades para el investigador, ya que este lenguaje o código es absolutamente numérico. Por este motivo, se han ideado lenguajes de utilización más cómoda y es la propia computadora la que, provista de un programa apropiado, traduce esos nuevos lenguajes a su propio código.



Los nuevos lenguajes se acercan a la forma humana de expresión; por ejemplo, el COBOL está adaptado a problemas de tipo administrativo mientras que el ALGOL, el BASIC y el FORTRAN son adecuados para la programación de cálculos matemáticos.

El lenguaje FORTRAN y el BASIC son los más utilizados en el campo científico. Los programas escritos en estos lenguajes están constituidos por proposiciones o sentencias individuales que deben cumplir una serie de normas análogas a las reglas gramaticales del idioma inglés (Organick, 1966).

#### PROGRAMAS DE COMPUTADORAS PARA APLICAR LA TAXONOMIA NUMERICA

Existen numerosos programas y sistemas de programas (= conjunto de programas diseñados para realizar operaciones dentro de un campo determinado) útiles en taxonomía numérica. Entre los programas más conocidos están las series de publicaciones dadas a conocer por el Kansas Geological Survey Computer Contributions (Bartcher, 1966; Bonham-Carter, 1967; Wahlstedt y Davis, 1968; Wishart, 1969b; Ondrick y Srivastava, 1970; Demirmen, 1969; Reymont, Ramden y Washstedt, 1969; Reymont y Ramden, 1970; McCammon y Wenniger, 1970).

Simultáneamente ha habido un gran avance en los sistemas de programas aplicables a la taxonomía numérica, los que pueden ser conseguidos fácilmente dirigiéndose a los autores (Lance y Williams, 1966; Tryon y Bailey, 1966; Moss, 1970; Wishart, 1972).

Todos estos sistemas poseen programas que comprenden los pasos elementales (confección de la matriz de similitud, análisis de agrupamientos y ordenación), pero difieren en las técnicas que ofrecen para llevarlos a cabo.

Entre estos sistemas, uno de los más difundidos es el NT-SYS ("Numerical Taxonomy System of Multivariate Statistical Programs") que ha sido ideado por F. James Rohlf, John Kishpaugh y David Kirk (1971). Este sistema es un conjunto de programas que permite realizar diversas operaciones sobre matrices, especialmente en relación con el análisis estadístico multivariado, con especial atención a los métodos útiles a la taxonomía numérica, tales como: análisis de agrupamientos y ordenación. Esos métodos comprenden una gran cantidad de técnicas. Debido a esta última característica, el NT-SYS es extremadamente valioso para cualquier centro de investigación que desee aplicar la taxonomía numérica en toda su extensión. Incluso contiene programas de aplicaciones estadísticas, útiles en biometría.

Existe una versión más simple del NT-SYS denominada MINT (mini-NT-SYS), establecida por los mismos autores con el objeto de ofrecer un sistema más modesto en cuanto a sus posibilidades técnicas.

Ambos sistemas, el NT-SYS y el MINT pueden solicitarse directamente del Dr. F. James Rohlf de la Universidad Estatal de Nueva York en Stony Brook, Estados Unidos.

Los datos de nuestro ejemplo de *Burmesia* fueron procesados en el Centro de Cómputos de la Universidad Nacional de la Plata (Centro de Estudios para el Procesamiento de la Información, CESPI) utilizando el NT-SYS en una IBM-4331.

En el ejemplo hipotético presentado en el capítulo 9 (Reconstrucción filogenética) se usaron programas elaborados por Farris (1970).



## VALOR Y LIMITACIONES DE LA TAXONOMIA NUMERICA

Después de exponer los distintos aspectos de la taxonomía numérica procede hacer una evaluación crítica de los mismos con el fin de delimitar el dominio de validez de este método.

En este capítulo se formularán consideraciones generales acerca de la importancia y las limitaciones de la taxonomía numérica. Examinaremos en primer término el feneticismo y sus aspectos filosóficos y en segundo término las técnicas numéricas.

## UNA EVALUACION DEL FENETICISMO

El mayor riesgo en la práctica del feneticismo radica en llevar el empirismo y el operacionismo a posiciones extremas. Según estas posiciones extremas todo aspecto teórico es superfluo hasta el punto de desechar la teoría como una especulación. Para ellas, la recopilación de datos es la única actividad científica válida.

Este punto de vista desconoce el hecho de que todo dato tiene sentido solamente dentro de un contexto teórico y que la acumulación de datos es infructuosa cuando no va acompañada de un análisis teórico (Bunge, 1961).

La parábola de Hull (1968) ilustra claramente los riesgos del empirismo y del operacionismo extremos: un granjero observó que sólo la mitad de sus árboles producían frutos; para aumentar su producción, comenzó a cortar los árboles "estériles". Tuvo éxito hasta que cortó el último árbol masculino; de allí en adelante, todos los otros árboles dejaron de producir.

Los términos teóricos son como los árboles masculinos: no son ni completamente empíricos ni completamente operacionales, pero son importantes para el progreso de la ciencia. El operacionismo y el empirismo son fructíferos sólo cuando no son extremos. Las versiones extremas del empirismo y del operacionismo son incompatibles con la teoría y las versiones moderadas imposibles sin ella.

Consideremos ahora los aspectos positivos del feneticismo, entre los cuales podemos señalar los siguientes:

- evita las definiciones imprecisas, por ejemplo: "las antenas de la especie A son largas";
- exige la interpretación empírica de ciertos conceptos como el de "homología";
- exige la comprobación empírica de las hipótesis;
- evita conceptos "fantasmas", por ejemplo, "esencias";
- al exigir un gran número de caracteres, obliga a acumular más información; e

- integra datos de una gran variedad de fuentes como la fisiología, la morfología, la química, promoviendo además relaciones interdisciplinarias dentro de la biología.

Sin duda, una de las contribuciones más importantes del feneticismo es el haber impulsado a la taxonomía hacia el reexamen de sus principios y la reestructuración de sus conceptos. Este reexamen no siempre fue el resultado de una coincidencia con los postulados del feneticismo, sino que muchos taxónomos fueron estimulados a explicitar sus principios cuando enfrentaron la controversia suscitada por esta doctrina.

#### UNA EVALUACION DE LAS TECNICAS NUMERICAS

La mayor crítica de que pueden ser objeto las técnicas numéricas no está dirigida a ellas mismas, sino a los investigadores que las aplican en la búsqueda de objetividad absoluta en el manejo de sus datos. Si se comparan las técnicas numéricas con la taxonomía clásica es evidente que el enfoque numérico es mucho más objetivo. Sin embargo, la objetividad absoluta es imposible pues la intervención de la subjetividad es ineludible, como en:

- la elección de las OTU;
- la determinación de homologías;
- la elección de los caracteres;
- la codificación de los caracteres;
- la determinación del número mínimo de OTU y caracteres empleados;
- la elección del coeficiente de similitud, y
- la elección de las técnicas de agrupamientos y ordenación.

El éxito o fracaso del uso de las técnicas numéricas depende, en gran medida, de las decisiones que adopte el investigador frente a la subjetividad.

En resumen, hay una región inalienable de la taxonomía (donde actúa esa subjetividad) en la cual las técnicas numéricas no pueden incorporarse. En los límites de esa región finaliza la taxonomía numérica y comienza el juicio del taxónomo.

Otro problema, aún sin resolver, es el de la significación estadística de las técnicas numéricas por su gran contenido empírico y por estar en cierto modo excluidas de las pruebas de significación. Esto está relacionado con la falta de una teoría estadística de la clasificación y de una definición matemática rigurosa del concepto de grupo ("cluster").

Entre las ventajas de las técnicas numéricas cabe señalar que:

- exigen precisión en los pasos taxonómicos y descripción de cada uno de ellos;
- requieren una clara definición de los caracteres;
- exigen la observación minuciosa de los organismos;
- sus resultados son sistemas generadores de hipótesis;

- permiten plantear nuevos problemas o replantear viejos problemas con un nuevo enfoque, y
- las clasificaciones obtenidas a partir de ellas son repetibles en gran parte del camino recorrido para construir las.

Debido a las virtudes de las técnicas numéricas, numerosos investigadores las han aplicado fuera de la clasificación biológica e incluso fuera de la biología.

A continuación presentaremos ejemplos extraídos de la extensa lista de aplicaciones de la taxonomía numérica fuera de la clasificación biológica: en biogeografía (Fisher, 1968); en ecología (Menni y Gosztonyi, 1982); en biología de poblaciones (Pahlen y colaboradores, 1980); en la clasificación biológica en un grupo fósil (Barnett, 1974); en estratigrafía (Damborenea y colaboradores, 1975); en antropología (Lorandi y colaboradores, 1979) en la medicina (Baron y Fraser, 1968); en psicología (Cattell, 1966); en geología (Marriam, 1966); en edafología (Etchevehere, 1971); en las artes y humanidades (Leed, 1966); en economía (Fisher y colaboradores, 1967) y en política y ciencias sociales (Alker, 1969).



## EPILOGO: UNA BUSQUEDA INCESANTE

¿Qué conclusiones generales pueden deducirse del desarrollo de la taxonomía numérica que hemos esbozado sólo en sus ideas fundamentales?

La taxonomía numérica es una de las principales protagonistas del vertiginoso renacimiento y transformación experimentados por la clasificación biológica en los últimos treinta años.

El enfoque propio de la taxonomía numérica comprende dos aspectos: uno filosófico, el feneticismo, y otro operativo, las técnicas numéricas.

El feneticismo sostiene que en la clasificación es preciso utilizar muchos caracteres y otorgar a cada uno de ellos el mismo peso o importancia sin tener en cuenta los acontecimientos filogenéticos.

Un análisis más profundo de esta doctrina demuestra que sus bases descansan sobre dos corrientes filosóficas del pensamiento científico: el empirismo y el operacionismo. El primero sostiene que el conocimiento sólo deriva de la experiencia sensible, mientras que el segundo otorga validez a un concepto científico sólo si pueden describirse las operaciones que conducen a establecerlo.

Las técnicas numéricas son el camino operativo para aplicar el feneticismo. Sin embargo, es posible aplicar las técnicas numéricas independientemente de los postulados feneticistas o practicar un feneticismo atemperado sin la ayuda de técnicas cuantitativas.

En la aplicación práctica de las técnicas numéricas, las operaciones se llevan a cabo en la siguiente secuencia: en un primer paso, se eligen los organismos a estudiar o unidades de estudio llamadas OTU. A continuación se escogen los caracteres que describen a esas unidades y se registran los estados de los caracteres presentes en los organismos. Con esta información se construye la matriz básica de datos (OTU x caracteres), a base de la cual, más un coeficiente, se calcula la similitud para cada par posible de OTU. De esta manera, se obtiene una matriz de similitud (OTU x OTU) que contiene los valores de similitud calculados. Utilizando la matriz de similitud y distintas técnicas (análisis de agrupamientos, ordenación, etc.) se logra la estructura taxonómica del grupo en estudio.

Todo este desenvolvimiento fue posible gracias a la aparición de las computadoras digitales que permiten procesar de manera eficiente y a gran velocidad la información proveniente de los numerosos caracteres utilizados.

La taxonomía numérica ha aportado al ámbito de la clasificación biológica precisión, observación minuciosa y una clara descripción de los pasos seguidos al establecer una clasificación. Asimismo, ha contribuido fundamentalmente y en forma notable al reexamen de los principios taxonómicos.

Hay dos consideraciones que hemos procurado transmitir en forma implícita en esta monografía, pero que deseamos explicitar en este epílogo.

- La taxonomía numérica no es un sucedáneo de la imaginación, sino un instrumento de ella.
- La taxonomía numérica no implica simplemente la habilidad de calcular, sino algo más importante: la capacidad de razonar cuantitativamente.

Es evidente que en la taxonomía numérica aún quedan cuestiones por resolver y que el problema de la clasificación biológica seguirá suscitando en el futuro respuesta de la vida más variada. Pero, sin lugar a dudas, la taxonomía numérica ha sido un aporte singular en la búsqueda incesante, por el hombre moderno, de una clave que le permita comprender la infinita complejidad de los seres vivos.



## AGRADECIMIENTOS

Los autores expresan su gratitud a los doctores Miguel O. Manceño, Susana E. Damborenea y Analía Lanteri, al profesor Jorge E. Bosch y a los licenciados María M. Cigliano y Jorge R. Casciotta por haber leído el manuscrito con espíritu crítico.

Asimismo, agradecen al técnico Víctor Hugo Calvetti, Miembro de la carrera del personal de apoyo a la investigación del Consejo Nacional de Investigaciones Científicas y Técnicas de la Argentina, el diseño y la realización gráfica de la portada y las figuras que ilustran esta monografía, y a la señorita Sheila Forziano el haber mecanografiado el texto.

Ninguna de estas personas es responsable en modo alguno de lo que esta monografía pueda tener de controvertible.

Finalmente, los autores hacen constar su reconocimiento al Consejo Nacional de Investigaciones Científicas y Técnicas de la Argentina, al cual pertenecen.



# BIBLIOGRAFIA

- (1) ADAMS, E.N. Consensus Technique and the Comparison of the Taxonomic Trees, *Syst. Zool.*, 21, 390 (1972).
- (2) ALKER Jr., H.R. Statistics and Politics: The Need for Causal Data Analysis. En: S.M. Lipser (ed.), Politics and the Social Sciences, Oxford University Press, Nueva York, N.Y., 328 págs. (1969).
- (3) ANONIMO. Trends, Priorities, and Needs in Systematic and Evolutionary Biology, *Syst. Zool.*, 23, 416 (1974).
- (4) BALL, G.H. Data Analysis in the Social Sciences: What About the Details? En: W.R. Rector (Chairman), Fall Joint Computer Conference, AFIPS Conference Proceedings, Spartan Books, Washington, D.C., vol. 27, part 1, 1100 págs. (1965).
- (5) BARNETT, R.S. An Application of Numerical Taxonomy to the Classification of the Nummulitidae (Foraminiferida), *J. Paleontol.*, 48, 1249 (1974).
- (6) BARON, D.N. y FRASER, P.M. Medical Applications of Taxonomic Methods, *Brit. Med. Bull.*, 24, 236 (1968).
- (7) BARTCHER, R.L. FORTRAN IV Program for Estimation of Cladistic Relationships Using the IBM 7040, *Kansas Geol. Surv. Computer Contrib.*, 6, 1 (1966).
- (8) BENZECRI, J.P. y colaboradores. L'Analyse des Données, Tomo 2: L'Analyses des Correspondances, Dunod, Paris, 632 págs. (1980).
- (9) BERTALANFFY, L. von. Perspectivas en la Teoría General de Sistemas. Alianza Universidad, Madrid, 166 págs. (1975).
- (10) BLACKITH, R.E. y REYMENT, R.A. Multivariate Morphometrics, Academic Press, Londres, 412 págs. (1971).
- (11) BLACKWELDER, R.E. y BOYDEN, A. The Nature of Systematics, *Syst. Zool.*, 1, 26 (1952).
- (12) BOCK, W.J. Philosophical Foundations of Classical Evolutionary Classification, *Syst. Zool.*, 22, 375 (1973).
- (13) BOCK, W.J. Principles and Methods of Comparative Analyses in Sociobiology. En: The Sociobiology Debate, A. Caplan (ed.), Harper & Row, Nueva York, N.Y., 514 págs. (1978).
- (14) BONHAM-CARTER, G.F. FORTRAN IV Program for Q-mode Cluster Analysis of Nonquantitative Data Using IBM 7090/7094 Computers, *Kansas Geol. Surv. Computer Contrib.*, 17, 1 (1967).
- (15) BORGMIEIER, T. Basic Questions of Systematics, *Syst. Zool.*, 6, 53 (1957).

- (16) BRUNDIN, L. Application of Phylogenetic Principles in Systematics and Evolutionary Theory, *Nobel Symp.*, 4, 473 (1968).
- (17) BUNGE, M. Causalidad. El Principio de Causalidad en la Ciencia Moderna, Eudeba, Buenos Aires, 403 págs. (1961).
- (18) BUNGE, M. La Investigación Científica. Ediciones Ariel, Barcelona, 955 págs. (1969).
- (19) CAIN, A.J. Deductive and Inductive Methods in Post-Linnaean Taxonomy, *Proc. Linn. Soc. Lond.*, 170th Session, 185, 217 (1959).
- (20) CAIN, A.J. The Use of Homology and Analogy in Evolutionary Theory. En: Methods of Inference from Animal to Human Behaviour, Von Cranach (ed.), Aldine, Chicago (1976).
- (21) CAIN, A.J. y HARRISON, G.A. An Analysis of the Taxonomist's Judgement of Affinity, *Proc. Zool. Soc. Lond.*, 131, 85 (1958).
- (22) CAMIN, J.H. y SOKAL, R.R. A Method for Deducing Branching Sequences in Phylogeny, *Evolution*, 19, 311 (1965).
- (23) CATTELL, R.B. Factor Analysis, Harper & Row, Nueva York, N.Y., 462 págs. (1952).
- (24) CATTELL, R.B. The Data Box: Its Ordering of Total Resources in Terms of Possible Relational Systems. En: R.B. Cattell (ed.), Handbook of Multivariate Experimental Psychology, Rand McNally, Chicago, 959 págs. (1966).
- (25) CLIFFORD, H.T. y STEPHENSON, W. An Introduction to Numerical Classification, Academic Press, Nueva York, N.Y., 229 págs. (1975).
- (26) COHEN, M. y NAGEL, E. Introducción a la Lógica y al Método Científico, Amorrortu Editores, Buenos Aires, Tomo I, 268 págs., Tomo II, 284 págs. (1971).
- (27) COLE, L.C. The Measurement of Interspecific Association, *Ecology*, 30, 411 (1949).
- (28) COLE, L.C. The Measurement of Partial Interspecific Association, *Ecology*, 38, 226 (1957).
- (29) COLLESS, D.H. The Phylogenetic Fallacy, *Syst. Zool.*, 16, 289 (1967).
- (30) COLLES, D.H. Congruence Between Morphometric and Allozyme Data for *Menidia* Species: A Reappraisal, *Syst. Zool.*, 29, 288 (1980).
- (31) COMAS, C.I., HUNZIKER, J.H. y CRISCI, J.V. Species Relationship in *Bulnesia* as Shown by Seed Protein Electrophoresis, *Biochem. System. Ecol.*, 7, 303 (1979).
- (32) CORMACK, R.M. A Review of Classification, *J. Roy. Statist. Soc. A*, 134, 321 (1971).
- (33) CRISCI, J.V. A Numerical-Taxonomic Study of the Subtribe Nassauviinae (Compositae, Mutisieae), *J. Arnold Arb.*, 55, 568 (1974).

- (34) CRISCI, J.V. Clasificación Biológica: Naturaleza, Objetivos, Fundamentos, Obra del Centenario del Museo de La Plata, *Bot.*, 3, 51 (1978).
- (35) CRISCI, J.V. Evolution in the Subtribe Nassauviinae (Compositae, Mutisieae): A Phylogenetic Reconstruction, *Taxon*, 29, 213 (1980).
- (36) CRISCI, J.V. La Especie: Realidad y Conceptos, *Symposia, VI Jornadas Argentinas de Zoología*, 21 (1981).
- (37) CRISCI, J. V. Parsimony in Evolutionary Theory: Law or Methodological Prescription? *J. Theoret. Biol.*, 97, 35 (1982).
- (38) CRISCI, J.V., HUNZIKER, J.H., PALACIOS, R.A. y NARANJO, C.A. A Numerical-Taxonomic Study of the Genus *Eulnesia* (Zygophyllaceae): Cluster Analysis, Ordination and Simulation of Evolutionary Trees, *Amer. J. Bot.*, 66, 133 (1979).
- (39) CRISCI, J.V. y STUESSY, T.F. Determining Primitive Character States for Phylogenetic Reconstruction, *Syst. Bot.*, 5, 112 (1980).
- (40) CHEETHAM, A.H. y HAZEL, J.E. Binary (Presence-Absence) Similarity Coefficients, *J. Paleontol.*, 43, 1130 (1969).
- (41) DAGNELIE, P. Contribution à l'Etude des Communautés végétales par l'Analyse factorielle, *Bull. Serv. Carte Phytogéogr.*, 8, 5, 7 y 93 (1960).
- (42) DAMBORENEA, S.E., MANCENIDO, M.O. y RICCARDI, A.C. Biofacies y Estratigrafía del Liásico de Piedra Pintada. Neuquén, Argentina, Actas Primer Congreso Argentino de Paleontología y Bioestratigrafía, 2, 173 (1975).
- (43) DAVIS, P.H. y HEYWOOD, V.H. Principles of Angiosperm Taxonomy, Van Nostrand, Nueva York, N.Y., 556 págs. (1963).
- (44) DE BEER, G. Homology, An Unsolved Problem, Oxford University Press, Nueva York, N.Y., 16 págs. (1971).
- (45) DE JONG, R. Some Tools for Evolutionary and Phylogenetic Studies, *Z. Zool. Syst. Evolut.-Forsch.*, 18, 1 (1980).
- (46) DEMIRMEN, F. Multivariate Procedures and FORTRAN IV Program for Evaluation and Improvement of Classifications, *Kansas Geol. Surv. Computer Contrib.*, 31, 1 (1969).
- (47) DUNCAN, T. y BAUM, B.R. Numerical Phenetics: its Uses in Botanical Systematics, *Ann. Rev. Ecol. Syst.*, 12, 387 (1981).
- (48) EADES, D.C. The Inappropriateness of the Correlation Coefficient as a Measure of Taxonomic Resemblance, *Syst. Zool.*, 14, 98 (1965).
- (49) EDMONDS, J.M. Numerical Taxonomic Studies on *Solanum* L. Section *Solanum* (Maurella), *Bot. J. Linn. Soc.*, 76, 27 (1978).
- (50) EDWARDS, A.W.F. y CAVALLI-SFORZA, L.L. Reconstruction of Evolutionary Trees. En: V.H. Heywood y J. McNeill (eds.), Phenetic and Phylogenetic Classification, 67, *Syst. Ass. Pub.*, 6, 1 (1964).

- (51) EHRLICH, P.R. y HOLM, R.W. Patterns and Populations, *Science*, 137, 652 (1962).
- (52) ELDREDGE, N. y CRACRAFT, J. Phylogenetics Patterns and the Evolutionary Process: Method and Theory in Comparative Biology, Columbia Univ. Press, Nueva York, N.Y. 349 págs. (1980).
- (53) ESTABROOK, G.F. Cladistic Methodology: A Discussion of the Theoretical Basis for the Induction of Evolutionary History, *Ann. Rev. Ecol. Syst.*, 3, 427 (1972).
- (54) ESTABROOK, G.F. Does Common Equal Primitive?, *Syst. Bot.*, 2, 36 (1977).
- (55) ESTABROOK, G.F. Some Concepts for the Estimation of Evolutionary Relationships in Systematic Botany, *Syst. Bot.*, 3, 146 (1978).
- (56) ETCHEVEHERE, P.H. Nuevas Perspectivas en el Campo de la Taxonomía y en Correlación de Suelos, VI Reunión Arq. de la Ciencia del Suelo, Córdoba, *Actas*, 2, 7 (1971).
- (57) FARINA, I.H. FORTRAN IV. Curso de Programación para Computadoras Digitales. Eudeba, Buenos Aires, 210 págs. (1972).
- (58) FARRIS, J.S. On the Cophenetic Correlation Coefficient, *Syst. Zool.*, 18, 279 (1969).
- (59) FARRIS, J.S. Methods for Computing Wagner Trees, *Syst. Zool.*, 19, 83 (1970).
- (60) FARRIS, J.S. The Hypothesis of Nonspecificity and Taxonomic Congruence, *Ann. Rev. Ecol. Syst.*, 2, 277 (1971).
- (61) FARRIS, J.S. On Comparing the Shapes of Taxonomic Trees, *Syst. Zool.*, 22, 50 (1973).
- (62) FISHER, D.R. A Study of Faunal Resemblance Using Numerical Taxonomy and Factor Analysis, *Syst. Zool.*, 17, 48 (1968).
- (63) FISHER, R.G., WILLIAMS, W.T. y LANCE, G.N. An Application of Techniques of Numerical Taxonomy to Company Information, *Econ. Rec.*, 43, 566 (1967).
- (64) FUNK, V.A. y STUESSEY, T.F. Cladistics for the Practicing Plant Taxonomist, *Syst. Bot.*, 3, 159 (1978).
- (65) GILMARTIN, A.J. Numerical Taxonomy-An Eclectic Viewpoint, *Taxon*, 16, 8 (1967).
- (66) GOODMAN, L.A. y KRUSKAL, W.H. Measures of Association for Cross Classifications, *J. Amer. Statist. Ass.*, 49, 732 (1954).
- (67) GOODMAN, L.A. y KRUSKAL, W.H. Measures of Association for cross Classifications, II. Further Discussion and References, *J. Amer. Statist. Ass.*, 54, 123 (1959).
- (68) GOODMAN, L.A. y KRUSKAL, W.H. Measures of Association for Cross Classification. III. Approximate Sampling Theory, *J. Amer. Statist. Ass.*, 58, 310 (1963).

- (69) GOWER, J.C. Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis, *Biometrika*, 53, 325 (1966).
- (70) GUEDES, M. La Méthode taxonomique d'Adanson, *Rev. Hist. Sci.*, 20, 361 (1967).
- (71) HARMAN, H.H. Modern Factor Analysis, University of Chicago Press, Chicago, 2a ed., 474 págs. (1967).
- (72) HARTIGAN, J.A. Clustering Algorithms, Wiley, Nueva York, N.Y., 351 págs. (1975).
- (73) HENNIG, W. Elementos de una Sistemática Filogenética, Eudeba, Buenos Aires, 353 págs. (1968).
- (74) HESSEN, J. Teoría del Conocimiento, Editorial Losada, Buenos Aires, 157 págs. (1938).
- (75) HEYWOOD, V.H. Plant Taxonomy Today. En: Modern Methods in Plant Taxonomy, V.H. Heywood (ed.), Academic Press, Londres, 312 págs. (1968).
- (76) HEYWOOD, V.H. Contemporary Objectives in Systematics. En: G.F. Estabrook (ed.), Proceedings of the Eighth International Conference on Numerical Taxonomy, W.H. Freeman, San Francisco, Ca., 429 págs. (1975).
- (77) HEYWOOD, V.H. y McNEILL, J. (eds.). Phenetic and Phylogenetic Classification, *Syst. Ass. Pub.*, 6, 1 (1964).
- (78) HOTELLING, H. Analysis of a Complex of Statistical Variables into Principal Components, *J. Educ. Psychol.*, 24, 417 y 498 (1933).
- (79) HSIAO, J. A. Test of the Nonspecificity Hypothesis Based on Chemical and Morphological Characters of the Different Plant-Parts of *Platanus*, *Taxon*, 24, 117 (1975).
- (80) HULL, D.L. The Effect of Essentialism on Taxonomy—Two Thousand Years of Stasis, *Brit. J. Phil. Sci.*, 15, 314, 16, 1 (1965).
- (81) HULL, D.L. The Operational Imperative: Sense and Nonsense in Operationism, *Syst. Zool.*, 17, 438 (1968).
- (82) HULL, D.L. Contemporary Systematic Philosophies, *Annual Rev. Ecol. Syst.*, 1, 19 (1970).
- (83) HUNZIKER, J.H. Evolución en *Bulnesia*: Un Enfoque Multidisciplinario, *Actas IV Congr. Latinoam. Genét.*, 2, 313 (1980).
- (84) INGLIS, W.G. The Purpose and Judgements of Biological Classifications, *Syst. Zool.*, 19, 240 (1970).
- (85) JANVIER, P., TASSY, P. y THOMAS, H. Le Cladisme, *La Recherche*, 117, 1396 (1980).
- (85a) JARDINE, N. y SIBSON, H. Mathematical Taxonomy, Wiley, Londres, 286 págs. (1971).
- (86) JOHNSON, L.A.S. Rainbow's End: The Quest for an Optimal Taxonomy, *Proc. Linn. Soc. New South Wales*, 93, 8 (1968).

- (87) KLUGE, A.G. Phylogenetic Relationships in the Lizard Family. Pygopodidae: An Evaluation of Theory, Methods and Data, Miscellaneous Publications Museum of Zoology, University of Michigan, 152, 1 (1976).
- (88) KLUGE, A.G. y FARRIS, J.S. Quantitative Phyletics and the Evolution of Anurans, *Syst. Zool.*, 18, 1 (1969).
- (89) KNELLER, G.F. Science as a Human Endeavor, Columbia Univ. Press, Nueva York, N.Y., 333 págs. (1978).
- (90) LANCE, G.N. y WILLIAMS, W.T. Computer Programs for Classification, Proc. ANCCAC Conference, Canberra, Paper 12, 3 (1966).
- (91) LEED, J. (ed.). The Computer and Literary Style: Introductory Essays and Studies, Kent State University Press, Kent, Ohio, 179 págs. (1966).
- (92) LE QUESNE, W.J. A Method of Selection of Characters in Numerical Taxonomy, *Syst. Zool.*, 18, 201 (1969).
- (92a) LEVIN, D.A. The Nature of Plant Species, *Science*, 204, 381 (1979).
- (93) LOCKHART, W.R. y LISTON, J. (eds.). Methods for Numerical Taxonomy, Amer. Soc. Microbiol., Bethesda, Md., 62 págs. (1970).
- (94) LORANDI, A.M., CRISCI, J.V., GONALDI, M.E. y CARAMAZANA, S.R. El Cambio Cultural en Santiago del Estero. Un Estudio de Taxonomía Numérica sobre Morfología de Bordes de Alfarería Ordinaria, *Relaciones*, 13, 85 (1979).
- (95) MAYR, E. Numerical Phenetics and Taxonomic Theory, *Syst. Zool.*, 14, 73 (1965).
- (96) MAYR, E. Especies Animales y Evolución, Universidad de Chile en coedición con Ediciones Ariel, S.A., Barcelona, 808 págs. (1968).
- (97) MAYR, E. Principles of Systematic Zoology, McGraw-Hill, Nueva York, N.Y., 428 págs. (1969).
- (98) MAYR, E. Cladistic Analysis or Cladistic Classification?, *Z. F. Zool. Syst. u. Evol.*, 12, 94 (1974).
- (99) MAYR, E. Biological Classification: Toward a Synthesis of Opposing Methodologies, *Science*, 214, 510 (1981).
- (100) McCAMMON, R.B. y WENNINGER, G. The Dendrograph, *Kansas Geol. Surv. Computer Contrib.*, 48, 1 (1970).
- (101) MEACHAM, C.A. A Manual Method for Character Compatibility Analysis, *Taxon*, 30, 591 (1981).
- (102) MENNI, R.C. y GOSZTONYI, A.E. Benthic and Semidermersal Fish Associations in the Argentine Sea, *Studies on Neotropical Fauna and Environment*, 17, 1 (1982).
- (103) MERRIAM, D.F. (ed.). Computer Applications in the Earth Sciences: Colloquium on Classification Procedures, *Kansas Geol. Surv. Computer Contrib.*, 7, 1 (1966).



- (104) MICHELOW, J. Introducción a la Computación, Monografía N° 21, Serie de Matemática, Programa de Monografías Científicas, OEA, Washington, D.C., ISBN-0-8270-1067-2, 88 págs. (1980).
- (105) MICHENER, C.D. Diverse Approaches to Systematics, *Evolut. Biol.*, 4, 1 (1970).
- (106) MICHENER, C.D. Discordant Evolution and the Classification of Allodapine Bees, *Syst. Zool.*, 6, 32 (1977).
- (107) MICHENER, C.D. y SOKAL, R.R. A Quantitative Approach to a Problem in Classification, *Evolution*, 11, 130 (1957).
- (108) MICKEVICH, M.F. Taxonomic Congruence, *Syst. Zool.*, 27, 143 (1978).
- (109) MICKEVICH, M.F. Taxonomic Congruence: Rohlf and Sokal's Misunderstanding, *Syst. Zool.*, 29, 162 (1980).
- (110) MICKEVICH, M.F. y JOHNSON, M.S. Congruence Between Morphological and Allozyme Data in Evolutionary Inference and Character Evolution, *Syst. Zool.*, 25, 260 (1976).
- (111) MINKOFF, E.C. The Effects on Classification of Slight Alterations in Numerical Technique, *Syst. Zool.*, 14, 196 (1965).
- (112) MORRISON, D.F. Multivariate Statistical Methods, McGraw-Hill, Nueva York, N.Y., 338 págs. (1967).
- (113) MOSS, W.W. Experiments with Various Techniques of Numerical Taxonomy, *Syst. Zool.*, 17, 31 (1968).
- (114) MOSS, W.W. NTPAK, a Manual for the Use of a Set of Numerical Taxonomy Programs, Adapted to and Written for the CDC 6400, University of California, Division of Entomology, Berkeley, Ca., 81 págs. (1970).
- (115) ONDRICK, C.W. y SRIVASTAVA, G.S. CORFAN-FORTRAN IV Computer Program for Correlation, Factor Analysis (Rand Q-mode) and Varimax Rotation, *Kansus Geol. Surv. Computer Contrib.*, 42, 1 (1970).
- (116) ORGANICK, E.I. A FORTRAN IV Primer, Addison-Wesley, Reading, Mass., 263 págs. (1966).
- (117) PAHLEN, von der, A.W., CRISCI, J.V., TELLERIAS POLO, W. y PEREZ LOPEZ, F. Clasificación de Poblaciones de Cebadilla Criolla (*Bromus unioloides*), y de Cebada Boliviana (*Hordeum vulgare*), *Actas IV Congr. Latinoam. Genét.*, 2, 207 (1980).
- (118) PETRIELLA, B. y CRISCI, J.V. Estudios Numéricos en Cycadales. II. Cycadales Actuales: Simulación de Árboles Evolutivos, Obra del Centenario del Museo de La Plata, 3, 151 (1978).
- (119) POGGIO, L. Cromatografía de Compuestos Fenólicos en el Género *Eulnesia* (Zygophyllaceae), *Bol. Soc. Arg. Bot.*, 19, 103 (1980).
- (120) POPPER, K.R. The Open Society and its Enemies, Routledge and Kegan Paul, Ltd., Londres, vol.1, 351 págs. (1950).
- (121) PRICE, D.J.S. Hacia una Ciencia de la Ciencia, Editorial Ariel, Barcelona, 181 págs. (1973).

- (122) PRIM, R.C. Shortest Connection Networks and Some Generalizations, *Bell Syst. Tech. J.*, **36**, 1389 (1957).
- (122a) RAVEN, P.H. Systematics and Plant Population Biology, *Syst. Bot.*, **1**, 284 (1976).
- (123) REYMENT, R.A., RAMDEN, H.A. y WAHLSTEDT, W.J. FORTRAN IV Program for the Generalized Statistical Distance and Analysis of Covariance Matrices for the CDC 3600 Computer, *Kansas Geol. Surv. Computer Contrib.*, **39**, 1 (1969).
- (124) REYMENT, R.A. y RAMDEN, H.A. FORTRAN IV Program for Canonical Variates Analysis for the CDC 3600 Computer, *Kansas Geol. Surv. Computer Contrib.*, **47**, 1 (1970).
- (125) ROBACK, S.S. y MOSS, W.W. Numerical Taxonomic Studies on the Congruence of Classifications for the Genera and Subgenera of Macropelopiini and Anatopyniini (Diptera: Chironomidae, Tanyptodinae), *Proc. Acad. Nat. Sci. Philadelphia*, **129**, 125 (1978).
- (126) ROHLF, F.J. A Numerical Taxonomic Study of the Genus *Ledes* (Diptera: Culicidae) with Emphasis on the Congruence of Larval and Adult Classifications, Ph.D. Thesis, University of Kansas, 98 págs. (1962).
- (127) ROHLF, F.J. A Randomization Test of the Nonspecificity Hypothesis in Numerical Taxonomy, *Taxon*, **14**, 262 (1965).
- (128) ROHLF, F.J. Sterograms in Numerical Taxonomy, *Syst. Zool.*, **17**, 246 (1968).
- (129) ROHLF, F.J. Adaptive Hierarchical Clustering Schemes, *Syst. Zool.*, **19**, 58 (1970).
- (130) ROHLF, F.J., KISHPAUGH, J. y KIRK, D. NT-SYS. Numerical Taxonomy System of Multivariate Statistical Programs, Tech. Rep. State University of New York at Stony Brook, Nueva York, N.Y., 87 págs. (1971).
- (131) ROHLF, F.J. y SOKAL, R.R. Coefficients of Correlation and Distance in Numerical Taxonomy, *Univ. Kansas Sci. Bull.*, **45**, 3 (1965).
- (132) ROHLF, F.J. y SOKAL, R.R. Comments on Taxonomic Congruence, *Syst. Zool.*, **29**, 97 (1980).
- (133) ROLLINS, R.C. On the Bases of Biological Classification, *Taxon*, **14**, 1 (1965).
- (134) ROMERO, F. Lógica e Introducción a la Problemática Filosófica, Editorial Losada, Buenos Aires, 302 págs. (1973).
- (135) ROSS, H.H. Biological Systematics, Addison-Wesley, Reading, Mass., 345 págs. (1974).
- (136) RUSE M. The Philosophy of Biology, Hutchinson University Library, Londres, 231 págs. (1973).
- (137) SCHLEE, D. Hennig's Principle of Phylogenetic Systematics, an "Intuitive" Statistico-Phenetic Taxonomy?, *Syst. Zool.*, **18**, 127 (1969).

- (138) SCHLEE, D. An Analysis of Numerical Phenetics, *Ent. Scand.*, 6, 1 (1975).
- (139) SEAL, H.L. Multivariate Statistical Analysis for Biologists, Methuen, Londres, 209 págs. (1964).
- (140) SIMPSON, G.G. Principles of Animal Taxonomy, Columbia University Press, Nueva York, N.Y., 247 págs. (1961).
- (141) SKERMAN, V.B.D. A Guide to the Identification of the Genera of Bacteria, Williams and Wilkins, Baltimore, Md., 2a. ed., 303 págs. (1967).
- (142) SNEATH, P.H.A. Classification of Microorganisms, Essays in Microbiology, Norris, J. R. y Richmond, M.A. (eds.), Wiley, Nueva York, N.Y., cap. 9, 1 (1978).
- (143) SNEATH, P.H.A. y SOKAL, R.R. Numerical Taxonomy. The Principles and Practice of Numerical Classification, Freeman, San Francisco, Ca., XV, 573 págs. (1973).
- (144) SOKAL, R.R. Distance as a Measure of Taxonomic Similarity, *Syst. Zool.*, 10, 70 (1961).
- (145) SOKAL, R.R. Numerical Taxonomy, *Sci. Amer.*, 215, 106 (1966).
- (146) SOKAL, R.R. Classification: Purposes, Principles, Progress, Prospects, *Science*, 185, 1115 (1974).
- (147) SOKAL, R.R. Numerical Phenetics: A Reply to Dieter Schlee, *Ent. Scand.*, 6, 209 (1975).
- (148) SOKAL, R.R. y CAMIN, J.H. The Two Taxonomies: Areas of Agreement and Conflict, *Syst. Zool.*, 14, 176 (1965).
- (149) SOKAL, R.R. y CROVELLO, T.J. The Biological Species Concept: A Critical Evaluation, *Amer. Natur.*, 104, 127 (1970).
- (150) SOKAL, R.R. y MICHENER, C.D. The Effects of Different Numerical Techniques on the Phenetic Classification of Bees of the *Ecophila* complex (Megachilidae), *Proc. Linn. Soc. Lond.*, 178, 59 (1967).
- (151) SOKAL, R.R. y ROHLF, F.J. The Comparison of Dendrograms by Objective Methods, *Taxon*, 11, 33 (1962).
- (152) SOKAL, R.R. y SNEATH, P.H.A. Principles of Numerical Taxonomy, Freeman, San Francisco, Ca., 359 págs. (1963).
- (153) SOLBRIG, O.T. The Phylogeny of *Gutierrezia*, An Eclectic Approach, *Brittonia*, 22, 217 (1970).
- (154) SOLBRIG, O.T. Polyphyletic Origin of Tetraploid Populations of *Gutierrezia sarothrae* (Compositae), *Madroño*, 21, 20 (1971).
- (155) SPENCE, N.A. y TAYLOR, P.J. Quantitative Methods in Regional Taxonomy, *Trapp. Geogr.*, 2, 1 (1970).
- (156) STEVENS, P.F. Evolutionary Polarity of Character States, *Ann. Rev. Ecol. Syst.*, 11, 333 (1980).

- (157) STUESSY, T.F. y CRISCI, J.V. Problems in the Determination of Evolutionary Directionality of Character States for Phylogenetic Reconstruction. En: *Cladistics Perspectives on the Reconstruction of Evolutionary History*, Duncan, T. y Stuessy, T.F. (eds.), Columbia Univ. Press, Nueva York, N.Y., (1983).
- (158) THOMPSON, W.R. The Philosophical Foundations of Systematics, *Can. Entomol.*, **84**, 1 (1952).
- (159) THOMPSON, W.R. Evolution and Taxonomy, *Studia Entomol.*, **5**, 549 (1962).
- (160) TRYON, R.C. y BAILEY, D.E. The BCTRY Computer System of Cluster and Factor Analysis, *Multivar. Behav. Res.*, **1**, 95 (1966).
- (161) WAGNER Jr., W.R. Problems in the Classification of Ferns, *Rev. Adv. Res.*, **1**, 841 (1961).
- (162) WAHLSTEDT, W.J. y DAVIS, J.C. FORTRAN IV Program for Computation and Display of Principal Components, *Kansas Geol. Surv. Computer Contrib.*, **21**, 1 (1968).
- (163) WALKER, M. El Pensamiento Científico, Editorial Grijalvo, México, 196 págs. (1968).
- (164) WARBURTON, F.E. The Purposes of Classifications, *Syst. Zool.*, **16**, 241 (1967).
- (165) WATROUS, L.E. y WHEELER, Q.D. The Out-Group Comparison Method of Character Analysis, *Syst. Zool.*, **30**, 1 (1981).
- (166) WHITTAKER, R.H. (ed.). Handbook of Vegetation Science, Part V, Dr. W. Junk b.v., Publishers, La Haya, Holanda, 737 págs. (1973).
- (167) WILEY, E.O. Phylogenetics. The Theory and Practice of Phylogenetic Systematics, Wiley, Nueva York, N.Y., XV, 439 págs. (1981).
- (168) WILLIAMS, W.T. Numbers, Taxonomy and Judgement. *Bol. Rev.*, **33**, 379 (1967).
- (169) WILLIAMS, W.T. y DALE, M.B. Fundamental Problems in Numerical Taxonomy, *Adv. Bot. Res.*, **2**, 35 (1965).
- (170) WILSON, E.O. A Consistency Test for Phylogenies Based on Contemporaneous Species, *Syst. Zool.*, **14**, 214 (1965).
- (171) WISHART, D. Mode Analysis, a Generalization of Nearest Neighbour which Reduces Chaining Effects. En: Cole, A.J. (ed.), Numerical Taxonomy. Proceedings of the Colloquium in Numerical Taxonomy Held at the University of St. Andrews, September 1968, Academic Press, Londres, 324 págs. (1969 a).
- (172) WISHART, D. FORTRAN II Programs for 8 Methods of Cluster Analysis (CLUSTAN I), *Kansas Geol. Surv. Computer Contrib.*, **38**, 1 (1969 b).
- (173) WISHART, D. A General Tripartite Clustering Method and Similarity Generating Function, *Statist. Div. Rep. N° R-31*, Civil Surv. Dep., Whitehall, Londres, 1 (1972).

Publicadas

**Serie de matemática**

- N° 1. La Revolución en las Matemáticas Escolares, por el Consejo Nacional de Maestros de Matemáticas de los Estados Unidos de América.
- N° 2. Espacios Vectoriales y Geometría Analítica, por Luis A. Santaló.
- N° 3. Estructuras Algebraicas I, por Enzo R. Gentile.
- N° 4. Historia de las Ideas Modernas en la Matemática, por José Babini.
- N° 5. Algebra Lineal, por Orlando E. Villamayor.
- N° 6. Algebra Linear e Geometría Euclídiana, por Alexandre Augusto Martins Rodrigues.
- N° 7. El Concepto de Número, por César A. Trejo.
- N° 8. Funciones de Variable Compleja, por José I. Nieto.
- N° 9. Introducción a la Topología General, por Juan Horváth.
- N° 10. Funções Reais, por Djairo G. de Figueiredo.
- N° 11. Probabilidad e Inferencia Estadística, por Luis A. Santaló.
- N° 12. Estructuras Algebraicas II (Algebra Lineal), por Enzo R. Gentile.
- N° 13. La Revolución en las Matemáticas Escolares (Segunda Fase), por Howard F. Fehr, John Camp y Howard Kellog.
- N° 14. Estructuras Algebraicas III (Grupos Finitos), por Horacio H. O'Brien.
- N° 15. Introducción a la Teoría de Grafos, por Fausto A. Toranzos.
- N° 16. Estructuras Algebraicas IV (Algebra Multilineal), por Artibano Micali y Orlando E. Villamayor.
- N° 17. Introdução à Análise Funcional: Espaços de Banach e Cálculo Diferencial, por Leopoldo Nachbin.
- N° 18. Introdução a la Integral de Lebesgue en la Recta, por Juan Antonio Gatica.
- N° 19. Introducción a los Espacios de Hilbert, por José I. Nieto.
- N° 20. Elementos de Biomatemática, por Alejandro B. Engel.
- N° 21. Introducción a la Computación, por Jaime Michelow.
- N° 22. Estructuras Algebraicas V (Teoría de Cuerpos), por Héctor A. Merklen.
- N° 23. Estructuras Algebraicas VI (Formas Cuadráticas), por Francisco M. Piscoya.
- N° 24. Estructuras Algebraicas VII (Estructuras de Algebras), por Artibano Micali.

**Serie de física**

- N° 1. Concepto Moderno del Núcleo, por D. Allan Bromley.
- N° 2. Panorama de la Astronomía Moderna, por Félix Cernuschi y Sayd Codina.
- N° 3. La Estructura Electrónica de los Sólidos, por Leopoldo M. Falicov.
- N° 4. Física de Partículas, por Igor Saavedra.
- N° 5. Experimento, Razonamiento y Creación en Física, por Félix Cernuschi.
- N° 6. Semiconductores, por George Bemski.
- N° 7. Aceleradores de Partículas, por Fernando Alba Andrade.
- N° 8. Física Cuántica, por Onofre Rojo y Harold V. McIntosh.
- N° 9. La Radiación Cósmica, por Gastón R. Mejía y Carlos Aguirre.

- N° 10. Astrofísica, por Carlos Jaschek y Mercedes C. de Jaschek.
- N° 11. Ondas, por Oscar J. Bressan y Enrique Gaviola.
- N° 12. El Láser, por Mario Garavaglia.
- N° 13. Teoría Estadística de la Materia, por Antonio E. Rodríguez y Roberto E. Caligaris.
- N° 14. Aplicações da Teoria de Grupos na Espectroscopia Raman e do Infra-Vermelho, por Jorge Humberto Nicola y Anildo Bristoti.

#### **Serie de química**

- N° 1. Cinética Química Elemental, por Harold Behrens LeBas.
- N° 2. Bioenergética, por Isaías Raw y Walter Colli.
- N° 3. Macromoléculas, por Alejandro Paladini y Moises Burachik.
- N° 4. Mecanismo de las Reacciones Orgánicas, por Jorge A. Brieux.
- N° 5. Elementos Encadenados, por Jacobo Gómez Lara.
- N° 6. Enseñanza de la Química Experimental, por Francisco Giral.
- N° 7. Fotoquímica de Gases, por Ralf-Dieter Penzhorn.
- N° 8. Introducción a la Geoquímica, por Félix González-Bonorino.
- N° 9. Resonancia Magnética Nuclear de Hidrógeno-1 y de Carbono-13, por Pedro Joseph-Nathan.
- N° 10. Cromatografía Líquida de Alta Presión, por Harold M. McNair y Benjamin Esquivel H.
- N° 11. Actividad Óptica, Dispersión Rotatoria Óptica y Dicroísmo Circular en Química Orgánica, por Pierre Crabbé.
- N° 12. Espectroscopia Infrarroja, por Jesús Morcillo Rubio.
- N° 13. Polarografía, por Alejandro J. Arvía y Jorge A. Bolzan.
- N° 14. Paramagnetismo Electrónico, por Juan A. McMillan.
- N° 15. Introducción a la Estereoquímica, por Juan A. Garbarino.
- N° 16. Cromatografía en Papel y en Capa Delgada, por Xorge A. Domínguez.
- N° 17. Introducción a la Espectrometría de Masa de Sustancias Orgánicas, por Otto R. Gottlieb y Raimundo Braz Filho.
- N° 18. Cinética Química, por Rodolfo V. Caneda.
- N° 19. Fuerzas Intermoleculares, por Mateo Díaz Peña.
- N° 20. Físico-Química de Superficies, por Tibor Rabockai.
- N° 21. Corrosión, por José R. Galvele.
- N° 22. Introducción a la Electroquímica, por Dionisio Posadas.
- N° 23. Cromatografía de Gases, por Harold M. McNair.
- N° 24. Cinética de Disolución de Medicamentos, por Edison Cid Cárcamo.
- N° 25. Introducción a la Química de Suelos, por Elemer Bornemisza.

#### **Serie de biología**

- N° 1. La Genética y la Revolución en las Ciencias Biológicas, por José Luis Reissig.
- N° 2. Bases Ecológicas de la Explotación Agropecuaria en la América Latina, por Guillermo Mann F.
- N° 3. La Taxonomía y la Revolución en las Ciencias Biológicas, por Elías R. de la Sota.
- N° 4. Principios Básicos para la Enseñanza de la Biología, por Oswaldo Frota-Pessoa.
- N° 5. A Vida da Célula, por Renato Basile.
- N° 6. Microorganismos, por J. M. Gutiérrez-Vázquez.
- N° 7. Principios Generales de Microbiología, por Norberto J. Palleroni.
- N° 8. Los Virus, por Enriqueta Pizarro-Suárez y Gamba.
- N° 9. Introducción a la Ecología del Bentos Marino, por Manuel Vegas Vélez.
- N° 10. Biosíntesis de Proteínas y el Código Genético, por Jorge E. Allende.

- N° 11. Fundamentos de Inmunología e Inmunoquímica, por Félix Córdoba Alva y Sergio Estrada-Parra.
- N° 12. Bacteriófagos, por Romilio Espejo T.
- N° 13. Biogeografía de América Latina, por Angel L. Cabrera y Abraham Willink.
- N° 14. Relación Hospedante-Parásito. Mecanismo de Patogenicidad de los Microorganismos, por Manuel Rodríguez Leiva.
- N° 15. Genética de Poblaciones Humanas, por Francisco Rothhammer.
- N° 16. Introducción a la Ecofisiología Vegetal, por Ernesto Medina.
- N° 17. Aspectos de Biología Celular y la Transformación Maligna, por Manuel Rieber.
- N° 18. Transporte a Través de la Membrana Celular, por P. J. Garrahan y A. F. Rega.
- N° 19. Duplicación Cromosómica y Heterocromatina a Nivel Molecular y Citológico, por Néstor O. Bianchi.
- N° 20. Citogenética Básica y Biología de los Cromosomas, por Francisco A. Sáez y Horacio Cardoso.
- N° 21. Ecología de Poblaciones Animales, por Jorge E. Rabinovich.
- N° 22. Metodología para el Estudio de la Vegetación, por Silvia D. Matteucci y Aída Colma.
- N° 23. Los Sistemas Ecológicos y la Humanidad, por Ariel E. Lugo y Gregory L. Morris.
- N° 24. A Germinação das Sementes, Por Luiz Gouvêa Labouriau.
- N° 25. Introducción a la Farmacocinética, por Edison Cid Cárcamo.
- N° 26. Introducción a la Teoría y Práctica de la Taxonomía Numérica, por Jorge Víctor Crisci y María Fernanda López Armengol.

#### En preparación

#### **Serie de matemática**

131

Ecuaciones en Derivadas Parciales, por Lorenzo Lara-Carrero.  
 Geometrías Finitas, por Oscar Barriaga.  
 Álgebra Elemental, por Leopoldo Nachbin.

#### **Serie de física**

Teoría de Fluidos en Equilibrio, por Antonio E. Rodríguez y Roberto E. Caligaris.  
 Geofísica, por Alvaro F. Espinosa.  
 Superconductividad, por Miguel Kiwi.  
 Efecto Mössbauer, por Jacques A. Dannon.  
 Fundamentos de Cristalografía Física, por Jaime Rodríguez Lara.  
 Introducción a la Espectroscopia Atómica, por Mario Garavaglia y Athos Giacchetti.  
 Aplicaciones Metrológicas del Láser, por Mario Garavaglia.  
 Teoría Cuántica del Impulso Angular, por Manuel de Llano y Mauricio Fortes.

#### **Serie de química**

Catálisis Homogénea, por Eduardo Humeres A.  
 Catálisis Heterogénea, por Sergio Droguett.  
 Físicoquímica de Interfases, por Francisco Javier Garfias.  
 Introducción a la Electrocatálisis, por Alejandro J. Arvía y María Cristina Giordano.  
 Química de Sólidos, por Julio César Bazán.  
 Química Bioinorgánica, por Henrique E. Toma.  
 Introducción al Estudio de los Productos Naturales, por Eduardo G. Gros.

## **Serie de biología**

**Etología:** El Estudio del Comportamiento Animal, por Raúl Vaz-Ferreira.

**Principios Básicos de la Contracción Muscular,** por Carlos Caputo.

**Clastogénesis y Contaminación Ambiental,** por Fernando Noel Dolout.

**Fotosíntesis,** por Rubén H. Vallejos.

**Cromosomas Humanos y de Primates,** por Máximo E. Drets y Héctor Seuanez.

**¿Qué es la Diferenciación Celular?,** por Roberto B. García y Susana Pereyra-Alfonso.

**Limnología Sanitaria. Estudio de la Pulución de Aguas Interiores,** por Samuel Murgel Branco.

**Aprovechamiento de Aguas Dulces y el Cultivo de Peces,** por Argentino Bonetto y Hugo P. Castillo.