

Evolución y Filogenia de Arthropoda

Sección I: Conceptos y métodos en el estudio de la filogenia

TRES DÉCADAS DE CLADISMO

Gonzalo Nieto Feliner

Real Jardín Botánico (CSIC),
Plaza. de Murillo, 2, 28014 Madrid

Resumen

A pesar de ser el método más ampliamente usado en reconstrucción filogenética, la Cladística no dispone de una síntesis conceptual que sea aceptada por las distintas corrientes que la componen. A partir del origen de esta escuela de Sistemática, el autor destaca algunos de los hitos que han marcado su desarrollo y diversificación, y comenta las desviaciones más significativas respecto a las ideas y métodos que en cada momento se consideraban mayoritarios dentro del Cladismo. Se señalan, además, algunos de los mitos atribuidos frecuentemente a la Cladística así como las controversias más debatidas en la actualidad.

Palabras clave: Cladística, Hennig, Sistemática.

Three decades of Cladistics

Abstract

Despite being the most widely used method in phylogeny reconstruction, Cladistics lacks a generally agreed conceptual synthesis. Starting from its foundation, the author gives an overview of the development and diversification of this school of Systematics, and comments the major departures from the mainstream ideas and practice. Some of the most pervasive myths attributed to Cladistics, as well as the issues that are most hotly debated in the last years, are also discussed.

Key words: Cladistics, Hennig, Systematics.

INTRODUCCIÓN

Pasar revista a lo que ha representado el Cladismo en la Biología me parece una tarea demasiado ambiciosa y ardua para poderse plasmar con un mínimo de coherencia y detalle en un solo artículo. Por ello, me voy a centrar en cuatro aspectos concretos que —inevitablemente, en un tema tan controvertido— enfocaré desde mi perspectiva personal, sin duda sesgada para otros autores. Comenzaré refiriéndome a la figura de Hennig (Dupuis 1984) y lo que han supuesto sus ideas, sobre todo aquellas que han resistido el paso del tiempo y la evolución y diversificación de una corriente inicialmente muy compacta (Hull 1988). Precisamente, el segundo aspecto que trato es éste: la “radiación” ideológica y metodológica del Cladismo o, al menos, alguna de sus manifestaciones más destacadas. De forma breve me refiero, a continuación, a algunas propiedades atribuidas a la Cladística que, cuando menos, son inexactas. Finalmente, comento alguna de las controversias actuales que considero más significativas dentro de este movimiento, así como alguno de los retos futuros.

WILLI HENNIG Y EL ORIGEN DEL CLADISMO

Cuando el entomólogo alemán Willi Hennig (1913-1976) publicó su *Grundzüge einer Theorie der Phylogenetischen Systematik* en 1950 faltaban casi 20 años para que sus ideas cristalizaran en la metodología e ideas, inicialmente bastante homogéneas, que hoy conocemos como Cladismo (Dupuis 1984). Por eso prefiero hablar de tres y no de cuatro décadas. Nadie, dentro ni fuera del Cladismo, duda que Hennig es el

padre de esta corriente. Buena prueba es el nombre que lleva la Sociedad que aglutina a sus adeptos [<http://www.nhm.ac.uk/hennig/>] y publica la revista *Cladistics*. También es ilustrativo el hecho de que cuando, en los años ochenta se discutían los méritos de dos corrientes dentro del Cladismo (los defensores de la parsimonia y los de la compatibilidad de caracteres, CCA), unos y otros pretendían interpretar mejor o ser herederos de las ideas de Hennig. Es posible, también, que en este consenso sobre la paternidad de la escuela haya jugado algún papel el que, a diferencia de muchos científicos cladistas y no cladistas, la suya no era una personalidad controvertida o difícil. Todo eso es cierto y, sin embargo, sólo con las ideas de Hennig la Cladística no podría estar donde está ahora: en la inmensa mayoría de los trabajos que se publican con algún ensayo de reconstrucción filogenética.

Por eso, tiene cierto interés el tratar de aclarar cuáles fueron sus contribuciones más influyentes y cuáles las que todavía perduran en la práctica de los filogeneticistas. Yo señalaría dos o tres ideas. La primera es la defensa de los grupos *monofiléticos*. Hennig (1968) no inventó el concepto de monofilia pero sí dejó clara la necesidad de que, para que su método tuviera coherencia, los grupos debían incluir ramas completas de un árbol filogenético. Esto se entiende cuando se representa con gráficos de teoría de conjuntos. En un diagrama de Venn (Fig. 1), para que las relaciones de parentesco representadas sean equivalentes a las que Hennig eligió, la intersección entre grupos debe ser cero o bien cualquier grupo debe poderse incluir totalmente en otro (Wiley et al., 1991).

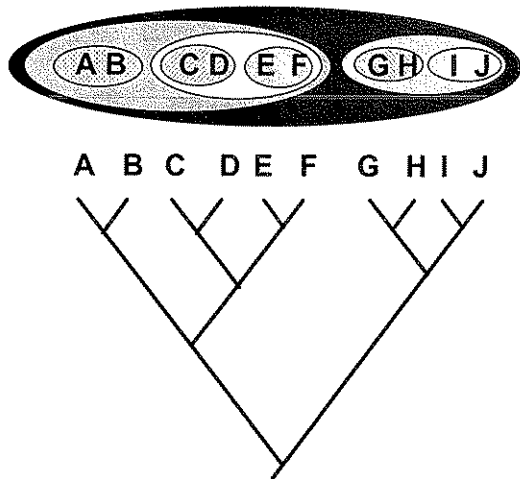


Fig. 1.- Divergencia de linajes y conjuntos encajados de sinapomorfias: equivalencia entre un cladograma y el diagrama de Venn correspondiente.

En realidad, la condición de monofilia, aunque más conocida, es una consecuencia lógica de otro de los avances conceptuales del sistema de Hennig: el concepto de grupo hermano. La evolución, al ser el origen de la jerarquía que observamos en la naturaleza, debe de alguna manera reflejarse en la clasificación que los taxónomos intentan hacer de los seres vivos. Este convencimiento parte ni más ni menos que de Darwin, quien descubre el vínculo escondido que los naturalistas habían venido buscando inconscientemente: las especies actuales se asemejan unas a otras porque descienden de un antepasado común. Sin embargo, sobre el modo de representar, de alguna manera, la filogenia en la clasificación, Darwin no nos da ninguna solución. Ello refleja la dificultad intrínseca de representar al mismo tiempo relaciones del tipo de antepasado-descendiente y patrones de ramificación (cladogénesis) en un sistema jerárquico. No han faltado autores posteriores que han insistido en la misma idea, entre otros Mayr (1942).

Lo que Hennig propone para solucionar este problema es ignorar cualquier aspecto de la filogenia relacionado con las relaciones antepasado-descendiente. En cambio, propone atender solamente a las relaciones de grupo hermano. Dos especies (o dos grupos, o una especie y un grupo) son hermanos si comparten un antepasado exclusivo de ambas, es decir, un antepasado más reciente del que cualquiera de ellos comparte con cualquier otro taxon. Este tipo de relación es menos precisa que una del tipo *la especie A es descendiente de la especie B*. En efecto, si B es el antepasado directo de A, entonces A y B también son grupos hermanos ya que comparten un antepasado común exclusivo, que es B. Por eso, la hipótesis de que B es el antepasado directo de A está incluida en la hipótesis de que A y B son hermanas. Por ser menos precisas las relaciones de grupo hermano, su representación —los cladogramas— es compatible con varios filogramas, que representen relaciones antepasado-descendiente (Forey et al., 1992). Dados tres taxones A B C, un cladograma que indique que A es hermano de B —expresado en notación parentética ((AB)C)— es compatible con 6 filogramas. ¿Qué razón hay para preferir una hipótesis más vaga? Son razones epistemológicas, ya que aun siendo la hipótesis de que A y B son hermanos más vaga que la otra, es la más explícita susceptible de ser refutada con las evidencias que los sistemáticos

manejan: distribuciones de caracteres. En cambio, una hipótesis que implique relaciones antepasado-descendiente del tipo de las representadas en los filogramas no puede refutarse sobre la base de distribuciones de caracteres (Engelmann & Wiley, 1977).

En el sistema lógico de Hennig, al ir formando grupos hermanos de forma aglomerativa, vamos construyendo grupos más inclusivos monofiléticos. Pero, ¿cómo sabemos que dos grupos o dos especies son hermanos? Hennig definió un concepto que, una vez más, ha trascendido ampliamente a sus seguidores pare ser aceptado incluso por los contrarios al cladismo: *sinapomorfia*. El entomólogo alemán señaló que, de entre las similitudes que un sistemático maneja, no todas son competentes para indicar o sugerir qué dos grupos son hermanos. De hecho, habla de una similitud primitiva y una similitud derivada. De los atributos que dos grupos comparten, solamente aquellos que son derivados, respecto a los que presentan otros taxones relacionados, son indicio de que efectivamente aquellos grupos son hermanos. Solamente este tipo de caracteres son sinapomórficos. Esto ha dado lugar a uno de los mitos más persistentes y del que luego hablaré: la necesidad de polarización previa.

Si bien el concepto de sinapomorfia fue un éxito y está plenamente vigente, el método por el cual Hennig trataba de estimar cuáles caracteres eran, en realidad, sinapomórficos y cuáles no es quizá el aspecto más débil de todo su sistema. Utilizaba una serie de criterios para decidir si un carácter era el derivado frente a sus alternativos (Wiley, 1981), criterios que han caído en desuso en su práctica totalidad. De ello, en mi opinión, lo más positivo fue lo que él llamaba *reciprocal illumination* aludiendo a un proceso iterativo de confrontación de unos caracteres con otros —cada uno con sus implicaciones— que conducía a identificar cuáles eran compatibles con otros, cuáles implicaban necesariamente una hipótesis distinta y, sobre todo, conducía a un re-análisis de alguno de ellos.

Es más, en mi opinión, una de las razones por las que el Cladismo no ha evolucionado como una corriente homogénea es la debilidad del método de Hennig en este preciso punto. Cuando las evidencias que usamos para hacer reconstrucción filogenética —los caracteres— son compatibles (encajan en una única hipótesis de parentesco) no hay ningún problema. Pero esto es, en la práctica, muy raro. El modo en el que decidimos qué hipótesis es preferible es un logro que se debe, fundamentalmente, a Farris (1983). Su mérito viene de la simplicidad, que es, en realidad, lo que propugna. El criterio que nos permite seleccionar entre hipótesis alternativas, que compiten para explicar unos datos, es el de la *parsimonia*, y ello no es más que una extensión de la famosa cuchilla de Ockham. Las hipótesis preferidas son aquellas que explican los datos disponibles de la forma más simple, esto es, implicando el menor número de apariciones independientes. Si varias especies poseen una novedad evolutiva —un carácter que creemos exclusivo de ellas e igual en todas—, es más simple suponer que se ha originado una sola vez, en su antepasado común más reciente, y no que se ha producido independientemente varias veces. El criterio metodológico de parsimonia selecciona aquella(s) hipótesis que explican los datos que observamos con el número mínimo de paralelismos, convergencias o reversiones. Este tipo de cambios —*homoplásicos*— resultantes, a posteriori, del análisis son en realidad hipótesis *ad hoc*, ya que inicialmente los consideramos homólogos y no tenemos evidencia independiente de que realmente no lo sean. Por ser *ad hoc*, es deseable minimizarlas, que es lo que hace el criterio de parsimonia (Farris, 1983; Kluge, 1984). Este criterio tan elegante y simple es, de hecho,

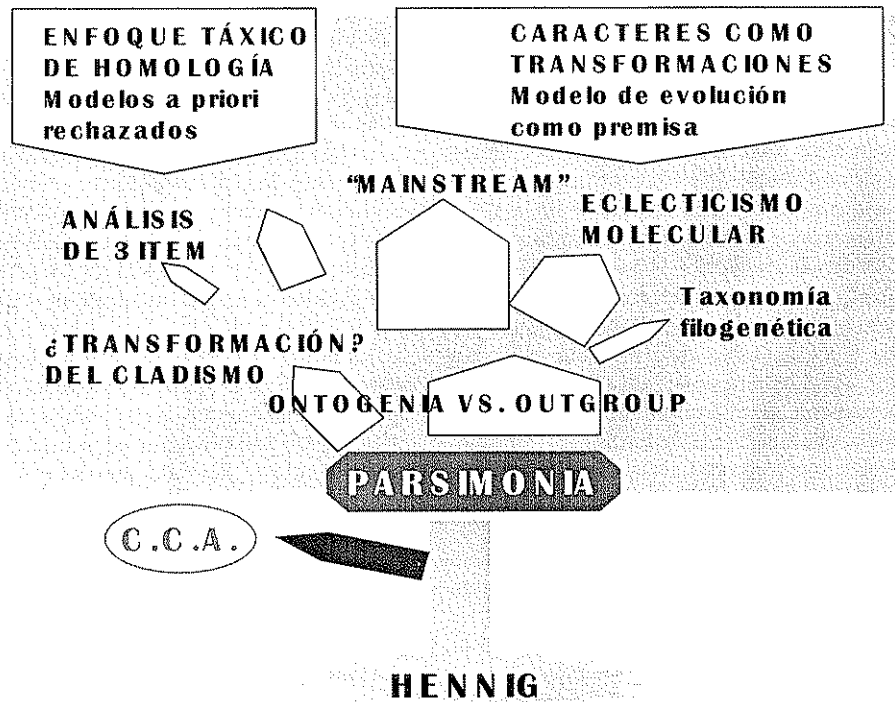


Fig. 2.- Esquema de las corrientes y controversias dentro del Cladismo en una secuencia más o menos cronológica. La dicotomía enfoque *táxico* frente a *transformacional* describe dos tendencias: la primera se basa sólo en distribuciones de caracteres para formular hipótesis sobre grupos mientras que otras propiedades, de grupos y caracteres, se “leen” a posteriori en el cladograma resultante. La segunda pone el énfasis en caracteres como atributos que se transforman y cuya hipótesis de evolución debe formularse a priori.

el responsable de que las ideas de Hennig se hayan implementado por parte de tantos investigadores. Por eso, cuando se comparan los distintos métodos de reconstrucción filogenética (Swofford et al., 1996), al cladístico suele denominársele como *de parsimonia*. También se suele hablar de forma impropia, principalmente en la jerga de la sistemática molecular, de *máxima parsimonia*. En efecto, no tiene sentido adjetivar la parsimonia ya que, sin necesidad de hacerlo, queda perfectamente claro cuál es el criterio seguido.

RADIACIÓN DEL CLADISMO

Quienquiera que haya tratado de enseñar lo que es la Cladística, y más si lo ha tratado de hacer con un cierto detalle —teórico y práctico— comprenderá que no es fácil hacer una síntesis conceptual cerrada, o al menos compacta, de las ideas de esta corriente, tan significativa en la historia de la Sistemática de la segunda mitad del siglo XX. Podría pensarse que casi cuatro décadas —si contamos a partir de la publicación del *Grundzuge* de Hennig— son tiempo suficiente para que una corriente científica haya madurado y, sobre todo, logrado una cierta síntesis conceptual de su cuerpo de doctrina. Téngase en cuenta, por ejemplo, que la llamada “Nueva Síntesis” o “Teoría Sintética de la Evolución”, surgida en los años treinta, se hallaba totalmente asentada en la década del setenta aun cuando hubiera algunas voces discordantes fuera de la misma.

No es éste el caso de la Cladística cuyo desarrollo, desde el descubrimiento de las ideas de Hennig (1966) ha sido marcadamente divergente. Es cierto que el impulso inicial correspondió a un grupo pequeño, compacto y muy activo de investigadores, principalmente del Museo de Historia Natural

de Nueva York (Hull, 1988). Esto sucedía a finales de los sesenta y comienzos de los setenta. Pero a finales de los setenta ya empezaban a surgir las primeras discrepancias y en el año 1986 había quien, en tono festivo, clasificaba a los cladistas sobre la base de sus ideas y enfoques en torno a la cladística (Carpenter, 1987).

Durante bastante tiempo el criterio de parsimonia no ha sido origen de gran controversia si exceptuamos la mantenida con los partidarios del análisis de compatibilidad de caracteres (Meacham & Estabrook, 1985; Farris & Kluge, 1979). Mucho más ruido originó otra corriente que se abría paso, dentro del cladismo, a finales de los setenta y que se conoce como transformación del Cladismo o *Pattern Cladistics* (Fig. 2). Ésta propugnaba una práctica sistemática de descubrimiento —de patrones— independiente de toda teoría causal (Platnick, 1985; Nelson & Platnick, 1981). En esto se asemejaba a otras corrientes fuera del cladismo, en concreto a la taxonomía numérica de los comienzos (Cain, 1959). No puede decirse que los cladistas transformados sean mayoría, entre otras cosas porque muchos creen que la metodología cladística estaría deslegitimizada (Scott-Ram, 1990) si no se asume, como premisa, un modelo darvinista de evolución. Sin embargo, algunos de sus argumentos distan mucho de ser descabellados. Afirman que han de recuperarse los patrones de diversidad de la naturaleza sin ninguna referencia a los procesos que han causado esos patrones —*no theory is necessary in order to discover patterns in nature*—. Además llaman la atención sobre la posible circularidad que supone el asumir como premisa una teoría causal: “para que la explicación causal de los patrones (o sea de los grupos que percibimos) sea convincente y eficiente, es mejor que el patrón no se perciba en términos del proceso explicativo (Fig. 2).

La denominación de cladistas transformados ha caído bastante en desuso y su heredera es una corriente que algunos llaman *taxic approach* (Kitching et al., 1998). La homología, salvo en sus comienzos engarzados con la anatomía comparada, siempre se ha definido por relación a un concepto no empírico (antepasado común). Las hipótesis sobre grupos (monofilia) también se definen sobre la base del mismo concepto no empírico. Para soslayar este problema, proponen manejar los caracteres como meras hipótesis sin ninguna referencia a antepasados comunes u otra causa (homología primaria, en el sentido de Pinna, 1991) en el análisis cladístico. Tras el análisis, las hipótesis sobre grupos resultantes se pueden interpretar como resultado de la evolución y, también a posteriori, las hipótesis sobre cambios entre caracteres pueden “leerse” en el cladograma: polaridad, series de transformación. Este enfoque difiere de aquel, más próximo a las ideas de Hennig, que pone el énfasis en las transformaciones entre caracteres, que hay que intentar descubrir en una fase previa al análisis cladístico y que, de hecho, condicionan ya de partida sus resultados.

Otra idea que ha surgido dentro del cladismo como alternativa a la parsimonia, y con no demasiada repercusión en los usuarios reales, es la del *Three-taxon statements analysis* o *three-item analysis* (Nelson & Platnick, 1991). Sus defensores codifican los caracteres de una forma muy simple. Todo se reduce a considerar y contabilizar simples hipótesis de relaciones del tipo de el taxon A está más relacionado con B que con C, sobre la base, por ejemplo de que A y B poseen un atributo ausente en C. Pero si ese atributo está también presente en D, hay que contabilizar, en principio, otras dos propuestas: (AD)C y (BD)C.

Otra corriente, muy minoritaria pretende sustituir el sistema de nomenclatura linneano por otro “filogenético”, de tal forma que lo que se nombrarían serían los clados y se harían sobre la base de conceptos filogenéticos como nodos o apomorfias (Queiroz & Gauthier, 1992). Esta idea ha sido recibida con indiferencia ya que no aportaría estabilidad a la nomenclatura, su uso sería complicado y, además, tendría la desventaja de ser muy dependiente de la teoría que lo sustenta.

Se puede, en definitiva, hablar del Cladismo como corriente falta de cohesión. Las causas de que ello sea así son varias, en mi opinión.

Las ideas de Hennig no estaban desarrolladas en el plano metodológico de forma totalmente satisfactoria. Los problemas que quedaron pendientes dejaban distintas vías de solución. Tampoco puede ignorarse la contribución a una cierta heterogeneidad que se deriva del hecho de que, en los últimos años, la metodología se ha convertido en la más usada para reconstrucción filogenética, por muchos taxónomos, muchos de los cuales están alejados de las controversias más teóricas (Kluge & Wolf, 1993). Ello puede introducir un cierto eclecticismo en una corriente que, desde su origen, ha dado mucha importancia a la legitimación epistemológica del método (Hull, 1983). Por otro, lado, la última fuente de evidencia incorporada a la Sistemática, los caracteres moleculares, por su gran importancia, ha ejercido una influencia importante en la Cladística en general, en la última década. Ello es así aunque no sea más que por desplazar el centro de gravedad de las controversias teóricas hacia aspectos específicos de los datos moleculares.

FALSOS MITOS SOBRE EL CLADISMO

No voy a referirme aquí a polémicas completamente superadas que se referían en realidad a convenciones metodológicas del

sistema de Hennig (Hull, 1988). Por ejemplo, el que cuando una especie ancestral especia —sufre cladogénesis— da lugar a dos especies hijas y la ancestral se considera extinta; tampoco a aquella de que los cladogramas debieran ser dicotómicos y la filogenia no tiene por qué ser así. En cambio, voy a referirme a tres mitos que han sido más persistentes.

La idea de que, con anterioridad al análisis, el investigador debe decidir cuál de los estados alternativos de un carácter es primitivo y cuál derivado está firmemente implantada en buena parte de los practicantes de la cladística. El origen y la explicación de este convencimiento son muy comprensibles pues están en el propio Hennig (1968). Cuando se decide, a priori, qué estados de un carácter se han transformado en cuáles otros —es decir, se polariza dicho carácter— y se registra la distribución de tales estados en los taxones que queremos analizar, de hecho, se puede ya formular una hipótesis de agrupación. Los taxones que comparten el estado que presumimos derivado formarían un grupo natural si realmente dicho estado fuera derivado, o sea, aparecido como novedad en el antepasado común más reciente o especie ancestral (*stem group*) de tales taxones. Esta estrategia, también llamada *hennigian argumentation*, es la que se sigue cuando se construyen cladogramas a mano, si bien cada vez esto es más raro y, desde luego, se limita al análisis de caracteres morfológicos.

La razón de que se cuestione esta estrategia es la falta de acuerdo sobre los criterios que nos permitirían polarizar los caracteres. Las polémicas entre los defensores del *outgroup comparison* (Brooks & Wiley, 1985), o comparación con el grupo externo, y del *ontogenetic criterion* (Nelson, 1985) fueron muy intensas en los años ochenta. Hennig había señalado varios criterios de polarización pero la mayoría —en realidad todos, menos los dos implicados en la polémica— se han ido abandonando por poco fiables (Forey et al., 1992). Baste, como ejemplo, mencionar el criterio de “común igual a primitivo”. Según éste, dentro de un grupo monofilético, cabría esperar que los estados de carácter primitivos (plesiomórficos) estuvieran más ampliamente distribuidos que el o los estados derivados (apomórficos). Es un criterio inadecuado porque asume que el proceso evolutivo tiende a conservar los caracteres plesiomórficos.

La solución al problema de la polarización ha sido formalizada recientemente (Nixon & Carpenter, 1993) y es el resultado de analizar las relaciones entre enraizamiento de cladogramas, polarización y comparación con el grupo externo. A diferencia de la polarización basada en el criterio ontogenético, considerada directa, la que se basa en el grupo externo se apoya en una filogenia de orden superior al de nuestro grupo de estudio y por ello se considera indirecta. La polarización de los caracteres y el *outgroup analysis* lo que persiguen, en definitiva, es atribuir estados de carácter al antepasado común más reciente del grupo de estudio o, dicho de otra forma, decidir en qué nodo se coloca la raíz del cladograma. Enraizar un cladograma es fundamental, pues al determinar la polaridad de los caracteres, condiciona cuáles son apomórficos y, por tanto, informativos en un contexto cladístico y en consecuencia qué grupos son propuestos como monofiléticos en la hipótesis representada en ese cladograma (Fig. 3). Por tanto, más que determinarse previamente, lo que ha de hacerse para conocer el sentido de la evolución de los caracteres es leerse a posteriori. ¿Dónde? En el cladograma óptimo, es decir, en la mejor hipótesis de relaciones con los datos disponibles. Ahora bien, si el problema es la posición de la raíz, ¿cómo se determina ésta? La propuesta es hacer un análisis no enraizado simultáneo. Esto significa que no cerramos la puerta a que, si los datos lo sugieren, se cuestione

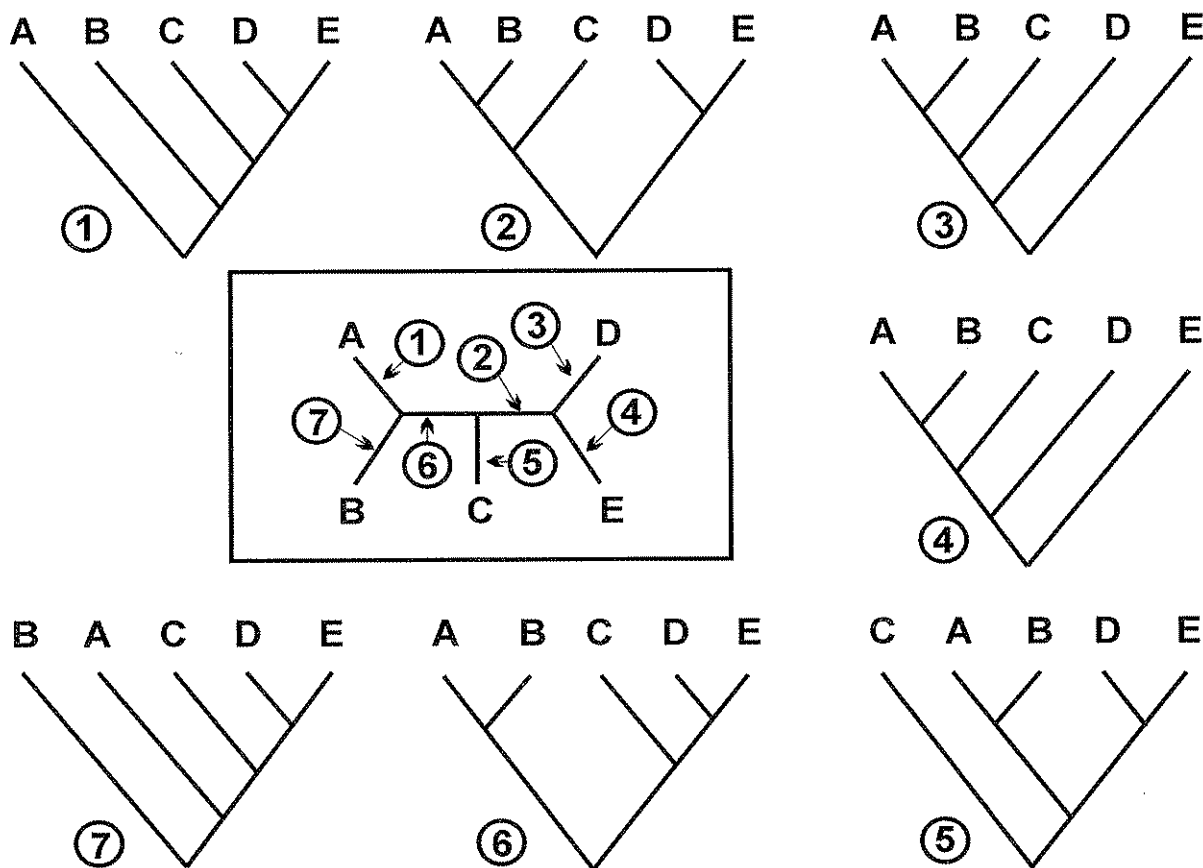


Fig. 3.- Siete hipótesis de relaciones diferentes, en un grupo de 5 taxones, resultantes de enraizar un cladograma no enraizado en las siete posibles posiciones.

la monofilia del grupo de estudio (ingroup). En este análisis debemos incluir más de un taxon que sabemos que está fuera del grupo. Una vez calculado el cladograma no enraizado —la red— más parsimoniosa, se enraza en el nodo que conecta outgroup con ingroup. De esta forma, lo que determinamos es la dirección de transformación global del árbol más que la dirección individual de los caracteres. Este procedimiento, que no necesita de polarización previa de caracteres, tiene una restricción: no se pueden aplicar modelos de evolución a los caracteres del tipo de los denominados Dollo y Camin-Sokal. Si impedimos apariciones paralelas del mismo estado de carácter o reversiones a un estado anterior, la longitud de un árbol no sería independiente de la posición de la raíz. Con lo cual, no habría una red que fuera la más parsimoniosa con independencia del enraizamiento y por tanto, no se podría hacer un análisis simultáneo.

Otra de las propiedades atribuidas a la cladística es que cualquier pequeña alteración en la matriz de datos puede cambiar drásticamente los resultados del análisis y, por tanto, la hipótesis de relaciones filogenéticas preferida. Esto puede ser cierto. Sin embargo, la solución es muy sencilla: no puede darse credibilidad a ramas (*clados*) que no estén firmemente apoyados por datos. En Cladística la única evidencia son las distribuciones de caracteres en los organismos estudiados. Para estimar cuáles ramas están más respaldadas por los datos hay varios métodos. Tal vez el más recomendado, aunque no el más sencillo, es el *Bremer support* o *decay index* (Bremer, 1994), consistente en averiguar en cuantos pasos (cambios o transformaciones en caracteres) hay que aumentar el árbol para que esa rama pueda desaparecer. O, dicho de otra forma, la diferencia en longitud entre los cladogramas más cortos con y sin la rama en cuestión. Si únicamente tenemos dos clado-

gramas igualmente parsimoniosos y en uno de ellos un determinado clado no aparece, el *Bremer support* de ese clado es 0, porque no es necesario aumentar la longitud del cladograma para que se pierda. Es decir, el respaldo para el clado por parte de los datos es bajo o, lo que es lo mismo, el grado de confianza que tenemos en que los taxones que componen ese clado formen realmente un grupo natural no es muy alto. Si por el contrario, necesitamos añadir bastantes pasos al cladograma más corto para que el clado no aparezca, dicho clado está sólidamente apoyado por los datos.

Al margen del *Bremer support*, los dos métodos más utilizados para estimar el apoyo de los grupos son el *Bootstrap* (Felsenstein, 1985) y el *Jackknife* (Farris et al., 1997). Ambos se basan en técnicas estadísticas de re-muestreo que alteran al azar todos o algunos de los caracteres de la matriz de datos. La búsqueda de los árboles óptimos para todas y cada una de las matrices re-muestreadas nos permite descubrir qué grupos constituyen más verosímilmente grupos naturales. En aquellos que aparecen en la mayoría de los árboles, "sin importarles" que la matriz se haya alterado ligera o más drásticamente, tenemos una confianza alta. Estos grupos con porcentaje de bootstrap alto no dependen solo de un carácter sino que están apoyados por tantos que aunque algunos de ellos se eliminen al alterar la matriz, siguen recibiendo el apoyo de otros. Como en cladística se asume que los caracteres son independientes, cuantos más apoyen un grupo más evidencia independiente hay de que es un grupo natural.

Hay otra creencia sobre el método cladístico, también muy extendida, y que no es correcta tal y como suele formularse (Goloboff, 1998). Se tiene la idea de que no puede hacerse un análisis cladístico —o que su fiabilidad es muy baja— a menos que incluyamos absolutamente todas las especies del

grupo. Sin embargo, el tipo de parentesco que el análisis cladístico persigue descubrir es relativo. Si con tres especies (A B C), A y B resultan estar más estrechamente emparentadas entre ellas —o sea, ((A B) C)—, la inclusión en el análisis de una cuarta especie D, no tiene por qué contradecir los resultados del primer análisis. Si se colocara junto a A —o sea (((A D) B) C)—, no alteraría la hipótesis relativa del primer análisis que dice que A y B están más estrechamente emparentadas entre sí que cualquiera de ellas con respecto a C. Esta consideración está muy relacionada con el *three-item analysis*, al que me he referido antes (Nelson & Platnick, 1991).

Es obvio que cuantos más taxones de un grupo monofilético estudiemos, más completa y fiable será una hipótesis de relaciones. Pero, en ausencia de tales condiciones, un primer análisis sin todos los taxones puede ser muy útil para una primera exploración de relaciones filogenéticas así como para dirigir la investigación futura. A pesar de esto, en un segundo análisis con D, es ciertamente posible que A resultara más próxima a C que a B o D —((A C) B) D)—, y esto sí estaría en contradicción con los resultados del primer análisis. Pero esta situación solo se da cuando las evidencias en que se basa ese primer análisis son débiles y cuando la muestra de taxones es muy poco representativa. Depende, en definitiva, de la calidad de los datos más que del método de análisis. Un temor por la fiabilidad de los resultados cuando se dispone de una muestra representativa y de un número de caracteres razonable, a la que le faltan por estudiar unos pocos taxones, no está justificado.

CONTROVERSIA ACTUALES Y RETOS FUTUROS

Total evidence vs. Taxonomic congruence

Esta polémica se ha avivado a raíz del aumento en la disponibilidad de datos surgida del empleo de técnicas moleculares. La congruencia entre caracteres es la mejor garantía —la única cuando se trata de datos moleculares— de que la hipótesis de relaciones que tenemos entre manos es buena (Patterson et al., 1993). Por ello, no es extraño que cada vez se nos demande a los sistemáticos más evidencias independientes, más matrices de otro tipo de datos que no sean morfológicos para ser confrontados con éstos. Cuando ya disponemos de datos de secuencias de un gen, se nos recomienda, con razón, que secuenciamos otro para ver si sugieren las mismas relaciones de parentesco entre los taxones que estudiamos. Por ello, es frecuente que estudiando un grupo de organismos podamos reunir caracteres potencialmente informativos de fuentes distintas y, en principio, independientes. La pregunta es: ¿cómo analizamos cladísticamente esos conjuntos de datos que provienen de distintas fuentes? ¿juntos o por separado? En el primer caso estamos procediendo de acuerdo con los partidarios de la *total evidence* (Kluge, 1989) o, como se puntualiza recientemente (Nixon & Carpenter, 1996), *simultaneous analysis*. En el segundo caso, después de hacer análisis cladísticos por separado de las distintas matrices de datos y por tanto generar cladogramas correspondientes, se buscan en éstos los puntos en común mediante árboles de consenso (Fig. 4). Esta práctica se denomina *taxonomic congruence* (Miyamoto & Fitch, 1995) o, por oposición al *simultaneous analysis*, *partitioned analysis*.

En principio, parecería lógico juntar todos los caracteres en una única matriz. Los análisis filogenéticos que aplican parsimonia, en definitiva, maximizan la homología. Esto es, favorecen aquella hipótesis en la que un mayor número de caracteres que presumíamos homólogos en nuestro estudio de

caracteres resultan serlo —aparecen una sola vez— tras el análisis. El mejor test de homología para tales hipótesis es la congruencia con otros caracteres. Por ello, cuantos más introduzcamos en una matriz, más riguroso será el test de homología. Además, en la práctica, cuando se sigue este procedimiento de *evidencia total*, suele alcanzarse una mayor resolución en los cladogramas, esto es, una hipótesis más explícita de relaciones, con menos ambigüedades.

Ahora bien, cuando los conjuntos de datos provienen de fuentes de evidencia muy dispares pueden contener una información esencialmente distinta, debida a que los caracteres de uno y otro tipo han sufrido procesos evolutivos distintos. Aunque ambos conjuntos de datos contengan información útil para estimar relaciones filogenéticas, pueden requerir una interpretación distinta y, si los mezclamos puede que estemos introduciendo ruido adicional, como consecuencia del desajuste entre dos señales filogenéticas distintas. El caso más claro es el de la confrontación de datos moleculares con datos morfológicos. La distinción entre árboles filogenéticos —de especies— y árboles de genes (Doyle, 1992; Maddison, 1997) es consecuencia del reconocimiento de estas diferencias. Una de ellas es la tasa de mutación en secuencias de determinados genes cuando se compara con la tasa de cambio evolutivo en caracteres morfológicos. Pero no es la única diferencia. Otro fenómeno que actúa sobre datos de secuencias de ADN es el llamado *lineage sorting* o *deep coalescence* (Pamilo & Nei, 1988). Se refiere al caso en que la persistencia de polimorfismos ancestrales a través de linajes que se van ramificando, unido a una pérdida al azar de parte de los linajes, provocaría una estimación errónea de relaciones.

En definitiva, se podría abordar este problema analizando los datos de distintos genes separadamente y aplicando distintos modelos de cambio evolutivo a cada uno, para intentar reconstruir sus historias evolutivas. Pero este enfoque ya no utiliza la parsimonia como criterio metodológico sino el de máxima verosimilitud (Felsenstein, 1981). Trata de calcular la probabilidad de obtener la distribución de caracteres observada en la práctica si dicho modelo de cambio evolutivo fuera cierto.

Pero no es necesario abandonar la cladística para preferir el analizar las matrices de datos separadamente. Éstas pueden contener niveles de homoplasia distintos. Si analizamos, con parsimonia, separadamente las dos matrices averiguamos esta diferencia y podemos tratar de interpretarla. En este sentido, hay una práctica intermedia que consiste en mezclar matrices de datos en una única, sólo cuando se dan unas ciertas condiciones. ¿Cuáles? Que las dos matrices no sean demasiado heterogéneas, demasiado incongruentes. Esto puede estimarse mediante diversos tests (cf. Kitching et al., 1998; Johnson & Soltis, 1998).

Grandes matrices de datos

Otra de las controversias actuales está también directamente relacionada con la evidencia molecular, por la cantidad de datos que ésta es capaz de proporcionar. Cuando, para intentar reconstruir la filogenia de un grupo, nos basamos en secuencias del ADN normalmente disponemos de varios cientos de caracteres, excepcionalmente de varios miles. Las matrices de unas decenas de caracteres morfológicos que sustituyeron a la argumentación hennigiana excluían toda posibilidad de búsqueda manual de los cladogramas más parsimoniosos. Algunos ya denominaron, peyorativamente o no, a esta práctica *numerical cladistics*.

Para manejar la cantidad de datos que uno o varios genes proporcionan, a nadie se le ocurre que pueda prescindir

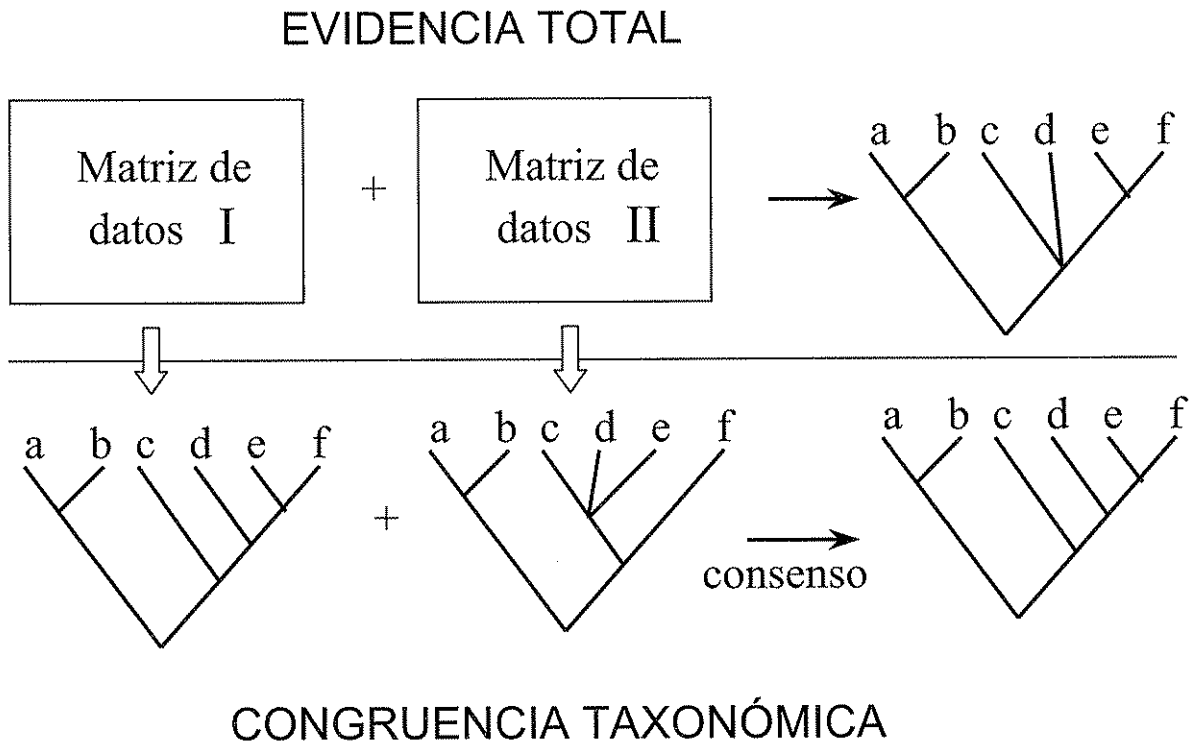


Fig. 4.- Cuando disponemos de más de un conjunto de datos diferentes para el mismo grupo de estudio hay dos procedimientos. En la *evidencia total* se juntan las dos matrices de datos en una sola y se busca el cladograma óptimo. En la *congruencia taxonómica* se analizan por separado ambas matrices y se calcula el cladograma de consenso de los cladogramas óptimos obtenidos a partir de cada una de las matrices.

se de programas de ordenador. En cambio, lo que se suscita es la eficiencia por parte de los programas para gestionar ese volumen de datos, así como las opciones de búsqueda de los árboles óptimos. Hay dos aspectos que deben tenerse en cuenta: el tiempo que el programa tarda en analizar los datos para encontrar el cladograma más parsimonioso y las garantías de encontrar éste último, y no alguno de los llamados subóptimos. Aunque se discute sobre la capacidad de uno u otro programa —PAUP, Hennig86, NONA— para producir unos resultados en tiempo mínimo (e.g. Nixon et al., 1998), este asunto podría resolverse con un estudio comparativo. Lo que sí merece más atención últimamente son las estrategias de búsqueda que minimicen el tiempo invertido así como los cálculos del grado de confianza que merecen las hipótesis de relaciones (Morgan, 1997). Las búsquedas pueden ser de dos clases, exactas o heurísticas. Las primeras garantizan que encuentran el cladograma más parsimonioso cualquiera que sea la cantidad de taxones y caracteres de la matriz. Pero, dado que por ejemplo el número de topologías enraizadas dicotómicas posibles para 10 taxones rebasa los 34 millones, estos métodos exactos pueden emplear tiempos de cálculo disparatados. Los heurísticos, en cambio, no garantizan encontrar el cladograma óptimo pero invierten mucho menos tiempo. Cuando nos enfrentamos a matrices tan grandes, la estrategia consiste en combinar distintas búsquedas heurísticas para que no se escape ninguna de las llamadas islas de máxima parsimonia (Maddison, 1991) y, así, tener una probabilidad alta de encontrar los cladogramas óptimos en un tiempo razonable.

Sin embargo, hay otro asunto prometedor surgido recientemente y que está relacionado con las grandes matrices de datos y con el apartado anterior sobre el análisis simultáneo o particionado de varias matrices. En 1993 se publicó un amplio estudio filogenético sobre angiospermas, basado en la

secuencia de un gen que codifica para la subunidad grande de la rubisco (*rbcL*) y que se encuentra en el ADN cloroplástico (Chase et al., 1993). Comprende más de 1400 pares de bases y por su tasa más bien baja de mutación es adecuado para analizar relaciones a nivel de familias. El trabajo se basaba en una matriz con 500 taxones y tuvo una repercusión amplia, apoyando algunas relaciones, previamente propuestas sobre la base de evidencias no moleculares, y sugiriendo otras nuevas.

En los últimos años varios grupos de investigación han trabajado para someter tales resultados al test de nuevos datos también moleculares. Otros, incluso, han re-analizado la misma matriz con estrategias de búsqueda distintas. Los genes que se han secuenciado, para ser confrontados con los resultados basados en *rbcL*, son el 18S (1855 pares de bases) perteneciente al ADN ribosómico, nuclear, y el *atbB* (1450 pares de bases) que forma parte del ADN cloroplástico (Soltis et al., 1998). Se hicieron varios análisis simultáneos de los tres genes —por pares o los tres conjuntamente— así como de los mismos por separado, aunque en un número de taxones algo menor que el primer estudio (190). El hecho más significativo que encontraron fue que el análisis simultáneo de los datos de los tres genes, es decir, uniendo las tres matrices en una sola que suma cerca de 5000 caracteres, superaba en todos los aspectos a los análisis por separado. En concreto, el tiempo de búsqueda se reducía muy sensiblemente para durar en torno a unos días. Ninguna de las matrices que combinaban dos genes había completado la búsqueda pasado el primer mes. Vale la pena mencionar que en el re-análisis de la matriz original de *rbcL* —de 500 caracteres— se habían invertido más de 11 meses de tiempo de funcionamiento de una CPU (Rice et al., 1997). Pero no sólo eso, además, con la matriz de 5000 caracteres se obtenía mayor resolución en clados internos y se obtenían valores de apoyo (bootstrap) altos o medianos para

un número significativamente mayor de ramas. Esto implica una menor ambigüedad en las hipótesis y una mayor eficiencia en la búsqueda de las mismas y representa un espaldarazo al enfoque de evidencia total. Por otro lado, contradice la conocida tendencia al aumento de homoplasia a medida que se añaden caracteres o taxones a una matriz.

Es un resultado prometedor porque implica que la señal filogenética que lleva cada uno de los genes, por un lado, se complementa. Es decir, que los genes que tienen una tasa de mutación más baja dan más información para resolver las relaciones de parentesco a niveles más basales —más antiguos— de la filogenia. Por el contrario, los de tasa más alta ayudan a resolver clados más terminales del cladograma de las angiospermas. Por otro lado, la congruencia con caracteres de los otros dos genes hace que afloren señales filogenéticas que estaban más o menos enmascaradas por ruido en las matrices individuales, es decir por variación que no se corresponde con la filogenia.

Enraizamiento del Árbol filogenético de los Seres Vivos

Este es otro reto científico importante, por la relevancia de la pregunta en sí misma y porque nuestra capacidad de darle respuesta está estrechamente relacionada con problemas metodológicos, y no sólo el del enraizamiento que he mencionado antes. Este problema está agravado por la dificultad de seleccionar objetos de comparación adecuados. Obviamente si los buscamos dentro del campo molecular, deberíamos atender a regiones del genoma que tengan una tasa de mutación muy baja y esto equivale a decir regiones codificantes para enzimas que intervengan en procesos muy vitales para un organismo (en el metabolismo, etc.). Se han secuenciado varios genes, entre otros del ARN ribosómico, para tratar de dilucidar las relaciones entre los grandes grupos como arqueobacterias (bacterias metanogénicas, halófilas, etc.), eubacterias (las verdaderas bacterias), eucariotas (protistas, plantas, hongos, animales, etc.) y los recientemente descubiertos eocitos. Sin embargo, no sólo se proponen topologías alternativas para las relaciones entre dichos grupos sino que se cuestiona la monofilia de alguno de ellos. Por otro lado, se ignora por completo la ubicación de los virus, en los cuales resulta todavía más difícil buscar secuencias homólogas. Téngase en cuenta que de los

criterios o estrategias para enraizar un árbol, podemos olvidarnos directamente del ontogenético. Una comparación con un grupo externo, ¿sería posible? Esta es una pregunta que no puedo responder. Con los datos disponibles un análisis simultáneo, en el sentido de Nixon & Carpenter (1996), tampoco parece posible. Porque ¿qué haríamos a la hora de colocar la raíz en el nodo que conectara al clado monofilético de la vida en el planeta con lo que es basal a dicho clado? Quizá, por el momento, tengamos que conformarnos con generar cladogramas no enraizados de genes, discutir los méritos relativos e implicaciones de las hipótesis alternativas resultantes de enraizarlos, y esperar.

El árbol filogenético de la vida en la tierra se puede ir construyendo y completando en sus distintos componentes o ramas. De hecho, hay una iniciativa que persigue este fin (*The tree of life* con la siguiente dirección Web:

<http://phylogeny.arizona.edu/tree/phylogeny.html>).

Pero para arrojar luz sobre las partes basales del Árbol y, especialmente, para lograr enraizarlo seguramente nos falta bastante camino.

Otro de los retos futuros, al menos en lo que se refiere a algunos grupos de seres vivos —entre ellos las plantas—, es el del tratamiento de las reticulaciones. Es decir, las consecuencias evolutivas de la hibridación. Este aspecto siempre se ha invocado por parte de los críticos de la Cladística (Cronquist, 1987). Con caracteres morfológicos no se ha avanzado mucho, al margen de concluir, con híbridos sintéticos F_1 , que este tipo de híbridos no tienen por qué causar grandes distorsiones en un análisis (McDade, 1992). En cambio, los marcadores moleculares pueden proporcionar información crítica aunque raras veces concluyente. Un caso, quizá poco representativo pero particularmente intrigante, es nuestro hallazgo de que mecanismos normalmente considerados problemáticos en reconstrucción filogenética —la hibridación y la homogeneización de secuencias multicopia— pueden, combinados, conducir a una estructura de datos jerarquizada (Fuertes et al., 1999). Sin embargo, el principal indicio de reticulación, en reconstrucción filogenética, son las discrepancias entre cladogramas basados en genes que se heredan de forma distinta (Rieseberg et al., 1996; Wendel & Doyle, 1998). Por ejemplo, entre genes nucleares heredados biparentalmente y otros, cloroplásticos o mitocondriales, heredados por vía materna.

BIBLIOGRAFÍA

- BREMER, K., 1994. Branch support and tree stability. *Cladistics*, **10**: 295-304.
- BROOKS, D. R. & WILEY, E. O., 1985. Theories and methods in different approaches to phylogenetic systematics. *Cladistics*, **1**: 1-11.
- CAIN, A.J., 1959. Deductive and inductive methods in post-Linnean taxonomy. *Proc. Linn. Soc. London*, **170**: 185-217.
- CARPENTER, J.A., 1987. Cladistics of cladists. *Cladistics*, **3**: 363-375.
- CHASE, M. W., SOLTIS, D. E., y 40 autores más, 1993. Phylogenetics of Seed Plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann. Missouri Bot. Garden*, **80**: 528-580.
- CRONQUIST, A., 1987. A botanical critique of Cladism. *Bot. Rev.*, **53**: 1-52.
- DOYLE, J. J., 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. *Syst. Bot.*, **17**: 144-163.
- DUPUIS, C., 1984. Willi Hennig's impact on taxonomic thought. *Ann. Rev. Ecol. Syst.*, **15**: 1-24.
- ENGELMANN, G. F. & WILEY, E. O., 1977. The place of ancestor-descendant relationships in phylogeny reconstruction. *Syst. Zool.*, **26**: 1-11.
- FARRIS, J. S., 1983. The logical basis of phylogenetic analysis. In: Platnick & Funk (eds.) *Advances in Cladistics*, vol. 2: 7-36.
- FARRIS, J. S. & KLUGE, A. G., 1979. A botanical clique. *Syst. Zool.*, **28**: 400-411.
- FARRIS, J. S., KÄLLERSJÖ, M., LIPSCOMB, D. & KLUGE, A. G., 1997. Parsimony jackknifing outperforms neighbor joining. *Cladistics*, **12**: 99-124.
- FELSENSTEIN, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.*, **17**: 368-376.
- FELSENSTEIN, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**: 783-791.
- FOREY, P. L., HUMPHRIES, C. J., KITCHING, I. L., SCOTLAND, R.W., SIEBERT, D. J. & WILLIAMS, D. M., 1992. *Cladistics. A practical course in Systematics*. Systematics Association Publications vol. 10. Clarendon Press, Oxford.
- FUERTES AGUILAR, J., ROSSELLÓ, J. A. & NIETO FELINER, G., 1999. Molecular evidence for the compilospecies model of reticulate evolution in *Armeria* (*Plumbaginaceae*). *Syst. Biol.*, **44** (en prensa).
- GOLOBOFF, P.A., 1998. *Principios básicos de Cladística*. Sociedad Argentina de Botánica. Buenos Aires.
- HENNIG, W., 1966. *Phylogenetic Systematics*. University of Illinois Press. Urbana, Illinois.
- HENNIG, W., 1968. *Elementos de una Sistemática Filogenética*. Editorial Universitaria de Buenos Aires.
- HULL, D. L., 1983. Karl Popper and Plato's metaphor. In: Platnick & Funk (eds.), *Advances in cladistics*, vol. 2: 177-189.
- HULL, D. L., 1988. *Science as a process. An evolutionary account of the Social and Conceptual Development of Science*. University of Chicago Press. Chicago & London.
- HUXLEY, T.H., 1874. On the classification of the Animal Kingdom. *Nature*, **11**: 101-102.
- JOHNSON, L. A. & SOLTIS, D. E. 1998. Assessing congruence: empirical examples from molecular data. In: SOLTIS, D. E. SOLTIS, P. S. & DOYLE, J. J. (eds.), *Molecular Systematics of Plants II. DNA sequencing*: 297-348. Kluwer Academic Publishers, Boston.
- KITCHING, I.L., FOREY, P.L., HUMPHRIES, C.J., & WILLIAMS, D.M., 1998. *Cladistics. Ed. 2. The theory and practice of parsimony analysis*. Systematics Association Publications vol. 11. Oxford University Press, Oxford.
- KLUGE, A.G., 1984. The relevance of parsimony to phylogenetic inference. In: DUNCAN, T. & STUESSY, T.F. (eds.), *Cladistics: perspectives on the reconstruction of evolutionary history*: 24-38. Columbia University Press, New York.
- KLUGE, A.G., 1989. A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boidae: Serpentes). *Syst. Zool.*, **38**: 7-25.
- KLUGE, A.G. & WOLF, A.J., 1993. Cladistics: what's in a word? *Cladistics*, **9**: 183-199.
- MADDISON, D.R., 1991. The discovery and importance of multiple islands of most-parsimonious trees. *Syst. Zool.*, **40**: 315-328.
- MADDISON, W.P., 1997. Gene trees in species trees. *Syst. Biol.*, **46**: 523-536.
- MAYR, E., 1942. *Systematics and the origin of species*. Cambridge University Press, Cambridge.
- MCDADE, L.A., 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution*, **46**: 1329-1346.
- MEACHAM, C.A. & FARRIS, J.S., 1985. Compatibility methods in systematics. *Ann. Rev. Ecol. Syst.*, **16**: 431-446.
- MIYAMOTO, M.M. & FITCH, W.M., 1995. Testing species phylogenies and phylogenetic methods with congruence. *Syst. Biol.*, **44**: 64-67.
- MORGAN, D.R., 1997. Decay analysis of large sets of phylogenetic data. *Taxon*, **46**: 509-517.
- NELSON, G., 1985. Outgroups and ontogeny. *Cladistics*, **1**: 29-45.
- NELSON, G. & PLATNICK, N., 1981. *Systematics and biogeography. Cladistics and vicariance*. Columbia University Press, New York.
- NELSON, G. & PLATNICK, N., 1991. Three-taxon statements: a more precise use of parsimony? *Cladistics*, **7**: 351-366.
- NIXON, K. C. & CARPENTER, J. M., 1993. On outgroups. *Cladistics*, **9**: 413-426.
- NIXON, K. C. & CARPENTER, J. M., 1996. On simultaneous analysis. *Cladistics*, **12**: 221-241.
- NIXON, K. C., DAVIS, J.A. & GOLOBOFF, P.A., 1998. Search Strategies for large datasets: an example using *rbcL*. *Amer. J. Bot.*, **85** (suppl.): 148.
- PAMILO, P. & NEI, M., 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.*, **5**: 568-583.
- PATTERSON, C., WILLIAMS, D. M., & HUMPHRIES, C. J., 1993. Congruence between molecular and morphological phylogenies. *Ann. Rev. Ecol. Syst.*, **24**: 153-188.
- PINNA, M.C.C. DE, 1991. Concepts and tests of homology in the cladistic paradigm. *Cladistics*, **7**: 367-394.
- PLATNICK, N. I., 1985. Phylosophy and the transformation of cladistics revisited. *Cladistics*, **1**: 87-94.
- QUEIROZ, K. & GAUTHIER, J., 1992. Phylogenetic taxonomy. *Ann. Rev. Ecol. Syst.*, **23**: 449-480.
- RICE, K.A., DONOGHUE, M.J. & OLMSTEAD, R.G., 1997. Analyzing large data sets: *rbcL* 500 revisited. *Syst. Biol.*, **46**: 554-563.
- RIESEBERG, L.H., WHITTON, J. & LINDER, C.R., 1996. Molecular marker incongruence in plant hybrid zones and phylogenetic trees. *Acta Bot. Neerl.* **45**: 243-262.
- SCOTT-RAM, N. R., 1990. *Transformed cladistics, taxonomy and evolution*. Cambridge University Press, Cambridge.
- SOLTIS, D.E., SOLTIS, P.E., MORT, M.E., CHASE, M.W., SAVOLAINEN, V., HOOT, S.B. & MORTON, C.M., 1998. Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for angiosperms. *Syst. Biol.*, **47**: 32-42.
- SWOFFORD, D.L., OLSEN, G.J., WADDELL, P.J. & HILLIS, D.M., 1996. Phylogenetic inference. In D.M.HILLIS, C. MORITZ & B.K. MABLE (eds.), *Molecular systematics ed. 2*: 407-514. Sinauer Assoc., Sunderland, Massachusetts.
- WENDEL, J. F. & DOYLE, J. J., 1998. Phylogenetic incongruence: Window into genome history and molecular evolution. In: SOLTIS, D. E. SOLTIS, P. S. & DOYLE, J. J. (eds.), *Molecular Systematics of Plants II. DNA sequencing*: 265-296. Kluwer Academic Publishers, Boston.
- WILEY, E. O., 1981., *Phylogenetics. The theory and practice of phylogenetic systematics*. John Wiley & sons, New York.
- WILEY, E.O., SIEGEL-CAUSEY, D., BROOKS, D.R. & FUNK, V.A., 1991. *The Compleat Cladist. A primer of phylogenetic procedures*. Univ. of Kansas, Museum Nat. Hist. special publication n° 19. Lawrence, Kansas.