

## CAPÍTULO 9

### CLADÍSTICA: MÉTODOS CUANTITATIVOS

CIGLIANO, M. MARTA, MARTA S. FERNÁNDEZ Y ANALÍA A. LANTERI



#### GENERALIDADES

La Cladística es un método de reconstrucción filogenética desarrollado a partir de las ideas del entomólogo alemán Willi Hennig y basado en el principio de Simplicidad o Parsimonia. Sin embargo, ha sido James Farris el autor que más ha contribuido al desarrollo numérico de la Cladística (Kluge & Farris, 1969; Farris, 1970, 1980, 1983, 1986; Farris *et al.*, 1982). En la práctica, este principio se aplica mediante algoritmos matemáticos computarizados, que tienen por objeto hallar el o los árboles más cortos (con menor número de pasos o cambios entre los estados de caracteres = mínima homoplasia) para un conjunto de datos determinado. Como generalmente resulta impracticable evaluar todos los árboles posibles para un set de datos (excepto que los taxones analizados no sean más de 10 o 20), se generan árboles cuya longitud es evaluada durante el procedimiento computacional y se retienen aquéllos más cortos (Goloboff, 1998). La obtención del árbol más corto (= óptimo) no es un procedimiento directo, sino que existen distintas estrategias de análisis, a partir de las cuales se obtendrán resultados más o menos confiables en cuanto a la posibilidad de haber obtenido el árbol más corto o todos los árboles más cortos.

En la figura 1a, se ilustra un cladograma con sus componentes. Términos tales como raíz, taxa terminales o nodos no fueron utilizados por Hennig (1968) sino que se asocian con los análisis de la Cladística cuantitativa o numérica. Asimismo en cladogramas publicados en la actualidad, que incluyen numerosos taxa y a veces varios miles de caracteres (cuando se emplean secuencias de ADN), las sinapomorfías y homoplasias no se indican sobre las ramas (ver cladogramas en Capítulo 8) sino que directamente se brindan valores que expresan el grado de soporte de dichas ramas (ver Capítulo 10). En Cladística se suelen usar los términos cladograma y árbol como sinónimos, aunque por lo general no se ilustran como árboles (con las longitudes de las ramas proporcionales a la cantidad de cambio acumulado).

**Raíz.** Es la base del cladograma.

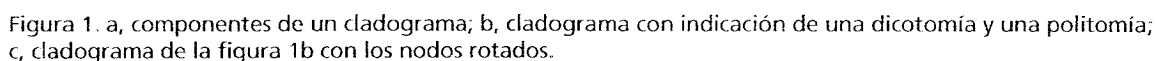
**Taxa terminales.** Son las unidades de estudio.

**Nodos internos.** Son los puntos de ramificación del cladograma. Los nodos que unen los taxa terminales pueden ser rotados sobre sí mismos expresando las mismas relaciones (Fig. 1 b-c)

**Ramas internas.** Son los segmentos que unen los nodos internos.

**Ramas externas.** Son los segmentos que unen los nodos internos y los *outgroups*.

En el cladograma de la figura 1a las relaciones entre los taxa son todas dicotómicas (cada nodo interno está conectado con otros dos nodos internos o taxones terminales); en la figura 1b, hay una politomía, ya que uno de los nodos internos está conectado con más de dos nodos internos o taxones terminales; en la figura 1c se han rotado los nodos del cladograma 1b, pero las relaciones entre los taxa no se alteran.



En Cladística cuantitativa se emplean como taxa terminales, además de los integrantes del *ingroup*, uno, dos o más grupos externos (*outgroups*), filogenéticamente próximos al *ingroup*. Esta es una diferencia importante con la argumentación hennigiana clásica, en que el *outgroup* se empleaba para polarizar los caracteres *a priori*, pero raramente se incluía en el análisis. En Cladística numérica la polaridad se infiere a partir del cladograma (Nixon & Carpenter, 1993), es decir *a posteriori* del análisis cladístico.

Un tipo de estrategia que se aplica con frecuencia en Cladística, es el análisis “simultáneo y no restringido” de *ingroup* y *outgroups* (Kitching *et al.*, 1998). Se denomina simultáneo porque los integrantes de ambos grupos son analizados al mismo tiempo, y no restringido o de parsimonia global, porque las relaciones del *outgroup* no están definidas antes del análisis. La ventaja es que mediante este tipo de análisis se ponen a prueba las hipótesis previas sobre la delimitación de *ingroup* y *outgroups*, de modo que podría suceder que al cabo del mismo, algunos de los miembros del *ingroup* se ubicaran fuera del grupo, o que algún *outgroup* resultara ser parte del *ingroup*.

En caso que los taxa terminales sean supraespecíficos, se han propuesto dos alternativas para su tratamiento (Yeates, 1995): el método del ejemplar (*exemplar*) y el del plan base (*groundplan*). Este último consiste en examinar todas las especies que conforman el taxón supraespecífico (familia, tribu, género, etc.) y deducir cual es la condición o estado ancestral del mismo. El método del ejemplar consiste en incluir como taxones terminales solo uno o algunos representantes de dicho taxón, por ejemplo la especie tipo de cada género (ver Capítulo 2). Cuanto mayor sea la diversidad dentro de un taxón terminal, tanto más especies "ejemplares" deberán incluirse en el análisis.

138

## ELECCIÓN DE UN CRITERIO DE OPTIMALIDAD O MODELO DE PARSIMONIA

En Cladística cuantitativa o numérica, la selección de caracteres, el establecimiento de homologías primarias y la codificación se realiza de modo similar al explicado en el Capítulo 8 referido a Cladística no numérica. Los programas de computación utilizados admiten dos opciones para tratar los datos: en forma no ordenada o no aditiva, y en forma ordenada o aditiva.

**Caracteres no ordenados o no aditivos:** Los estados no pueden ser ordenados en una secuencia de grados de atributos, por ejemplo, el color de ojos azul, rojo y amarillo; o los nucleótidos de las bases nitrogenadas Adenina, Timina, Citosina, y Guanina, cuando se trabaja con secuencias de ADN (ver Capítulos 4 y 10).

**Caracteres ordenados o aditivos:** Los estados se pueden ordenar en una secuencia determinada, sobre la base de una magnitud creciente o decreciente en la cualidad estudiada (morfocline o estados del desarrollo ontogenético). El reconocimiento de un orden implica la postulación de una hipótesis sobre la secuencia de cambio entre los diferentes estados del carácter o serie de transformación, sobre la base de su similitud. Dicha secuencia puede ser lineal o ramificada (ver Capítulo 8). Algunos programas de computación como el PAUP, permiten ingresar secuencias ramificadas.

La decisión de analizar los datos multiestados como ordenados tiene una gran incidencia en los resultados del análisis cladístico, y comporta la elección de un "criterio de optimalidad o modelo de parsimonia", es decir la elección de un modelo de transformación de los caracteres. Si bien la parsimonia busca minimizar el número de pasos en las transformaciones entre estados de caracteres, los distintos modelos difieren en las restricciones que imponen a dichas transformaciones (Forey *et al.*, 1992; Lipscomb, 1994). Los modelos empleados con mayor frecuencia en el análisis de datos morfológicos son los de Wagner (= Farris) y de Fitch; y para secuencias de ADN, el de Fitch o el de parsimonia generalizada (Kitching *et al.*, 1998; Schuh, 2000). Cuando los datos multiestado se analizan de forma ordenada, se está aplicando parsimonia de Wagner (= Farris), pero si los datos se analizan como no ordenados, el modelo que se aplica es el de Fitch, por esa razón, la longitud e inclusive la topología de los árboles puede variar según los caracteres se analicen de una u otra forma. Los modelos de Wagner y de Fitch admiten que durante su transformación, los estados de caracteres homoplásicos se expresen como paralelismos o como reversiones, otros modelos como el de Dollo, no admite los paralelismos (Kitching *et al.*, 1998).

### Parsimonia de Wagner ó de Farris (Wagner, 1961; Kluge & Farris, 1969; Farris, 1970)

Se emplea cuando se analizan caracteres binarios y multiestados ordenados. Las probabilidades de cambios en ambos sentidos son iguales y se suma un paso por cada transformación de un estado a otro, de modo que cambios 0-1 y 1-2 tienen un costo 1, pero cambios 0-2 tienen un costo 2. Para un carácter con tres estados ordenados, las transformaciones posibles son tres:

0 → 1 → 2      2 → 1 → 0      0 ← 1 → 2

### Parsimonia de Fitch (Fitch, 1971)

Se aplica para caracteres binarios (en este caso no difiere de Wagner) o multiestados desordenados, tales como los datos de secuencias de ácidos nucleicos. Cada paso puede derivarse de otro en cualquier orden y se contabiliza siempre un paso por cada transformación de modo que cambios 0-2 o 0-1 tienen un costo de 1. Por ejemplo, para un carácter con tres estados habrá 9 transformaciones posibles.

0 → 1 → 2      0 → 2 → 1      1 ← 0 → 2  
1 → 2 → 0      1 → 0 → 2      0 ← 1 → 2  
2 → 0 → 1      2 → 1 → 0      1 ← 2 → 0

## CODIFICACIÓN DE DATOS FALTANTES E INAPLICABLES. TAXA CON CARACTERES POLIMÓRFICOS

Varios especialistas han discutido algunos problemas derivados de la codificación de los caracteres empleados en análisis numéricos, por ejemplo cómo deben codificarse los datos inciertos (faltantes e inaplicables). A estos datos muchas veces se los suele codificar de igual forma (con un signo de pregunta "?"), pero no deberían ser tratados de este modo pues la naturaleza de los mismos es diferente (Platnick *et al.*, 1991).

**Datos faltantes o no observables.** Esto puede deberse al deficiente estado de conservación de algunos de los ejemplares en estudio, a que un determinado estado de desarrollo (e.g. larva, huevo) o alguno de los sexos (macho o hembra) aún no han sido hallados o estudiados; a que alguna posición nucleotídica no pudo secuenciarse correctamente, etc. El signo "?" en tales casos representa cualquiera de los estados existentes en los restantes taxa. Por ejemplo, si se tratara de un carácter binario, el ? podría representar **1 o 0**. Cuando se construye el árbol ambas posibilidades son evaluadas en la búsqueda del cladograma más corto (Lipscomb, 1998).

Taxa con numerosos datos faltantes pueden dificultar la búsqueda del árbol más corto y en ciertos casos conviene excluirlos del análisis (Lipscomb, 1998). Esto ocurre frecuentemente cuando se analizan datos en Paleontología.

**Datos inaplicables o inexistentes.** El carácter no existe en el taxón en estudio. Es relevante para resolver una parte del cladograma, pero irrelevante o inaplicable para resolver otra parte (Lipscomb, 1998). Por ejemplo, si se desea analizar la filogenia de los hexápodos, el carácter tipo de alas es inaplicable a los apterigotas. El "?" no representa ninguno de los estados existentes en los restantes taxa. Si se tratara de un carácter binario, el "?" no representaría **ni 1 ni 0**. En este caso se recomienda corregir la codificación a fin de evitar los datos inaplicables, por ejemplo, incorporando la ausencia como un estado más, al principio de la serie de transformación. Una corrección posible en el caso de los hexápodos sería la siguiente: alas ausentes (0), alas membranosas (1), alas anteriores membranosas y posteriores transformadas en balancines (2), alas anteriores transformadas en hemiélitros y posteriores membranosas (3), alas anteriores transformadas en élitros y posteriores membranosas (4).

También se suelen codificar erróneamente con un "?" los **caracteres polimórficos** (con más de un estado del mismo carácter en cada taxón). En este caso el "?" representa **1 y 0** (Platnick *et al.*, 1991). Nixon & Davis (1991) criticaron este tipo de codificación ya que no tiene en cuenta la variación en los taxa terminales durante el análisis cladístico. El problema se puede resolver dividiendo los taxa con caracteres polimórficos en distintas unidades terminales, del modo que se indicó previamente en este capítulo, bajo el título "Elección de taxa terminales". Otra estrategia, es usar un programa computarizado, como el NONA (Goloboff, 1996) o el PAUP (Swofford, 1999), que permiten ingresar los distintos estados presentes en un taxa polimórfico y analizar dichos taxa.

Como se ha dicho previamente, la Cladística cuantitativa no requiere ninguna información *a priori* acerca de la polaridad de los caracteres. Las distribuciones de los caracteres en los taxa es la única información necesaria para establecer las hipótesis de relación en el grupo estudiado. Esto es así, dado que la mayoría de los algoritmos computarizados para obtener cladogramas construyen primeramente redes (= *networks*), es decir árboles no enraizados, y al ubicar la raíz, generalmente entre el grupo interno y el o los grupos externos, se determina la polaridad de los caracteres (Nixon & Carpenter, 1993).

## BÚSQUEDA DE LOS ÁRBOLES DE LONGITUD MÍNIMA

Los algoritmos de Parsimonia implementados para obtener cladogramas, mediante el uso de programas de computación como el Hennig 86 (Farris, 1988), NONA y Pee-Wee (Goloboff, 1996), PAUP (Swofford, 1999) y TNT (Goloboff *et al.*, 2003) construyen árboles no enraizados (*networks*),

procedimiento que reduce el número de cálculos a realizar, ya que para cada árbol no enraizado, existen varios árboles con raíz posibles (Fig. 2). El número de pasos del árbol o la parsimonia global de los caracteres no depende de la posición de la raíz, sin embargo la ubicación de la raíz determina, además de la polaridad de los caracteres, la topología final del árbol (Nixon & Carpenter, 1993).

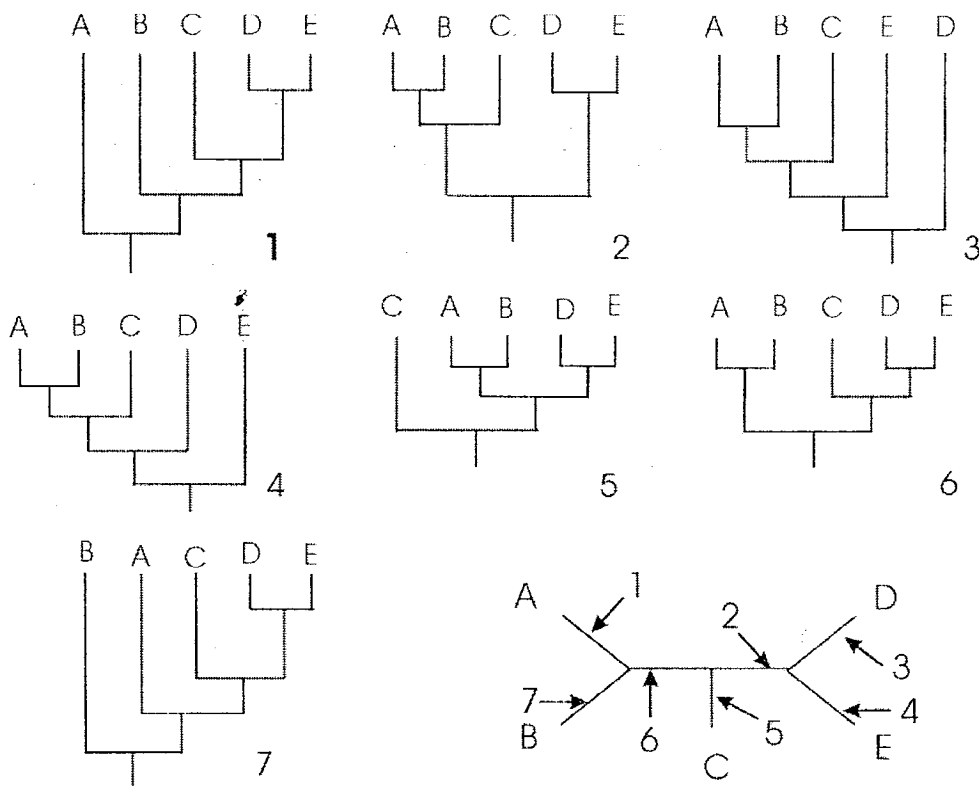


Figura 2. Árbol no enraizado de cinco taxones (A hasta E) y los árboles con raíz derivados del mismo (1-7). Las flechas con números 1 a 7 indican las diferentes posiciones en que se puede insertar la raíz (modificado de Nieto Feliner, 1999).

El primer algoritmo para construir árboles no enraizados corresponde a Kluge & Farris (1969) y Farris (1970) y se conoce como algoritmo de Wagner. Este algoritmo genera árboles óptimos en ausencia de homoplasia, sin embargo, para matrices de más de 15-20 taxa y datos con considerable homoplasia, es probable que no permita encontrar el árbol más corto o todos los árboles más cortos (Goloboff, 1998).

#### Algoritmo de Wagner (Fig. 3)

Procede de la siguiente manera:

a. Conecta el primer taxón terminal que aparece en la matriz de datos (generalmente coincide con el *outgroup*) con el taxón más próximo (es el que se diferencia en un menor número de caracteres). En el ejemplo de la figura 3 el *outgroup* OG se conecta con A, ya que éste difiere sólo en dos caracteres, en tanto que B y C difieren en tres y cuatro caracteres, respectivamente.

b. Busca el siguiente taxón más próximo y lo inserta en el punto que produce un menor incremento en la longitud del árbol (en el ejemplo se debe incorporar B). En caso de empate suele ingresar el taxón que aparece primero en el orden de la matriz de datos. En el punto de unión se ubica el primer nodo (nodo 1), para el cual será preciso calcular los estados de los caracteres. El estado de cada carácter del nodo, será igual a la mediana (valor medio) entre los estados registrados para dicho carácter, en los tres taxa a los cuales se ha conectado (mediana entre OG, A y B).

c. El próximo taxón (en el ejemplo es C) se incorporará al segmento que comporte el menor incremento en longitud del árbol y en el punto de unión se ubicará el nodo 2. Para saber a qué segmento del *network* deberá conectarse C, es preciso calcular todas las opciones posibles, es

decir, las distancias entre C y los segmentos OG-Nodo 1, Nodo 1-A, Nodo 1-B. Una vez realizados estos cálculos, se elegirá la opción que produzca menor incremento en longitud. En el ejemplo de la figura 3 la conexión de C que menos incrementa la longitud del *network* es con el Nodo1-B.

d. Para incorporar los restantes taxones, si los hubiera, se procede de igual modo que en el paso anterior, solo que al incrementarse el número de taxones y de caracteres, aumenta también el número de cálculos que es preciso realizar.

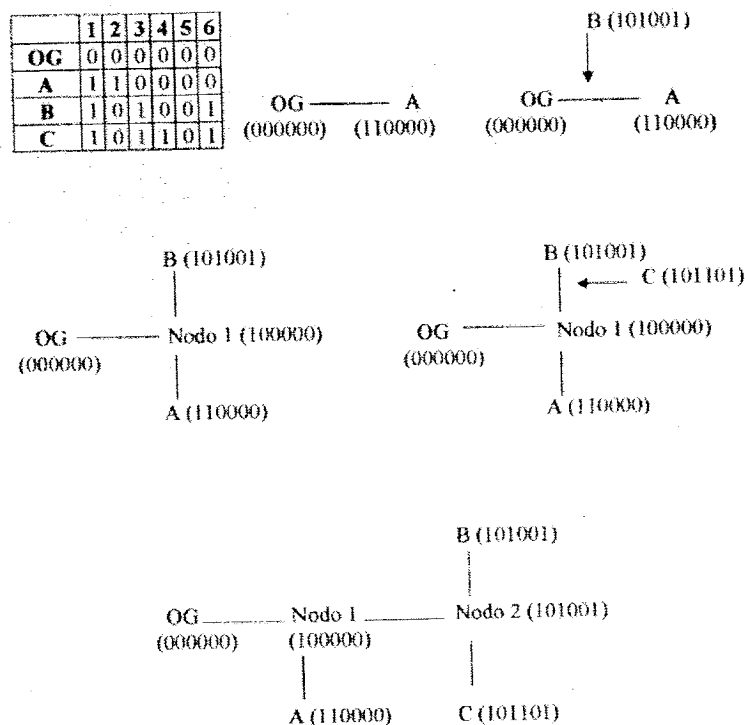


Figura 3 Matriz de datos y pasos para la construcción de un árbol de Wagner de cuatro taxones (A-B-C y un *outgroup*).

El algoritmo de Wagner permite encontrar árboles más cortos que si se generaran al azar y por lo tanto constituye un buen punto de partida para la aplicación de otras estrategias de búsqueda más exhaustivas. Existen distintas opciones en lo que respecta a las secuencias de incorporación de los taxa al árbol de Wagner (Forey *et al.*, 1992): incorporación según su ubicación u orden en la matriz de datos, al azar, o de acuerdo con los valores de distancia con respecto al *outgroup* (Fig. 3). En algunos casos, a partir de distintas opciones de secuencia de adición de los taxa al árbol, se pueden obtener distintas topologías.

Dado que el algoritmo de Wagner no garantiza que se haya encontrado el árbol más corto o todos los árboles más cortos, se han propuesto otras alternativas de búsqueda, aunque muchas de ellas parten del citado algoritmo. Existen dos tipos de búsquedas principales, *exactas* y *heurísticas*.

#### Búsquedas exactas

Garantizan que se han hallado todos los árboles de longitud mínima, pero esto sólo es posible cuando se analizan matrices de menos de 20-25 taxa. Existen dos tipos de algoritmos exactos, los exhaustivos y el algoritmo de *branch and bound*= *bandb* (Hendy & Penny, 1982; Swofford & Olsen, 1990).

La búsqueda exhaustiva o enumeración implícita (Goloboff, 1998; Lipscomb, 1998) consiste en revisar todos los cladogramas totalmente resueltos posibles y elegir el más corto. Dado que el número de árboles posibles crece muy rápidamente, este algoritmo es solo aplicable para sets de datos con menos de 11 taxones (Forey *et al.*, 1992).

El algoritmo de *bandb* se usa para matrices que no superen los 20-25 taxones y permite llegar a un resultado exacto sin calcular todos los árboles posibles. Para ello construye primero un árbol mediante el algoritmo de Wagner y considera que su longitud será el límite superior (*upper bound*)

que no podrán exceder otros árboles hallados. El algoritmo de *bandb* va incorporando taxones secuencialmente y evaluando la longitud de los árboles parciales generados, pero abandona los caminos que resultan en una longitud mayor que la del árbol de Wagner de referencia asegurando de este modo la recuperación de todos los árboles más cortos (Swofford & Olsen, 1990).

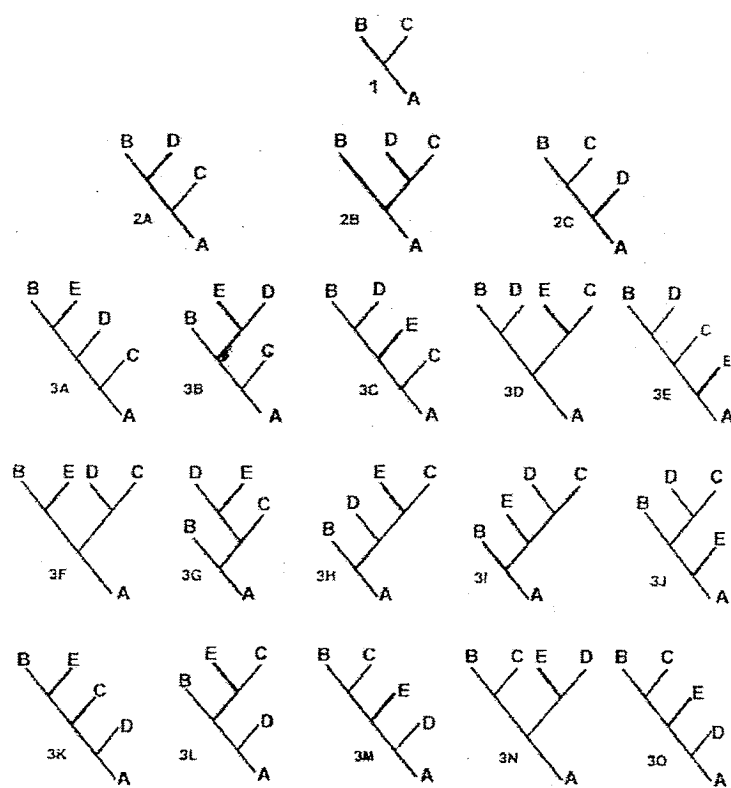


Figura 4. Ilustración de la estrategia de búsqueda *branch and bound* para la obtención del cladograma más parsimonioso entre cinco taxones (A-E). Dado un cladograma inicial de tres taxones (A,B,C), D podrá incorporarse en tres posiciones diferentes (2A, 2B, 2C). Dados tres cladogramas de cuatro taxones E podrá incorporarse en 15 posiciones diferentes (3A hasta 3O), pero el algoritmo abandonará los caminos que resulten de un árbol más largo que el de referencia (tomado de Swofford & Olsen, 1990).

### Búsquedas heurísticas

Se utilizan en el análisis de matrices de más de 20- 25 taxa y proporcionan soluciones aproximadas, que no garantizan el hallazgo de todos los árboles de mínima longitud. Los métodos heurísticos buscan árboles más cortos a base de prueba y error, usando como punto de partida uno o más árboles iniciales. El programa de cómputo aplica un procedimiento de permutación de ramas (*branch swapping*) que consiste en mover las ramas del árbol inicial a diferentes partes del árbol, contar los pasos y guardar los árboles de igual longitud o los más cortos. El movimiento de ramas se repite hasta evaluar todas las combinaciones posibles en los árboles guardados en memoria y se detiene cuando ya no hay árboles de igual longitud o más cortos, o cuando no alcanza la memoria disponible. La búsqueda de árboles requiere de una estrategia definida por el usuario.

Las formas más comunes de permutación de ramas (Fig. 5) son NNI (*Nearest Neighbour Interchange*), SPR (*Subtree Pruning and Regrafting*) y TBR (*Tree Bisection Reconnection*) siendo la última la más exhaustiva (Swofford & Olsen, 1990). En NNI los árboles se dividen en subárboles con una rama libre cada uno, que intercambian ramas con los vecinos; en SPR uno de los subárboles conserva una rama libre y el otro no, de modo que la rama libre se conecta en diferentes puntos al otro subárbol. En TBR ninguno de los subárboles tiene ramas libres, y cada uno se conecta a través de sus ramas internas, a las ramas internas del otro subárbol. Para una matriz de 40 taxones se realizan aproximadamente 5000 reacomodamientos SPR y 25.000 TBR.

Un problema que surge con el uso de los algoritmos de *branch swapping* es la formación de islas de árboles (= grupos de topologías de árboles), que son conjuntos de árboles igualmente parsimoniosos separados entre sí por un solo *swap* o una sola permutación de sus ramas (Swofford & Olsen, 1990). Las permutaciones de ramas permiten encontrar las distintas topologías presentes en una misma isla, pero por definición, no permitirán hallar las topologías presentes en otras islas. En consecuencia, si hubiera varias islas con topologías muy diferentes, lo

más recomendable es partir de más de un árbol de Wagner inicial, e.g. diez, obtenidos por secuencias de adición al azar, y aplicar a ellos permutación de ramas. Cuando la permutación de ramas se aplica a un sólo árbol de Wagner inicial, sólo podrán hallarse los árboles óptimos, si éstos pertenecen a la misma isla.

También surgen problemas para encontrar el o los árboles mas cortos cuando se analizan matrices de datos muy grandes (por ejemplo la matriz Zilla de 500 taxones de plantas). Para estos casos se han propuesto estrategias de búsqueda tradicionales más eficientes (Schuh, 2000) y nuevas estrategias como la denominada *Parsimony ratchet*, desarrollada por Nixon (1999 a) e implementada en el programa TNT (Goloboff *et al.*, 2003).

Las estrategias de búsqueda tradicionales eficientes se llevan a cabo de la siguiente manera:

- Se realizan varios análisis de TBR a partir de diferentes árboles de Wagner construidos por secuencias de adición al azar, reteniendo relativamente pocos árboles en cada búsqueda (a mayor número de terminales menor cantidad de árboles retenidos).
- Se retienen los árboles más cortos y se realizan TBR más exhaustivos, reteniendo en este paso un mayor número de árboles en la memoria.
- Se repiten los pasos anteriores varias veces.

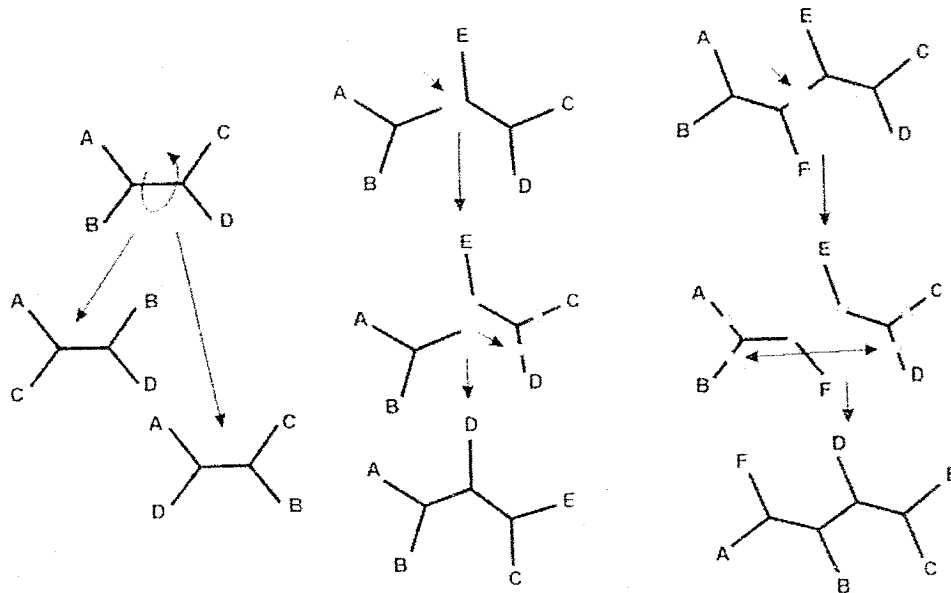


Figura 5. Esquemas de distintos procedimientos de permutaciones de ramas. A la izquierda NNI (*Nearest Neighbour Interchange*), en el centro SPR (*Subtree Pruning and Regrafting*), a la derecha TBR (*Tree Bisection Reconnection*). El último tipo de permutación es el más exhaustivo y uno de los más utilizados.

## OPTIMIZACIÓN DE CARACTERES

Optimizar un carácter en un árbol dado, significa asignar un estado definitivo a cada nodo interno del árbol, de modo que el número de transformaciones entre los estados de los caracteres sea mínimo (Schuh, 2000). Una vez optimizados los caracteres, es posible calcular su número de pasos y la longitud de todo el árbol. Para optimizar un carácter es preciso realizar un procedimiento de lectura desde los taxones terminales del cladograma hacia la raíz, asignando a los nodos internos los estados compartidos por los taxones hermanos. En el caso de registrarse ambigüedades, éstas se resuelven mediante una lectura hacia arriba, en que a cada nodo interno se le asigna el estado que reporte menor costo de transformación, de acuerdo con el modelo de parsimonia elegido. Si se optimiza con Wagner o Fitch, que aceptan tanto los paralelismos como las reversiones, hay dos opciones: ACCTRAN (*Accelerated transformation*) o FAST, que optimiza cerca de la raíz y favorece las reversiones, y DELTRAN (*Delay transformation*) o SLOW, donde los cambios se producen más lejos de la raíz, y se da prioridad a los paralelismos (Fig. 6).



La opción UNAMBIGUOUS de los programas WINCLADA, NONA y TNT, representa sólo los clados soportados bajo ambos tipos de optimización (FAST y SLOW) dado que es posible que aparezcan clados soportados por un solo tipo de optimización (Fig. 7), y si ésta cambia, dichos clados o ramas colapsan.

Reglas para la lectura hacia abajo:

1. Si los taxones hermanos comparten el mismo estado de carácter, o uno de los estados, el nodo tendrá dicho estado.



2. Si los taxones hermanos tienen estados de caracteres diferentes, el nodo tendrá ambos (el estado es ambiguo), y la ambigüedad se resolverá en la lectura hacia arriba según la opción elegida sea FAST (Fig. 6, izquierda) o SLOW (Fig. 6, derecha).

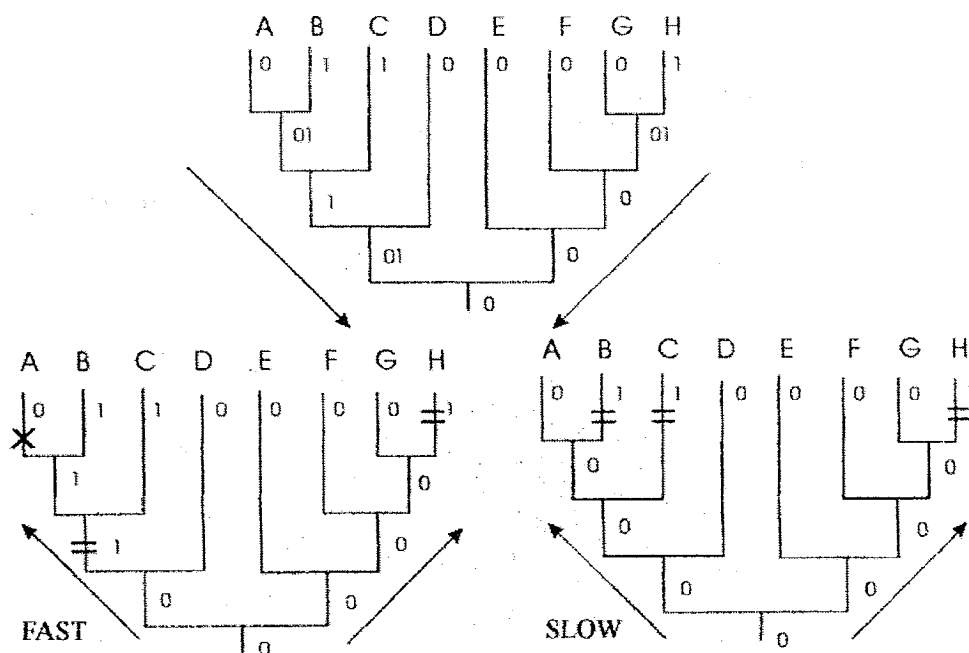
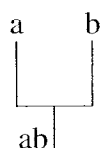


Figura 6. Optimización de un carácter doble estado, según parsimonia de Wagner (en este caso igual a Fitch). En el primer cladograma se muestra la lectura hacia abajo, lo que determina ambigüedad en la asignación de estados a algunos nodos. En los otros dos cladogramas se muestra la resolución de dichas ambigüedades mediante la lectura hacia arriba, y bajo las opciones FAST (prioriza reversiones) y SLOW (prioriza paralelismos).

	1	2	3	4	5	6	7
A	0	0	0	0	0	0	0
B	0	0	0	1	1	0	0
C	0	1	0	0	1	0	1
D	0	1	1	0	1	1	1
E	1	1	1	0	1	1	0

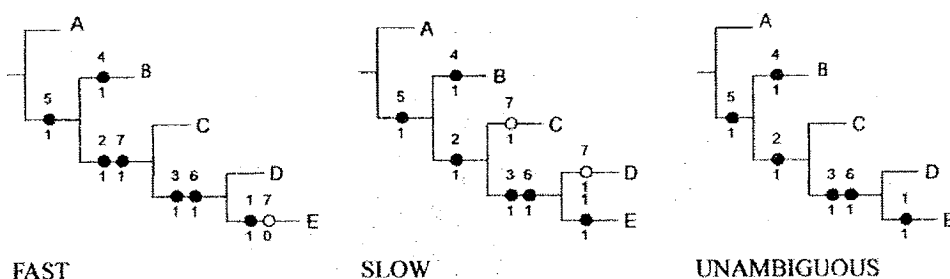


Figura 7. Matriz de datos de cinco taxones (A-E) y siete caracteres, y optimización de los caracteres según Parsimonia de Wagner (en este caso igual a Fitch), obtenida mediante el programa WINCLADA (Nixon, 1999 b). El carácter 7 cambia su distribución de acuerdo con la optimización elegida: Opción FAST; opción SLOW y opción UNAMBIGUOUS (en esta última no aparece dicho carácter por ser ambigua la optimización).

## PARÁMETROS DEL ÁRBOL

### Longitud del árbol

La longitud de un árbol es el parámetro más simple para evaluar el grado de homoplasia o ajuste de sus caracteres. Se calcula como la suma del número de transformaciones de todos los caracteres, multiplicado por el peso de dichos caracteres. El peso es generalmente 1, para todos los caracteres, al principio del análisis cladístico. Otra forma de medir el grado de ajuste de los caracteres a un árbol dado, es mediante la aplicación de los índices de consistencia (CI) y de retención (RI).

### Índice de consistencia (Kluge & Farris, 1969)

El valor de este índice varía entre 0 y 1. Se aproxima a 1 cuando no se producen cambios o transformaciones "extras" en los caracteres, y tiende a cero a medida que aumentan los paralelismos y las reversiones. Se calcula para cada carácter en forma individual y para todo el árbol:

Para cada carácter  
 $ci = m/s$

Para el cladograma  
 $CI = M/S$

**m:** cantidad de cambios (= pasos) mínimos posibles entre los estados de un carácter (= rango del carácter).

**s:** número de cambios registrados para cada carácter, en un árbol dado.

**M:** sumatoria de "m" de todos los caracteres (= sumatoria de los rangos).

**S:** sumatoria de "s", o de los cambios registrados en todos los caracteres.

Los valores de CI son influenciados por las autapomorfías de los caracteres y los caracteres invariantes, razón por la cual conviene calcularlo excluyendo los mismos.

### Índice de retención (Farris, 1989)

Fue propuesto con posterioridad al índice de consistencia, para representar la relación entre sinapomorfías aparentes y sinapomorfías reales (o su complemento, homoplasia relativa con respecto a homoplasia máxima posible). En ausencia de homoplasia es igual a 1. No está influido por las autapomorfías o los caracteres invariantes.

Para cada carácter  
 $ri = (g - s) / (g - m)$

Para el cladograma  
 $RI = (G - S) / (G - M)$

**g:** número máximo de pasos o transformaciones posibles de un carácter. Equivalente a una politomía total.

**G:** sumatoria de "g" de todos los caracteres.

#### Índice de consistencia re-escalado (Farris, 1989)

Resulta de multiplicar *ci* por *ri* (*rci*) o *CI* por *RI* (*RCI*). El *rci* se utiliza para asignar pesos diferenciales a los caracteres, de acuerdo con su grado de homoplasia, durante el procedimiento de pesado sucesivo.

**Tabla I.** Matriz de datos y cálculo de los índices de consistencia (*ci*) y retención (*ri*) para los caracteres 6 y 8, según su ajuste al cladograma de la figura 8. Dichos caracteres presentan el mismo valor de *ci* (0.5), pero difieren en el *ri* (0.5 y 0) (modificado de Wiley *et al.*, 1991).

	1	2	3	4	5	6	7	8
OG	0	0	0	0	0	0	0	0
A	1	0	0	0	0	1	0	1
B	1	1	1	0	1	0	1	0
C	1	0	1	1	1	0	0	0
D	1	1	1	0	1	1	1	0
E	1	1	1	1	1	1	1	1

Carácter 6

Carácter 8

$$ci = 1/2 = 0.5$$

$$ci = 1/2 = 0.5$$

$$ri = 3-2/3-1 = 0.5$$

$$ri = 2-2/2-1 = 0$$

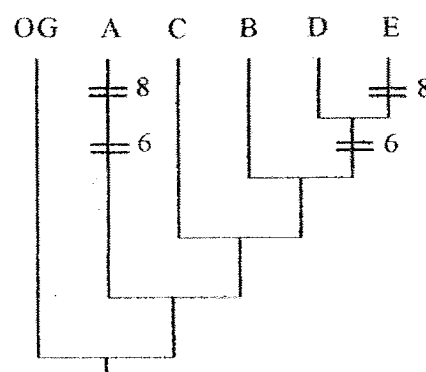


Figura 8. Cladograma donde se observa la distribución paralela de los caracteres 6 y 8. El carácter 6 registra menos homoplasia que el 8, pues una de las ocurrencias paralelas justifica la sinapomorfía aparente de los taxa D y E; el carácter 8, en cambio, no justifica ningún agrupamiento.

## PESADO DE CARACTERES

La asignación de pesos diferenciales a los caracteres, *a priori* del análisis cladístico, no resulta recomendable, debido a que es un procedimiento arbitrario. Por lo contrario, se acepta la asignación de pesos en función del comportamiento de los caracteres en el análisis cladístico, y de este modo se puede reducir el número de árboles igualmente parsimoniosos que en muchos casos se obtienen como resultado. Se han implementado distintos procedimientos de pesado, que se aplican durante o *a posteriori* del análisis. Uno de los más utilizados y que está implementado en el programa Hennig86 es el "pesado sucesivo de caracteres" (Farris, 1969; Carpenter, 1988, 1994), otro es el "Análisis de parsimonia bajo pesos implicados" (Goloboff, 1993) implementado en el programa Pee-Wee.

El pesado sucesivo opera del siguiente modo, una vez obtenidos los árboles más parsimoniosos a partir de una matriz donde todos los caracteres tienen igual peso, se les asigna mayor peso a aquellos caracteres que registran valores de índice de consistencia re-escalado más elevados, y se vuelven a calcular los cladogramas. Por ejemplo, el programa Hennig86 suele asignarles un peso de 10 a los caracteres con *rci* = 1 (no producen ninguna homoplasia); a aquellos caracteres con *rci* = 0 (máxima homoplasia) no les asigna peso; y a los caracteres con valores de homoplasia intermedios, les suele asignar pesos entre 3 y 5. El procedimiento de pesado sucesivo es iterativo, de forma tal que una vez pesados y recalculados los árboles de mínima longitud, el procedimiento se repite hasta que se estabiliza el número de árboles obtenidos y sus índices de consistencia y retención no varían. Un inconveniente posible del pesado sucesivo es que al cambiar los pesos iniciales, se puede llegar a una solución estable diferente (Goloboff, 1998).

## PROGRAMAS DE COMPUTACIÓN EN CLADÍSTICA

Los programas de computación utilizados en Cladística pueden servir para llevar a cabo las siguientes tareas de un análisis cladístico: edición de la matriz de datos, alineación de las secuencias nucleotídicas (ver Capítulo 10), obtención de cladogramas por parsimonia o *Maximum Likelihood* (ver Capítulo 10), y análisis y edición de cladogramas con alta calidad gráfica. A continuación se brinda información acerca de los programas de uso más frecuente.

### Programas que aplican algoritmos de simplicidad para la construcción de cladogramas

Existen varios programas para obtener árboles por parsimonia (Crisci *et al.*, 1994), algunos de los más utilizados son Hennig86 (Farris, 1988), NONA y Pee-Wee (Goloboff, 1996), PAUP (Swofford, 1999) y TNT (Goloboff, *et al.* 2003). Estos programas han sido desarrollados para computadoras IBM-PC compatibles, exceptuando el PAUP que fue desarrollado para Macintosh, pero tiene una versión beta para usar en sistema IBM. En general, el NONA y TNT son los que calculan árboles con mayor rapidez.

Hennig86 (Farris, 1988) es un programa de análisis muy rápido, desarrollado en MS-DOS, pero que no permite leer directamente datos de secuencias nucleotídicas (los estados de caracteres expresados como A, T, C y G deben ser codificados con valores numéricos, por ejemplo 0, 1, 2, 3), y no puede procesar matrices mayores de 128 taxa y 900 caracteres. En consecuencia es un programa útil para analizar matrices no muy grandes y con datos morfológicos. Para mayor información consultar el sitio [www.cladistics.org](http://www.cladistics.org).

NONA (Goloboff, 1996) es un programa desarrollado en MS-DOS, que tiene una variedad de opciones para encontrar el árbol más parsimonioso y para evaluar el soporte de sus ramas. Puede leer matrices con más de 128 taxa y 900 caracteres, con datos de secuencias nucleotídicas. Usa comandos similares al Hennig86 y es más fácil de utilizar en conjunto con WINCLADA. Una versión demo de este programa está disponible en el sitio [www.cladistics.com](http://www.cladistics.com).

Pee-Wee (Goloboff, 1996) es similar al NONA, pero se diferencia porque permite calcular árboles aplicando el procedimiento de pesos implicados. Una versión demo de este programa está disponible en el sitio [www.cladistics.com](http://www.cladistics.com).

PAUP (Swofford, 1999) además de realizar el análisis cladístico de los datos permite editar la matriz de datos. Contiene una vasta selección de opciones para el análisis tanto de datos morfológicos como de secuencias nucleotídicas, incluyendo la obtención de árboles a partir del cálculo de *Maximum Likelihood* (ver Capítulo 10). En general este programa demora más tiempo que el NONA en obtener los resultados. Información sobre este programa puede obtenerse en el sitio [www.sinauer.com](http://www.sinauer.com).

TNT (Goloboff *et al.*, 2003) a diferencia de los restantes programas permite obtener árboles, a partir de matrices de datos muy grandes (e.g. 300-500 taxa), en tiempos razonables, y también se aplica para el análisis de dichos árboles. Una versión demo está disponible en el sitio [www.cladistics.com](http://www.cladistics.com).

### Programas auxiliares para la edición de matrices y análisis de cladogramas

WINCLADA (Nixon, 1999 b) no realiza por sí solo todos los cálculos inherentes al análisis cladístico y es utilizado principalmente para editar y combinar matrices, visualizar y reordenar árboles, optimizar caracteres y mapearlos directamente sobre el árbol, e imprimir cladogramas con alta calidad gráfica. Trabaja bajo el entorno de Windows y permite leer matrices de datos con formato MS-DOS de Hennig86, NONA y PeeWee; además, provee una serie de menús que actúan como interfase para realizar el procesamiento de los datos o ciertos subprocesos mediante dichos programas. WINCLADA está disponible en el sitio [www.cladistics.com](http://www.cladistics.com).

MacClade (Maddison & Maddison, 2001) permite realizar tareas similares a WINCLADA pero con archivos obtenidos a partir de PAUP, en computadoras Macintosh. Información sobre este programa puede obtenerse en el sitio: [www.sinauer.com](http://www.sinauer.com).

# Ejercitaciones

## Ejercitaciones

### EJERCICIO 1

En la figura 9 se ilustra un árbol no enraizado de cuatro taxones (A-D) y la matriz de datos correspondiente. Dicha matriz incluye cuatro caracteres doble-estado, cuya polaridad no ha sido establecida *a priori*, de modo que los estados "a" y "b" representarán los códigos "0" y "1" o viceversa, dependiendo de la posición de la raíz.

1. Construya todos los árboles con raíz posibles, utilizando para ello cada una de las ramas terminales del *network* de la figura 9.
2. ¿Algunos de esos árboles presenta ramas no soportadas por caracteres? En este caso ¿cómo representaría las relaciones entre dichos taxa?
3. Entre todas las opciones de enraizamiento posibles ¿cuál le parece mejor, de acuerdo con la evidencia que brindan los caracteres?

	1	2	3	4
A	a	a	a	b
B	a	b	a	a
C	a	b	b	a
D	b	b	b	a

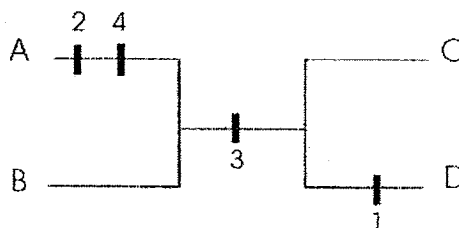


Figura 9. Matriz de datos para cuatro taxones (A-D) y el árbol no enraizado correspondiente.

### EJERCICIO 2

Sobre la base de los cladogramas obtenidos en el capítulo 8 para las especies hipotéticas de Brooks *et al.* (1984), y para los géneros de Saurocephalidae, realice los siguientes procedimientos:

*Sin computadoras*

1. Calcule la longitud y el índice de consistencia de los árboles.
2. Calcule los índices de consistencia y retención para algunos de sus caracteres.

*Con computadoras*

3. Grabe las matrices de datos construidas para las especies hipotéticas de Brooks *et al.* y para los géneros de Saurocephalidae, con el formato adecuado para ser analizadas mediante un algoritmo de parsimonia computarizado.
4. Obtenga el cladograma de mayor simplicidad y observe la distribución de los caracteres sobre el árbol. Distinga sinapomorfias, autapomorfias y homoplasias.

### EJERCICIO 3

El género *Galapaganus* (Coleoptera: Curculionidae), distribuido en las islas Galápagos y en las costas de Ecuador y Perú, incluye 15 especies, reunidas en dos grupos prin-

cipales, el grupo *femoratus* con las especies *G. femoratus* y *G. howdenae*, y el grupo *darwini*, con las restantes especies (Lanteri 1992). En la figura 10 se representa uno de los cladogramas óptimos obtenidos por Lanteri (1992) sobre la base de una matriz de datos morfológicos. Los caracteres multiestado 1, 14, 17 y 30 fueron analizados como no ordenados y los restantes como ordenados. Sobre la base de este resultado responda las siguientes preguntas:

*Sin computadoras*

1. ¿Qué grupo de especies se halla mejor soportado por sinapomorfías? ¿el grupo *femoratus* o el grupo *darwini*?
2. ¿Qué especies se hallan mejor justificadas por la evidencia de los caracteres empleados?
3. ¿Cuál es el número mínimo de colonizaciones independientes del archipiélago de Galápagos, a partir del continente, que podría postularse sobre la base del cladograma ilustrado?

*Con computadoras*

4. Analice la matriz de datos de Lanteri (1992) mediante distintas estrategias de búsqueda (exactas y heurísticas) y cambiando las opciones de "ordenados" y "no ordenados" de los caracteres multiestado. ¿Cuántos árboles óptimos obtuvo en cada caso? ¿Cuáles son los índices de consistencia y retención obtenidos para los mismos?
5. En los casos que haya obtenido más de un árbol óptimo, aplique el procedimiento de pesado sucesivo. ¿Se reduce el número de árboles? ¿Qué cambio observa en los valores de los índices de consistencia (CI) y de retención (RI) de dichos árboles? Compare los valores de estos índices, al calcularlos con y sin autapomorfías incluidas.
6. Analice los valores de los índices de consistencia (ci) y retención (ri) de los caracteres, antes y después de aplicado el procedimiento de pesado sucesivo.

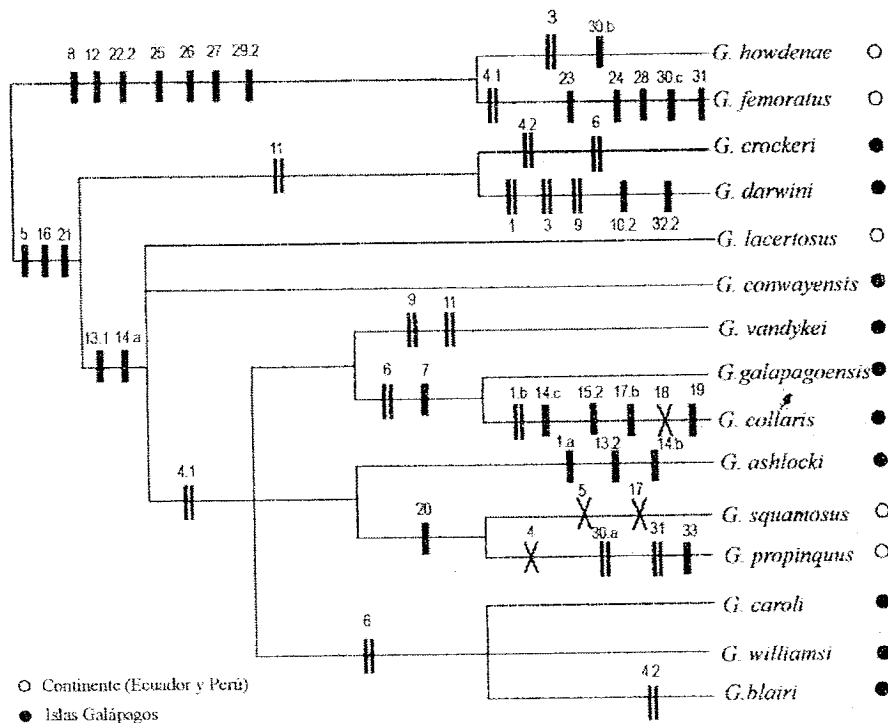


Figura 10. Cladograma de las especies de gorgojos del género *Galapaganus* (Coleoptera: Curculionidae) según Lanteri (1992). Aquellas especies indicadas con un círculo blanco se distribuyen en el continente (Perú y Ecuador) y las señaladas con un círculo negro, en el archipiélago de Galápagos.

## EJERCICIO 4

El género de gorgojos *Erycideus* (Coleoptera: Curculionidae) se halla distribuido en el continente americano, desde la Argentina hasta Arizona en los Estados Unidos de América. Según Lanteri (1995) dicho género incluye 16 especies: *Erycideus bahiensis*, *E. hancocki*, *E. schoenherri*, *E. nigropunctatus*, *E. sedecimpunctatus*, y *E. argentinensis* se distribuyen en Sudamérica; *E. yucatanus* (endémica de la provincia de Yucatán), *E. roseiventris*, *E. quadripunctatus*, *E. cupreolus*, *E. modestus*, *E. viridans*, *E. duodecimpunctatus* y *E. forreri*, se distribuyen en el sur de México y América Central; y *E. lautus* y *E. placidus* se hallan restringidas al noroeste de México y suroeste de USA (Arizona). El cladograma ilustrado en la figura 11 se obtuvo sobre la base de 35 caracteres morfológicos, considerando a los multiestado 14, 16 y 25, como no aditivos y a los restantes como aditivos. Sobre la base de estos datos responda las siguientes preguntas:

### Sin computadoras

1. ¿Las especies sudamericanas de *Erycideus* se separan en el cladograma óptimo de la figura 11, de las restantes distribuidas en México, América Central y USA? ¿Observa una dirección del cambio evolutivo con sentido geográfico norte-sur o sur-norte?
2. ¿Cuáles son los grupos monofiléticos del cladograma mejor soportados por sinapomorfias?

### Con computadoras

3. Analice la matriz de datos morfológicos del trabajo de Lanteri (1995) aplicando distintas estrategias de búsqueda (exactas y heurísticas) y considerando a los caracteres multiestado como ordenados y no ordenados. ¿Cuántos árboles obtuvo en cada caso? ¿Cuáles son los valores de los índices de consistencia (CI) y retención (RI)?
4. El cladograma de *Erycideus* fue enraizado con el género *Lamprocyphus* como *outgroup*. Cambie la raíz del árbol, colocando alguna de las especies del *ingroup* en la base. ¿Qué cambios observa en la topología del cladograma? ¿Cambian los valores de CI y de RI con respecto al árbol original?

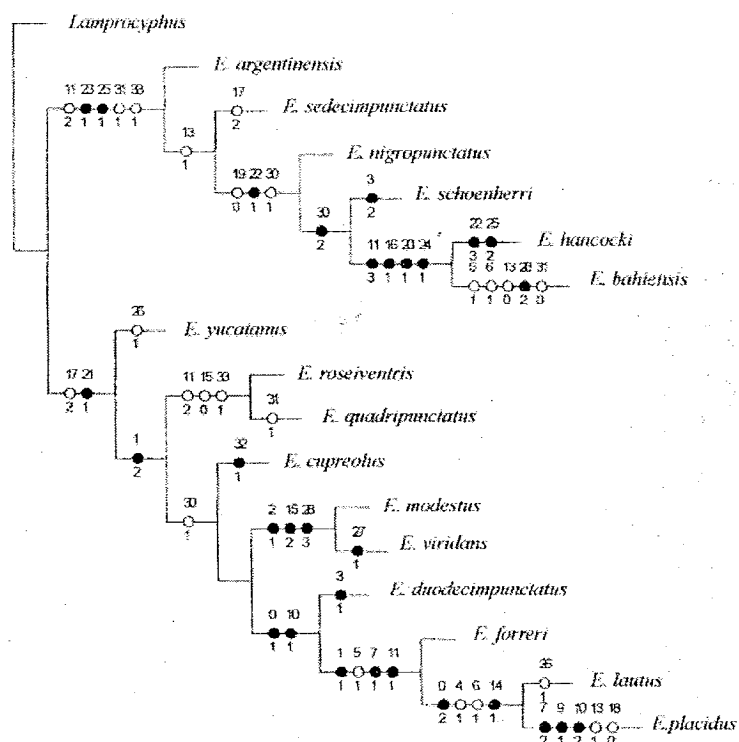


Figura 11. Cladograma más parsimonioso correspondiente a las especies de *Erycideus* (Coleoptera: Curculionidae) según Lanteri (1995). Puntos negros: apomorfias; puntos blancos: homoplasias. Los caracteres están numerados desde 0.

## EJERCICIO 5

La familia Tristiridae (Orthoptera) comprende 18 géneros endémicos de América del Sur y su monofilia está definida por características del complejo fálico. El análisis filogenético llevado a cabo por Cigliano (1989) dio como resultado el cladograma que se ilustra en la figura 12. Los 29 caracteres morfológicos utilizados corresponden a la morfología externa (caracteres 0 a 10), genitalia de la hembra (carácter 11), y genitalia de los machos (caracteres 12 a 28). Sobre la base de esta información responda las siguientes preguntas:

### Sin computadoras

1. ¿Qué caracteres justifican o brindan mayor soporte a los principales grupos del cladograma? ¿los de la morfología externa o los de la genitalia?
2. ¿Considera Usted que la diferenciación de los caracteres genitales de los machos ha precedido a la diferenciación de otros caracteres? Justifique su respuesta.

### Con computadoras

3. Sobre la base de la matriz de datos publicada por Cigliano (1989) obtenga un cladograma óptimo, analizando los caracteres 14, 16 y 25 como no ordenados.
4. Mediante el uso del programa WINCLADA visualice la transformación de los caracteres, empleando distintas opciones de optimización: FAST, SLOW y UNAMBIGUOUS. ¿Qué caracteres registran cambios en su evolución al emplear las distintas opciones de optimización?

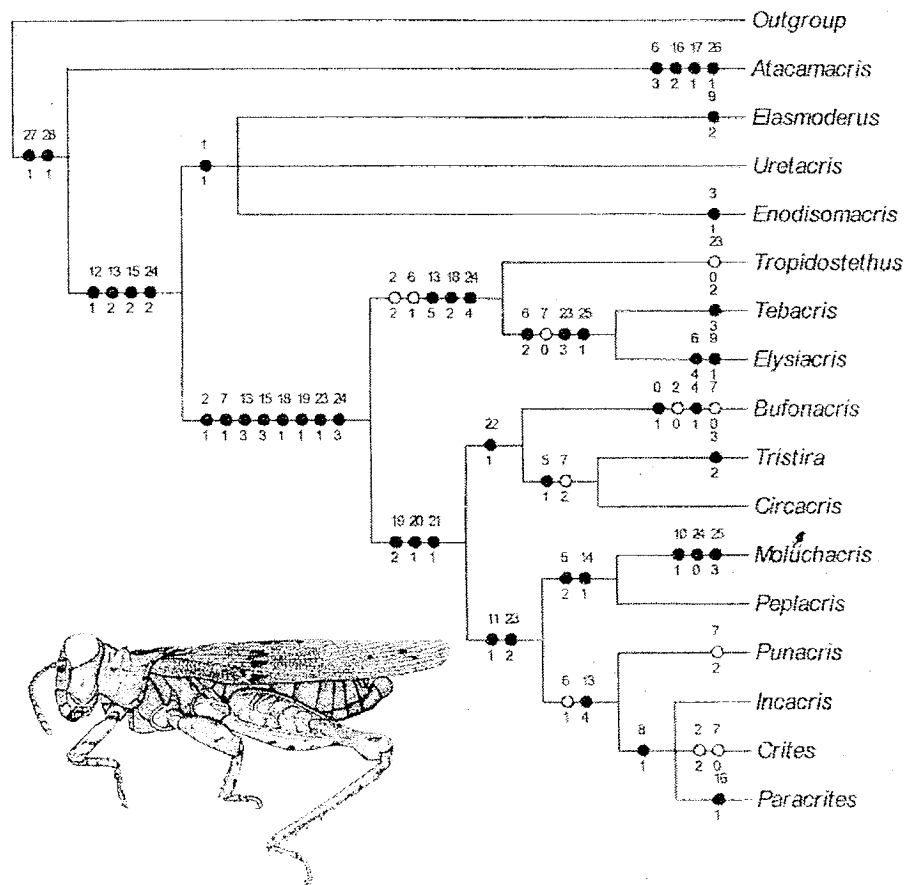


Figura 12. Cladograma de la familia Tristiridae (Orthoptera) según Cigliano (1989). Círculos negros: apomorfias; círculos blancos: homoplasias. Los caracteres están numerados desde 0.



## EJERCICIO 6

En la figura 13 se ilustra un cladograma correspondiente a 12 géneros de lagartos marinos mesozoicos, en el cual se indican los estados de un carácter doble-estado.

1. Optimice el carácter, mediante Parsimonia de Wagner, utilizando las opciones ACCTRAN y DELTRAN.

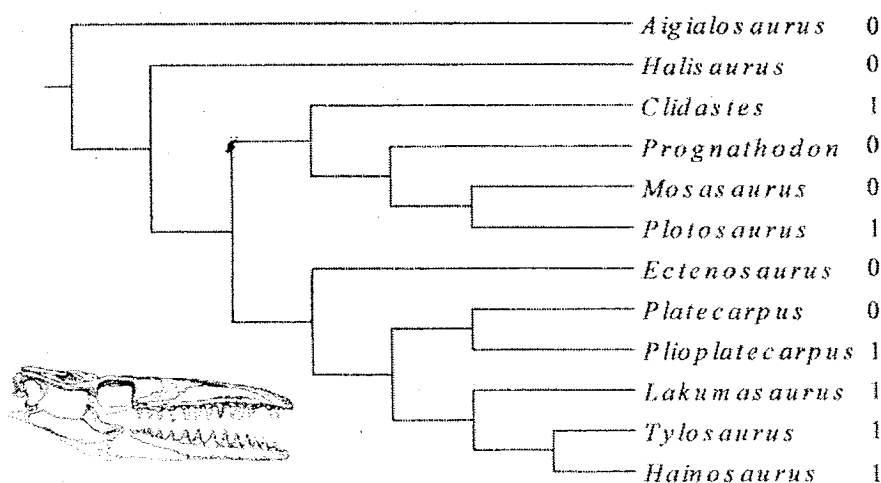


Figura 13. Cladograma correspondiente a 12 géneros de lagartos marinos mesozoicos (mosasaurios). Los códigos 1-0 son los estados de un carácter presente en los taxones terminales (modificado de Novas et al., 2002).

## BIBLIOGRAFÍA CITADA

- BROOKS, R. F., J. N. CAIRO, T. R. PLATT & H. H. PRITCHARD. 1984. *Principles and methods of Phylogenetic Systematics. A cladistic workbook*. Univ. Kansas, Museum of Natural History, Special Publication Nº 12.
- CARPENTER, J. M. 1988. Choosing among multiples equally parsimonious cladograms. *Cladistics* 4: 291-296.
- CARPENTER, J. M. 1994. Successive weighting, reliability and evidence. *Cladistics* 10: 215-220.
- CIGLIANO, M. M. 1989. A cladistic analysis of the Tristiridae (Orthoptera, Acrididae). *Cladistics* 5: 379-393.
- CRISCI, J. V., A. A. LANTERI & E. ORTIZ JAUREGUIZAR. 1994. Programas de computación en Sistemática y Biogeografía Histórica: revisión crítica y criterios para su selección. Pp. 207-228. En: Llorente Bousquets, J. & I. Luna Vega (comp.). *Taxonomía Biológica*, UNAM, Fondo de Cultura Económica, México.
- FARRIS, J. S. 1969. A successive approximations approach to character weighting. *Systematic Zoology* 18:374-385.
- FARRIS, J. S. 1970. Methods for computing Wagner trees. *Systematic Zoology* 19: 83-92.
- FARRIS, J. S. 1980. The efficient diagnoses of the phylogenetic system. *Systematic Zoology* 29: 386-401.
- FARRIS, J. S. 1983. The logical basis of phylogenetic analysis. Pp. 7-36. En: Platnick N. I. & V. Funk (eds.). *Advances in Cladistics*. New York Botanical Garden, New York.
- FARRIS, J. S. 1986. Synapomorphy, parsimony, and evidence. *Taxon* 35: 298-315.
- FARRIS, J. S. 1988. Hennig86 reference. Versión 5.1. Published by the author.
- FARRIS, J. S. 1989. The retention index and the rescaled consistency index. *Cladistics* 5: 417-419.
- FARRIS, J. S., A. G. KLUGE & M. F. MICKEVICH. 1982. Phylogenetic analysis, the monothetic group

- method, and myobatrachid frogs. *Systematic Zoology* 31: 317-327.
- FITCH, N. M. 1971. Towards defining the course of evolution: minimal change for a specific tree topology. *Systematic Zoology* 20: 406-416.
- FOREY, P.L., C. L. HUMPHRIES, I. J. KITCHING, R.W. SCOTLAND, D. J. SIEBERT & D.M. WILLIAMS. 1992. *Cladistics. A practical course in systematics*. Clarendon Press, Oxford.
- GOLOBOFF, P. 1993. Methods for faster parsimony analysis. *Cladistics* 12: 199-220.
- GOLOBOFF, P. 1996. Pec-Wee, NONA, SPA and Phast programs and documentation. Willi Hennig Society Web-Site: <http://www.vims.edu/~mes/hennig/hennig.html>.
- GOLOBOFF, P. 1998. *Principios básicos de la cladística*. Sociedad Argentina de Botánica, Buenos Aires.
- GOLOBOFF, P.; K. NIXON & J. FARRIS. 2003. (TNT), Tree analysis using New Technology. Published by the authors, Tucumán, Argentina.
- HENDY, M. D & D. PENNY. 1982. Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Bioscience* 59: 271-290.
- HENNIG, W. 1968. *Elementos de una sistemática filogenética*. EUDEBA, Buenos Aires.
- KITCHING, I.J., P.L. FOREY, C.J. HUMPHRIES & D.M. WILLIAMS. 1998. *Cladistics. The theory and practice of parsimony analysis*. Second Edition. The Systematics Association Publication No. 11. Oxford University Press Inc., New York.
- KLUGE, A.G. & J.S. FARRIS 1969. Quantitative phyletics and the evolution of anurans. *Systematic Zoology* 18:1-32.
- LANTERI, A. A. 1992. Systematics, cladistics and biogeography of a new weevil genus, *Galapaganus* (Coleoptera: Curculionidae) from Galápagos islands, and coasts of Ecuador and Perú. *Transactions of the American Entomological Society* 118:227-267.
- LANTERI, A. A. 1995. Systematic revision of *Eriocydeus* Pascoe (Coleoptera: Curculionidae). *Entomologica scandinavica* 26(4): 393-424.
- LIPSCOMB, D. 1994. *Cladistic analysis using Hennig86*. George Washington University, Washington D. C.
- LIPSCOMB, D. 1998. *Basics of cladistic analysis*. George Washington University, Washington D. C.
- MADDISON, W.P. & D.R. MADDISON. 2001. *MacClade*. Interactive analysis of phylogeny and character evolution. Sinauer Associates, Inc., Sunderland.
- NIETO FELINER, G. 1999. Tres décadas de Cladismo. *Boletín de la Sociedad Entomológica Aragonesa* 26:85-93.
- NIXON, K.C. 1999 a. The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics* 15(4): 407-414.
- NIXON, K. C. 1999 b. WINCLADA, Ithaca, Nueva York, publicado por el autor.
- NIXON, K.C. & J. M. CARPENTER. 1993. On outgroups. *Cladistics* 9: 413-426.
- NIXON, K.C. & J. I. DAVIS. 1991. Polymorphic taxa, missing values and cladistic analysis. *Cladistics* 7: 233-241.
- NOVAS, F., M. FERNÁNDEZ, Z. CASPARINI, J.M. LIRIO, H. NÚÑEZ & P. PUERTA. 2002. *Lakumasaurus antarcticus*, n.gen. et sp., a new mosasaurus (Reptilia, Squamata) from the Upper Cretaceous of Antarctica. *Ameghiniana* 39 (2):245-249.
- PLATNICK, N. I., C. E. GRISWOLD & J. A. CODDINGTON. 1991. On missing entries in cladistic analysis. *Cladistics* 7: 337-343.
- SCHUH, R. T. 2000. *Biological Systematics. Principles and applications*. Cornell Univ. Press, Ithaca.
- SWOFFORD, D. L. 1999. PAUP. Phylogenetic Analysis Using Parsimony. Version 4.0. Illinois, Saunders.
- SWOFFORD, D. L. & G. J. OLSEN. 1990. Phylogeny reconstruction. Pp. 411-501. En: Hillis, D.M. & C. Moritz (eds.). *Molecular Systematics*. Sinauer, Sunderland.
- WAGNER, W.H. 1961. Problems in classification of ferns. *Recent Advances in Botany* 1:841-844.
- WILEY, E.O, D. SIEGEL-CAUSEY, D.R. BROOKS & V.A. FUNK. 1991. *The complete cladist. A primer of phylogenetic procedures*. The University of Kansas, Museum of Natural History, Special Publication 19.
- YEATES, D. K. 1995. Groundplans and exemplars: paths to the tree of life. *Cladistics* 11:343-357.