

Caracteres Bioquímicos y Nucleares en los Métodos de la Sistemática Moderna*

ADRIAN NIETO***
JORGE LLORENTE **

INTRODUCCION

Desde la aparición del libro clásico *The New Systematics* de Huxley (1940) se ha avanzado significativamente en la Taxonomía Biológica. Dicho libro trata, entre otros puntos, sobre la posibilidad de que varios conjuntos de caracteres "no morfológicos" pudieran ayudar a superar diversos problemas de discriminación efectiva de las especies (aún las "cripticas"), del análisis filogenético y la clasificación de organismos recientes. Se han publicado desde entonces —durante casi cinco décadas— muchos libros y ensayos sobre Quimiotaxonomía, Serotaxonomía, Cariotaxonomía y otros tópicos afines, como el uso de las secuencias de proteínas y de ácidos nucleicos en la Taxonomía (Alston y Turner, 1963; Blackwelder, 1967; Bendz y Santesson, 1974; Bisby, Vaughan y Wright, 1980; Wiley, 1981; Goodman, 1982 y Oxford y Rollinson, 1983).

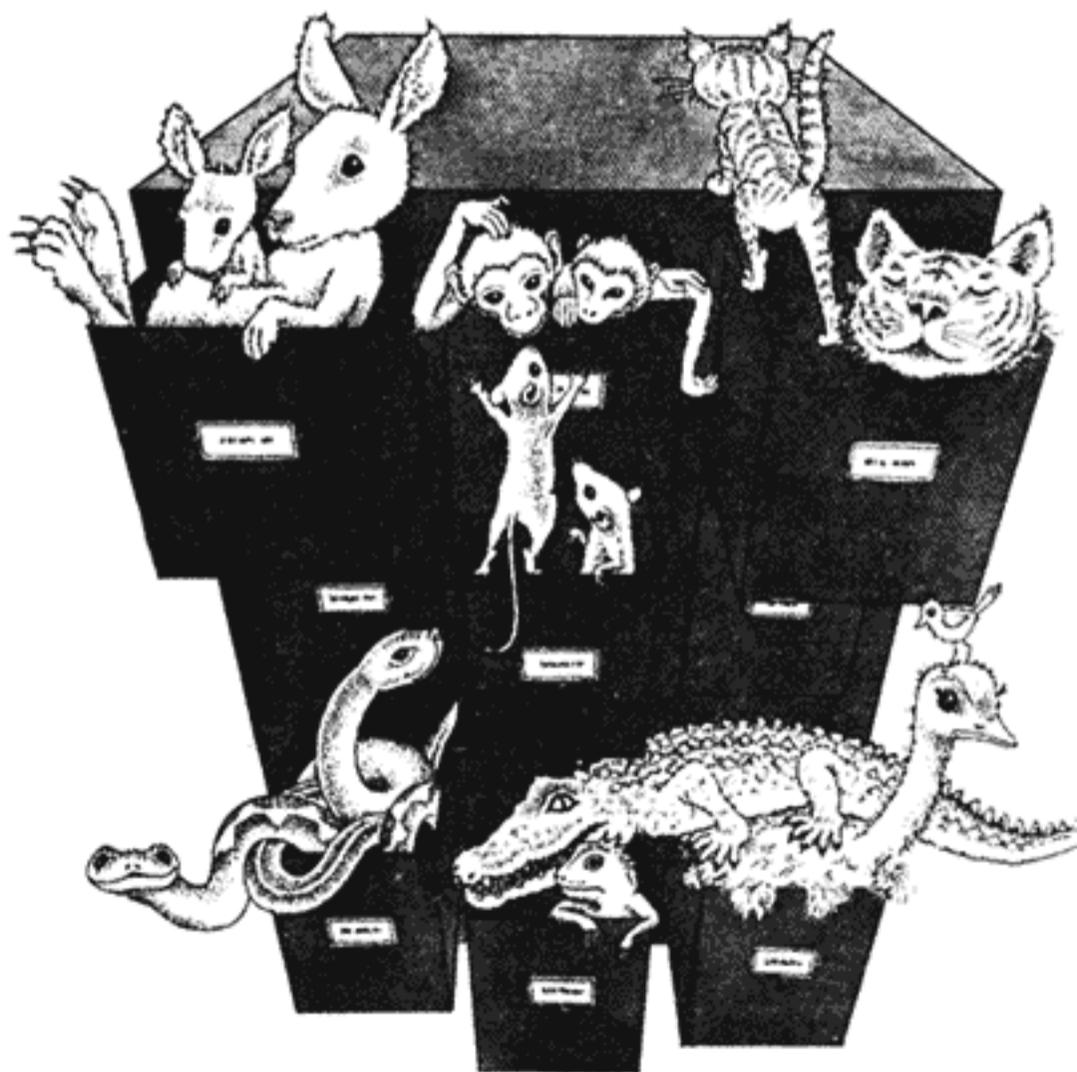
Hubo un tiempo en que el uso de la Inmunología y el estudio de los cromosomas fueron vistos como elementos potencialmente convincentes para el reconocimiento de afinidades entre los taxa; sin embargo, su uso no ha tenido el impacto que se esperaba, a pesar del éxito obtenido en algunos aspectos. Durante las últimas dos décadas el conocimiento de los ácidos nucleicos y de la estructura de las proteínas vino a dar un nuevo impulso a este enfoque de la Sistemática; éste aún no se agota, aunque cada día se comprende mejor que no por llegar más cerca del gen se logra necesariamente la "mejor" o "más adecuada" Sistemática de los organismos (Wiley, 1981).

Algunos fenómenos genético-reproductivos y evolutivos han planteado diversos problemas a los que no se han encontrado soluciones plenamente satisfactorias; éstos tratan sobre la interpretación y utilización apropiada de ciertos caracteres en Filogenia de organismos. Por otro lado, diversos

• Este trabajo se inició en los seminarios del Curso de Posgrado Sistemática Zoológica Contemporánea I que dirige el coautor.

* ENEP Zaragoza, UNAM.

** Museo de Zoología, Facultad de Ciencias UNAM.



fenómenos de Genética Molecular descubiertos en años recientes, pueden hacer cambiar algunas de nuestras concepciones sobre la interpretación de la similitud para proponer o descubrir relaciones genealógicas entre los organismos (*v. gr.* genes que no se expresan por periodos largos de tiempo en varios linajes y que son heredables (simplesiomorfias no reconocidas)).

Algunos de los principales objetivos de la Sistemática son el reconocimiento y la explicación de las relaciones de parentesco (Genealogía), de la similitud (relaciones patristicas) y de la distribución de los caracteres entre las especies y los taxa supraespecíficos. Estos objetivos se realizan a través del análisis de los caracteres con base en el estudio de la posición filogenética de los taxa (Wiley, 1981).

La comparación y la evaluación de las diferencias y similitudes morfológicas del fenotipo constituyen las fuentes de datos tradicionales y más frecuentemente utilizadas en la práctica taxonómica; pero si bien proveen de una aproximación legítima, relativamente económica y fácil de estudiar, esos datos no ofrecen una información completa y, en muchos de los casos, tampoco efectiva para el análisis en Filogenia. Esto último se aplica especialmente al caso de la sistemática de especies en las que las discontinuidades morfológicas no son una expresión necesaria suficiente de los mecanismos de aislamiento reproductivo y, en última instancia, ni de independencia evolutiva, que son los factores principales a reconocer (ver Wiley, 1978 y 1981, para una explicación más detallada). De acuerdo con Reig (1983) los faneromorfismos intraespecíficos (manifestaciones morfológicas muy distintas en una misma especie) y los muchos casos de sinmorfismo interespecífico (especies crípticas) así lo confirman.

Durante la década de 1960 convergieron varias ideas en Sistemática: similitudes metodológicas y conceptuales en Bioquímica y Cladística, generándose ideas equivalentes Numérico-Cladísticas, a las que Patterson (1981) llamó Sistemática Analítica; las mismas analogías se presentan en la Biogeografía Histórica. En el caso de la Sistemática Bioquímica, Fitch y Margoliash (1967) fueron los iniciadores. Estos autores desarrollaron métodos de reconstrucción filogenética a partir de datos moleculares, independientemente de los cladistas. Kluge y Farris (1969) adoptaron también los métodos numéricos, acoplándolos a los conceptos cladistas para la reconstrucción de la historia evolutiva.

Finalmente muchos taxónomos comenzaron a seguir las ideas de Hennig (1968) también conocidas como Sistemática Filogenética. Se puede considerar que las características en común de las tres tendencias o enfoques de la Sistemática son las siguientes: 1. Todos los taxa o unidades empíricas son considerados ramas terminales en cualquiera de sus dendrogramas. 2. El principio para producir los dendrogramas se basa en el criterio de mayor parsimonia en la distribución de homologías (caracteres, componentes o sinapomorfias). Las tres escuelas de pensamiento insistieron en la propiedad de homología y en el criterio metodológico de parsimonia. Por otra parte, criticaron a la "Sistemática Narrativa" o tradicional porque permitía cualquier explicación y se inmunizaba contra la crítica (Humphries, 1982).

En este artículo se comentan principalmente las técnicas de obtención de caracteres "no morfológicos"; sin embargo, en muchos casos estas técnicas son inseparables de los métodos de análisis (*v. gr.* en la evidencia inmunológica o hibridación de ADN); por otra parte, tales métodos pueden ser aplicables tanto a los caracteres morfológicos como a los bioquímicos. "Similitud total" de los fenetistas *versus* criterio de posición filogenética —por medio del establecimiento de las relaciones del grupo hermano— de los cladistas, son posiciones metodológicas y filosóficas de las que los caracteres bioquímicos y nucleares no se pueden sustraer. Respecto a ello conviene señalar que, para muchos sistematas bioquímicos, el fundamento básico del problema está en el acceso a la información genética; piensan que la expresión fenotípica, cualquiera que ésta sea, es de importancia secundaria en comparación con la información de los ácidos nucleicos.

Otras teorías biológicas, en relación con el concepto de epigenotipo, han demostrado los posibles errores en este tipo de juicios. Algunos sistematas han respondido de manera sencilla que nos ocupamos de clasificar organismos y no macromoléculas; por ello los taxónomos echan mano de todo tipo de caracteres, que de acuerdo a su comportamiento en la evolución, sólo son útiles en un determinado nivel de jerarquía. No obstante estas aparentes limitaciones, los caracteres bioquímicos son de enorme valor para la sistemática de cualquier taxón y pueden ser analizados a través de una gran variedad de métodos cladistas, gradistas y fenetistas. La tendencia moderna es más bien ecléctica, ya que algunos caracteres bioquímicos son de gran valor para el estudio de grupos en determinados niveles de la jerarquía taxonómica. En otros grupos o en otros niveles de la jerarquía, los caracteres morfológicos continúan siendo valiosos y superiores a los caracteres bioquímicos; de cualquier modo, no se trata ahora —después de una polémica de varias décadas— de encontrar cuáles son los "mejores" caracteres en la Sistemática, sino de reconocer y justipreciar el valor de ellos para cada grupo y nivel taxonómicos y que nos permitan así comparaciones más precisas de los patrones temporales y espaciales de las relaciones entre los organismos. Entre mayor sea la congruencia en los conjuntos de evidencia analizados por métodos robustos, mayor será la aproximación a la verdadera filogenia.

A continuación se exponen técnicas no morfológicas que han venido a complementar y superar muchas de las limitaciones del pensamiento morfológico clásico y que en varios casos pueden considerarse, en la concepción filogenética de Hennig (1968), integrantes del método holomorfológico. Se presentan en este trabajo aspectos sistemáticos del uso de los cromosomas, las proteínas, los lípidos y los ácidos nucleicos.

CROMOSOMAS

Diversos cambios en el número, tamaño y forma de los cromosomas han ocurrido en la evolución orgánica a través de fusiones, fisiones, translocaciones, inversiones, duplicaciones y deleciones cromosómicas (Dobzhansky *et al.*, 1977). Durante las décadas de los cincuenta y sesenta se determinaron el número y la morfología gruesa (tamaño, forma, longitud de los brazos, proporciones entre las longitudes de los brazos, posición del centrómero y otros caracteres) de los cromosomas de algunos taxa. Hubo un interés creciente en la citología cromosómica por parte de zoólogos y botánicos, y un incremento del uso de los cariotipos para inferir relaciones filogenéticas (Atchley, 1972 y Duellman, 1985). Ya se han efectuado revisiones de los datos que comprenden el número y la forma de los cromosomas para los animales (Robinson, 1971; White, 1973 y Wiley, 1981) y también se han discutido los trabajos basados en los cromosomas de las plantas (Darlington, 1958; Stebbins, 1971 y Radford *et al.*, 1974).

Sin embargo, durante las décadas citadas, resultó bastante especulativo el sólo uso de los números y las morfologías cromosómicas gruesas para el análisis de las relaciones genealógicas; ello se debió, principalmente, a la imposibilidad de determinar las homologías cromosómicas en diferentes genomas, así como al desconocimiento de los mecanismos que intervienen en los cambios de las morfologías cromosómicas y la dirección de aneuploidía (Atchley, 1972). Estos problemas tuvieron como consecuencia, en los casos en que las homologías cromosómicas de diferentes genomas no podían determinarse, que el grado de exactitud de los resultados fuese inversamente proporcional al mayor número cromosómico.

Durante este periodo, los únicos métodos para obtener información "exacta" sobre las homologías de los cromosomas eran la hibridación experimental y en algunos casos, la autorradiografía (útil para señalar patrones de marcaje sincrónicos o asincrónicos donde han ocurrido translocaciones) y aunque muchos biólogos aplicaron correctamente la aproximación citogenética de la hibridación artificial y el análisis cromosómico subsecuente



—con el fin de obtener evidencias de diferenciación genética entre los taxa— muchos otros "sistematas cariotípicos" se dieron por satisfechos con la comparación de diferentes taxa sobre la base de sus números y morfologías cromosómicas; con ello especularon sobre relaciones genealógicas y divergencia genética, a pesar de los problemas señalados por Atchley (*op. cit.*).

Existen diversos ejemplos de tales trabajos (Nadler, 1966; Gorman, 1970 y Bogart, 1970) donde se usaron solamente los números y las morfologías cromosómicas gruesas para sacar conclusiones detalladas acerca de la divergencia genética en varios grupos de vertebrados. Los estudios cromosómicos se estancaron en este nivel hasta que el desarrollo de nuevas técnicas permitió la identificación de unidades homólogas en los cromosomas (Duellman, 1985).

Sin embargo, los problemas antes expuestos se refieren, básicamente, al uso de los cariotipos compuestos por cromosomas somáticos típicos y no se extienden al uso de los cromosomas politénicos gigantes (Atchley, 1972). Estos cromosomas de los Diptera exhiben secuencias de bandas individualmente reconocibles que permiten la detección de las homólogas cromosómicas y de inversiones u otros rearrreglos cromosómicos; muchas veces las afinidades genealógicas pueden determinarse a partir de estos rearrreglos, sobre todo aquéllos que involucran "inversiones superpuestas" (Dobzhansky *et al.*, 1977 y Ayala y Kiger, 1984).

Tales inversiones son inversiones cromosómicas que incluyen parte de segmentos cromosómicos previamente invertidos. Por ejemplo, supóngase que las secuencias de las bandas en tres diferentes cromosomas homólogos son ABCDEFGH, AEDCBFGH y AEDFBGHI. La segunda secuencia puede surgir de la primera, o la primera de la segunda, por una sola inversión que comprendiera al segmento BCDE; de manera similar la tercera secuencia puede surgir de la segunda o viceversa por una sola inversión del segmento CBF. Sin embargo la tercera secuencia no puede surgir por una sola inversión de la primera, ni la primera de la tercera. La segunda secuencia representa un paso intermedio necesario entre la primera y la tercera; si no hay otra información disponible no hay modo de saber cuál de las tres secuencias pudo haber sido la ancestral, no obstante sólo hay tres series de transformación posibles para inferir la genealogía: $1 \rightarrow 2 \rightarrow 3$, $3 \rightarrow 2 \rightarrow 1$, $6 \ 1 \rightarrow 2 \rightarrow 3$.

Al término de los años 60 varias genealogías habían sido ya propuestas basándose en las inversiones superpuestas en los organismos con cromosomas politénicos, tales como *Drosophila*, las moscas negras (Simuliidae) y algunos mosquitos (Culicidae); Dobzhansky *et al.* (1977) ilustraron algunos ejemplos notables de tales genealogías (*v.gr.* la propuesta para *Drosophila*; Carson y Kaneshiro, 1976).

Volviendo a los cromosomas somáticos típicos, en los inicios de la década de 1970 por fin se desarrollaron técnicas para identificar unidades homólogas de cromosomas de diferentes genomas. Estas técnicas tienen diferencialmente los cromosomas, revelando regiones claras y oscuras alternas (bandas) a lo largo de los mismos. Dos técnicas de bandeado usadas comúnmente son el método de giemsa (de bandas G; Pardue y Gall, 1970) y el método de "mostaza-quinacrina" (de bandas C; Arrighi y Hsu, 1971); además de una técnica que permite la identificación de caracteres subestructurales de los cromosomas, que es la resolución de regiones organizadoras nucleolares, desarrollada por Goodpasture y Bloom (1975).

Es posible un trabajo más detallado mediante el examen de inversiones, deleciones y translocaciones en los cromosomas bandeados y de las características subestructurales de los cromosomas (Wiley, 1981). Utilizando la morfología gruesa y los caracteres subestructurales de los cromosomas se han podido desarrollar varios trabajos aplicando los principios de la cladística; por ejemplo, en las genealogías de las tortugas criptodiras (Bickham y Carr, 1983), de los ratones peromiscinos (Yates, Baker y Barnett, 1979; Stangl y Baker, 1984) y de los murciélagos glosofagíneos (Haiduk y Baker, 1982). Con todo, son pocas las genealogías "cromosómicas" enmarcadas explícitamente en las bases de la Sistemática filogenética *sensu* Hennig (Wiley, 1981).

A manera de ejemplo de tales análisis filogenéticos se ofrece el siguiente: Yates, Baker y Barnett (1979) publicaron un estudio de tres géneros de ratones peromiscinos; en él usaron comparaciones de grupo externo ("outgroup") para intentar establecer el cariotipo primitivo para el género *Peromyscus*, examinando los cariotipos de miembros de los géneros *Baiomys* y *Neotomodon*. En el primer recuadro se ofrece un resumen de dicho análisis:

En realidad, el análisis de Yates, Baker y Barnett (1979) es mucho más detallado y completo de lo que aquí se ha mostrado, pero el interés de mencionar este ejemplo es demostrar que los datos cariológicos son susceptibles de análisis filogenético, similar al que usualmente se realiza con datos de morfología externa y consecuentes con las ideas de Holomorfofilia y análisis filogenético de Hennig (1968) y de seguidores como Wiley (1981).

Actualmente se sabe que la especiación puede estar acompañada por rearrreglos cromosómicos en una de las dos especies involucradas (White, 1978). Los rearrreglos pueden promover o asegurar la independencia de los linajes y producir la inviabilidad o la esterilidad diferencial de los individuos híbridos entre los tipos cromosómicos, pero también pueden permanecer como polimorfismos o fijarse en uno o más linajes como un subproducto del proceso de especiación, más que ser su principal causa (Hall y Selander, 1973 y Wiley, 1981).

De esta manera los datos cromosómicos ofrecen al sistemata un conjunto de caracteres que le permiten analizar y sentar bases muy importantes para la definición de especies, el proceso de especiación y el análisis filogenético para grupos en los que la similitud morfológica clásica hace extremadamente difícil o imposible el análisis. Por esta razón, en muchos grupos (especialmente en plantas) los datos cromosómicos se han convertido en una parte de los datos básicos que deben ser mencionados en un trabajo sistemático estándar, aunque las clases de datos cromosómicos dependen del nivel de análisis realizado en el grupo y de la factibilidad de que a los cromosomas puedan aplicárseles las técnicas disponibles.

Por otra parte, las relaciones filogenéticas reconstruidas en diversos grupos —sobre la base de datos de rearrreglos cromosómicos—, descansan en la suposición de que el origen de cada nuevo tipo de rearrreglo es un evento único en la evolución de cada linaje (Farris, 1978 y Wiley, 1981), tanto en el caso de los cromosomas politénicos como en el caso de los cromosomas típicos. Sin embargo, se ha demostrado que la premisa de unicidad por sí sola es



insuficiente para determinar la reconstrucción de un árbol filogenético, ya que no siempre da por resultado un árbol filogenético único, en virtud de que la inversión, como cualquier otro carácter, debe pasar a través de una etapa de polimorfismo (Farris, 1978).

Por ejemplo:

Supóngase que cualquier sitio de inversión de un cromosoma individual puede mostrar sólo uno de los dos tipos de inversión alternativos, A ó D, de los cuales A es el tipo ancestral y D el derivado. La evolución de tal sitio ocurre cuando un cromosoma que muestra el nuevo tipo D surge en un individuo perteneciente a una población original homocigótica para A. La población tiene entonces una condición heterocigótica llamada H. Una vez en esta condición, una población puede pasar a la fijación del sitio considerado en cualquiera de dos formas: volviéndose una población homocigótica para el tipo de inversión A o para el tipo D. El postulado de unicidad requiere que se consideren solamente aquellos árboles filogenéticos en que se postule sólo una transición desde la condición homocigótica A hasta la condición heterocigótica H para cualquier sitio. Las transiciones desde la condición H hasta cualquiera de las condiciones A ó D no involucra el origen de ningún tipo nuevo de inversión y, así, no son cubiertas por el postulado de unicidad. Puesto que son poblaciones y no individuos los que figuran en las ramas de un árbol filogenético, la inclusión de la condición H en el modelo biológico del problema es necesaria por razones biológicas. El hecho es que el tipo de inversión ancestral A puede ser "salvado" en una población heterocigótica y reemerger en alguna población descendiente en una condición homocigótica; esto hace posible que un árbol filogenético de cualquier forma sea siempre congruente con el postulado de unicidad. Por ejemplo, todos los árboles de la figura siguiente son consistentes con los datos hipotéticos correspondientes, dadas las poblaciones ancestrales heterocigóticas adecuadamente (como es necesario). Así, si la inversión apomórfica mostrará o no relaciones de grupo hermano entre X e Y depende de ciertas probabilidades que son semejantes en sus propiedades a aquéllas de los datos morfológicos (Farris, 1978).

Se ha sugerido que los árboles filogenéticos que disminuyen el número de transiciones desde la condición heterocigótica hasta las condiciones derivadas son la mejor solución al problema bajo el principio metodológico de parsimonia (Farris, 1978). Afortunadamente la mayoría de los investigadores parecen haber seguido esta pauta con los datos de inversiones cromosómicas,

aun cuando sus trabajos no estaban inscritos en la Sistemática filogenética *sensu* Hennig. Esto significa que:

- a) Los datos cromosómicos se manejan en forma similar a otros caracteres.
- b) Se puede tener cierta confianza en que las hipótesis filogenéticas publicadas previamente basadas en datos cromosómicos se correlacionan fuertemente con el análisis filogenético de otros caracteres.
- c) Cuando la similitud morfológica evita o impide el análisis filogenético en este nivel, pueden usarse los datos cromosómicos para producir hipótesis genealógicas y sentar las bases de clasificaciones filogenéticas (Farris, 1978 y Wiley, 1981).

Uno de los problemas no resueltos es la evaluación y comparación de las tasas de evolución cromosómica (divergencia evolutiva-relaciones patristicas). En algunos grupos, tales como las tortugas, la evolución cromosómica parece haber procedido muy lentamente, mientras que en otros, especialmente los mamíferos, la tasa de evolución ha sido mucho más rápida (Duellman, 1985). Las tasas son altamente variables aun dentro de algunos grupos de mamíferos; por ejemplo, se ha encontrado sólo un rearreglo cromosómico en seis géneros de murciélagos glosófagos, pero se han reconocido más de 100 rearreglos en otros seis géneros (Haiduk y Baker, 1982). Al parecer, la evolución cromosómica no procede de manera regular en el tiempo; como consecuencia de ello, la correlación entre la distancia cromosómica y la distancia taxonómica no es universal (Reig, 1983), tal como ocurre generalmente con los caracteres morfológicos, en donde una evaluación fenética —por similitud total— no coincide con una evaluación genealógica, que es lo que requerimos. No obstante, esto no quiere decir que las diferencias en las tasas de cambio no nos permita la reconstrucción de relaciones genealógicas, sino más bien no nos permite inferir tiempos de divergencia entre los taxa.

Por otra parte, como en cualquier carácter de origen diploide, es necesario enfatizar que los datos cariológicos no permiten establecer límites y relaciones entre especies de taxa que presentan el fenómeno de la uniformidad cromosómica intra-taxón; el polimorfismo cromosómico poblacional también representa una limitación de los datos cariotípicos (Reig, 1983).

Otro problema es la localización de los genes en los cromosomas (sólo aceptablemente establecida en pocas especies, como en algunas de *Drosophila*); los experimentos de hibridación, como los que se han realizado con ranas (Wright *et al.*, 1983) están empezando a proporcionar datos sobre marcadores genéticos que permitirán la localización de los *loci* para ciertos genes en otros grupos. De cualquier modo este aspecto no está dentro de los objetivos de la Cariotaxonomía.

PROTEINAS

Las proteínas son macromoléculas poliméricas formadas por residuos de la hidrólisis de aminoácidos (los mismos 20 aminoácidos se encuentran en todos los organismos). Las proteínas pueden estar compuestas por uno, dos o más polipéptidos; cada polipéptido es codificado por un gen. La secuencia de aminoácidos en un polipéptido (estructura primaria) está determinada por la secuencia de nucleótidos en el ADN que lo codifica; el número de diferencias entre dos polipéptidos o proteínas homólogas frecuentemente refleja el número de diferencias entre sus genes correspondientes. La diferenciación genética entre las especies y, por lo tanto, sus probables relaciones filogenéticas, pueden inferirse a partir del grado de diferenciación en la estructura primaria de sus proteínas (Dobzhansky *et al.*, 1977).

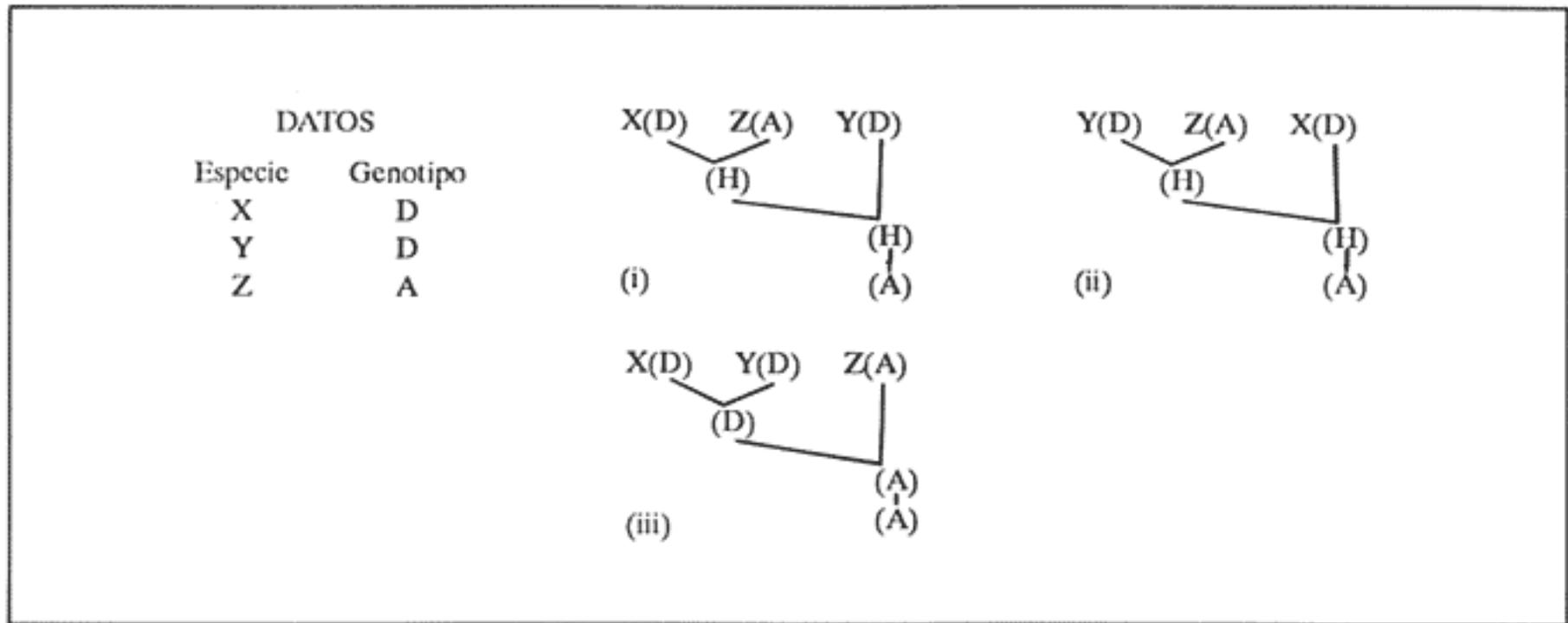
Las estimaciones del grado de similitud entre proteínas específicas de diferentes taxa, pueden hacerse mediante la comparación de las composiciones y secuencias de aminoácidos de las proteínas seleccionadas para el análisis. *Grosso modo* hay dos formas básicas de examinar los cambios en dicha composición: 1. La aproximación directa mediante la determinación de la

VARIOS TRABAJOS PREVIOS sugerían que la tendencia evolutiva en **Peromyscus** era hacia el incremento de la cantidad de heterocromatina y del número de cromosomas con dos brazos (mediante inversiones pericéntricas y adiciones de brazos cortos heterocromáticos); si esto era correcto, entonces, tanto **Peromyscus** como el grupo hermano de **Peromyscus** deberían haber tenido un cariotipo primitivo con todos (o casi todos) los cromosomas de tipo acrocéntrico y sólo heterocromatina centromérica. Para probar esta predicción era necesario realizar comparaciones de grupo externo examinando los cariotipos de otros géneros que se supusieran estrechamente relacionados con **Peromyscus**. Si los patrones de bandeo de los cromosomas acrocéntricos de las especies de estos géneros difirieran de sus contrapartes correspondientes de dos brazos en **Peromyscus** por una inversión pericéntrica y sólo exhibieran heterocromatina en las regiones centroméricas, estas serían fuertes evidencias de un cariotipo esencialmente acrocéntrico primitivo para **Peromyscus**.

Se seleccionaron para las comparaciones a **Neotomodon**, por ser considerado el pariente más cercano a **Peromyscus**, y a **Baiomys**, como un pariente más lejano; ambos géneros tienen un número cromosómico diploide de 48, al igual que **Peromyscus**. Se encontró que el cariotipo de **Baiomys taylori** era muy similar al cariotipo de **Peromyscus critinus**, que previamente había sido sugerido como el más primitivo para el género **Peromyscus**; ambos cariotipos diferían sólo por inversiones pericéntricas en los pares 1, 22 y 23, con dos brazos en **Peromyscus** y acrocéntricos en **Bayomis**. El alto grado de similitud entre los cariotipos podía explicarse por:

- a) una estrecha relación filogenética entre estas especies. (Pero esta explicación fue rechazada porque varias evidencias morfológicas la contradecían).
- b) Convergencia. Esta explicación también se rechazó porque requería eventos convergentes independientes (inversiones pericéntricas) en tres pares de cromosomas homólogos.
- c) Simplesiomorfías. Los autores interpretaron la similitud cariotípica de esta manera. Si las características cariotípicas compartidas por **P. critinus** y **B. taylori** eran simplesiomórficas, entonces se podía sospechar que estos elementos estaban presentes en el cariotipo primitivo del género **Peromyscus**; previamente se había sugerido que el cariotipo primitivo del género estaba constituido por cromosomas de dos brazos en los pares 1, 22 y 23 y que los demás cromosomas debían haber sido acrocéntricos, con excepción de los pares 2, 3, 6 y 9, cuya morfología no podía ser sujeta a hipótesis todavía. Los cromosomas en los pares 2, 3, 6 y 9 son acrocéntricos en **B. taylori**, lo cual sugirió que estos pares también eran acrocéntricos en el linaje pre-**Peromyscus**. Lo anterior a su vez implicaba que algunas especies de **Peromyscus** que presentaban cromosomas con dos brazos en los pares 2, 3, 6 y 9 poseían este carácter como una sinapomorfía, [ver abajo]. Por otra parte, todos los cromosomas de **B. taylori** (48) fueron acrocéntricos, y sólo presentaron heterocromatina centromérica; ésto apoya la hipótesis de que el cariotipo primitivo del género **Peromyscus** consistía principalmente de cromosomas acrocéntricos, con los pares 22 y 23 con dos brazos y el par 1 de morfología desconocida; además, dicho cariotipo debió contener sólo heterocromatina centromérica.

En cuanto a **Neotomodon**, los especímenes examinados de **N. alstoni** ($2n=48$) tuvieron 30 cromosomas acrocéntricos y 16 con dos brazos, con sólo heterocromatina centromérica; sin embargo, se encontraron inversiones pericéntricas sinapomórficas en los pares cromosómicos 2, 3, 6 y 9 entre **N. alstoni** y **P. melanotis**, **P. maniculatus**, **P. polionotus**, **P. gossypinus** y **P. floridanus**, lo que indicaba que estas especies de **Peromyscus** y **N. alstoni** compartieron un ancestro común después de las inversiones de los pares de cromosomas 2, 3, 6 y 9. Con estas evidencias dichos autores concluyeron que **Neotomodon** y **Peromyscus** son congénéricos.



secuencia de aminoácidos de la misma y 2. La aproximación indirecta mediante la inmunología. Una tercera forma de análisis es por electroforesis, que puede considerarse una aproximación indirecta.

Secuencia de Aminoácidos de las Proteínas

Un resumen ilustrado del procedimiento común para establecer la secuencia de aminoácidos de un polipéptido o una proteína, puede hallarse en Dobzhansky *et al.* (1977). La primera proteína de la cual se obtuvo la secuencia de aminoácidos o estructura primaria fue la insulina, que consta de 51 aminoácidos. Durante los primeros años de la década de 1950 se demostró que las secuencias de aminoácidos de las insulinas de vaca, cerdo, oveja, caballo y ballena de esperma (cachalote), son idénticas, excepto por reemplazamientos en tres aminoácidos consecutivos. Los procedimientos para establecer la secuencia primaria de las proteínas son extremadamente laboriosos; a pesar de esto, más de 500 secuencias completas o parciales se conocían ya en 1972, debido a los esfuerzos de muchos investigadores en diversos laboratorios, y muchas secuencias más se han determinado año con año (Dobzhansky, *et al.*, *op. cit.*).

El grado de diferencia o distancia entre cualquier par de proteínas homólogas de secuencia conocida puede medirse en dos formas: la primera es simplemente contar el número de diferencias de aminoácidos entre dos secuencias dadas, y la segunda, que da más información que la primera, hace uso del código genético. El reemplazamiento de un aminoácido por otro puede requerir una, dos o tres sustituciones nucleotídicas en el triplete de ADN correspondiente. Por ejemplo, en la posición 66 el citocromo c del hombre tiene isoleucina, mientras que en el mono rhesus hay treonina en tal posición; el número mínimo de diferencias nucleotídicas entre los codones que codifican estos dos aminoácidos es uno. Por otro lado, en la posición 20 el hombre tiene metionina y el caballo tiene glutamina; los codones correspondientes difieren por no menos de dos nucleótidos.

El número mínimo de diferencias de nucleótidos se determina para cada reemplazamiento de aminoácidos entre dos proteínas y todas esas diferencias se suman. Esto da el número mínimo de cambios de nucleótidos que deben haber ocurrido en la evolución por separado de los dos "genotipos" que codifican las proteínas comparadas (llamado también distancia de mutación). La distancia de mutación entre dos genes frecuentemente es mayor que el número de diferencias de aminoácidos entre las proteínas correspondientes; además, el número efectivo de diferencias nucleotídicas entre dos genes puede ser mayor que su distancia de mutación, puesto que ésta mide sólo el número mínimo de diferencias nucleotídicas.

Un ejemplo clásico de la proposición de filogenias mediante el uso de las secuencias de aminoácidos en las proteínas, es la comparación del citocromo c realizada para 20 especies por Fitch y Margoliash (1967). Las secuencias de aminoácidos originales empleadas por Fitch y Margoliash pueden ser utilizadas para aplicar el método habitual de inferencia filogenética citado por Wiley (1981); este autor ha ejemplificado tal uso: se parte de las secuencias de

aminoácidos de las posiciones 9 a 20 de los citocromos c del hombre, del mono rhesus y del caballo, y se incluyen grupos externos compuestos de "todos los otros mamíferos" y "todos los otros metazoarios". Los citocromos c son idénticos excepto para las posiciones 19 y 20; aquí, los humanos (*Homo sapiens*) y los monos rhesus (*Macaca mulata*) comparten isoleucina y metionina, mientras que los caballos (*Equus*) y los grupos externos comparten valina y glutamina. Es más aceptable (aplicando el principio de parsimonia) considerar la presencia de isoleucina y metionina (en las posiciones 19 y 20, respectivamente) como sinapomórfica para los humanos y los monos rhesus (Primates), que considerar la presencia de valina o glutamina como sinapomórfica, debido a que el taxón Primates como grupo monofilético es más congruente con otros muchos datos que un grupo monofilético compuesto de caballos, ballenas grises y canguros (entre otros mamíferos). Puede obtenerse una corroboración adicional a esta conclusión con otros grupos externos: 4 especies de aves, una tortuga, una rana, un toro y un atún tienen todos valina y glutamina en las posiciones 19 y 20, respectivamente; por lo cual se puede concluir que se trata de dos caracteres simplesiomórficos, de acuerdo al criterio de posición filogenética basado en la regla de grupo externo (ver Hennig, 1968; Wiley, 1981; Llorente, 1986).

Sin embargo, uno de los principales problemas de trabajar manualmente las relaciones filogenéticas de los taxa (cuando éstas son evidenciadas por las secuencias de las proteínas) es el enorme número de comparaciones que uno tendría que hacer; considérese que en el ejemplo anterior solamente el citocromo c tiene de 103 a 112 aminoácidos en su cadena. Por esta razón, los llamados "sistematas bioquímicos" han desarrollado algoritmos de computadora, construidos sobre el principio de parsimonia máxima, que agrupan a los taxa formando dendrogramas a partir de matrices de distancias genéticas entre ellos (o distancias de mutación). Tales dendrogramas son comúnmente interpretados como una representación de las relaciones filogenéticas de las especies; y, de hecho, estos algoritmos se aplican también a otros datos, como los que proceden de las técnicas inmunológicas o de las técnicas de hibridación de ADN-ADN (nuclear o mitocondrial). Ver recuadro 2.

Dobzhansky *et al.* (1977) señalaron que los cambios evolutivos en las proteínas pueden comprender no sólo reemplazamientos de aminoácidos, sino también adiciones o deleciones de uno o más aminoácidos en distintas posiciones; las proteínas homólogas pueden diferir, por tanto, no sólo en sus secuencias de aminoácidos, sino también en la longitud de sus polipéptidos. Las cadenas alfa y beta de la hemoglobina humana son homólogas (codificadas por genes que surgieron por duplicación ancestral); con todo, la cadena alfa consta de 141 aminoácidos, mientras que la cadena beta de 146. Para determinar el grado de similitud entre proteínas homólogas debe tomarse en cuenta la posible existencia de adiciones y deleciones. La similitud entre las cadenas alfa y beta de la hemoglobina puede aumentarse si las secuencias se alinean sobre 148 posiciones y se supone que existen discontinuidades en las posiciones 2, 48, y 56 a 60 de la cadena alfa, y en las posiciones 19 y 20 de la cadena beta; así, cada discontinuidad puede asignarse a una deleción en una cadena o a una adición en la otra.

La presencia de adiciones y deleciones de aminoácidos en proteínas homólogas origina un problema: una proteína típica consta de aproximadamente 100 o más aminoácidos; es probable que cada uno de los 20 aminoácidos comunes se encuentre varias veces en cualquier proteína dada. Si se toma la libertad de poner discontinuidades en cualquier posición, seguramente cualquier par de proteínas tendrá aminoácidos idénticos en varios sitios, ya sea que las proteínas sean homólogas o no (similitud estructural). Se han ideado varios métodos para decidir si las similitudes en la composición de aminoácidos entre dos proteínas se deben a coincidencia accidental o a homología. La demostración de que dos proteínas están relacionadas había sido intentada, hasta 1970, usando dos criterios diferentes: 1. Mostrar que sus estructuras y funciones bioquímicas eran muy similares y 2. Mostrar que las estructuras genéticas subyacentes a las proteínas son más similares que lo que podría esperarse por azar (Fitch, 1970). Sin embargo, dos proteínas pueden parecer muy similares porque descienden —con divergencia— de un gen ancestral común (es decir, son homólogas) o porque descienden —con convergencia— de genes ancestrales separados (es decir, son análogas); las restricciones impuestas por una aptitud funcional pueden causar la suficiente convergencia como para producir una aparente relación genética. Por lo tanto, la demostración de que dos secuencias actuales son significativamente similares, ya sea por criterios químicos o genéticos, deja aún sin decisión la cuestión de si su similitud se debe a un proceso convergente o es lo que queda de un proceso divergente (Fitch, *op. cit.*).

Fitch (1970) imaginó un método mediante el cual es posible determinar si dos proteínas comparadas tienen un ancestro común o son de origen independiente; tomó como ejemplo al citocromo c de los hongos y de los metazoarios, planteando el problema de la siguiente manera: para demostrar que los citocromos c de los hongos y de los metazoos son en realidad homólogos, basta con sólo demostrar que las secuencias de los aminoácidos de los citocromos c ancestrales de ambos grupos eran más semejantes entre sí que lo que son los citocromos c de los representantes actuales de estos dos grupos (demostrar divergencia por similitud genética ancestral a partir de un ancestro común). El razonamiento en el método de Fitch es el siguiente:

1. Las secuencias de aminoácidos de las proteínas ancestrales de las actualmente comparadas pueden ser reconstruidas (es decir, determinadas con bastante aproximación) por un método previamente descrito por Fitch y Margoliash (1967). Este método supone evolución divergente.
2. Una vez que se lleva a cabo el proceso de reconstrucción de las secuencias polipeptídicas ancestrales, se puede calcular la probabilidad de que una posición nucleotídica elegida al azar en una secuencia reconstruida, sea de tipo convergente, si se supone que las secuencias de partida no tienen ninguna relación entre sí.
3. Esta probabilidad multiplicada por el número total de posiciones nucleotídicas a ser examinadas, es el número de posiciones que uno esperaría fuesen de tipo convergente en la secuencia reconstruida.
4. Posteriormente se reconstruyen las secuencias ancestrales y se obtiene su distancia de mutación.
5. Si se encuentra un exceso —estadísticamente significativo— de posiciones de tipo convergente en las secuencias reconstruidas respecto a lo predicho (distancia de mutación observada entre las secuencias reconstruidas mayor de lo esperado), se puede concluir que los dos grupos de genes son probablemente de origen independiente.
6. Si, por otro lado, se encuentra un exceso estadísticamente significativo, de posiciones de tipo divergente en las secuencias reconstruidas respecto a lo predicho (siendo la distancia de mutación entre éstas menor que lo esperado), se puede concluir que los dos grupos de genes tuvieron probablemente un ancestro común.

Mediante este método, Fitch (1970) demostró que un conjunto de 16 secuencias de aminoácidos elaboradas al azar, no estaban relacionadas, y con

otro conjunto de 16 proteínas reales, pero presuntamente no relacionadas, obtuvo un resultado similar. Asimismo, el método mostró correctamente una relación de convergencia en un conjunto de 24 proteínas modelo que estaba compuesto de dos grupos de evolución independiente que convergieron hacia la misma estructura química y, finalmente, mostró correctamente una relación de divergencia en un conjunto de 24 citocromos c compuesto de 5 hongos y 19 metazoarios. Por otra parte, Fitch (*op. cit.*) reconoció dos tipos de genes homólogos: 1. Genes ortólogos, que son los descendientes de un gen ancestral presente en el último ancestro común de las especies que poseen dichos genes (y que comienzan su evolución separada en el momento en que las especies correspondientes tienen su último ancestro común) y que, por tanto, su evolución refleja la evolución de las especies en las cuales se hallan; y 2. Genes parálogos, que son los genes homólogos resultantes de la duplicación de un gen ancestral (y que evolucionan independientemente unos de los otros desde el momento en que la duplicación tiene lugar). Estos genes evolucionan independientemente durante la historia de la especiación del grupo examinado. Las relaciones filogenéticas entre las especies se deben determinar comparando las estructuras primarias de proteínas ortólogas, no parálogas (Dobzhansky *et al.*, 1977; Wiley, 1981), y pueden hacerse inferencias erróneas si los genes o proteínas que se piensa que son ortólogos en realidad son parálogos. Una revisión de varios métodos propuestos para distinguir proteínas homólogas de análogas puede hallarse en Fitch (1973).

Cuanto se ha expuesto aquí acerca de la comparación de las secuencias de aminoácidos de las proteínas, puede resumirse de la siguiente manera: si las homologías entre las proteínas son determinadas correctamente y se usan algoritmos de matrices adecuados para el análisis de los datos resultantes de las secuencias, tales comparaciones son una fuente de datos filogenéticos, válida y muy útil (Wiley, 1981). Otra característica de estas comparaciones es que requieren de un procedimiento muy laborioso y costoso para su investigación, por lo que el número de secuencias conocidas es relativamente limitado (Dobzhansky *et al.*, 1977).

La filogenia de Fitch y Margoliash (1967) basada en las secuencias del citocromo c, aunque corresponde más o menos bien con la filogenia de los organismos determinada a partir del registro fósil y de otras fuentes, también presenta algunas discordancias (muy pocas). Sin embargo, a pesar de estas relaciones "erróneas" (en función de la evidencia más apoyada y reconocida), es notable que el estudio de una sola proteína represente con mediana exactitud la filogenia de una muestra muy diversa de 20 organismos, que son muy divergentes desde hace mucho tiempo para otros caracteres.

Los citocromos c son proteínas de evolución muy lenta, y su conservacionismo hace posible establecer el grado de relación entre organismos emparentados lejanamente, aunque también hace imposible el establecer relaciones entre organismos estrechamente emparentados. Afortunadamente hay proteínas de tasa evolutiva diferente y las relaciones entre organismos muy emparentados puede establecerse estudiando proteínas de evolución rápida, como las anhidrasas carbónicas y los fibrinopéptidos en los mamíferos (Dobzhansky *et al.*, 1977). Una idea de la diversidad de proteínas que habían sido secuenciadas en animales y plantas hasta 1973, puede obtenerse revisando los trabajos presentados en el Symposium "Evolutionary significance of proteins" realizado en Boulder, Colorado. Varias de las presentaciones correspondientes aparecieron en *Systematic Zoology* en 1974 (4): 507-608.

Convendría examinar la secuenciación de proteínas de evolución más rápida en grupos vivientes muy diversificados, cuyos caracteres morfológicos y filogenia se conocen bien, con el fin de efectuar comparaciones entre los métodos referidos, analizando tanto las genealogías obtenidas por ambos métodos como las distancias "genéticas" y "fenéticas". Si de este grupo supuestamente "modelo" se tuvieran fósiles y un conocimiento adecuado de su biogeografía histórica, sería muy interesante analizar de modo integral los distintos componentes de su historia evolutiva (relaciones cladísticas, patristicas, cronísticas y espaciales). Los resultados de todo esto serían significativos para estudios en Especiación, Clasificación y Biogeografía. No obstante, con el advenimiento de las técnicas de secuenciación de ácidos nucleicos, la de proteínas ha venido cayendo en desuso, sobretodo por resultar ahora totalmente impráctica para estos fines.

Immunología

Siguiendo a Dobzhansky *et al.* (1977), la comparación de las proteínas usando las técnicas inmunológicas se lleva a cabo de la siguiente manera: una proteína, por ejemplo la albúmina, se purifica a partir de un animal, que puede ser el hombre; la proteína purificada se inyecta en un mamífero como el conejo, lo que ocasiona que éste produzca una reacción inmunológica y elabore anticuerpos contra la proteína extraña (antígeno). Los anticuerpos producidos por el conejo inmunizado reaccionarán por tanto no sólo contra el antígeno específico usado (la albúmina humana en el ejemplo), sino también contra otras proteínas relacionadas (tales como las albúminas de otros primates). Entre mayor sea la similitud entre la proteína usada para inmunizar al conejo y la proteína examinada, mayor será la extensión de la reacción inmunológica. El grado de disimilitud entre la albúmina del hombre (o cualquier proteína usada en la inmunización original) y las albúminas de diferentes especies, se expresa como "distancia inmunológica".

Un método inmunológico sensible y eficiente que requiere sólo pequeñas cantidades de proteína purificada, es el de la fijación de microcomplemento. El "complemento" es una serie de proteínas que actúan secuencialmente y que se halla en el suero de los vertebrados. Cuando el complemento se añade a los antígenos y anticuerpos, bajo condiciones experimentales adecuadas, se fija dentro de la red tridimensional de los complejos antígeno-anticuerpo.

La extensión de la reacción antígeno-anticuerpo se mide determinando la cantidad de complemento fijada en la reacción. Se adicionan glóbulos rojos adecuadamente preparados ("sensibilizados") y cualquier complemento no fijado en los complejos antígeno-anticuerpo queda disponible para lisar las células. La cantidad de células lisadas se mide espectrofotométricamente. El número de glóbulos rojos lisados es proporcional a la cantidad de complemento libre. Entre mayor sea la cantidad de células lisadas menor será la extensión de la reacción antígeno- anticuerpo.

La fijación de microcomplemento empleando varias proteínas ha sido usada para determinar distancias inmunológicas en una variedad de organismos, incluyendo anfibios, reptiles, aves y mamíferos (Sarich y Wilson, 1967; Goodman y Moore, 1971; Wallace, King y Wilson, 1973; Maxson y Wilson, 1975; Wilson, Carlson y White, 1977; Gorman, Buth y Wyles, 1980). Dobzhansky *et al.* (1977) ilustraron algunos ejemplos del uso de esta técnica.

Como hay una buena correlación entre las distancias inmunológicas y el número de diferencias de aminoácidos entre las proteínas ortólogas de dos especies (Champion *et al.* 1974; Wilson, 1975), las técnicas inmunológicas producen una estimación indirecta del número global de sustituciones de aminoácidos que han ocurrido desde que las especies divergieron de su ancestro común; sin embargo, Farris, Kluge y Micevich (1982) sugieren que esto no es así (ver más adelante). Debido a que estos cambios se pueden considerar estocásticos, uno puede construir un árbol filogenético para las especies comparadas (Wiley, 1981).

Las técnicas inmunológicas no llevan por sí mismas a un análisis cladista clásico *sensu* Hennig (Wiley, 1981), ya que no comprenden comparaciones de grupo externo de cada complejo de caracteres (proteínas) para el reconocimiento formal de proteínas ortólogas primitivas (plesiomórficas) y derivadas (apomórficas). No obstante, el análisis puede ser intuitivamente cladístico, en el sentido de que cualquier cambio que ocurra —después de la separación desde un ancestro común— representa innovaciones evolutivas únicas y, así, la similitud total puede ser una medida relativa de similitud sinapomórfica (Wiley, 1981). Sin embargo, deben hacerse dos suposiciones tácitas:

1. Las similitudes no homólogas no son de importancia significativa en el análisis. Esta suposición debe permanecer siempre como tal, puesto que las sustituciones de aminoácidos efectivas no pueden determinarse a partir de los resultados.
2. El algoritmo usado agrupa sobre la base de las similitudes sinapomórficas. Los datos inmunológicos sólo pueden analizarse usando métodos fenéticos de matrices; tema tratado en el recuadro 2.

Farris, Kluge y Micevich (1982) criticaron el uso de los datos inmunológicos para determinar distancias entre los taxa y señalaron un error común en el uso del algoritmo de Fitch y Margoliash (1967). Este algoritmo, para hallar los árboles del mínimo porcentaje de desviación estándar (% DS), es un procedimiento de ensayo y error; un árbol inicial, obtenido por agrupamiento fenético (de acuerdo a la similitud), es sujeto a una serie de pequeños rearrreglos, conservando aquellos cambios que reducen el % DS. Rara vez el árbol de % DS mínimo es obtenido fácilmente y los árboles que se obtienen sin seguir el procedimiento cabalmente no son árboles en el sentido estricto de Fitch y Margoliash (como se menciona en muchos trabajos), pues éstos serán bastante similares a aquéllos producidos por el agrupamiento fenético inicial; además, es bien conocido que tales agrupamientos sólo pueden ser interpretados genealógicamente — si es que se puede — cuando las tasas de divergencia evolutiva son invariables o casi lo son, ya que a menos que se suponga que la tasa de divergencia es uniforme (o casi), no hay razón para pensar que los taxa mutuamente más similares estén también genealógicamente más relacionados. Al parecer, estos errores son comunes en el análisis filogenético a partir de matrices de distancias moleculares; Farris, Kluge y Micevich (*op. cit.*) han señalado que muchos "sistematas inmunológicos" toman como un hecho bien establecido la supuesta naturaleza uniforme de la evolución molecular.

Históricamente, el problema en particular se puede atribuir a Sarich y Wilson (1967), quienes propusieron que la evolución de la albúmina procedía de una manera semejante a un reloj; desde entonces, varios investigadores han insistido en que la evolución de las proteínas procede con una tasa aproximadamente constante y que la comparación de esta tasa con fósiles de edades conocidas y eventos geológicos fechados da una base sólida para datar momentos secuenciales de divergencias de linajes de organismos vivos. Evidentemente, aquellas agrupaciones de taxa mutuamente más similares por agrupamientos fenéticos, produce árboles consistentes con la hipótesis del reloj molecular. Pero este concepto no es un principio *a priori*, sino un descubrimiento empírico, y las evidencias de tal descubrimiento apuntan hacia la misma conclusión: si bien la similitud molecular está generalmente bien correlacionada con la antigüedad de la ancestría común, la correlación ciertamente no es perfecta (Farris, Kluge y Micevich, 1982); esto implica la necesidad de usar un método que permita que las excepciones sean reconocidas como tales en la reconstrucción



filogenética. Mientras el agrupamiento fenético actúa siempre como si la similitud implicara relación y no puede usarse legítimamente para la inferencia genealógica, el método de Fitch-Margoliash (al disminuir el % DS) —si se usa adecuadamente— es capaz de agrupar sin presuponer una homogeneidad de las tasas evolutivas, y en este punto, por lo menos, su uso está bien fundado.

Otra tendencia común de los "sistematas inmunológicos", señalada por Farris, Kluge y Mickevich (1982), es que cuando las evidencias morfológicas entran en conflicto con las evidencias moleculares, se tienden a desechar las primeras; muchos de estos conflictos son aparentes y se deben al mal procesamiento de los datos moleculares. En el futuro debería recordarse que la calibración original del "reloj molecular" se basó necesariamente en los tiempos de divergencia que a su vez se derivaron de las filogenias basadas en la morfología. Por otra parte, los autores citados señalaron que muchos "sistematas inmunológicos" han trabajado sobre la premisa de que las distancias inmunológicas miden las diferencias entre las secuencias de aminoácidos de las proteínas, pero ellos mismos (Farris, Kluge y Mickevich, 1979) señalaron previamente que la correlación entre las distancias inmunológicas y las secuencias de aminoácidos generalmente no podría ser tan estrecha, ya que las últimas son métricas, mientras que las primeras no lo son; esto también puede implicar que las distancias inmunológicas no puedan ser totalmente semejantes a un reloj. Farris y sus colegas reconocieron, en conclusión, que el considerar otras evidencias como irrelevantes o patentemente erróneas a la "luz de la inmunología", tratar las distancias inmunológicas meramente como semejantes a un reloj (a pesar de las evidencias en contra) y llegar a filogenias "indisputables" sin siquiera obtener el árbol de mejor ajuste, son actitudes más bien típicas que excepcionales entre los "sistematas inmunológicos".

Wiley (1981) consideró que las técnicas inmunológicas que usan algoritmos adecuados para el tratamiento de los resultados (p. ej., el de Farris, 1972), son una fuente válida de datos filogenéticos. Tales técnicas son mucho menos laboriosas y costosas que obtener la secuencia de aminoácidos de las proteínas, y son también aplicables a divergencias muy antiguas, a diferencia de las técnicas electroforéticas. No obstante, otro problema detectado en el análisis de los datos de las técnicas inmunológicas, es el de ciertas discrepancias en los resultados obtenidos con sólo una y con diferentes proteínas. Por ejemplo, se ha encontrado en diferentes trabajos que mientras las lisozimas del hombre y el orangután presentan mayor similitud que las lisozimas del hombre y del gorila, las albúminas muestran la relación inversa. Adicionalmente, las distancias inmunológicas entre el hombre y el babuino, cuando se usa suero antihumano, es de 126; mientras que la misma distancia, cuando se usa suero antibabuino es de sólo 66. Pero estas incongruencias también se dan en otros caracteres, incluso los morfológicos, por lo que no son razón suficiente para desechar las técnicas inmunológicas de los estudios filogenéticos (Dobzhansky *et al.*, 1977). Más bien serían una razón para desechar los métodos fenéticos con los que se analizan resultados de cualquier técnica (por sofisticada que sea) en los estudios filogenéticos.

Electroforésis

Otra de las técnicas usadas para estimar las similitudes y diferencias entre las proteínas de los organismos es la electroforesis. Las proteínas que difieren en su carga eléctrica neta pueden separarse mediante técnicas de electroforesis en gel; no todos los cambios de aminoácidos en la estructura primaria pueden detectarse con esta técnica, sino sólo los que determinen diferencias en la carga eléctrica de las proteínas y ocasionalmente, aquéllos que cambien la conformación de las mismas. Por esta razón y también debido a la redundancia del código genético (varios tripletes para algunos aminoácidos), no todas las diferencias alélicas pueden detectarse por la electroforesis de las proteínas correspondientes, subestimándose así la variación genética en un 60 %; sin embargo, las proteínas que se pueden distinguir mediante la electroforesis son siempre producto de diferentes alelos. Un aspecto que es importante anotar, es que la electroforesis queda limitada a proteínas para las cuales existen métodos histoquímicos de purificación o reconocimiento.

Con esta limitación significativa, la electroforesis de las proteínas es un medio de estimar diferenciación genética entre diferentes poblaciones de organismos. La técnica supone a menudo que un número de genes (o

proteínas) pueden ser seleccionados para estudio sin saber (*a priori*) si son variables o cuan variables son; también se supone que una muestra moderada de *loci* representa el genoma de los organismos estudiados. Entonces, las estimaciones de variación genética obtenidas a partir del estudio de unos pocos *loci* (o proteínas) pueden ser extrapoladas a la totalidad del genoma.

La electroforesis en gel para fines taxonómicos es un método medianamente simple de medir la diferenciación genética entre las especies; cuando se estudian varios taxa puede construirse un dendrograma que refleje la diferenciación genética entre ellos. La aplicación original de la electroforesis comprendió estimaciones de variabilidad genética en poblaciones de *Drosophila pseudoobscura* (Hubby y Lewontin, 1966), pero después fue aplicada al estudio de la variación genética entre poblaciones intraespecíficas e interespecíficas. En años recientes, la electroforesis en gel se ha usado para estimar la diferenciación genética entre especies estrechamente relacionadas en varios grupos de organismos, desde plantas anuales, pasando a través de insectos y otros invertebrados, hasta peces, anfibios, reptiles, aves y mamíferos (Dobzhansky *et al.*, 1977). Estos estudios han dado mucha información relevante y en muchos casos crítica, acerca de las relaciones filogenéticas (cladísticas y patristicas).

Los datos obtenidos por electroforesis son frecuencias génicas y genotípicas para los alelos detectables electroforéticamente. Las frecuencias génicas y genotípicas observadas en dos especies diferentes se transforman en medidas de distancia genética, usando cualquier método estadístico seleccionado entre cierta variedad. Dos estadísticos usados ampliamente son el de "distancia genética", *D*, y el de "identidad genética", *I*, desarrollados por Nei (1972). La distancia genética puede estimarse también empleando el método desarrollado por Rogers (1972). Recientemente, los métodos de Nei y de Rogers han sufrido algunas modificaciones (ver Duellman, 1985).

A partir de los valores de distancias genéticas de Nei o de Rogers pueden construirse "árboles filogenéticos" (fenogramas); con tales valores se forma una matriz que integra las comparaciones entre los pares de taxa examinados; ésta puede procesarse por medio de distintos algoritmos. Los algoritmos que no hacen implícita la suposición de una uniformidad de las tasas evolutivas son preferibles para la obtención de filogenias (Farris, 1979; Wiley, 1981; ver el recuadro de algoritmos en este artículo).

Los datos electroforéticos también son susceptibles de tratarse de una manera "cladista clásica" (criterio de posición filogenética por medio de la regla del grupo externo), como ha sido delineada por Wiley (1981). Por ejemplo, Mickevich y Johnson (1976) analizaron datos electroforéticos codificando los alelos de cada gen como ausentes o presentes en cada población y calcularon después el árbol de distancia mínima de Wagner, en vez de usar métodos matriciales. Baverstock *et al.* (1979) también consideraron la presencia o ausencia de los alelos de cada gen en las diferentes poblaciones y usaron además grupos externos, lo que les permitió determinar la apomorfia o plesiomorfia relativas de los alelos; usaron pocos datos, pero su análisis es interesante. De esta manera, los datos electroforéticos no están restringidos en su análisis a los métodos matriciales.

Los estudios electroforéticos son particularmente útiles para establecer filogenias de taxa que exhiben poca o ninguna diferenciación morfológica (grupos de especies crípticas, especies de un género o géneros muy relacionados) y en general son más sensibles para comparaciones de especies de divergencia reciente. Sin embargo, no son efectivos para comparar organismos que están evolutivamente muy separados, porque es muy probable que sean diferentes electroforéticamente en todos o la mayoría de los *loci* (Ayala y Kiger, 1984). Las técnicas son relativamente "baratas" y simples en comparación con otras técnicas de estudio de las proteínas (Dobzhansky *et al.*, 1977).

Siempre se debe tener en cuenta que en las técnicas electroforéticas, dos formas de una proteína se juzgan iguales si tienen migraciones idénticas (aunque podrían no serlo); mientras que siempre que dos proteínas migran a diferentes velocidades se sabe que son estructuralmente diferentes, aunque no se sabe si difieren en uno, dos o más reemplazamientos de aminoácidos. Las restricciones y problemas a los que están sujetos para su análisis los datos electroforéticos en la reconstrucción filogenética fueron resumidos por Wiley

LOS ALGORITMOS SON procedimientos diseñados para realizar un método; tales diseños de procedimientos pueden hacerse para ser procesados por computadora. Existen varios algoritmos concordantes con las ideas filogenéticas; se agrupan bajo dos enfoques: algoritmos de parsimonia y algoritmos de compatibilidad (Wiley, 1980). Existen también otros algoritmos no concordantes con ideas filogenéticas, que son algoritmos de acuerdo a la idea de similitud total de los fenetistas, quienes siguen la hipótesis de la no especificidad (Ehrlich y Ehrlich, 1967).

Uno de los algoritmos más conocidos y empleados por los sistematas bioquímicos es el desarrollado por Fitch y Margoliash (1967). Fue mediante ese algoritmo que estos autores realizaron la comparación de los citocromos c de las 20 especies aludidas anteriormente. Gracias al algoritmo lograron analizar completamente las secuencias de aminoácidos de las 20 especies y proponer un árbol filogenético para las mismas. Un algoritmo de aceptación general entre los cladistas es el de Farris (1972). Sin embargo existen muchos algoritmos diferentes, tanto de orientación feneticista (similitud total) como cladista (relaciones de grupo hermano por sinapomorfias). Considerando que los algoritmos de matrices se usan para procesar y analizar los datos que producen las diversas técnicas discutidas en este trabajo, este aspecto requiere especial atención. Fitch (1984) publicó una serie de consideraciones sobre los problemas, peligros y perspectivas potenciales de los algoritmos de matrices; a continuación se resumen algunos de los puntos más importantes de su trabajo.

Fitch mostró que la construcción de un árbol filogenético sobre la única base del principio de parsimonia (como en el ejemplo del citocromo c) es un procedimiento superior a cualquiera basado en matrices, ya que se acerca más a los cambios históricos ocurridos en la evolución de los taxa. Esto se debe a que se pierde una parte de la información contenida en la matriz carácter vs taxón, cuando ésta se simplifica por alguna clase de proceso algebraico promediador del conjunto de números que contiene, al seguir un método feneticista. Pero, si bien la parsimonia es mucho mejor, las técnicas de matrices se siguen usando porque algunos datos sólo se obtienen de esa manera, como en el caso de las distancias inmunológicas y los datos de las técnicas de hibridación de ADN's; o bien, como ya se ha mencionado, porque sería muy tardado analizarlos a mano, como en el caso de las secuencias de aminoácidos.

Fitch (1984) señaló también las causas principales por las que los algoritmos fallan o pueden fallar: una de éstas es la ocurrencia de fenómenos homoplásticos (paralelismos, convergencias y reversiones); cuando esto se presenta, causa problemas al usar los algoritmos, ya que la mayoría de éstos tienen como característica alguna clase de suposición implícita de una tasa de evolución uniforme y progresiva (el método de distancia de Wagner desarrollado por Farris en 1972, es una excepción); de

este modo, tienden a agrupar como taxa hermanos a aquéllos cuyas diferencias son menores que las que hay entre otros pares de taxa, esto es, por cantidad de divergencia. En consecuencia, este error ocurre con mayor probabilidad en los casos en que se tienen tasas evolutivas con desigualdades sustanciales; ésta es la razón por la que el análisis cladista —cuando se puede hacer—, da con mayor probabilidad un resultado más satisfactorio. Fitch demostró que los métodos feneticistas (por ejemplo el UPGMA) dan por resultado una topología errónea de los árboles cuando los datos incluyen paralelismos u homoplasias en los mismos casos en que otros métodos como el suyo propio (es decir, el de Fitch y Margoliash, 1967) y el de Farris (1972) resultan en la topología "correcta").

Pero en el método mismo de Fitch y Margoliash (1967), aunque no se presupone explícitamente una uniformidad de las tasas evolutivas, la elección de qué uniones hacer y el orden en que éstas ocurran son función de las diferencias relativas; ésto equivale o se hace análogo a una suposición implícita de una igualdad relativa de las tasas evolutivas, lo cual conduce en los árboles a errores similares a los que causan las homoplasias que Fitch (1984) discute.

Por otra parte, recuérdese que los algoritmos de matrices se han construido sobre la base del principio de parsimonia máxima, por ello disminuyen el número de sustituciones de aminoácidos necesarias para explicar los datos de las secuencias; dicho en una forma más general, dados los estadios de carácter de los taxa y la topología del árbol, tratan de hallar el número mínimo de cambios de estadio. Sin embargo, Fitch (1984) también señaló algunos aspectos muy discutibles relativos a la parsimonia máxima: uno de ellos es que la adición de un nuevo carácter al análisis puede cambiar la topología del árbol más parsimonioso; otro es el que la adición de un nuevo taxón al análisis también puede alterar la estructura del árbol más parsimonioso, lo que origina problemas si se quiere añadir un taxón primitivo e hipotético para "enraizar" los diagramas (convirtiéndolos en árboles) o bien definir los estadios primitivos; finalmente, un tercer punto es que el árbol más parsimonioso puede ser inconsistente con el árbol de máxima compatibilidad de caracteres.

Asimismo, Fitch mostró algunos puntos de interés relativos a los métodos de matrices usando ejemplos artificiales. En el primero de ellos, él demuestra que mediante la elección adecuada de los caracteres a combinar y considerar en conjunto, uno puede crear deliberadamente conjuntos de datos que hagan que diferentes métodos produzcan diferentes árboles, usando cada método su propio criterio. Mediante otro ejemplo, Fitch demuestra que en un caso en que sólo hay tres árboles posibles, no enraizados, como solución, cada uno de tres métodos (el de Fitch y Margoliash, 1967; el de Farris, 1972 y el procedimiento de parsimonia de Fitch, 1984) produce un árbol diferente como "el

mejor". Demuestra también en otro ejemplo que los datos pueden ser aditivos (que son supuestamente idóneos para los métodos), llevando al "mejor árbol" con una topología que no es la más parsimoniosa. Una breve discusión de los algoritmos puede consultarse en Wiley (1980); comentarios sobre los resultados del uso de diversos tipos de algoritmos se presentan en el trabajo de Hillis (1987).

Finalmente, Fitch (1984) citó el trabajo de Baum (1984), quien presentó un conjunto de datos para cultivos de *Avena*, cuya historia evolutiva era "suficientemente" conocida a partir de los registros de sus cruces, y los analizó por medio de la mayoría de los métodos generalmente disponibles; ninguno dio la filogenia correcta, a pesar de que sólo consideró un poco más de 100 años durante los que se habían mantenido cruces.

Lo anteriormente expuesto no quiere decir, desde luego, que los métodos de matrices o las técnicas que producen matrices no deban emplearse; sin embargo, sí implica que estos métodos no son una solución universal a los problemas sistemáticos, pues tienen limitaciones y se debe ser cauteloso y muy crítico en su uso. No hay un método que tenga todas las bondades y tampoco uno que no tenga defectos (Fitch, 1984). Pero al construir árboles sobre la base del principio de parsimonia, algunos son mejores que otros (Wiley, 1981) y algunos agrupan con base en similitudes plesiomórficas más que apomórficas y producen diagramas de ramificación no genealógicos (Mickey, 1978).

La congruencia (similitud y diferencia en las agrupaciones) entre las clasificaciones de los mismos organismos basadas en diferentes caracteres, mide la estabilidad de una clasificación conforme se consideran nuevas fuentes de información, ésto siguiendo a Mickey (1978). Este autor comparó la estabilidad de 5 métodos feneticistas y 3 métodos filogenético-numéricos, hallando que todos los métodos feneticistas tienen —esencialmente— el mismo comportamiento inestable y que su estabilidad es demostrablemente más sensible a la cantidad de evolución homoplástica que el método de Wagner, que fue el más estable de los métodos filogenéticos. Concluyó que dado que la metodología filogenética es más estable que la agrupación fenética debe usarse la Sistemática Cladista como sistema de referencia general para la Biología.

Puede señalarse, a manera de conclusión acerca de estos aspectos, que la estrecha correspondencia entre los dendrogramas producidos al usar las técnicas que producen matrices de datos y las hipótesis filogenéticas reconocidas, apoyan la idea de que estas técnicas si usan algoritmos de agrupación adecuados (p. ej., el de Fitch y Margoliash, 1967, o el de Farris, 1972), son una fuente importante de datos y de análisis filogenético (Futuyma, 1986); una lista de referencias sobre la materia ha sido dada por Wiley (1980; 1981) y Hillis (1987).

(1981) del modo siguiente: 1. Dos diferentes *loci* genéticos pueden codificar enzimas de movilidad electroforética idéntica; entonces puede presentarse la posibilidad de que entre mayor sea el número de especies con que se trate pero más lejana sea su relación filogenética, la probabilidad de ocurrencia de homologías falsas (homoplasias) pueda ser mayor. Aquí, nuevamente los métodos más directos para determinar homologías y analogías en las proteínas requieren el conocimiento de las secuencias de aminoácidos de las mismas, pero estos métodos generalmente no están disponibles. Por ello, se hace la suposición implícita de que los electromorfos son homólogos. 2. La electroforesis detecta sólo aquellas sustituciones de aminoácidos que alteran la movilidad electroforética de las proteínas y subestima constantemente las diferencias efectivas entre las especies. Así, se debe hacer implícita la suposición de que la probabilidad de tener mutaciones que afecten la movilidad electroforética es igual en cada población y también suponer que estas mutaciones son una muestra al azar de los cambios genéticos totales en la enzima. 3. Las técnicas electroforéticas están restringidas a proteínas solubles en agua, codificadas por genes estructurales.

Kitching (1985) resumió los trabajos previos sobre técnicas de electroforesis de enzimas en Lepidoptera y estudió el caso de los Danainae (Rhopalocera: Nymphalidae), enfatizando el caso de *Danaus plexippus*, la monarca migratoria. Encontró que el uso de tales técnicas es de gran utilidad para casos de relaciones filogenéticas infragenéricas y de muy poco valor para casos de géneros, subtribus y tribus dentro de la subfamilia que estudió. No obstante, según Geiger (1981) y Pashley (1983), citados por Kitching, para Pieridae y Tortricidae respectivamente, la utilización de tales técnicas a nivel supragenérico conduce a patrones congruentes con aquellos derivados a partir de técnicas morfológicas tradicionales. Para una discusión sobre las desventajas y limitaciones de las técnicas electroforéticas en la detección de la variación de enzimas dentro y entre especies es recomendable consultar el trabajo de Johnson (1977).

Wiley (1981) también señaló algunas conclusiones que se han derivado de los estudios electroforéticos, advirtiendo que muchos de estos estudios han sido más bien fenéticos que filogenéticos: 1. Las poblaciones coespecíficas son generalmente más similares electroforéticamente que las poblaciones de diferentes especies. Generalizando esto, se han obtenido intervalos empíricos de diferencia dentro de los que suelen caer las distancias intraespecíficas y las interespecíficas, al usar tanto el coeficiente de Nei como el de Rogers. Por otra parte, aunque hay evidencia de correlación entre la similitud morfológica y la similitud electroforética, algunos estudios han demostrado un desfase pronunciado entre la similitud morfológica y la electroforética. Por esta razón, la electroforesis no es la panacea para tomar decisiones al nivel de especie (Wiley, 1981). 2. Las similitudes y diferencias en los electromorfos no son de un tipo diferente al de otras fuentes de datos para la evaluación de hipótesis filogenéticas. Los datos electroforéticos pueden ser analizados por métodos filogenéticos al igual que los caracteres morfológicos, llevando a la misma filogenia.

Algunos autores han tratado de datar los eventos de especiación con base en los datos electroforéticos (a semejanza de lo que se ha hecho con los datos inmunológicos), calibrando aquéllos con éstos (Wilson, Carlson y White, 1977). La crítica realizada por Farris *et al.* (1982) al mal uso de los datos inmunológicos por algunos "sistematas inmunológicos" es aplicable en este caso también al mal uso de los datos electroforéticos y aún más, ya que éstos se han calibrado con datos inmunológicos, calibrados a su vez con base en la premisa de que el "reloj molecular" es un hecho. Farris (1979) señaló que muchos "sistematas bioquímicos" parecen pensar que los datos moleculares son tan maravillosos que revelan la "verdadera filogenia", sean analizados correctamente o no.

Finalmente, Reig (1983) señaló que las técnicas electroforéticas sólo permiten cuantificar distancias genéticas sobre la base de frecuencias de genes que codifican proteínas, pero que la especiación puede tener lugar con una cantidad muy variable de diferenciación al nivel de dichos genes. Los cambios en las secuencias de ADN que intervienen en la regulación génica parecen ser el mecanismo más importante en la diferenciación de las especies y no son detectados por este método. La especiación es un evento que puede resultar de la diferenciación genética, pero también puede ser independiente de ella;



puede haber gran divergencia, grandes cambios genéticos sin especiación y, también, puede presentarse poca divergencia, muy pocos cambios genéticos y haber especiación. Estos hechos han conducido a plantear finalmente que cualquier tipo de diferencia o similitud debe interpretarse, para efectos de estudios genealógicos, a la luz del conocimiento de los procesos biológicos que intervienen en cada fuente de caracteres (morfológicos, genéticos, bioquímicos, etc.). La cladogénesis puede ser independiente (White, 1978), paralela (gradualismo filético) o causada por procesos anagenéticos (Wiley, 1981; LLorente, 1986).

OTRAS MOLECULAS Y METABOLITOS SECUNDARIOS

La presencia o la ausencia de varias moléculas en grupos de seres vivos ha servido también en Taxonomía; principalmente en la clasificación de las plantas. El uso de metabolitos secundarios (alcaloides, glicósidos, terpenos, flavonoides u otras micromoléculas) ha tenido ventajas, desventajas, potencialidades y limitaciones. Un defecto fundamental de estos enfoques micromoleculares (Takhtajan en Bendz y Santesson, 1974) es el hecho de que varias moléculas no afines genealógicamente —por ser resultantes de evolución convergente— dificultan la interpretación de las similitudes en directos taxa bajo análisis, aunque lo mismo sea aplicable a los caracteres morfológicos. No obstante, dicho enfoque parece ser muy útil a nivel genérico en las plantas y de limitado valor en grupos superiores en las mismas. Estos hechos, el acceso creciente a la información genética y el reconocimiento de que la cantidad de información evolutiva se incrementa de modo notable con el conocimiento de proteínas y ácidos nucleicos, es lo que ha venido haciendo a un lado el estudio de las micromoléculas como fuente de información relevante en estudios filogenéticos y clasificatorios.

Desde la década pasada las críticas al uso de las micromoléculas en Taxonomía han sido severas (ver trabajos en Bendz y Santesson, 1974), pero se destaca en esa crítica la diferencia entre las técnicas de obtención, fuente o tipo de caracteres, y los métodos de análisis y sus filosofías de clasificación biológica subyacentes en las escuelas de pensamiento sistemático (fenetismo, gradismo y cladismo). A pesar de ello, los estudios de rutas biosintéticas (el proceso, no sólo las micromoléculas) o bien, nuevos estudios de micromoléculas en animales, de gran valor por su relación microestructural, parecen prometer información muy valiosa y de gran significación en tales investigaciones evolutivas con propósitos sistemáticos; éste es el caso del estudio reciente de los lípidos en insectos que a continuación se resume.

Lípidos

Cane (1983) señaló algunos de los problemas que se encuentran comúnmente al inferir relaciones filogenéticas usando datos bioquímicos: 1. la determinación de las homologías y 2. el basar las hipótesis filogenéticas en la presencia o ausencia de una sola y única molécula hasta el grado de excluir datos bioquímicos más comunes pero menos diagnósticos (problemas que también ocurren con otras fuentes de caracteres, como los mismos datos morfológicos); y señaló también que los problemas encontrados al inferir relaciones filogenéticas usando datos bioquímicos pueden ser superables, especialmente si se aplica el análisis cladístico a combinaciones discretas de moléculas que pueden ser cuantificadas, caracterizadas estructuralmente, relacionadas biosintéticamente y elegidas sin influencias preconcebidas. Cane (*op. cit.*) destacó que los lípidos exócrinos de los insectos son bastante adecuados para la tarea de reconstrucción filogenética de esos grupos. Sus argumentos son los siguientes: los lípidos pueden ser definidos como un conjunto heterogéneo de sustancias químicas que son relativamente solubles en solventes orgánicos, pero que no se mezclan en agua; son de distribución universal en la naturaleza, e incluyen algunos de los llamados "metabolitos secundarios", pigmentos, semioquímicos (p. ej., feromonas), grasas, ceras y otros. Las secreciones exócrinas de los insectos están compuestas típicamente por mezclas de lípidos, cuyos constituyentes generalmente pueden separarse por medio de cromatografía y cuantificarse.

La elección de lípidos exócrinos para los análisis filogenéticos no necesita ser influenciada más que por la cantidad mínima absoluta de cualquier compuesto requerida para su identificación exitosa (frecuentemente meros nanogramos). Esta es una ventaja notable sobre las inevitables cuestiones de elección de caracteres que plagan a los morfólogos (Cane, 1983).

La obtención, separación-aislamiento e identificación de los lípidos frecuentemente es directa y reproducible; además, los lípidos exócrinos de los insectos a menudo están compartimentalizados convenientemente en glándulas, lo cual ayuda a su disección y disminuye las oportunidades de contaminación. Los esquemas biosintéticos de los lípidos de los insectos están avanzando rápidamente y ofrecen un criterio similar a la ontogenia para estimar o evaluar hipótesis filogenéticas.

Puesto que los lípidos exócrinos son secretados comúnmente como mezclas de constituyentes cuyos componentes pueden ser cuantificados con precisión y relacionados unos con otros biosintéticamente, las filogenias pueden inferirse usando tanto estructuras moleculares novedosas como otras combinaciones compartidas.

La homología puede establecerse tanto para el órgano secretor como para la ruta biosintética de los lípidos analizados. Las funciones ecológicas y/o fisiológicas de los lípidos exócrinos muchas veces pueden discernirse, lo cual provee al investigador de hipótesis *a priori* para anticipar convergencias, radiaciones, paralelismos y otros fenómenos evolutivos.

Finalmente, los lípidos exócrinos ofrecen un conjunto de datos independiente que pueden complementar la información morfológica y ecológica para apoyar o refutar las clasificaciones existentes. Cane (1983) concluyó que los lípidos exócrinos poseen ventajas múltiples para un análisis "quimiosistemático" que no son compartidas en conjunto por ninguna otra clase de sustancias. Aunque no lo sugirió explícitamente, otra ventaja aparente de los lípidos es la gran cantidad de caracteres que se pueden analizar simultáneamente.

En general, la técnica para obtener los datos lipídicos es la siguiente (Cane, 1983): primero se remueven las glándulas respectivas de los insectos; de ellas se extraen los lípidos con algún solvente orgánico y posteriormente se separan, identifican y cuantifican, usando cromatografía de gases y el espectrómetro de masas. Los datos pueden ser manejados por métodos filogenéticos comunes, analizándolos tanto manualmente como con algoritmos de computadora; son también posibles las comparaciones de grupos externos para distinguir las apomorfias y plesiomorfias y, finalmente reconocer las sinapomorfias.

Cane (1983) pudo usar hasta 101 lípidos diferentes para efectuar un análisis cladístico de las relaciones entre algunos grupos de abejas. Su principal problema fue que la cromatografía de gases/espectrómetro de masas tienen un límite de detección, que —aunque no arbitrario— subdivide al continuo cuantitativo en las categorías "detectado" y "no detectado". Para datos micromoleculares, la "ausencia" de un compuesto particular es evidenciada sólo por su falta de detección y considérese que la "pérdida" de un carácter (como en una reversión) puede significar solamente que el lípido se ha vuelto muy diluido en la secreción como para ser todavía detectado en el análisis; ésta y otras razones sugirieron que usar un código cuantitativo en el análisis por computadora sería ventajoso sobre el uso de un código de presencia/ausencia.

Como conclusión preliminar, se puede decir aquí que la "quimiosistemática" parece tener un futuro prometedor si resultasen verdaderas y generalizables todas sus ventajas; pero esto debe ser sustentado por muchas investigaciones más en grupos taxonómicos variados, ya que estas afirmaciones son conocidas desde Huxley (1940). Estos métodos y técnicas son bastante recientes, por lo que hay que esperar aún para definir con mayor confiabilidad sus defectos y bondades, para poder así detectar sus potencialidades y limitaciones.

ACIDOS NUCLEICOS

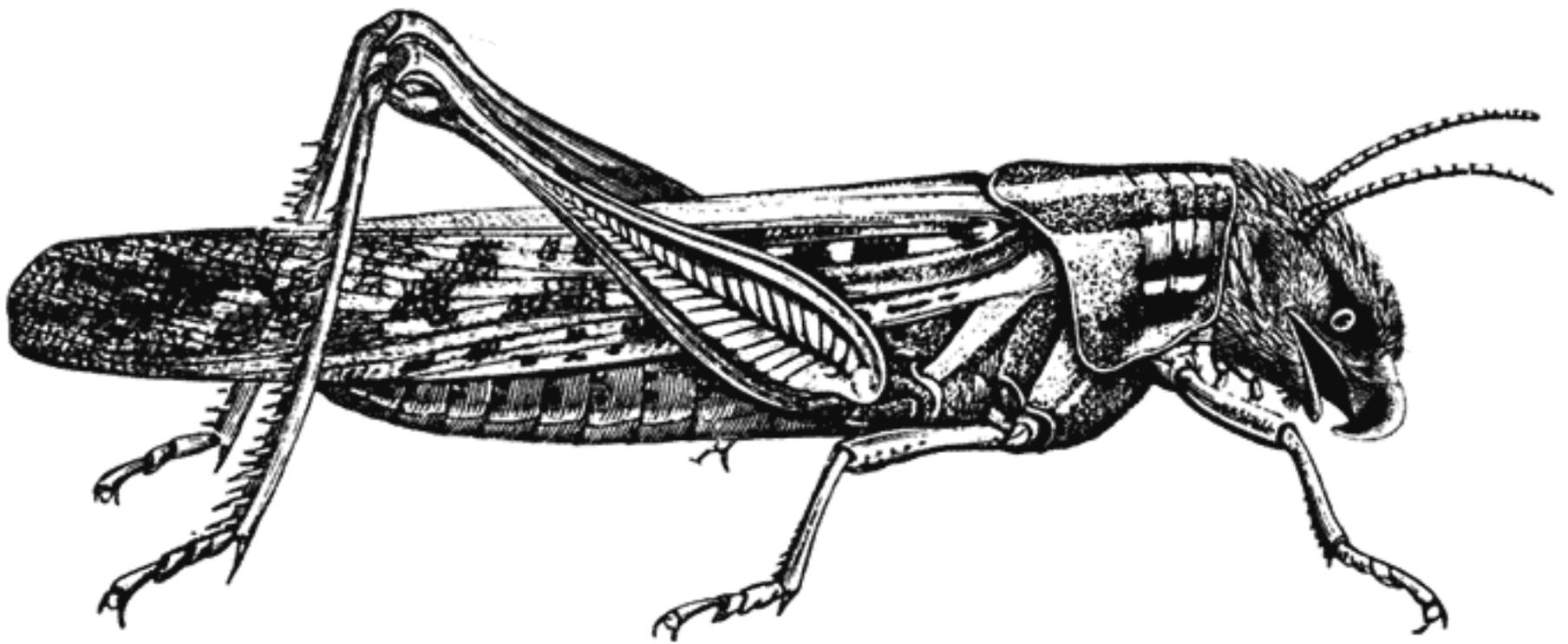
Cinética de reasociación y estabilidad térmica del ADN híbrido

Las técnicas de hibridación del ADN se usan para determinar cuanta similitud tienen entre sí dos cadenas sencillas y "complementarias" de ADN. Las técnicas aprovechan tres características del ADN: 1. Una doble hélice puede ser desnaturalizada por calentamiento en una serie de cadenas sencillas de nucleótidos 2. Partes de la doble hélice pueden ser renaturalizadas incubando una mezcla de cadenas sencillas 3. El grado de renaturalización es directamente proporcional a la similitud (complementariedad) de las cadenas y esto puede medirse calentando la cadena híbrida para determinar su punto de disociación (Wiley, 1981).

Un excelente resumen de la técnica de hibridación del ADN es el de Dobzhansky *et al.* (1977). Siguiendo a Wiley (1981), las bases del análisis son las siguientes:

1. Se marca radiativamente el ADN de un organismo.
2. Se desnaturaliza el ADN de este organismo y se obtienen cadenas de aproximadamente la misma longitud o se obtienen cadenas de ADN "único" (la mejor y más reciente de las técnicas para el trabajo sistemático; ver adelante).
3. Se producen los ADN híbridos mezclando las cadenas del primer organismo con las de un segundo organismo.
4. Se elimina todo el ADN "libre"; es decir, el que no formó ADN híbrido.
5. Se determina el punto de disociación del ADN híbrido y la cantidad de ADN del segundo organismo que ha formado cadenas dobles con el ADN del primer organismo, midiendo la cantidad de radiactividad emitida por la doble cadena híbrida durante el calentamiento.

Cuanto mayor sea la temperatura necesaria para desnaturalizar la cadena de ADN híbrido, más similar será el ADN de las dos especies. En otras palabras,



entre más estable térmicamente sea el ADN híbrido, más similares serán las fibras de las dos especies.

Pueden obtenerse dos medidas distintas a partir de la diferenciación del ADN por medio de la técnica de hibridación: 1. la extensión de la reacción de hibridación y 2. la proporción de pares de nucleótidos que son complementarios en las moléculas híbridas. Para obtener la segunda medida, el ADN híbrido se calienta incrementando la temperatura a una tasa de 1°C cada pocos minutos; el ADN dúplex gradualmente se disocia en fibras sencillas que se recolectan a intervalos regulares y la proporción de ADN dúplex disociado en cada temperatura se grafica. La proporción de nucleótidos no complementarios se determina comparando las curvas de disociación de los dúplex del ADN híbrido y el ADN testigo, este último formado por la reasociación de cadenas sencillas del ADN de un solo organismo. El parámetro crítico llamado estabilidad térmica (TS) es la temperatura a la cual el 50% del ADN dúplex se ha disociado. La diferencia (delta TS) entre el TS del ADN híbrido y el ADN testigo es aproximadamente proporcional al número de nucleótidos no apareados en el ADN híbrido. En la actualidad se supone que la correspondencia es de aproximadamente 1°C delta TS = 1% de nucleótidos no apareados.

Algunos ejemplos de filogenias obtenidas a partir de datos de hibridación de ADN pueden consultarse en Dobzhansky *et al.* (1977). Aunque muchas filogenias semejantes no están inscritas siguiendo un análisis cladista, Wiley (1981) señaló que las relaciones filogenéticas reconocidas por medio de esa técnica son suficientemente similares a las relaciones genealógicas bien establecidas (pero basadas en otros caracteres) como para inspirar confianza en los resultados obtenidos de tales análisis.

Wiley (1981) advirtió que los primeros trabajos de hibridación de ADN se llevaron a cabo sin separar el ADN en secuencias únicas y repetitivas; el problema era que uno nunca sabía realmente con cuáles de ellas se estaba trabajando, por ello los resultados no eran confiables. Entonces se tenían esperanzas en los resultados obtenidos usando sólo segmentos únicos, como se hace en la actualidad.

Desde 1974, Sibley y Ahlquist (1983) han realizado la hibridación del ADN de más de 1,000 especies de aves, representando todos los órdenes y 163 de las 174 familias de las mismas; estos autores han presentado, como una conclusión de sus estudios, una serie de consideraciones sobre las cualidades y potencialidades de los datos de hibridación de ADN-ADN en la filogenia y en la clasificación de los organismos. Su trabajo comprende una descripción cabal de la técnica de hibridación del ADN.

Sibley y Ahlquist señalaron las razones y ventajas de efectuar el análisis usando solamente las secuencias de ADN no repetidas, eliminando las copias de secuencias repetidas (éstas secuencias pueden consistir desde unas pocas hasta centenares de miles de copias); ya que la tasa de reasociación es función del número de copias, las secuencias repetidas contribuyen a desproporcionar las medidas precisas de comparación de la radiactividad, aparentando mayor

similitud. Esto distorsionaría la equivalencia de los datos y el análisis, porque la fracción de secuencias repetidas es variable aun dentro de una especie; no puede suponerse que las diferencias observadas entre las secuencias repetidas de dos especies hayan ocurrido desde el momento de divergencia de las mismas. Por otro lado, las secuencias no repetidas compartidas entre dos especies deben ser descendientes de la misma secuencia ancestral común que estuvo presente en el ancestro común más reciente (como los genes parálogos y ortólogos, respectivamente).

Los mismos autores discutieron también el problema del análisis de los datos de hibridación del ADN: considerando la relación entre la temperatura de disociación de un ADN híbrido y el porcentaje de apareamiento de bases en el mismo, definieron la "distancia genética" como una medida de esa relación. Así, un valor de delta TS es una distancia genética promedio; pero la conversión de medidas de cualquier clase —en filogenias— requiere de un procedimiento para obtener una agrupación jerárquica de los taxa; ellos usaron una versión modificada del método de agrupación por unión promedio (*average linkage*), que agrupa esencialmente por similitud "total" (en el sentido fenético). Su argumento de base para el empleo de este método es que "la única razón para la 'cercanía' es la verdadera homología de los caracteres, pero la distancia puede resultar del fallo al identificar las homologías existentes. Los datos de hibridación del ADN son especialmente compatibles con este punto de vista porque no hay razón conocida que no sea homología entre las secuencias de nucleótidos, que explique la estabilidad térmica del ADN híbrido".

Sibley y Ahlquist (1983) dieron argumentos en favor de la exactitud con la cual se miden las diferencias en los puntos de disociación de los ADN híbridos y ofrecieron ejemplos de que estas medidas son: 1. confiables (medido por la reproducibilidad de los parámetros de las curvas de disociación de los híbridos ADN-ADN: híbridos idénticos mostraron tener propiedades casi idénticas, con desviaciones bajas atribuibles a error experimental), 2. sensibles (pueden discriminar entre especies estrechamente relacionadas y morfológicamente muy similares, e incluso pueden detectar polimorfismos genéticos), y 3. recíprocas (la distancia evolutiva del taxón A al taxón B es la misma que de B a A).

Con respecto al problema de la homología, estos autores consideran que los conceptos y definiciones de genes parálogos y ortólogos de Fitch (1970) no afectan la interpretación de los datos de hibridación del ADN, ya que en el sentido más general, todos los genes son homólogos porque es probable que todos ellos, en todos los organismos, se hayan originado por duplicación y divergencia génica a partir de las secuencias de nucleótidos preexistentes que acompañaron el origen de la vida. Así, el grado de complementariedad por el cual se identifican los genes homólogos, declina gradualmente con el tiempo. No obstante, a menudo pueden reconocerse homólogos entre genes que comenzaron su divergencia hace cientos de millones de años. De este modo, genes "totalmente nuevos" no surgen de la nada, sino que son producto de la duplicación y divergencia de genes preexistentes. Lo anterior es muy importante para la interpretación de los datos de hibridación del ADN y para

el concepto de homología en las secuencias de ADN, de acuerdo a los autores referidos.

Sibley y Ahlquist (1983) argumentaron con diversos ejemplos la uniformidad promedio en las tasas de evolución del ADN; ellos encontraron evidencia de que la tasa de sustituciones de nucleótidos, medida en todo el genoma y en el tiempo, es la misma en todos los linajes de aves. Esto implicaría que los valores de delta TS podrían tomarse como medidas relativas del tiempo de divergencia evolutiva. Así, las tasas de divergencia serían medibles por esta técnica y los valores de distancias resultantes podrían ser usados para reconstruir las filogenias de los organismos vivos en términos de tiempo relativo. Esta escala de tiempo relativo podría convertirse a una escala absoluta mediante la calibración de los valores de ADN-ADN contra fuentes externas de datos: fósiles y eventos geológicos o geofísicos. Sibley y Ahlquist (*op. cit.*) propusieron que un delta TS de 1.0 es igual a aproximadamente 4.5 millones de años.

Estos autores han ido aún más lejos al abarcar aspectos de clasificación, señalando que gracias a la uniformidad de la tasa promedio de evolución genética, las relaciones genéticas y genealógicas son idénticas y que los rangos absolutos de las categorías taxonómicas podrían basarse sobre sus edades de origen (de hecho, proponen una escala para este efecto con base en sus resultados en las aves). Desde luego, toda la clasificación podría basarse también en estos datos, ya que miden directamente la divergencia genotípica y pueden calibrarse contra el tiempo absoluto, alcanzando una clasificación cladista ortodoxa e idealizada (*sensu* Hennig, 1968; Wiley, 1981), ya que la uniformidad de la tasa promedio de evolución del ADN significaría que la filogenia está estrechamente relacionada con el tiempo; esto satisfaría el criterio hennigiano de rangos de taxa basados en el tiempo, en el cual las categorías de igual edad de origen son equivalentes y los grupos hermanos corresponden a iguales rangos o categorías taxonómicas.

Evidentemente Sibley y Ahlquist son ardientes partidarios del método de hibridación del ADN. Su exposición de las grandes ventajas del método parece convincente excepto por dos puntos: 1. Arguyen que no hay ninguna otra razón que pueda explicar la hibridación y estabilidad térmica del ADN, que no sea la homología entre las fibras sencillas hibridadas; y de este modo parecen ignorar las posibilidades de considerar homoplasias en el método. Es decir, también pueden darse homoplasias a nivel de las secuencias de nucleótidos que expliquen la estabilidad térmica del ADN híbrido. Su argumento de que los conceptos y definiciones de genes parálogos y ortólogos no afecta

la interpretación de los datos de hibridación del ADN tampoco resulta muy convincente. 2. Señalan haber encontrado tasas evolutivas constantes a nivel del ADN en todas las aves estudiadas y de ello desprenden sus argumentos para: a) emplear métodos fenéticos en el análisis de sus datos y b) datar la historia de la especiación con los rangos de los taxa basados en su edad de origen. Quedaría por verse en el futuro si esto es cierto para el caso de las aves; pero ya anteriormente Dobzhansky *et al.* (1977) señalaron algunas fuertes discrepancias en las tasas de evolución de nucleótidos por unidad de tiempo en varios linajes diferentes, aun considerando errores experimentales.

Dobzhansky *et al.* (1977) señalaron algunas cualidades de estas técnicas, que hay que atender para usarlas y evaluarlas en estudios de historia evolutiva: 1. Son muy laboriosas y caras por la necesidad de marcar el ADN con un isótopo radiactivo, 2. Los resultados son muy sensibles a las condiciones experimentales usadas, y 3. Los resultados inconsistentes no son raros, aun en un conjunto único de experimentos, usando procedimientos idénticos.

Secuencias de ADN Mitocondrial

Awise, Lansman y Slade (1979) introdujeron al análisis de las poblaciones naturales una técnica molecular que incluye el uso de endonucleasas de restricción para comparar secuencias de ADN mitocondrial (ADNmt). Para ese entonces era claro que el ADNmt de los animales se estaba volviendo la "pieza" mejor conocida del genoma eucariote y que también debido a sus propiedades estaba adquiriendo un gran interés para los genetistas de poblaciones y los sistematas. Estas propiedades, expuestas detalladamente por Wilson *et al.* (1985), son las siguientes: 1. El ADN mitocondrial llegó a ser la unidad mejor conocida del ADN eucariote porque es más fácil de purificar que cualquier segmento específico de ADN nuclear; esto se debe a una densidad de flotación poco frecuente, un alto número de copias y su presencia en un organelo diferente al núcleo. 2. El ADNmt es fácil de caracterizar porque es pequeño y carece de muchos de los rasgos complejos del ADN nuclear (tales como intrones y secuencias repetitivas). 3. El ADNmt está distribuido universalmente en el reino animal y es notablemente uniforme en contenido génico. Los ADN mitocondriales de todos los organismos multicelulares probados y de algunos protozoarios parecen tener el mismo conjunto de 37 genes, que especifican 22 ARNt, 13 ARNm y 2 ARNr; estos genes están dispuestos apretadamente en alrededor de 15 kilobases de ADN de doble cadena; su arreglo es muy estable y sólo se han hallado cambios en el orden de sus genes en comparaciones entre diferentes phyla. 4. Los tipos de cambio evolutivo que sufre el ADNmt son relativamente simples, son principalmente sustituciones de bases y adiciones mutacionales de longitud que se acumulan predominantemente en las pequeñas regiones no codificadoras. 5. El ADNmt tiene una tasa de cambio mucho más rápida que el ADN nuclear. Algunos trabajos han indicado que el ADNmt acumula mutaciones a una velocidad casi 10 veces mayor (o mucho mayor todavía) que el ADN nuclear, lo que hace que las comparaciones usando ADNmt sean mucho más sensibles que las que usan el ADN nuclear. 6. Hasta donde se sabe, el ADNmt se hereda sólo maternalmente y todos los organismos son haploides con respecto al número de tipos de ADNmt transmitidos a la siguiente generación (aunque poliploides con respecto al número de copias de ADNmt por célula). 7. De acuerdo a lo que se conoce, la recombinación no ocurre en el ADNmt (excepto, quizá, en el asa de desplazamiento). Esta propiedad, aunada a la anterior, sugieren que el ADNmt puede ser una herramienta genealógica muy poderosa (ver adelante).

Desde sus inicios, para comparar las secuencias de ADNmt se introdujo el uso de endonucleasas de restricción (Awise, Lansman y Slade, 1979); éstas rompen en fragmentos específicos las moléculas de ADN, aisladas a partir de diversos organismos; luego los fragmentos son separados por su tamaño mediante electroforesis en gel. Las endonucleasas de restricción reconocen secuencias de nucleótidos particulares, específicas, de unos pocos nucleótidos de extensión: 4, 5 ó 6 pares de bases de largo, rompiendo los enlaces fosfodiéster dentro de la secuencia, uno en cada fibra del dúplex. Si los nucleótidos en la secuencia de reconocimiento no están protegidos por metilación, una enzima de restricción rompe consistentemente esta secuencia en el ADN de cualquier organismo. Así, las diferencias en los tamaños de los fragmentos obtenidos digiriendo los ADN homólogos con una enzima



dada, reflejan diferencias exactamente en los sitios de reconocimiento en las secuencias.

Las enzimas de restricción son particularmente valiosas para comparar los ADNmt estrechamente relacionados. Hay varios métodos que hacen uso de estas enzimas (Wilson *et al.*, 1985). Un primer método de gran poder de resolución, usa un conjunto de aproximadamente 10 enzimas, capaces de romper en secciones de 4 bases; así, cada una de tales enzimas reconoce una secuencia específica de cuatro bases y corta el ADNmt dondequiera que haya un sitio de reconocimiento; los fragmentos de ADNmt resultantes se hacen detectables uniendo fosfato radiactivo a sus extremos. Una enzima típica de éstas corta el ADNmt de los vertebrados en aproximadamente 25 fragmentos, la mayoría de los cuales son separables unos de otros por electroforesis en un largo gel de poliacrilamida. Los patrones de fragmentación, hechos visibles con una película de rayos X, se comparan entonces para diferentes ADNmt que se han tratado con la misma enzima. Las diferencias entre los patrones de fragmentos se deben generalmente a sustituciones de bases que pueden generar pérdida de sitios de reconocimiento, aunque algunas veces son el resultado de adiciones. La técnica que usa el marcaje terminal y las enzimas que rompen secuencias de cuatro bases es capaz de distinguir entre ADNmt que difieren en menos de 0.05% de sus secuencias de bases —cuando se usa un conjunto de 10 enzimas semejantes. También existen métodos que usan enzimas que reconocen secuencias de 6 u 8 bases, pero son de menor poder de resolución debido al tamaño mayor de los fragmentos resultantes. Un ejemplo del uso de este método puede hallarse en *Avise et al.* (1979); estos autores analizaron los patrones de fragmentos producidos por 6 endonucleasas de restricción que actuaron sobre los ADNmt de 23 muestras de tres especies de *Peromyscus* (ver adelante).

Un segundo método de comparación de los ADNmt incorpora la obtención de mapas de restricción; las localizaciones relativas de los sitios pueden determinarse comparando los fragmentos producidos, digiriendo el ADNmt con mezclas de dos enzimas de 6 bases, con aquellos producidos por cada enzima sola. Los mapas de restricción resultantes mejoran la exactitud de las estimaciones del grado de divergencia de las secuencias y de la construcción de árboles genealógicos.

Un tercer método de comparación de los ADN mitocondriales es el de secuenciamiento de regiones clonadas de los mismos. Los métodos del ADN recombinante permiten la clonación de regiones específicas del genoma mitocondrial a partir de individuos y de especies estrechamente relacionadas. El secuenciamiento de estas piezas clonadas de ADNmt da indicios de la naturaleza de los cambios evolutivos que tienen lugar en él; la técnica ha demostrado que una gran proporción de las sustituciones de bases que tienen lugar en las secuencias del ADN son transiciones silentes en los codones, es decir, que no causan sustituciones de los aminoácidos codificados. El secuenciamiento en secuencia de regiones codificadoras clonadas puede usarse para evaluar la precisión de las estimaciones de divergencia entre secuencias hechas a partir de comparaciones de mapas de restricción de los ADNmt completos. Este método también facilita probar y enraizar árboles intraespecíficos para ADNmt. Pero aunque las técnicas de restricción son valiosas para construir tales árboles, tienen una debilidad potencial, ya que la pérdida paralela de un sitio puede considerarse como un estado derivado compartido; las secuencias de fragmentos no tienen esta limitación y pueden revelar si las mutaciones que causaron la pérdida del sitio en dos ADNmt diferentes, se deben al mismo cambio de base.

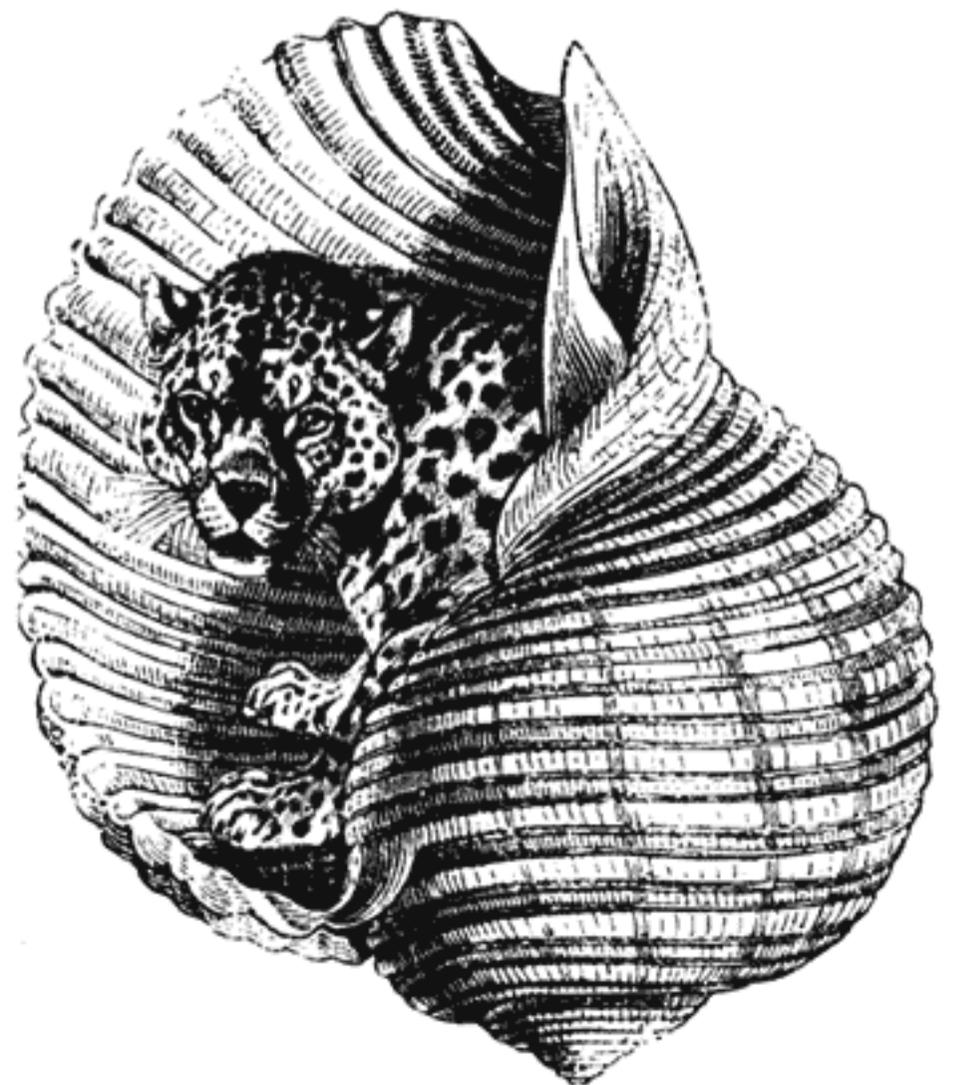
Si las secuencias contienen regiones codificadoras, surge otra ventaja: al enfocar la atención sobre los sitios silentes (los sitios donde ocurren las sustituciones silentes, las cuales se acumulan rápidamente), uno puede trabajar sobre las ramificaciones recientes del árbol: enfocando la atención sobre las sustituciones de acumulación más lenta que causan reemplazamientos de aminoácidos, uno puede trabajar sobre tiempos antiguos de divergencia y enraizar el árbol.

La razón final para el entusiasmo por la obtención de las secuencias de bases de fragmentos clonados de ADNmt, es que tales fragmentos son probablemente las únicas formas de ADNmt que podrían recobrase de los

restos orgánicos de animales muertos mucho tiempo atrás. Estos restos, que abundan en museos, tumbas, suelo congelado y ámbar, representan poblaciones antiguas y algunas veces ancestrales. Debido al alto número de copias de ADNmt por célula, la probabilidad de que estos restos contengan fragmentos clonables de ADNmt es miles de veces mayor que la probabilidad de que exista un fragmento de cualquier pieza específica de ADN nuclear, del que existe una sola copia (Wilson *et al.*, 1985).

Avise et al. (1979) usaron los patrones de fragmentos generados por los rompimientos de las endonucleasas de restricción, a fin de calcular la divergencia entre secuencias para cada par de sus 23 muestras de 3 especies de *Peromyscus* (análisis cuantitativo). Las estimaciones de divergencias entre las secuencias de nucleótidos de todos los pares de muestras —usando diversas enzimas— generan matrices; la información contenida en tales matrices de distancias puede ser analizada por cualquiera de los procedimientos existentes para estimar dendrogramas y árboles filogenéticos (Fitch y Margoliash, 1967; Farris, 1972; ver recuadro sobre algoritmos en este trabajo). Los autores citados también analizaron los mismos datos desde una aproximación diferente —de naturaleza cualitativa— que enfatiza las similitudes en los patrones de rompimiento o digestión, más que sus diferencias. La suposición básica de este análisis es que los complejos fenotipos del patrón de digestión no pueden ser el resultado de evolución convergente a partir de fenotipos no relacionados. Así, los individuos o unidades familiares que comparten un patrón multibandeado después de la digestión del ADNmt, con una o más enzimas, pueden asignarse sin ambigüedades a un origen evolutivo común. Por ejemplo, puesto que todas las muestras examinadas de *Peromyscus polionotus* compartieron un patrón de digestión común de dos fragmentos producidos por la enzima BstEI, dichos autores consideraron que todas ellas pertenecían a un solo "linaje matriarcal". Evidentemente, este tipo de análisis puede hacerse dentro del marco conceptual del cladismo y con su metodología usual, ya que implica la presencia o ausencia de caracteres y evita así los problemas que conllevan otras técnicas bioquímicas, las cuales producen medidas de distancia abstractas (Duellman, 1985).

Avise et al. (1979) señalaron que la principal fuente de error en los cálculos del análisis cuantitativo es la posibilidad de que fragmentos no homólogos de peso molecular similar puedan ser tomados como idénticos. Esta probabilidad es mayor cuando se comparan muestras de secuencias ampliamente diferentes o cuando se incluyen en el análisis enzimas que producen un gran número de fragmentos de peso molecular similar. A pesar de esto, ambos tipos de



análisis, el cuantitativo y el cualitativo, les llevaron a panoramas generales de las relaciones evolutivas entre las poblaciones de *Peromyscus* que concordaban con las relaciones obtenidas previamente a partir de datos morfológicos y genéticos. Sin embargo, ellos concluyeron que la aproximación cualitativa sería más útil y poderosa porque no requiere extrapolaciones estadísticas ni suposiciones débiles para establecer la ancestría común de o entre las muestras comparadas.

Un hecho que parece apoyar el punto de vista anterior es el siguiente: DeBry y Slade (1985) demostraron que existe una gran asimetría entre las probabilidades de ganar un nuevo sitio y las de perder un sitio ya existente en el ADNmt: la pérdida de un sitio es un hecho mucho más común. Ya que la parsimonia de Wagner supone que las probabilidades de cambios hacia adelante y de reversiones son iguales, es una forma muy deficiente de apreciar las relaciones partiendo de estos datos. Por otro lado, estos autores señalaron que la parsimonia de Dollo supone que la probabilidad de una pérdida es mucho mayor que la de una ganancia, encajando, por tanto, correctamente con los datos de los mapas de restricción.

Se ha calculado la tasa de divergencia promedio en la molécula de ADNmt a lo largo del tiempo, estableciéndose que ésta es de aproximadamente 2% por millón de años en muchos grupos de organismos; también se han hecho intentos por fechar los eventos de especiación y de divergencia sobre la base de tal tasa (Avice *et al.*, 1979), suponiendo, como era de esperar, que la evolución del ADNmt se comporta de manera semejante a un reloj. Pero Wilson *et al.* (1985) han señalado que aunque no se conocen grandes alejamientos de dicha tasa para la molécula de ADNmt como un todo, en una región específica de ella se ha observado lo que se podría considerar, quizá, como el mayor alejamiento del comportamiento semejante a un reloj encontrado en estudios de evolución molecular de los vertebrados.

Los autores anteriores han apuntado que para evaluar objetivamente las potencialidades de las comparaciones del ADNmt, se requieren aún más estudios comparativos del mismo, tanto con nuevas técnicas como con las ya existentes; también han indicado que todo el poder de resolución de la técnica estará disponible sólo cuando haya una manera de obtener rápidamente la secuencia del ADNmt.

Por su parte, Avice *et al.* (1979) mencionaron las siguientes características para el análisis del ADNmt usando enzimas de restricción (características que hacen poderosa a tal técnica en el análisis de la estructura de poblaciones y de las relaciones filogenéticas de especies estrechamente relacionadas):

1. Cada fenotipo de restricción complejo es único. Las posibilidades de que un fenotipo idéntico surgiera por convergencia a partir de fenotipos no relacionados son extremadamente remotas.
2. Los fenotipos se transmiten intactos, inalterados por la recombinación durante la reproducción sexual, de tal manera que cualquier cambio en las secuencias sólo puede surgir por mutación.
3. Las mutaciones que se fijan en un individuo dan por resultado un nuevo fenotipo que puede ser relacionado con su progenitor sin ambigüedades.
4. La tasa de aparición de nuevos fenotipos por mutación es significativa, pero no incontrolablemente alta. Así, la relación entre los individuos de una población es aparente a pesar de la presencia de polimorfismo fácilmente detectable.
5. Los fenotipos son fácil y rápidamente obtenibles. Preparaciones de 10 fenotipos pueden hacerse en un tiempo razonablemente corto, aunque es un procedimiento laborioso.

Un estudio reciente que ilustra el enorme valor y ventajas de esta técnica, es descrito por Cann (1987) en la investigación acerca del origen del hombre. La autora apoya con sus resultados la hipótesis de un origen africano para *Homo sapiens*.



Wiley (1981) concluyó, refiriéndose a otras técnicas bioquímicas tales como las electroforesis de proteínas, la inmunología y otras, que dado que éstas han originado genealogías muy semejantes a aquéllas basados en hipótesis filogenéticas bien documentadas (análisis cladistas), pueden considerarse fuentes válidas de datos filogenéticos, siempre que estos sean analizados correctamente y bajo las suposiciones referidas a lo largo del trabajo. Esta conclusión puede hacerse extensiva a las comparaciones de secuencias de ADN mitocondrial; no obstante, son menos las suposiciones en este caso, pero la técnica es más cara que la electroforesis.

Secuencias de ARN Ribosomal

El estudio de las relaciones filogenéticas entre los organismos también se ha efectuado usando el ácido ribonucleico (ARN); en especial, el ácido nucleico ribosomal (ARNr). El uso de este ácido se ha generalizado más o menos simultáneamente al uso del ADN, a partir de finales de la década pasada, época en la que se desarrollaron las técnicas necesarias para obtener la secuencia de nucleótidos de ambos ácidos nucleicos (Woese, 1981).

Los argumentos que comúnmente se esgrimen a favor del uso del ARNr son - en buena medida - semejantes a los que se han dado en favor del uso del ADN: principalmente son argumentos que enfatizan las deficiencias y limitaciones de las técnicas tradicionales para el reconocimiento de caracteres "significativos" en las relaciones filogenéticas entre los organismos, sobre todo en el caso de grupos muy antiguos. Así, Field *et al.* (1988) señalaron que, si bien hay cierto acuerdo en los agrupamientos de taxa en clases y phyla, hay grandes dificultades en la determinación de las relaciones filogenéticas entre los grupos más lejanamente relacionados (phyla y reinos), y que es éste uno de los grandes retos en la Sistemática Biológica.

Lo anterior es debido a que los diferentes phyla tienen "tipos estructurales" bien diferentes, con características muchas veces restringidas a un solo phylum (autapomorfias), lo que da por resultado que haya muy pocos caracteres para agrupar los phyla en unidades mayores o establecer sus relaciones filogenéticas (a través de sinapomorfias); por otra parte, cuando hay caracteres comunes a más de un phylum, su homología es comúnmente incierta. Tampoco existe un registro fósil que muestre la continuidad histórica de caracteres que pudieran usarse para establecer relaciones entre los phyla y, finalmente, los datos embriológicos son de difícil interpretación porque los procesos de desarrollo no son suficientemente entendidos; no debe dejarse de lado que los procesos de especiación y sus patrones, así como fenómenos evolutivos hasta ahora no completamente comprendidos en estos niveles (*e.g.* endosimbiosis de organelos) impiden un tratamiento de absoluta equivalencia al realizado en organismos pluricelulares.

Todo lo anteriormente mencionado es la causa de que los árboles filogenéticos de estos taxa, basados en la anatomía, la embriología y la paleontología sean muy especulativos y controvertidos. Además, normalmente las relaciones filogenéticas se establecen agrupando organismos que comparten características homólogas derivadas, innovaciones que deberían reflejar una ancestría común.

Sin embargo, a nivel de phyla, frecuentemente es difícil determinar la dirección del cambio evolutivo y saber cuáles son los caracteres derivados, ya que también ha habido evolución convergente, lo cual ha oscurecido tanto el establecimiento de las verdaderas homologías como el de la dirección de cambio (polaridad evolutiva).

Woese (1981) dio argumentos similares a los anteriores para un caso todavía más extremo: el de la comprensión de las relaciones filogenéticas entre las bacterias. Durante todo un siglo los microbiólogos han tratado de establecer las relaciones entre las bacterias, investigando su gran diversidad de formas, sistemas fisiológicos y ecologías. La variedad de las bacterias es básicamente una variedad dentro de la simplicidad y proporciona por ello poca información acerca de sus relaciones filogenéticas. Por ejemplo, bastoncillos, esferas y espirales, que son las formas bacteriales típicas, son formas de "evolución fácil" y han evolucionado muchas veces; lo mismo puede decirse de algunos aspectos de la bioquímica bacteriana. Aunque es posible que hayan caracteres bacterianos que puedan ser buenos indicadores filogenéticos, hasta el momento es imposible decir de antemano cuáles lo son y cuáles no. Y aunque existen microfósiles de bacterias de hasta 3 500 millones de años de antigüedad, estos fósiles no son estructuras muy informativas: poco puede inferirse de la impresión de una pequeña esfera o bastoncillo.

Recientemente Hori y Osawa (1987:445) señalaron: "Puesto que los cambios evolutivos de tales caracteres (fisiológicos y morfológicos) son muy complicados y la tasa de cambio es variable en diferentes grupos de organismos o en diferentes periodos evolutivos, no puede tenerse mucha confianza en tales sistemas (es decir, en los sistemas de clasificación actuales del mundo vivo, basados en características fisiológicas y morfológicas...)"

Por otra parte, también se han dado argumentos y consideraciones teóricas respecto al valor de los ácidos nucleicos en general y en particular, del ARNr *per se* para el estudio de las relaciones filogenéticas entre los organismos:

1. Woese (1981) argumentó que todas las entidades que en el inicio de la vida pudieron autorreplicarse, tuvieron necesariamente sistemas para el mantenimiento y propagación de la información genética y para traducir dicha información en proteínas y que, por ello, las grandes moléculas actuales que intervienen en estos procesos (los ácidos nucleicos) deben tener sus orígenes en los mismos albores de la evolución de la célula (aún antes de la existencia de los procariontes), por lo que uno esperaría que estas moléculas tuvieran las propiedades de un "marcador filogenético"; dicho de otro modo, las células son un registro de su pasado en términos de las secuencias de nucleótidos de sus ácidos nucleicos, o bien de las secuencias de aminoácidos de sus proteínas (debido a la relación colineal entre un gene y su producto, ya sea éste una proteína o uno de los varios tipos de ARN). Sin embargo entre linajes muy antiguos (por ejemplo de bacterias), aun proteínas como el citocromo c no están distribuidas universalmente, no son estrictamente constantes en su función y, por lo tanto, no son del todo comparables y pueden presentar grandes diferencias de un linaje a otro (lo que no ocurre entre eucariotas), por lo que las filogenias bacterianas resultantes de la comparación de proteínas, son incompletas e inciertas.

2. Woese (1981) y Field *et al.* (1988) establecieron, refiriéndose al problema de la homología, que lo que hace de los genes registros notables del pasado evolutivo es que el "espacio" de secuencia genética es enorme, a tal grado que en todo el curso de la evolución sólo se ha producido una pequeña parte de todas las secuencias genéticas posibles, lo que permite inferir que si dos genes de una cadena de ácido nucleico que contenga un número significativo de nucleótidos son semejantes, esto sólo puede implicar que tuvieron un ancestro común y que son por tanto homólogos (criterio similar al que siguen Sibley y Ahlquist en el caso del ADN nuclear: homólogos estructurales siempre indican ancestría común).

Basándose en los argumentos anteriores, Woese (1981) dio por sentado que el análisis de las secuencias genéticas brinda dos tipos de información: puede revelar relaciones genealógicas y es a la vez un registro de características ancestrales (las comparaciones de secuencias homólogas hace posible reconstruir con cierta exactitud la versión original del gene); también, considera que

tales secuencias pueden medir tiempo evolutivo (funcionando como un "reloj molecular"); pero aquí es de notar que no hay acuerdo entre todos los investigadores que han hecho uso del ARNr: para Field *et al.* (1988) es definitivo el hecho de que no puede suponerse que exista constancia en la tasa de cambio de por lo menos un tipo de ARNr: el ARNr 18S (ver abajo), mientras que Hori y Osawa (1987) consideran bien establecido que la tasa de cambio en cualquier ADN o ARN es constante.

Adicionalmente, se han señalado algunas ventajas prácticas del uso del ARNr en especial (Woese, 1981, y Field *et al.*, 1988):

1. El ARNr es fácil de aislar en cantidades adecuadas, ya que una bacteria típica tiene de 10 000 a 20 000 ribosomas.
2. El ARNr parece haber mantenido una función constante durante espacios de tiempo evolutivo muy extensos en los diferentes linajes y así, es enteramente comparable.
3. Por lo menos en alguna parte de su secuencia el ARNr ha cambiado con la suficiente lentitud como para que la secuencia ancestral no haya sido modificada totalmente.
4. El uso del ARNr garantiza que las secuencias comparadas representan genes transcritos a ARNm y no genes menores o inactivos.

Existen tres clases de moléculas de ARNr. En bacterias existe uno "grande" (ARNr 23S), uno "pequeño" (ARNr 16S) y uno "muy pequeño" (ARNr 5S). Las designaciones 23, 16 y 5 S se refieren a la velocidad con que las diferentes moléculas se sedimentan en una centrifuga (medida en unidades Svedberg); son por tanto una medida indirecta del tamaño molecular. El ARNr 23S tiene aproximadamente 2900 nucleótidos de largo; el 16S, aproximadamente 1540; y el 5S, sólo 120. En eucariotas los tamaños son similares: 25-28S, 18S y 5S.

Se han usado para la determinación de las relaciones genealógicas entre los taxones de los tipos de ARNr señalados; esto por distintas razones: Woese (1981) reconoció que por su tamaño el ARNr 5S resultaba más fácil de caracterizar —era más fácil determinar su secuencia de nucleótidos— que la de los otros ARNr, pero señaló que no era el más exacto como indicador de relaciones genealógicas, principalmente por razones estadísticas (el ARNr 5S algunas veces exhibe grandes diferencias anómalas en las secuencias de un taxón a otro y por su pequeño tamaño, tiene asociados errores estándar mayores en la construcción de dendrogramas filogenéticos que los ARNr más largos); por otro lado, el ARNr 23S es el doble de largo y más difícil de caracterizar que el ARNr 16S, por lo que consideró que este último era el más adecuado para el caso. Sin embargo, elegir al ARNr 16S como el más adecuado obliga a considerar sólo las secuencias más conservadoras de éste para el análisis, pues son las más útiles para la investigación de relaciones genealógicas muy antiguas, ya que para organismos muy lejanamente emparentados no es posible determinar la "homología" entre nucleótidos en las partes de mayor tasa de cambio de la molécula (Woese, 1981; Field *et al.*, 1988). De hecho, aún conociendo completamente las secuencias de nucleótidos de las moléculas 16S o 18S, sólo partes de ellas son útiles en la determinación de las relaciones genealógicas. Aunque esto no fue considerado problemático por Woese (1981) ni por Field *et al.* (1988), Hori y Osawa señalaron un problema a considerar en el análisis exclusivo de las regiones conservadoras, como se verá abajo, al tratar de las limitaciones de estas técnicas.

Hori y Osawa (1987), por otra parte, consideraron que el ARNr 5S es particularmente adecuado para el análisis de las relaciones genealógicas entre organismos muy distantemente relacionados en virtud de dos atributos: 1. su baja tasa de sustitución de nucleótidos (promedio + error estándar = 0.18 ± 0.05 sustituciones/sitio de nucleótido/10 E9 años) y 2. su similitud básica de estructura entre todos los organismos, así como una longitud casi igual (de aproximadamente 120 nucleótidos en todos ellos), lo que permite alinear fácilmente sus secuencias para la construcción de los dendrogramas filogenéticos.



Una vez decidido el tipo de ARN más adecuado para el estudio, la técnica subsiguiente es similar para cualquiera de ellos: el primer paso es obviamente el aislamiento del ARN y la obtención de su secuencia de nucleótidos, ya sea de la molécula entera o de sus regiones conservadoras; después, las regiones de estructuras primarias y secundarias homólogas en las secuencias se alinean, asignando discontinuidades o huecos para compensar en las regiones donde hay variación en la longitud de las secuencias. Normalmente, las regiones de "homología cuestionable" o donde hay grandes discrepancias en las longitudes de las secuencias, se excluyen del análisis para reducir los errores residuales sistemáticos y azarosos, favoreciendo de este modo que los errores en la determinación de las secuencias o en los alineamientos no contribuyan mucho a la incertidumbre en el orden de ramificación del dendrograma filogenético. Después se determinan las "distancias evolutivas" entre cada par de secuencias; es decir, el número promedio de mutaciones puntuales fijas o sustituciones de bases por sitio de nucleótido que separan cada par de secuencias (que han ocurrido desde la separación de las dos secuencias, teóricamente). Para ello existen fórmulas y factores de corrección asociados a éstas, los cuales consideran las mutaciones paralelas y superpuestas y las discontinuidades o huecos de uno o más nucleótidos de longitud hallados al alinear las secuencias (ver Hori y Osawa, 1987).

Con las distancias "evolutivas" obtenidas de esta forma entre cada par de organismos comparados, se forma una matriz, a partir de la cual se obtiene el dendrograma filogenético usando métodos de agrupación fenéticos, como el UPGMA de Sneath y Sokal (1973) u otros métodos estadísticos semejantes empleados en la construcción de dendrogramas.

Son muchos los trabajos taxonómicos que ya se han realizado usando el ARN. Sólo hasta 1981, los ARN de casi 200 especies de bacterias y eucariotas habían sido caracterizados (Woese, 1981); quizá los resultados más espectaculares obtenidos con esta técnica son la demostración de que la secuencia de nucleótidos del ARN en los cloroplastos está específicamente relacionada con las secuencias de ARN en las cianobacterias y de que el ARN en las mitocondrias de las plantas parece ser de tipo bacteriano, lo que ha sido tomado —por muchos investigadores— como prueba del origen simbiótico de esos organelos celulares a partir de procariontes.

También se podría señalar la ubicación de las "arquibacterias" (metanógenas, termoacidófilas y halófilas extremas) como un nuevo reino en la historia y el orden natural de la vida (Woese, 1981; ver también Hori y Osawa, 1987), basado en sus secuencias de ARN. Lo que reveló la existencia de las arquibacterias en este caso, fueron diferencias genéticas puramente cuantitativas que no decían nada de las diferencias cualitativas — fenotípicas— que pudiera haber entre las arquibacterias y las eubacterias (el resto de las bacterias conocidas); si la separación de las arquibacterias como un nuevo reino era justificada, la inspección detallada de éstas debería probar que eran tan diferentes de las eubacterias como cualquiera de las dos lo es de los eucariotes; tal inspección reveló que las arquibacterias son en verdad organismos atípicos que presentan características únicas, no encontradas en las demás bacterias; por ejemplo: en los lípidos que forman su pared celular, en su ARN de transferencia, en la estructura de su ARN polimerasa y en otras características más; esto sin que las arquibacterias hayan sido examinadas a fondo todavía (Woese, 1981). Así, el análisis del ARN en estos grupos ha resultado de valor predictivo, lo que parecería un argumento más a su favor.

Otros ejemplos de grupos en los que se han obtenido dendrogramas filogenéticos basados en las secuencias del ARN 5S son: varios grupos de eubacterias, protozoarios, hongos, plantas "superiores", meso y metazoarios, organelos celulares y, finalmente, una generalización para todos los organismos vivos (Hori y Osawa, 1987).

Algunos autores que han trabajado con las secuencias del ARN han señalado algunas de las limitaciones que existen en la construcción de dendrogramas filogenéticos a partir de ellas. Hori y Osawa consideran 4 limitaciones fundamentales (enfocadas principalmente al uso del rARN 5S):

1. La tasa de sustitución de nucleótidos en la molécula de ARN puede variar en cierto grado en diferentes grupos de organismos.
2. Los genes que codifican los ARN son miembros de familias multigénicas, por lo que no es improbable que tales ARN en los diferentes organismos comparados sean derivados de genes parálogos (los genes que codifican el ARN 5S no existen en una copia única en ningún organismo); aunque sus secuencias son generalmente similares —mostrando sólo algunas sustituciones de nucleótidos— se ha documentado la existencia de poblaciones con un ARN heterogéneo en algunas especies animales, lo que implica que a veces los genes del ARN 5S pueden diversificarse considerablemente entre organismos. Así, entre invertebrados y vertebrados podrían estarse comparando ARN 5S derivados de diferentes genes que se separaron dentro de su ancestro común y que han evolucionado independientemente.
3. El contenido de las bases nitrogenadas, guanina y citosina (G+C), en el ADN genómico se ha diversificado, yendo del 25 al 75% entre bacterias y ciertos grupos de eucariotas. Recientemente se encontró que el contenido de G+C del ARN 5S refleja más o menos el contenido de G+C en el ADN genómico de las eubacterias, mientras que en los eucariotas la presión de mutación que opera para alterar el contenido de G+C genómico no parece alterar el contenido de G+C en el ARN 5S. No se sabe realmente como tal presión de mutación sobre el ADN genómico afecta la tasa de sustitución nucleotídica en el ARN, pero cabe suponer que podría influenciarla en algunos grupos, y desviar las comparaciones intergrupales.
4. Las incertidumbres anteriores se pueden aplicar a todos los ARN. Los ARN más grandes (16s, 18s, 23s) son mejores que el ARN 5S para reducir los errores estándar en la construcción de los dendrogramas filogenéticos. Sin embargo, ciertas discrepancias en la posición filogenética de las arquibacterias entre dendrogramas obtenidos usando ARN 5S y 16S son atribuidas por Hori y Osawa a la influencia de la adición de un gran número de nucleótidos al ARN 16S que parece haber ocurrido después de la emergencia de los eucariotas (la longitud del ARN 16S de los procariontes es de aproximadamente 1473-1567 nucleótidos, mientras que en el caso de los eucariotas es de 1771-2305); tal adición podría haber influenciado la tasa de sustitución nucleotídica en algunos grupos, lo que alteraría el

análisis aunque sólo se consideren las "regiones conservadas" de los ARNr en el mismo. En contraste, la longitud del ARNr 5S parece ser más o menos constante en todos los organismos y Hori y Osawa suponen que tiene una tasa de cambio constante en todos ellos.

Puede observarse que las limitaciones 1, 3 y 4 convergen hacia un punto: la tasa de sustitución de nucleótidos probablemente no es la misma en todos los grupos de organismos. Ello puede traer consigo problemas al tratar de resolver las relaciones genealógicas entre diferentes grupos. Por ejemplo, Field *et al.* (1988) obtuvieron un dendrograma filogenético para todos los grupos principales de animales, basado en las secuencias del ARNr 18S. En tal dendrograma incluyeron 4 representantes de los artrópodos: un "miriápodo" (*Spiroboles* sp), un "cangrejo cacerola" (*Limulus* sp), un "camarón de las salinas" (*Artemia* sp) y una "mosca de la fruta" (*Drosophila* sp). La secuencia de *Limulus* es de cambio lento o *slow clock* (esto es, ha acumulado —relativamente— pocos cambios en su secuencia, desde su divergencia a partir de un ancestro común), mientras que otras secuencias son de cambio rápido (*fast clock*). Incluir especies de cambio rápido en una filogenia introduce un error sistemático; estas especies frecuentemente parecen divergir en el dendrograma más profundamente de lo que se supondría. *Limulus* y los otros tres artrópodos generalmente forman un sólo grupo cuando se incluyen todos en el dendrograma; pero en algunos dendrogramas, *Limulus* parece más cercano a los protostomados eucelomados, mientras que los otros tres artrópodos forman un grupo cerca de la base del dendrograma. Surgen así dos posibilidades: o la verdadera posición del grupo de los artrópodos es la rama más profunda del grupo de los protostomados, posición tomada por *Limulus* cuando se deja a los otros artrópodos fuera del dendrograma, o los artrópodos son polifiléticos, estando los quelicerados emparentados con los protostomados.

La limitación anterior también surge del método de análisis de la información obtenida a partir de la secuenciación del ARNr; esto es, de las limitaciones inherentes a los métodos matriciales para la construcción de dendrogramas filogenéticos; pero si bien es cierto que algunos caracteres sólo

pueden ser analizados de esta forma (por ejemplo, los análisis inmunológicos en su versión antigua), los datos de las secuencias de nucleótidos también son susceptibles de un análisis cladista, que es mucho más robusto que uno fenetista en la investigación de las relaciones genealógicas, como se mencionó previamente en este trabajo, en los apartados que tratan de los algoritmos y del análisis comparado del ADN nuclear (ver recuadro). En este contexto debe recordarse también que si, debido a la naturaleza de los datos obtenidos, el análisis comparativo sólo puede realizarse con métodos matriciales, existen algoritmos de agrupación para la construcción de dendrogramas filogenéticos (como los de Fitch y Margoliash, 1967 y Farris, 1972) que han demostrado ser mejores que aquellos que agrupan con base en la similitud total de lo comparado (como el UPGMA arriba citado).

Debe hacerse énfasis en que las "distancias evolutivas" son sólo indicadores de disimilitud total entre organismos pues expresan las diferencias entre dos secuencias nucleotídicas; por otra parte, el problema de conocer relaciones genealógicas es más complejo que el de clasificar secuencias exclusivamente por su semejanza; tal semejanza o disimilitud, genotípica o fenotípica, no equivale siempre a una estrecha relación genealógica (Wiley, 1981). Ello debería ser considerado siempre al pensar en obtener dendrogramas filogenéticos con metodologías feneticistas o simplemente al interpretar los análisis semejantes que aparecen en la literatura.

Finalmente, Field *et al.* (1988) señalaron que no existen medidas simples de confiabilidad para la posición de los vértices de ramificación de alguna rama particular en un dendrograma filogenético obtenido por los métodos arriba descritos. Tales métodos permiten calcular un error estadístico (error estándar) para el cálculo de cada "distancia evolutiva", pero la conversión de estos errores en cada par de secuencias a un error total para todos los vértices de ramificación es compleja, aunque los dendrogramas sean simples. Generalmente se considera "robusta" la posición de una secuencia dada sobre el dendrograma si es relativamente insensible a la elección de los nucleótidos usados en el análisis y es independiente de los organismos incluidos en el dendrograma. Y por último, si bien el análisis del ARNr puede ser de utilidad para analizar filogenias muy antiguas, no lo es para filogenias mucho más recientes.

CONCLUSIONES

Reig (1983) señaló que, en general, las filogenias macromoleculares tienen mayor grado de verosimilitud que las basadas sobre datos morfológicos o cromosómicos, ya que el cambio evolutivo a nivel morfológico es sumamente complejo y procede a diferentes ritmos (lo que determina que, por regla, no se correlacione linealmente con el tiempo evolutivo), mientras que el cambio cromosómico puede o no estar asociado a la divergencia evolutiva y puede proceder a tasas muy diferentes aun entre linajes relacionados de un mismo taxón; por el contrario, los cambios macromoleculares detectados en las proteínas o los ácidos nucleicos presentan un patrón evolutivo muy simple y constante y su tasa de cambio es proporcional a los tiempos de las sucesivas divergencias evolutivas.

A lo largo del presente trabajo se han ofrecido diversos argumentos y opiniones que contradicen los señalamientos de Reig, los cuales a su vez contrastan con las conclusiones que sobre el tema ha expuesto Wiley (1981): las técnicas bioquímicas no son la panacea para la reconstrucción filogenética, pero proporcionan datos útiles que es posible incorporar a los estudios filogenéticos cuando se está consciente de que tales datos son susceptibles de ser analizados de la misma manera que otros tipos de datos. El viejo adagio de que "entre más cerca del gene, más cerca de la verdad" simplemente no funciona. Entre más cerca del gene, hay más posibilidades de admitir homoplasias que son estructuralmente idénticas, pero de origen distinto, ya que no tienen características químicas o morfológicas que permitan el reconocimiento diferencial entre homólogos estructurales y filogenéticos.

Dados estos problemas, los datos bioquímicos deben analizarse cuidadosamente con métodos que sean relativamente robustos para tratar los problemas



de la no-homología. Técnicas tales como los algoritmos del árbol de Wagner usado por Mücke y Johnson (1976) o el método de parsimonia máxima de Fitch (1971), parecen aportar soluciones filogenéticas (*sensu* Hennig) para datos bioquímicos complejos. Los métodos fenéticos están bastante supeditados a los problemas de la no-homología como para dar respuestas satisfactorias.

De acuerdo a Cadle (1988) hay tres puntos básicos de controversia en la Sistemática Molecular: 1. La capacidad de los datos moleculares en el estudio de las relaciones evolutivas entre los taxa 2. La utilidad de los datos moleculares en la consideración de las relaciones cronísticas de una filogenia dada, y 3. Los métodos a usar en la reconstrucción de relaciones genealógicas a partir de datos moleculares, sean en la forma de secuencias o de medidas indirectas del parecido molecular. Este autor señaló lo siguiente respecto a esos puntos: los datos moleculares permiten proponer hipótesis de relaciones filogenéticas que pueden desarrollarse independientemente de aquellas derivadas de la morfología; esto conduce a un examen de congruencia entre los resultados obtenidos de los análisis de ambos conjuntos de evidencia, del mismo modo como ocurre con los datos ontogenéticos al examinar relaciones evolutivas. Los datos moleculares o los morfológicos, por sí solos, no son la respuesta a cada problema filogenético; existe una única historia filogenética, de modo que conjuntos de evidencia congruentes entre sí apuntan al descubrimiento de ese patrón histórico. En abstracto, no existe superioridad de un conjunto de caracteres sobre el otro; por otra parte, es de esperarse una gran complementariedad entre ambos tipos de estudios en la filogenia y clasificación de los seres vivos.

Las críticas a las bases lógicas subyacentes a los métodos de análisis de distancias, continúan siendo muy sólidas, por lo cual los resultados de la interpretación de las distancias en filogenia, con base en esos métodos, seguirán siendo cuestionadas por muchos autores. Hay dos problemas interrelacionados en la crítica al uso del concepto de reloj molecular: 1. la existencia real de los relojes y 2. su calibración, esto es, si en realidad existen; pues aunque la naturaleza dependiente del tiempo para la evolución molecular sea un hecho, su regularidad debe ser empíricamente probada en cada caso, calibrando el "reloj" en cada estirpe o linaje (Cadle, 1988). Por otra parte, es de esperarse que en un mismo linaje haya variación en las tasas de cambio del ADN nuclear y el ADN mitocondrial.

Hillis (1987) admitió que durante las dos últimas décadas se presentó un gran debate sobre los enfoques morfológico y molecular, pero que —actualmente— las investigaciones moleculares en la Sistemática son indispensables y se realizan de la misma forma en varios grupos taxonómicos. Continúa el desacuerdo en algunos aspectos, pero en otros hay cierto consenso; en cuanto a los conjuntos de caracteres (morfológicos y moleculares), éstos tienen tanto ventajas como desventajas, potencialidades y limitaciones. Lo anterior ocurre debido a la naturaleza de las técnicas de obtención u observación y sus resultados, así como a los métodos de análisis que se les aplican. Los datos morfológicos o moleculares parten de suposiciones, en ambos casos discutibles, y todavía debemos esperar una creciente polémica; en ambos casos, los fenómenos epigenéticos relacionados con la diferenciación en el organismo y sus expresiones fenotípicas prometen añadir una mayor complejidad a esta problemática.

Hillis (1987) enumeró algunas ventajas para ambos conjuntos de caracteres —moleculares y morfológicos—; aquí se resumen: en el primer caso destacó que los datos moleculares pueden ser mayores en número que los morfológicos *v. gr.* ADN + ARN. Los datos moleculares tienen mayores posibilidades de aplicación y comparación que los morfológicos, ya que estos últimos dependen de las influencias ambientales para su expresión, por lo que hay menor posibilidad de confusión en el estudio de tales caracteres moleculares (ADN + ARN + proteínas); además, los caracteres morfológicos pueden ser incomparables de una clase o de un phylum a otros. En el segundo caso advirtió que para los estudios morfológicos hay muestras muy numerosas y accesibles en museos, reconociendo que un número considerable de especies sólo son conocidas por el tipo o a través de pocos ejemplares, o bien es prohibitiva su recolección para estudios bioquímicos por su rareza o peligro de extinción. Los estudios morfológicos pueden utilizar la evidencia fósil, no así —en general— los estudios moleculares; el criterio de grupo externo para determinar posición

genealógica tiene el auxilio o la complementariedad del método ontogenético, de un modo mucho más accesible, en el caso de los datos morfológicos. Finalmente, los problemas para financiar los estudios sistemáticos favorecen a los estudios morfológicos, pues aun cuando hay técnicas caras en morfología, a menudo son baratas, mientras que en el caso de los caracteres moleculares éstas son normalmente más caras, tanto por el equipo como por los reactivos; además del mantenimiento del laboratorio y la necesidad de muestras frescas o "crioconservadas".

El desarrollo de métodos de análisis que admitan diferencias en las tasas de cambio, a la vez que permitan pruebas estadísticas de congruencia o consenso —entre conjuntos de datos obtenidos independientemente (morfológicos y moleculares)— quizá será prometedor en el futuro. Ciertas incongruencias aparentes pueden ser complementarias en la reconstrucción filogenética de un taxón dado.

Un punto central de esta problemática seguirá siendo el alto costo de los estudios; por caros que éstos resulten, sin embargo, continuarán efectuándose en aquellos taxa en los que las técnicas y los métodos de análisis puedan ser examinados y contrastados. Las técnicas de obtención de datos moleculares están progresando a pasos agigantados y sus resultados continúan siendo espectaculares, pero sólo los estudios combinando varios enfoques (Anatomía Comparada, estudios moleculares comparados y el uso del método ontogenético) permitirán una mayor certeza y cohesión de datos, hipótesis y teorías en la Historia Evolutiva.

Es evidente que muchas de las técnicas aquí mencionadas son muy poco accesibles o no lo son en absoluto para la mayoría de los taxónomos en México y de muchos países en el mundo. Precisamente, por esta razón, deben enfatizarse dos puntos: 1) Las fuentes de datos tradicionales (morfología y desarrollo ontogenético) siguen siendo relativamente accesibles y válidas y 2) Más que las propias técnicas de obtención de datos comparados, son importantes los conceptos subyacentes a su análisis. Entonces, para hacer Taxonomía de buen nivel, es totalmente prioritaria una sólida formación teórica y conceptual del taxónomo, en comparación al conocimiento de las técnicas sofisticadas disponibles, sobre todo cuando se está en las primeras etapas en la Taxonomía de muchos grupos. Tal formación, en principio, no es inaccesible en nuestro país y obtenerla debe ser un requisito indispensable para todo taxónomo profesional.

AGRADECIMIENTOS

Numerosos borradores de distintas versiones del original fueron amablemente corregidos y criticados por Luisa Alba Lois, Jesús Manuel León Cázares, Alfonso Delgado, Alfonso Neri García Aldrete, Jaime Martínez Medellín, Juan Carlos Morales Muciño, Juan Pedro Lacleste, Daniel Piñero Dalmau, Alfonso Torre Blanco y Víctor Valdés. Sin las valiosas aportaciones, consejos y literatura adicional que nos proporcionaron algunos de los colegas citados, este trabajo no hubiera podido llegar a término, por lo que expresamos nuestro agradecimiento a todos ellos. Para el desarrollo del trabajo se contó con el apoyo de la Facultad de Ciencias de la UNAM, su programa PSPA y el CONACYT.

LITERATURA CITADA

- Alston, R.E. y B.L. Turner. 1963. *Biochemical Systematics*. Prentice Hall, Englewood Cliffs, New Jersey.
- Arrighi, F.E. y T.C. Hsu. 1971. Localization of heterochromatin in human chromosomes. *Cytogenetics* 10: 81-86.
- Atchley, W.R. 1972. The chromosome karyotype in estimation of lineage relationships. *Syst. Zool.* 21: 199-209.
- Ayala, F.J. y J.A. Kiger, Jr. 1984. *Genética moderna*. Fondo Educativo Interamericano, Barcelona, España.



- Avise, J.C., R.A. Lansman y R.O. Shade. 1979. The use of restriction endonucleases to measure mitochondrial ADN sequence relatedness in natural populations. I. Population structure and evolution in the genus *Peromyscus*. *Genetics* 92: 279-295.
- Baum, B.R. 1984. Application of compatibility and parsimony methods at the infraespecific and generic levels in Poaceae. En *Cladistics: Perspectives on the reconstruction of evolutionary history* (Duncan, T. y T.F. Stuessy, eds). Columbia University Press, New York. 312 pp.
- Baverstock, P.R., S.R. Cole, B.J. Richardson y C.H.S. Watts. 1979. Electrophoresis and cladistics. *Syst. Zool.* 28: 214-219.
- Bendz, G. y J. Santesson (eds.). 1974 *Chemistry in botanical classification*. Academic Press. New York. 340 pp.
- Bickham, J.W. y J.L. Carr. 1983. Taxonomy and phylogeny of the higher categories of cryptodiran turtles based on cladistic analysis of chromosomal data. *Copeia* 1983: 918-932.
- Bisby, F.A., J.G. Vaughan y C.A. Wright (Eds.) 1980. *Chemosystematics: Principles and practice*. Academic Press, Nueva York.
- Blackwelder, R.E. 1967. *Taxonomy: a text and reference book*. John Wiley & Sons. New York. 698 pp.
- Bogart, J.P. 1970. Systematic problems in the amphibian family Leptodactylidae (Anura) as indicated by karyotype analysis. *Cytogenetics* 9: 369-383.
- Cadle, J.E. 1988. Phylogenetic relationships among advanced snakes: a molecular perspective. *University California publications (Zoology)* 119: 1-77.
- Cane, J.H. 1983. Preliminary chemosystematics of the Andrenidae and exocine lipid evolution of the short-tongued bees (Hymenoptera: Apoidea). *Syst. Zool.* 32: 417-430.
- Carr, R.L. 1987. In search of eve. *The Sciences*: 30-37.
- Carson, H.L. y K. Y. Kaneshiro. 1976. *Drosophila* of Hawaii: Systematics and ecological genetics. *Ann. Rev. Ecol. Syst.* 7: 311-345.
- Champion, A.B., E.M. Prager, D. Wachter y A.C. Wilson. 1974. Microcomplement fixation. En Wright (ed.), *Biochemical and immunological taxonomy of animals*. Academic Press, London. pp 397-416.
- Darlington, C.D. 1958. *The evolution of genetic systems*. Basic Books, Inc., New York.
- DeBry, R.W. y N.A. Slade. 1985. Cladistic analysis of restriction endonuclease cleavage maps within a maximum-likelihood framework. *Syst. Zool.* 34: 21-34.
- Dobzhansky, T., F.J. Ayala, G.L. Stebbins y J.W. Valentine. 1977. *Evolution*. W.H. Freeman and Co., San Francisco, California.
- Duellman, W.E. 1985. Systematic Zoology: slicing the Gordian Knot with Ockham's Razor. *Amer. Zool.* 21: 199-209.
- Ehrlich, P.R. y A.H. Ehrlich. 1967. The phenetic relationships of the butterflies. I. Adult Taxonomy and the nonspecificity hypothesis. *Syst. Zool.* 16 (4): 301-317.
- Farris, J.S. 1972. Estimating phylogenetic trees from distance matrices. *Amer. Natur.* 106: 645-668.
- Farris, J.S. 1978. Inferring phylogenetic trees from chromosome inversion data. *Syst. Zool.* 27: 275-284.
- Farris, J.S. 1979. Biosystematics in Agriculture (revisión). *Syst. Zool.* 28: 239-243.
- Farris, J.S., A.G. Kluge y M.F. Mickevich. 1979. Paraphyly of the *Rana boylei* species group. *Syst. Zool.* 28: 627-634.
- Farris, J.S., A.G. Kluge y M.F. Mickevich. 1982. Immunological distance and the phylogenetic relationships of the *Rana boylei* species group. *Syst. Zool.* 31: 479-491.
- Field, K.G., G.J. Olsen, D.J. Lane, S.J. Giovannoni, M.T. Ghiselin, E.C. Raff N.R. Pace y R.A. Raff. 1988. Molecular phylogeny of the animal kingdom. *Science* 239: 748-753.
- Fitch, W.M. 1970. Distinguishing homologous from analogous proteins. *Syst. Zool.* 19: 99-113.
- Fitch, W.M. 1971. Toward defining the course of evolution: Minimum change for a specific tree topology. *Syst. Zool.* 20: 406-416.
- Fitch, W.M. 1973. Aspects of molecular evolution. *Ann. Rev. Genet.* 7: 343-380.
- Fitch, W.M. 1984. Cladistic and other methods: Problems, pitfalls, and potentials. En T. Duncan and T.F. Stuessy (eds.), *Cladistics: Perspectives on the reconstruction of phylogenetic history*, pp. 221-252. Columbia Univ. Press, New York.
- Fitch, W.M. y E. Margoliash. 1967. The construction of phylogenetic trees. *Science* 155: 1200-1203.
- Futuyma, D. 1986. *Evolutionary Biology*. Sinauer, Sunderland, Massachusetts.
- Geiger, H. 1981. Enzyme electrophoretic studies on the genetic relationships of pierid butterflies (Lepidoptera: Pieridae). I. European taxa. *Jour. Res. Lep.* 19: 181-195.
- Goodman, M. (Ed.) 1982. *Macromolecular sequences in Systematic and Evolutionary Biology*. Plenum Press, New York. 418 pp.
- Goodman, M. y G.W. Moore. 1971. Immunodiffusion systematics of the primates I. The Catarrhini. *Syst. Zool.* 20: 19-62.
- Goodpasture, C. y S.E. Bloom. 1975. Visualization of nucleolar organizer regions in mammalian chromosomes using silver staining. *Chromosoma* 53: 37-50.

- Gorman, G.C. 1970. Chromosomes and the systematics of the family Teiidae (Sauria, Reptilia). *Copeia* 1970: 369-383.
- Gorman, G.C., D.G. Buth y J.S. Wyles. 1980. *Anolis* lizards of the eastern Caribbean: A case study in evolution. III. A cladistic analysis of albumin immunological data, and the definition of species groups. *Syst. Zool.* 29: 143-158.
- Haiduk, M.W. y R.J. Baker. 1982. Cladistical analysis of G-banded chromosomes of nectar feeding bats (Glossophaginae: Phyllostomidae). *Syst. Zool.* 31: 252-265.
- Hall, W.P. y R.K. Selander. 1973. Hybridization of karyotypically differentiated populations in the *Sceloporus grammicus* complex (Iguanidae). *Evolution* 27: 226-242.
- Hennig, W. 1968. *Elementos de una Sistemática Filogenética*. EUDEBA, Argentina. 317 pp.
- Hillis, D.M. 1987. Molecular versus morphological approaches to systematics. *Ann. Rev. Ecol. Syst.* 18: 23-42.
- Hori, H. y S. Osawa. 1987. Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences. *Mol. Biol. Evol.* 4 (5): 445-472.
- Hubby, J.L. y R.C. Lewontin. 1966. A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics* 54: 577-594.
- Humpries, C.J. 1982. Vicariance Biogeography in Mesoamerica. *Ann. Miss. Bot. Gard.* 69: 444-463.
- Huxley, J.S. (ed.) 1940. *The new Systematics*. Oxford University Press. London.
- Johnson, G.B. 1977. Assessing electrophoretic similarity: the problem of hidden heterogeneity. *Ann. Rev. Ecol. Syst.* 8: 309-328.
- Kitching, I.J. 1985. Allozyme variation in the milkweed butterflies (Lepidoptera: Danainae). *Zool. Jour. Linn. Soc.* 86: 367-389 + 1 fig.
- Kluge, A. y J.S. Farris 1969. Quantitative phyletics and the evolution of the anurans. *Syst. Zool.* 18: 1-32.
- Llorente, J. 1986. Algunas ideas de la Teoría Sistemática Contemporánea: conceptos en cladismo. *Ciencias No. Especial 1*. pp. 74-87. (Reimpreso en este número)
- Maxson, L.R. y A.C. Wilson. 1975. Albumin evolution and organismal evolution in tree frogs (Hylidae). *Syst. Zool.* 24: 1-15.
- Mickevich, M.F. 1978. Taxonomic congruence. *Syst. Zool.* 27: 143-158.
- Mickevich, M.F. y M.S. Johnson. 1976. Congruence between morphological and allozyme data in evolutionary inference and character evolution. *Syst. Zool.* 25: 260-270.
- Nadler, C.F. 1966. Chromosomes and systematics of the American ground squirrels of the subgenus *Spermophilus*. *Jour. Mammal.* 47: 579-596.
- Nei, M. 1972. Genetic distance between populations. *Amer. Natur.* 106: 283-292.
- Oxford, G.S. y D. Rollinson. 1983. *Protein polymorphism: adaptative and taxonomic significance*. Academic Press, Nueva York.
- Pardue, M.L. y J.G. Gall. 1970. Chromosomal localization of mouse satellite DNA. *Science* 168: 1356-1358.
- Pashley, D.P. 1983. Biosystematic study in the Tortricidae (Lepidoptera), with a note on evolutionary rates of enzymes. *Ann. Ent. Soc. Amer.* 76: 139-148.
- Patterson, C. 1981. Methods of paleobiogeography. 446-489. En G. Nelson y D.E. Rosen (Eds.). *Vicariance Biogeography: A critique*. Columbia University Press. New York.
- Radford, A.E., W.C. Dickison, J.R. Massey, C.R. Bell, et al. 1974. *Vascular plant systematics*. Harper and Row, New York.
- Reig, O.A. 1983. Métodos no morfológicos en la sistemática animal contemporánea. *Informe final del IX Congreso Latinoamericano de Zoología, Perú*. Oct. de 1983, p. 242 (resumen).
- Robinson, R. 1971. *Lepidoptera Genetics*. Pergamon Press, Oxford. 687 pp.
- Rogers, J.S. 1972. Measures of genetic similarity and genetic distance. *Univ. Texas Publ.* 7213: 145-153.
- Sarich, V.M. y A.C. Wilson. 1967. Immunological time scale for hominid evolution. *Science* 158: 1200-1203.
- Sibley, C.G. y J.E. Ahlquist. 1983. Phylogeny and classification of birds based on the data of DNA-DNA hybridization. En R.F. Johnston (ed.), *Current ornithology*, Vol. 1, pp. 249-292. Plenum Press, New York.
- Sneath, P.H.A. y R.R. Sokal. 1973. *Numerical taxonomy. The principles of numerical classification*. Freeman, San Francisco. 573 pp.
- Stangl, F.B., Jr. y R.J. Baker. 1984. Evolutionary relationships in *Peromyscus*: Congruence in chromosomal, genic, and classical data sets. *J. Mammal.* 65: 643-654.
- Stebbins, G.L., Jr. 1971. *Chromosomal evolution in higher plants*. Addison-Wesley Publishing Co., Inc. Reading, Massachusetts.
- Wallace, D.G., M.C. King y A.C. Wilson. 1973. Albumin differences among ranid frogs: taxonomic and phylogenetic implication. *Syst. Zool.* 22: 1-13.
- White, M.J.D. 1973. *Animal cytology and evolution* (3rd Ed.) Cambridge University Press.
- White, M.J.D. 1978. *Modes of speciation*. W.H. Freeman and Co., San Francisco, California.
- Wiley, E.O. 1978. The evolutionary species concept reconsidered. *Syst. Zool.* 27: 17-26.
- Wiley, E.O. 1980. Phylogenetic Systematics and Vicariance Biogeography. *Syst. Bot.* 5 (2): 194-220.
- Wiley, E.O. 1981. *Phylogenetics*. John Wiley and Sons, New York.
- Wilson, A. 1975. Relative rates of evolution of organisms and genes. *Stadler Genet. Symp.* 7: 117-134.
- Wilson, A.C., R.L. Cann, S.M. Carr, M. George, et al. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. Jour. Linn. Soc.* 26: 375-400.
- Wilson, A.C., S.S. Carlson y T.J. White. 1977. Biochemical evolution. *Ann. Rev. Biochem.* 46: 573-639.
- Woese, C.R. 1981. Archaeobacteria. *Sci. Am.* 244: 94-107.
- Wright, D.A., C.M. Richards, J.S. Frost, A.M. Camozzi y B.J. Kunz. 1983. Genetic mapping in amphibians. Isozymes: Current topics. *Biol. Med. Res.* 10: 287-311.
- Yates, T.L., R.J. Baker y R.K. Barnett. 1979. Phylogenetic analysis of karyological variation in three genera of peromyscine. *Syst. Zool.* 28: 40-48.

Ilustraciones: Cosas de familia de Rafael López Castro y Felipe Garrido