

## REVISIÓN DE LOS MÉTODOS FILOGENÉTICOS DE SECUENCIAS: UNA CRÍTICA A LA TEORÍA DE KIMURA SOBRE EL NEUTRALISMO DE LA EVOLUCIÓN MOLECULAR

ENRIQUE MELÉNDEZ-HEVIA\*, RAQUEL R. RAPOSÓ†, RUTH MELÉNDEZ\* Y HÉCTOR CABEZAS\*

Universidad de La Laguna, \* Departamento de Bioquímica, Facultad de Biología/ † Departamento de Química Física, Facultad de Química/ 38206 Tenerife, Canary Islands, Spain.

### ABSTRACT

The approaches for phylogenetic studies based on protein and nucleic acid sequence analysis have yielded a remarkable advance in our knowledge of biodiversity in the last 30 years. However its correct use has many problems, so data processing is a hard task. The general hypothesis on which all these methods are based is the so-called neutral theory of molecular evolution, presented by Kimura. In this work we present a review of these methods and of their main problems, and present a criticism to Kimura's theory, based on the results of our research in the last years. A conclusion of this work is that the correct interpretation of molecular phylogeny data demonstrates that arthropods and annelids are artificial groups with no phylogenetic consistency.

### RESUMEN

Los métodos para estudiar la filogenia basados en el análisis de secuencias de proteínas y ácidos nucleicos han producido un avance considerable en nuestro conocimiento sobre la biodiversidad en los últimos 30 años. Sin embargo, su correcto uso tiene muchos problemas, por lo que el tratamiento de los datos es complicado. La hipótesis general en la que están basados estos métodos es la llamada *teoría del neutralismo de la evolución molecular* formulada por Kimura. En este trabajo hacemos una revisión de estos métodos y de sus principales problemas, y presentamos una crítica a la teoría de Kimura, basada en los resultados de nuestra investigación en los últimos años. Una conclusión de este trabajo es que la correcta interpretación de los datos de filogenia molecular demuestra que los artrópodos y los anélidos son grupos artificiales sin consistencia filogenética.

## INTRODUCCIÓN

La construcción del árbol filogenético es uno de los problemas más trascendentales de la Biología, comparable en cierto modo con lo que fue para la Química la construcción de la Tabla periódica. Esta tarea no ha perdido actualidad desde las primeras consideraciones de Darwin en 1859 y los primeros intentos de Haeckel en 1966 [3, 42]. La filogenia es el estudio de las relaciones de parentesco entre los distintos grupos de seres vivos. El criterio general para investigar es el establecimiento de *homologías*. Una homología—o marcador filogenético—es una característica familiar, propia del grupo, que no depende de las condiciones ambientales o del desarrollo y diversificación de una rama, sino que se mantiene inalterable en la línea desde su aparición. El problema más importante de la investigación filogenética es la búsqueda de buenas homologías. En general, las mejores son las que carecen de valor selectivo. Como veremos, el mundo molecular es un buen campo para encontrar homologías muy valiosas, pero su búsqueda y la correcta interpretación de esos datos están llenas de obstáculos.

En los últimos años, con el desarrollo de nuevos métodos experimentales ha resurgido la filogenia molecular: el estudio de relaciones filogenéticas basado en las secuencias de nucleótidos en ácidos nucleicos, como años antes lo habían sido las de aminoácidos en proteínas. Estos métodos están basados en que estas secuencias cambian a lo largo del tiempo con velocidad constante; esa velocidad es distinta para cada proteína o ácido nucleico de acuerdo con ciertas características de su papel en la célula pero, en principio, sin diferencias entre especies de distintos grupos taxonómicos, siempre que la molécula en cuestión desempeñe el mismo papel en todos ellos.

Entre los años 1974 y 1983 [22] Kimura formuló la “teoría del neutralismo en la evolución molecular”, la cual se ha interpretado como la razón de esa velocidad constante. El uso extenso que están teniendo estos métodos en los últimos años ha demostrado que existen muchos problemas para su uso correcto. Algunos de estos problemas son prácticos y pueden resolverse analizando más muestras, pero otros son teóricos y su solución puede ser mucho más difícil o imposible. En este trabajo hacemos una revisión de esos problemas y presentamos una crítica a la tesis de Kimura basada en nuestros resultados de los últimos años sobre la optimización del metabolismo. En la última sección presentamos también unas conclusiones sobre la situación filogenética de los Artrópodos y de los Anélidos mostrando que ambos grupos son artificiales sin fundamento filogenético.

Es importante tener un *reloj evolutivo* para medir el tiempo de los acontecimientos que han ocurrido a lo largo de la evolución. Las primeras herramientas que se usaron en este sentido fueron los relojes paleontológicos, que están basados en el cambio morfológico de una especie producido continuamente en el tiempo, como el ejemplo mostrado en la figura 1. Después se desarrolló la datación de los períodos geológicos midiendo la desintegración de elementos radiactivos. El tercer paso fueron los relojes moleculares. A partir de los primeros trabajos de Zuckerkandl y Pauling [44], y de Margoliash [28], comparando secuencias de aminoácidos de hemoglobina y de citocromo *c*, Zuckerkandl y Pauling [45] propusieron que la velocidad de variación de la secuencia de aminoácidos sería aproximadamente constante para cada proteína a lo largo de su evolución, la misma en cada rama filogenética divergente, aunque distinta para cada proteína. Esto abrió la puerta de la filogenia molecular. A partir de entonces, durante los últimos 30 años, nuestro conocimiento de la filogenia ha progresado mucho con el desarrollo y la aplicación de esta idea, estudiando secuencias de proteínas y ácidos nucleicos como marcadores moleculares de la evolución. Véanse revisiones en [10, 31, 39 y 43].

Supongamos que una proteína tal como el citocromo *c* varía a velocidad constante mediante cambios aleatorios en su secuencia de aminoácidos. En principio cualquier varia-

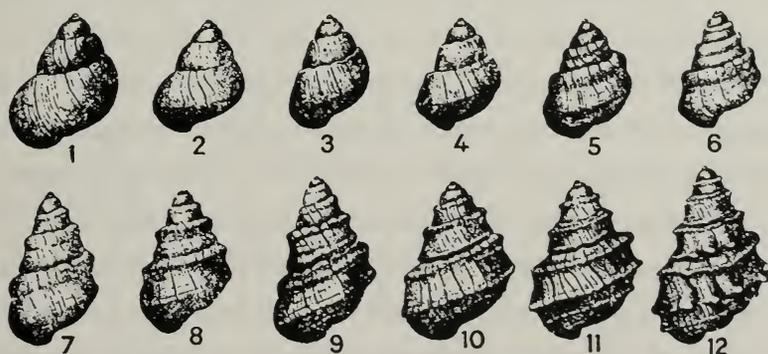


Figura 1.—Serie de diferentes especies fósiles del Gasterópodo *Vivipara* que han vivido en Hungría desde el Terciario (65 Ma) hasta la actualidad. La serie es una interesante demostración paleontológica de la evolución biológica, y además es un buen reloj paleontológico. Como su variación ha sido continua, cada forma es característica de un corto período de tiempo [15].

ción que se produzca en su secuencia puede ser virtualmente compartida por todos los miembros de la población, dado que hay libertad de intercambio de material genético entre todos ellos. En realidad una especie genéticamente definida debe ser homogénea para tales variaciones. La proposición anterior significa que las variaciones en la secuencia de una proteína se fijan en toda la población a una velocidad constante. Supongamos ahora que algunos individuos quedan aislados genéticamente produciéndose una nueva rama evolutiva divergente; la proteína continuará variando a la misma velocidad en cada grupo, pero sus cambios (aleatorios) en cada comunidad genéticamente aislada serán diferentes, de forma que cuanto más distante en el tiempo haya sido la divergencia es de esperar que haya más diferencias entre ellos. Este principio es la base de la filogenia molecular. La figura 2 muestra un ejemplo de estos resultados: el primer árbol filogenético molecular que se construyó, basado en la secuencia del citocromo c.

El principal problema de estos métodos es que la constancia de la velocidad es una propiedad que no siempre se cumple. Cuando ocurre esto—cuando la misma proteína cambia más deprisa en una especie que en otra—puede haber un conflicto importante pues una especie, o el grupo que representa, podría quedar muy deslocalizado de su posición filogenética real. Sarich y Wilson [40] han propuesto el *test de velocidad relativa* para resolver estos problemas. Este método no requiere un conocimiento previo del tiempo de divergencia, sino el uso de una tercera especie, externa al grupo, como referencia. Únicamente se necesita saber que tal especie se separó de las dos estudiadas antes que ellas entre sí. Esto se usa generalmente en la actualidad, y es muy útil para algunos estudios. Por ejemplo, se pueden estudiar las relaciones entre familias de mamíferos tomando un pájaro como referencia. Obviamente la especie de referencia debe ser representativa de la velocidad media del reloj; no puede ser una especie de reloj rápido ni lento. El problema cuando se investigan las relaciones entre grandes grupos es asegurarse que la especie (o mejor, varias especies) de referencia sea realmente externa al problema que se estudia. Esto es relativamente fácil entre grupos próximos, pero cuando se analizan grupos potencialmente muy separados, la elección de la especie de referencia puede ser un problema importante.

Supongamos que estamos estudiando con RNA las divergencias entre dos Mamíferos de familias distintas, y tomamos un pájaro como especie externa (pues sabemos bien por otros métodos que la rama que condujo a las Aves se separó de la que llevó a los Mamíferos antes de que las familias de Mamíferos se separasen entre sí). Supongamos ahora que ve-

mos más diferencias en la secuencia del RNA entre los dos Mamíferos que entre uno de ellos y el pájaro. Entonces, está claro que la molécula ha cambiado más deprisa en un Mamífero que en el otro, y se concluye que uno de ellos es una “especie de reloj rápido” o el otro una “especie de reloj lento”. Esto es normal y ocurre con cierta frecuencia; el fundamento teórico de este efecto y la forma de corregirlo se explican más abajo. Supongamos ahora que estudiásemos las relaciones entre un murciélago y ciertas aves (con la hipótesis errónea de que el murciélago es una especie relacionada directamente con las Aves) y tomásemos un caballo como especie externa de referencia. Como el caballo y el murciélago (ambos Mamíferos) son parientes próximos, observaríamos más diferencias entre el murciélago y cualquier ave que entre el murciélago y el caballo. Si persistiésemos en la hipótesis equivocada de que el murciélago es pariente más próximo de las aves que del caballo, entonces el murciélago sería calificado como una *especie de reloj rápido*.

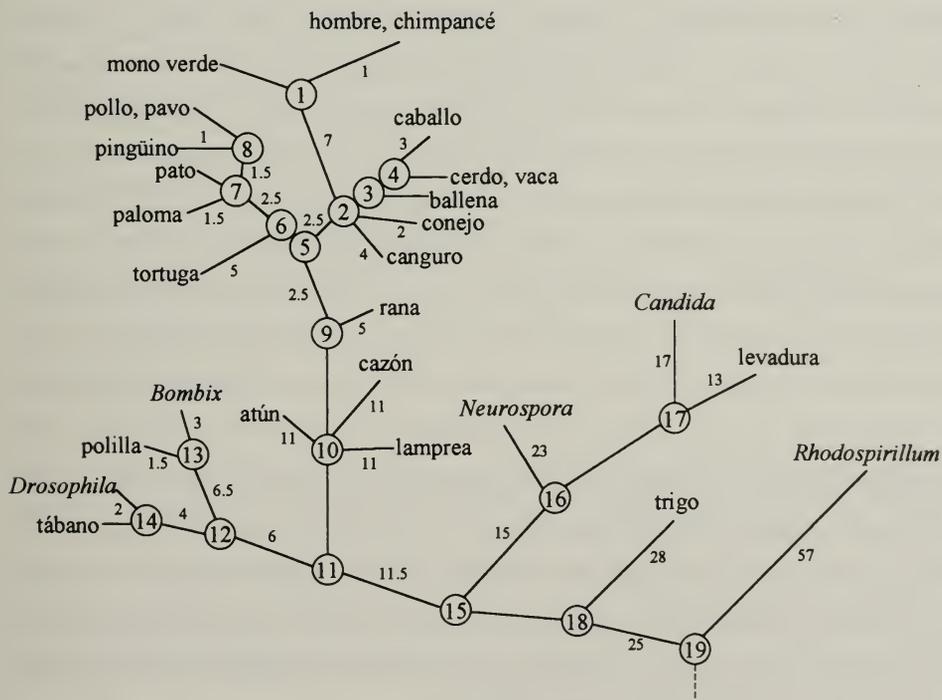


Figura 2.—Árbol filogenético basado en el citocromo c. Las distancias filogenéticas expresadas en cada rama se han calculado por las diferencias de aminoácidos entre cada dos especies, como se muestra en la tabla I. Los círculos numerados indican los puntos de radiación evolutiva [10].

Esto ha ocurrido con varios estudios que se han hecho sobre la filogenia de los Artrópodos: Boore *et al.* [6], tratando de investigar si los Artrópodos eran un grupo homogéneo, tomaron dos Nematodos como puntos de referencia, suponiendo que su divergencia de la línea de los Artrópodos (lo que es admitir *a priori* que los artrópodos constituyen una línea) debió ocurrir mucho antes que la de los artrópodos entre sí. Dos años más tarde, el grupo de Lake [2] usando Equinodermos y Cnidarios como puntos de referencia llegaron a la conclusión opuesta: los Nematodos estarían realmente *dentro* del grupo de los Artrópodos (¡si esto fuese cierto, Boore *et al.* habrían tomado como especies externas de referencia unas que pertenecerían al mismo grupo! sin embargo, ninguno de los dos grupos de investigadores interpretó correctamente sus resultados).\*

Este principio es la base para determinar la velocidad evolutiva de una proteína o de un ácido nucleico. La tabla I muestra las diferencias en las secuencias de aminoácidos del citocromo c entre 20 especies de varias líneas filogenéticas (16 animales, una planta y tres hongos). El árbol filogenético de la figura 2 se ha construido con estos datos. Además los mismos datos se pueden usar para determinar la velocidad evolutiva de la proteína, como se muestra en la figura 3. Cada punto de la gráfica se determina por la diferencia de aminoácidos entre dos especies y el tiempo transcurrido desde la divergencia entre sus dos líneas. Se supone que ese tiempo se conoce previamente por el registro fósil, o eventualmente por cualquier otra fuente. La tabla II muestra las velocidades evolutivas de varias proteínas determinadas con este método [10, 31, 43]. Ahora, una vez conocida la velocidad de evolución de una proteína, la curva puede usarse para determinar la edad de la divergencia entre otros grupos. Además, como puede verse, unas proteínas evolucionan más deprisa que otras, lo que es de esperar, de acuerdo con el razonamiento expuesto arriba. Así podemos disponer de relojes diferentes para estudiar problemas más amplios o más concretos. Por ejemplo, una molécula rápida, como el péptido C de la Insulina que evoluciona muy deprisa será apropiada para estudiar relaciones filogenéticas cercanas que se han producido en períodos cortos de tiempo, tales como familias de mamíferos, mientras que una proteína lenta, como el citocromo c, será buena para estudiar relaciones lejanas, como la revisión completa de todos los metazoos.

---

\* Este tratamiento poco riguroso de los datos ha llevado a estos autores [2] a proponer el “superphylum” *Ecdisozoa* que englobaría Artrópodos, Nematodos y otros grupos que tienen en común el fenómeno de la muda. En realidad este grupo no tiene ningún fundamento morfológico ni molecular.

Tabla I

Diferencias en la secuencia del citocromo c [33,34]. Como puede verse, cuanto mayor es la distancia filogenética entre dos especies mayor es el número de aminoácidos distintos. Este tipo de datos se usa para construir árboles filogenéticos como el de la figura 1, y para calcular velocidades evolutivas (ver la figura. 3 y la tabla II).

	hombre	mono Rhesus	caballo	burro	cerdo, vaca, oveja	perro	ballena	conejo	canguro	pollo, pato	pingüino	pato de Pekin	serpiente de cascabel	tortuga	sapo	atún	<i>Drosophila</i>	mariposa de seda	trigo	<i>Neurospora</i>	levadura	<i>Candida</i>
<i>Candida</i>	51	51	51	50	50	49	50	50	51	51	50	51	51	53	51	48	47	47	50	42	27	0
levadura	45	45	46	45	45	45	45	45	46	46	45	46	47	49	47	47	45	47	47	41	0	
<i>Neurospora</i>	48	47	46	46	46	46	46	46	49	47	48	46	47	49	49	48	41	47	54	0		
Trigo	43	43	46	45	45	44	44	44	47	46	46	46	46	46	48	49	45	45	0			
Mariposa de seda	31	30	29	28	27	25	27	26	28	28	27	27	31	28	29	32	14	0				
<i>Drosophila</i>	27	26	22	22	22	21	22	21	24	23	24	22	29	24	22	24	0					
atún	21	21	19	18	17	18	17	17	18	17	18	17	26	18	15	0						
sapo	18	17	14	13	11	12	11	11	13	11	12	11	24	10	0							
Tortuga	15	14	11	10	9	9	8	9	11	8	8	7	22	0								
Serpiente cascabel	14	15	22	21	20	21	19	18	21	19	20	17	0									
pato de Pekin	11	10	10	9	8	8	7	6	10	3	3	0										
pingüino	13	12	12	11	10	10	9	8	10	2	0											
pollo, pato	13	12	11	10	9	10	9	8	12	0												
canguro	10	11	7	8	6	7	6	6	0													
conejo	9	8	6	5	4	5	2	0														
ballena	10	9	5	4	2	3	0															
perro	11	10	6	5	3	0																
cerdo, vaca, oveja	10	9	3	2	0																	
burro	11	10	1	0																		
caballo	12	11	0																			
mono Rhesus	1	0																				
hombre	0																					

La principal hipótesis en la que están basados estos métodos es que las macromoléculas evolucionan a velocidad constante. Sin embargo esto tiene varios problemas, como lo demuestra este ejemplo: la separación de las líneas evolutivas que condujeron a las aves y a los mamíferos ocurrió en el Pérmico, hace unos 275 millones de años (Ma), mientras que la separación entre las diferentes líneas de insectos ocurrió después, en el Triásico (hace unos 225 Ma). Como la divergencia entre los vertebrados ocurrió antes que la divergencia entre los Insectos, es de esperar que las diferencias de la secuencia de aminoácidos entre dos órdenes de Insectos, como Lepidópteros y Dípteros fuese menor que las que pueda haber entre dos Vertebrados distantes, tales como Aves y Mamíferos, pero esto no es así siempre:

La velocidad media estimada de variación de secuencia en el citocromo c es de un aminoácido de cada 100 cambiado cada 20,16 millones de años (véase la figura 3 y la tabla II). El citocromo c tiene 104 aminoácidos, lo que hace que su velocidad evolutiva global sea:

$$v_{\text{evolución}} = \frac{\text{aa que cambian}}{\text{tiempo}} = \frac{\left(\frac{104}{100}\right) \text{aa}}{20,16 \text{My}} = \frac{1 \text{aa}}{19,38 \text{Ma}}$$

es decir, un aminoácido cambia en el total de la molécula de proteína cada 19,38 Ma en cualquier rama filogenética. Como los órdenes de insectos se separaron hace 225 Ma es de esperar que la diferencia entre el citocromo c de una mosca y de una mariposa sea de

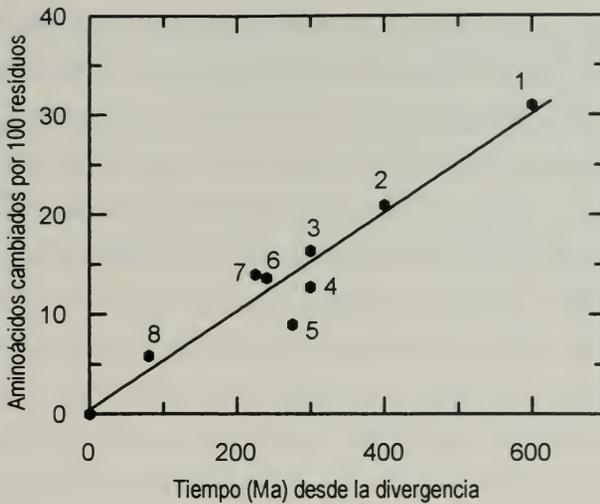
$$\text{aa que cambian} = v_{\text{evolución}} \cdot \text{tiempo} = \frac{1 \text{aa}}{19,38 \text{Ma}} \cdot 225 \text{Ma} = 11,61 \text{ aminoácidos}$$

por la misma razón, entre un pájaro y un mamífero debía haber  $275/19,38 = 14,19$  aminoácidos diferentes. Sin embargo, los datos experimentales no concuerdan con las predicciones de la teoría: hay 14 aminoácidos diferentes entre la mosca del vinagre (*Drosophila*) y la mariposa de la seda (*Bombix*)—ciertamente más de 11,61; hay 8 aminoácidos diferentes

**Tabla II**

Las distintas proteínas evolucionan a velocidades diferentes. Estos datos se han determinado a partir de las pendientes de curvas como la de la figura 3. El período evolutivo es el tiempo en millones de años (Ma) requerido para producir una diferencia de un 1% en la secuencia de aminoácidos. Las proteínas lentas como las histonas tienen períodos evolutivos muy largos. Por el contrario, los fibrinopéptidos evolucionan muy deprisa. El citocromo c es una proteína de velocidad media, y así útil para estudios filogenéticos amplios (véase la figura 2) [1, 31, 43].

Proteína	Período evolutivo (Ma)
Histona H4	400
Histona H3	300
Histona H2B	60
Colágeno $\alpha$ -1	36
Citocromo c	20
Lactato deshidrogenasa (B)	19
Insulina (péptidos A and B)	14
Mioglobina	6
Anhidrase carbónica B	4
Albúmina de suero	3
Insulina (péptido C)	1.9
Fibrinopéptido A	1.7
Fibrinopéptido B	1.1



**Figura 3.—Velocidad de evolución de una proteína (citocromo c).**— Cada punto corresponde a la comparación entre dos especies: diferencias en la secuencia de aminoácidos frente al tiempo de divergencia. En general puede admitirse que hay una relación lineal entre estas dos variables, aunque hay puntos que se desvían notablemente (véase el texto). La pendiente de esta recta es 0,0496, que significa el tanto por ciento de aminoácidos de la proteína que han cambiado cada millón de años. El valor inverso es el período evolutivo (20.16 Ma para que se produzca un 1% de cambio). Los datos de diferencias de aminoácidos se han tomado de la tabla I, y los datos de tiempo se han tomado del registro fósil, para las siguientes divergencias: 1, vertebrados/insectos; 2, anfibios/peces; 3, 4 y 5, mamíferos/reptiles-aves; 6, reptiles/aves; 7, insectos entre sí; 8, mamíferos entre sí. Los puntos que están por encima de la línea recta están promovidos por especies de reloj rápido, y los que están por debajo de la línea, por especies de reloj lento. El punto 5 es el reloj lento comentado en el texto. Datos moleculares tomados de [1, 10, 31] y más puntos de la tabla I, y los datos de tiempo de [29]. Las velocidades de evolución de proteínas de la tabla II se han determinado con gráficas como ésta.

entre el perro y el pato de Pekin, y 10 entre el perro y el pingüino o el pato (bastante menos de los 14,9 que le corresponderían). El reloj del citocromo c ha ido más deprisa de lo normal en los insectos y más despacio en los vertebrados.

Este caso demuestra pues, que la misma proteína puede actuar eventualmente como un reloj rápido o lento en distintas líneas evolutivas. Sin embargo, muchas proteínas exhiben una velocidad constante de evolución que se puede determinar estadísticamente. Estos estudios deben tener muchos datos a fin de que sean estadísticamente significativos. En la figura 3, los puntos 3, 4 y 5 describen el mismo efecto: la divergencia entre las dos ramas de vertebrados que condujeron a mamíferos y reptiles, respectivamente. Cada punto representa un caso particular de diferencia entre dos especies.

No siempre es fácil saber si una especie es un reloj rápido o lento. El problema es que la especie de referencia debe ser establecida previamente como externa a los grupos (separada de ellos antes de su diversificación), es decir, que su posición filogenética debe ser aceptada *a priori*. Cuando esto no se conoce con certeza, antes de afirmar que una especie es un reloj rápido o lento se debe analizar una muestra muy amplia de muchas especies. Este criterio, sin embargo, no se ha aplicado con rigor. Por ejemplo, Field et al. [11] analizaron cuatro artrópodos (un insecto, un miriápodo, un crustáceo y un quelicerado). Sus resultados demostraron un parentesco entre ellos mucho más lejano de lo que cabría esperar si los artrópodos fuesen un grupo filogenéticamente homogéneo. Entonces decidieron que tres de ellos eran especies de reloj rápido y los descartaron del estudio. Por motivos similares, el grupo de Lake [2] descarta tres Nematodos de cuatro analizados, etc.

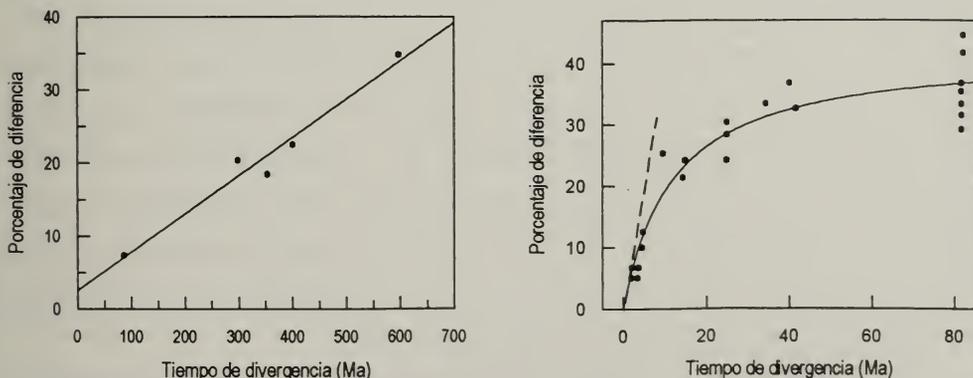
La velocidad de variación de una macromolécula depende de dos variables: la probabilidad de que ocurra cada variación, y las posibilidades de sea aceptada. Consideremos la primera: lo cierto es que no todos los cambios pueden ocurrir con la misma probabilidad. Por razones de mecanismos químicos, ciertos cambios ocurren con mayor facilidad. Este problema ha sido estudiado profundamente para el caso de la sustitución de nucleótidos en los ácidos nucleicos. El primer método que se usó, denominado *modelo de un parámetro*, fue propuesto por Jukes y Cantor [20]: se supone que todas las variaciones ocurren al azar entre los cuatro nucleótidos con la misma probabilidad, de manera que cualquier sustitución tiene la misma probabilidad de ocurrir. Esta suposición no es realista, ya que los cambios que sólo implican purinas o pirimidinas entre sí, esto es, las *transiciones* (cambios  $A \leftrightarrow G$  o  $C \leftrightarrow T$ ) están mecanísticamente favorecidas frente a las *transversiones* (cambios purina  $\leftrightarrow$  pirimidina  $A \leftrightarrow C$ ,  $A \leftrightarrow T$ , ó  $G \leftrightarrow C$ ,  $G \leftrightarrow T$ ), ya que las primeras ocurren por el mecanismo químico simple de tautomerización, mientras que las segundas requieren un mecanismo mucho más complejo.

Con objeto de corregir este problema, Kimura [21] propuso el *modelo de dos parámetros* considerando independientemente *transiciones* ( $pur \leftrightarrow pur$ ,  $pyr \leftrightarrow pyr$ ) y *transversiones* ( $pur \leftrightarrow pyr$ ). Desde su formulación, esta corrección se considera absolutamente necesaria en los estudios de filogenia molecular [19, 25].

Este problema es mucho más complicado cuando se estudian secuencias de proteínas, ya que ahí no hay sólo dos velocidades diferentes, sino prácticamente 19! (factorial de 19). En efecto, suponiendo que cualquier cambio de un aminoácido por otro pueda ser acepta-

do (lo que es otro problema diferente, discutido arriba), cada uno de esos cambios tiene una probabilidad diferente: unos pueden ocurrir cambiando sólo una base, otros necesitan que cambien dos bases, y otros necesitan que cambie todo el triplete. Algunas de estas sustituciones de bases pueden ocurrir con mayor facilidad que otras, como acabamos de ver. Un cambio de un aminoácido por otro que tiene que ocurrir en tres pasos consecutivos transcurre a través de posiciones intermedias cuya aceptación puede ser muy diferente de la que pueda tener la situación final. De forma que cuando observamos las modificaciones que se han producido en una proteína a lo largo del tiempo, estamos viendo realmente una sombra borrosa de la variación real que ha ocurrido en su gen. En general, la variación de un ácido nucleico es más rápida y más estable que la de una proteína, y así también es estadísticamente más significativa. Esto, unido a que hoy día el análisis de secuencias de ácidos nucleicos es más fácil que el de proteínas, ha hecho que aquéllos hayan reemplazado a éstas en el campo de la filogenia molecular.

Hay proteínas que evolucionan más deprisa que otras, como muestran los datos de la tabla II. Entre los ácidos nucleicos hay también diferencias importantes. Por ejemplo, El DNA mitocondrial y el RNA ribosómico, dos tipos de ácidos nucleicos muy usados para estos estudios, son herramientas muy diferentes (véase la figura 4). El DNA mitocondrial es una molécula que evoluciona muy deprisa, y así es útil para estudiar relaciones entre especies que se han separado recientemente, no hace más de 15 Ma (por ejemplo, los hu-



**Figura 4.**—Velocidades de evolución de dos ácidos nucleicos. El rRNA evoluciona lentamente (tiene un periodo evolutivo de 18 Ma para que se produzca una diferencia de un 1% en la secuencia) y muestra una variación muy regular con el tiempo. El DNA mitocondrial evoluciona muy deprisa y además su velocidad cambia rápidamente.

manos y los monos). Esta rapidez—del orden de un 25% de divergencia entre especies separadas desde hace sólo 15 Ma—se debe probablemente a la carencia del sistema de reparación del DNA en la mitocondria [39]. Por el contrario, el RNA ribosómico (por ejemplo, el rRNA 18S, o su gen llamado “rDNA 18S”) cambia mucho más lentamente: hay sólo un 33% de diferencia entre especies separadas desde hace 600 Ma. El rRNA es una buena herramienta para estudiar relaciones filogenéticas amplias, y en efecto, se ha usado para revisar las relaciones entre metazoos (artrópodos, moluscos, poliquetos, etc) y bacterias. Varios trabajos basados en rRNA 18S están comentados más abajo.

### ¿TIENE UNA BASE TEÓRICA LA TESIS DE KIMURA?

La figura 3 muestra que se puede admitir en principio que el citocromo c evoluciona a velocidad constante—aunque también hay desviaciones notables de esa regla—. En la tabla II hay más ejemplos. ¿Por qué muchas proteínas evolucionan a velocidad constante, y por qué hay diferencias de velocidad entre unas y otras? En general se acepta que la probabilidad de mutación es esencialmente la misma para cada sitio en la secuencia de una proteína y que las diferencias de velocidad de evolución se deben a diferencias en la probabilidad de que las mutaciones se fijen [43-47]. Este hecho está basado en el concepto de restricciones funcionales: cualquier proteína tiene partes altamente comprometidas con su función donde los cambios son difícilmente aceptados, y otras cuyo papel no es tan crucial y donde hay más holgura para aceptar mutaciones. La proporción entre los sitios no funcionales y funcionales condicionará la proporción de mutaciones que pueden fijarse, determinando la velocidad de evolución. Uno de los casos que mejor ilustran este efecto es la insulina, una proteína hormonal. La insulina se sintetiza en forma de una molécula más grande, denominada proinsulina. Durante la maduración la proinsulina se rompe en tres trozos, denominados respectivamente péptidos A, B y C. El péptido C es necesario para que el proceso de maduración se haga correctamente, pero la hormona funcional está formada sólo por los péptidos A y B. Como puede verse en la tabla I, la velocidad de evolución de los péptidos A y B de la insulina es lenta (1% de cambio cada 14 Ma), mientras que la del péptido C es muy rápida (1% de cambio cada 1,9 Ma).

La teoría neutralista de la evolución molecular propuesta por el genetista de poblaciones Motoo Kimura en 1974, y revisada por él mismo en profundidad en 1983 [22] pretende dar una explicación a la velocidad constante de variación de las secuencias de proteínas. Esta teoría afirma la mayor parte de los cambios producidos a nivel molecular (secuencias

de aminoácidos en proteínas, o de bases en ácidos nucleicos) son neutrales, es decir, que son cambios que no tienen valor selectivo. Sin embargo Kimura no ha dado realmente una explicación teórica a su tesis; se limita a presentar una colección de datos empíricos y a sacar ciertas conclusiones de ellos, sin un razonamiento teórico convincente que la apoye.

Hemos visto que existen desviaciones importantes de la velocidad constante de cada proteína. ¿A qué se debe este efecto? Según Goodman [12, 13] es muy probable que la velocidad de variación se acelere después de la duplicación de un gen (mecanismo básico para crear proteínas nuevas), ya que eso promovería nuevas adaptaciones que podrían desencadenar una radiación adaptativa. Un buen ejemplo de este efecto se ha observado con las extremadamente altas velocidades de sustitución de aminoácidos que siguieron a la duplicación de los genes en la separación de las hemoglobinas  $\alpha$  y  $\beta$ . [9].

En nuestro grupo hemos estado trabajando durante los últimos 15 años en la optimización del metabolismo y en la filogenia basada en datos bioquímicos, y los resultados obtenidos nos dan una base teórica para comprender mejor este hecho. Hemos demostrado que varios aspectos del metabolismo han estado sometidos a un proceso de optimización. Esto incluye diseños de rutas metabólicas, tales como el ciclo de las pentosas-fosfato [32-34], el ciclo de Calvin [32] y la glicólisis [18, 35]. Sin embargo, no todo el metabolismo tiene un diseño optimizable. Por ejemplo el diseño del ciclo de Krebs es un problema de solución única que no puede optimizarse porque no hay otras posibilidades [36]. Los mecanismos de optimización han operado también en la estructura molecular. Nosotros lo hemos demostrado para la estructura de la molécula de glucógeno [30, 37] y también hemos descrito un caso de metabolismo no optimizado—*paleometabolismo*—en la ostra (*Ostrea edulis*), una paleoespecie viva documentada en el registro fósil desde el Triásico (hace 280 Ma) [38]. La optimización de la molécula de glucógeno demuestra que la estructura de otras macromoléculas mucho más complejas, más decisivas para la vida y con más posibilidades de cambio, tienen que ser también optimizables. Es lógico suponer que la optimización de una macromolécula mucho más complicada como una enzima sea mucho más costosa y que pueda modificarse para adaptarse a condiciones ambientales diferentes, lo que puede implicar procesos de reoptimización cuando el grupo se diversifica ocupando nuevos ambientes. Los trabajos del grupo de Heinrich [16, 17] sobre la optimización enzimática apoyan esta idea.

Supongamos que una proteína está optimizada. Supongamos para simplificar que esa proteína está dividida en dos partes diferenciadas que llamaremos *zona funcional* y *zona no funcional*, respectivamente. Podemos admitir que la selección natural no permite ningún cambio en la zona funcional—pues ésta degeneraría—, pero los permite todos en la no funcional. Como la probabilidad de que *ocurra* una mutación es la misma en cualquier sitio, los cambios ocurrirán permanentemente en la zona no funcional. Esto determina una velocidad constante de variación que será mayor o menor según la proporción de estas dos zonas en el total de la proteína. Esto es igualmente aplicable a ácidos nucleicos.

Admitamos que la evolución de una proteína es un proceso de optimización de su estructura, de acuerdo con nuestros resultados comentados arriba, y veamos cómo puede transcurrir este proceso. Al principio de su historia, cuando la proteína está lejos de su estructura óptima, la probabilidad de que un cambio al azar pueda mejorarla es grande, de manera que podrán ser aceptadas muchas mutaciones, y así su velocidad de variación será muy alta. Esa velocidad irá decreciendo a medida que la proteína se aproxima a su estructura óptima a base de acumular mutaciones favorables, de forma que cuando al fin se alcance la estructura óptima, sólo los cambios neutros (que no modifican su función) serán aceptados. La proporción de cambios neutros aceptables será específica de cada proteína, según sea el porcentaje de aminoácidos estrechamente comprometidos con su función, lo que determinará que cada proteína, una vez alcanzada su estructura óptima, tenga una velocidad constante específica de variación.

Llamemos  $a$  a la zona funcional de la proteína, y  $b$  a la zona no funcional. La velocidad total de variación de la proteína  $v_T$ , será la suma de las velocidades de las dos partes:

$$v_T = v_a + v_b$$

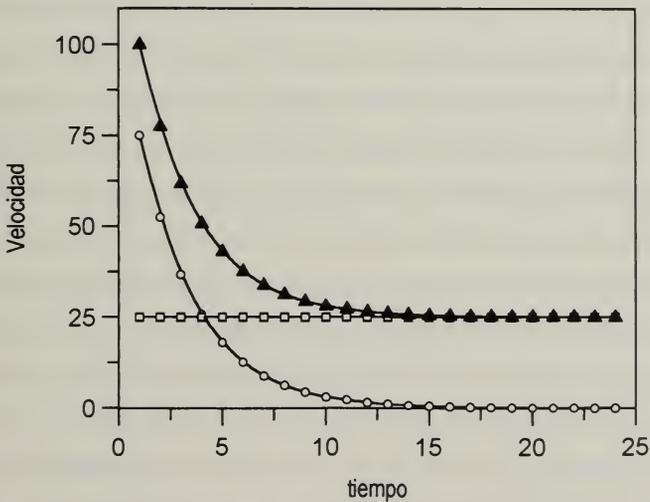
En la figura 5 se representa la variación de la velocidad total a lo largo del tiempo durante la optimización de la proteína. Puede verse que la propiedad del neutralismo se cumple únicamente cuando la proteína está muy optimizada y la parte funcional ya no evoluciona.

Este razonamiento demuestra que la teoría neutralista de la evolución molecular de Kimura no es consistente, ya que carece de soporte teórico y, en cuanto a su cumplimiento empírico, describe un fenómeno que sólo ocurre cuando la proteína está optimizada. Una idea parecida ha sido propuesta por Zuckerkandl [48]; según este autor, la velocidad constante de variación de una proteína se debe a que su estructura fluctúa en torno a un estado subóptimo con respecto a propiedades generales, tales como solubilidad y carga;

entonces, como todas las propiedades no pueden optimizarse simultáneamente, hay una competencia continua para optimizar cada propiedad. El resultado acaba en una serie ilimitada de cambios, incluso sin que varíe el ambiente: estas variaciones se pueden considerar como un ruido evolutivo, aunque basado en un principio de selección.

No puede decirse, como afirma Kimura, que la mayor parte de los cambios que se producen a nivel molecular son neutrales. Cuando la proteína está optimizada, los cambios neutrales son en realidad los únicos que se producen, pero durante el proceso de optimización, el porcentaje de cambios neutrales y selectivos depende de dos variables: del porcentaje de aminoácidos estrechamente comprometidos con la función, y de lo cerca que estemos de la estructura óptima.

Hay, sin embargo, algo de cierto en la tesis de Kimura: Es mucho más fácil que los cambios neutrales se produzcan a nivel molecular que a nivel macroscópico, ya que un cambio molecular neutro no trasciende y pasa absolutamente inadvertido, mientras que cualquier



**Figura 5. Efecto de la optimización de una proteína sobre su velocidad de evolución.** La velocidad de evolución de una proteína (▲) es la suma de la velocidad de evolución de la parte funcional (○) y la no funcional (□). Al principio, cuando la proteína está poco optimizada, su parte funcional evoluciona muy deprisa, y esta velocidad va disminuyendo hasta llegar a ser nula cuando se alcanza la estructura óptima. Por el contrario, la velocidad en la zona no funcional es independiente del grado de optimización de la proteína, por lo que es siempre constante.

cambio morfológico, por inocuo que pueda parecer en un principio, es fácil que acabe teniendo un significado. Un cambio de color, unos lunares en una posición determinada, etc, son rasgos que pueden adquirir fácilmente un valor selectivo. Hasta ahora nadie parece haber demostrado que exista una diferencia morfológica sin valor selectivo. Pero que los cambios neutrales puedan ocurrir con más facilidad a nivel molecular, no significa que sean los que predominan o hayan predominado siempre a ese nivel. La tesis de Kimura podría formularse más ajustadamente diciendo simplemente que cuando una proteína está optimizada, como no puede mejorarse, los únicos cambios que se aceptan son los que no estropean su función; pero eso es tan obvio que se explica sólo, sin teoría.

### CAMBIOS EN LA VELOCIDAD DE EVOLUCIÓN

Podría pensarse que tras una larga historia evolutiva todas las proteínas deberían estar ya optimizadas, y así todas estarían variando a velocidad constante—como acabamos de ver, la filogenia molecular está basada en esta hipótesis—. Es cierto que lo lógico es esperar que esa hipótesis se cumpla actualmente en muchos casos después de una larga evolución, pero suponer que se ha estado cumpliendo desde hace muchos millones de años puede producir un error grave a la hora de deducir relaciones filogenéticas. Hay varios problemas: las condiciones de optimización pueden haber sido muy diferentes en los distintos grupos, y además pueden haber cambiado a lo largo del tiempo. Cada molécula tendrá también una diferente sensibilidad a cambiar su velocidad de variación; algunos cambios ambientales pueden promover una estructura óptima muy diferente, condicionando que su velocidad de variación pueda cambiar durante un período discreto de tiempo; después podría recuperarse su velocidad inicial, o podría cambiar el tamaño de la zona funcional, lo que llevaría a cambiar su velocidad permanentemente. El resultado es en cualquier caso el mismo: la velocidad de evolución puede cambiar por diversos motivos.

Por ejemplo, la hemoglobina y el citocromo c tienen que ser proteínas sensibles a cambiar sus estructuras como consecuencia de cambios en la presión de oxígeno en el ambiente. Como la presión externa de oxígeno debe condicionar su concentración dentro de la célula, un cambio de aquélla ha de modificar la afinidad química de la cadena respiratoria, y el citocromo c debe ajustarse a esas nuevas condiciones. Es de esperar pues, que cuando se produce una amplia radiación de un grupo—una expansión de especies para ocupar nuevos nichos con ambientes más o menos aeróbicos—la velocidad de evolución

de las proteínas relacionadas con el transporte y el metabolismo del oxígeno pueda tener incrementos dramáticos como consecuencia de un proceso de reoptimización. Este fenómeno será diferente en especies distintas de la misma familia, o en familias distintas derivadas del mismo grupo. El mismo concepto puede aplicarse a otras proteínas relacionadas con otras variables ambientales. La elección de la molécula con la que se va a hacer el estudio no es trivial, hay muchas variables diferentes que deben considerarse en cada caso.

Pero ¿qué pueden significar los relojes rápidos en el RNA ribosómico? Es fácil comprender que se dé el efecto de reloj rápido en proteínas tales como la hemoglobina o el citocromo c, donde acabamos de ver que existe una relación obvia entre su estructura óptima y las condiciones ambientales. Sin embargo, no es fácil comprender que ocurra este efecto con el rRNA 18S, donde no existe esta relación. En realidad, el RNA ribosómico no es únicamente una molécula lenta, sino que su velocidad de variación es muy estable.

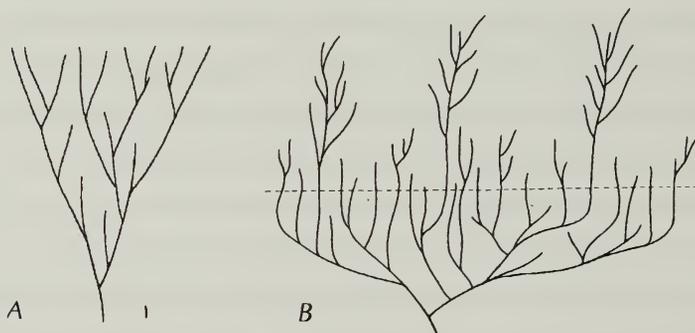
### MÁS PROBLEMAS

Hay muchos más problemas. A continuación señalaremos algunos de los más importantes.

**El efecto del tiempo de generación.**—En ciertas moléculas se han descrito velocidades de variación más altas en monos que en humanos, y también velocidades más altas en roedores que en primates. Estas diferencias de velocidad se explican por el efecto, propuesto por Kohne (1970) [23], del *tiempo de generación* (tiempo medio en la especie que transcurre desde que un individuo nace hasta que se reproduce): el tiempo de generación en roedores es mucho más corto que en primates, y entre éstos, el tiempo de generación de la especie humana es más largo que el de otros, de forma que el número de replicaciones del DNA por año podría ser mucho más alto en roedores que en el hombre (cf. Li & Graur [25]). En los artrópodos hay especies de baja velocidad de generación en todos los grupos principales (quelicerados, insectos, miriápodos y crustáceos), y muchas veces no se conoce con exactitud esta velocidad, de manera que este problema sólo puede ser resuelto analizando muchas muestras distintas. Otra vez vemos que la mejor forma de resolver los problemas de la estadística es con más estadística; formalmente hablando, la estadística debe usarse con rigor: el tamaño de la muestra debe ser grande, y especialmente cuando los resultados parecen incoherentes.

**El tiempo que tarda en fijarse una mutación.**—Este es un problema complejo. Para comprenderlo consideremos la diferencia entre selección natural y selección artificial. A dife-

rencia de la primera, en la selección artificial hay un plan preconcebido y ejecutado personalmente por el cuidador, el cual selecciona los progenitores de cada generación a fin de acentuar un carácter buscado, o eliminar otro indeseable. El hecho de seleccionar drásticamente los progenitores determina que todos los descendientes tendrán ese carácter, y así elimina el largo período de fijación de la mutación en la población. Por eso la selección artificial es mucho más rápida que la selección natural, y usualmente se consigue el objetivo buscado en unas pocas generaciones. En la selección natural el tiempo de fijación de un gen depende de la magnitud de la población. En una especie distribuida por todo un continente, con libertad de intercambio genético en todo el territorio, una mutación puede tardar muchos miles o millones de años en fijarse, mientras que en una población pequeña, aislada geográficamente, el tiempo de fijación puede ser muy corto. Las especies evolucionan mucho más deprisa en una isla que en el continente. Una proteína optimizada con una zona no funcional grande, donde se aceptan muchas mutaciones sin dificultad, en una población muy amplia, producirá inevitablemente una dispersión neutra (polimorfismo neutral) muy grande; ciertos individuos tendrán unos aminoácidos cambiados distintos de los que tengan otros sin que haya diferencias funcionales entre ellos; véase un ejemplo en los puntos de 80 Ma en la evolución del DNA mitocondrial (fig. 4). Esto origina una proteína con mucha variación individual, donde se aceptan mutaciones a una velo-



**Figura 6. Nuevas Ideas sobre la forma del árbol filogenético.** A, Forma clásica del árbol como puede verse en muchos libros de texto, con la forma típica de un cono de diversidad creciente. B, Forma propuesta por Gould [14]: un árbol 'horizontal' que parece más bien un arbusto después de una poda selectiva, con la mayor diversificación al principio, y una posterior eliminación de muchas ramas. Esta idea está fuertemente soportada por las dos grandes explosiones de biodiversidad (el Precámbrico de Ediacara, y el Cámbrico medio de Burgess Shale). La línea discontinua representa la época de Burgess Shale, con la más alta diversidad conocida.

cidad mayor que a la que se fijan, originando mucha dispersión de puntos en las gráficas.

**La forma del árbol filogenético.**—Hay otro problema importante que no se ha mencionado antes en la literatura en relación con la filogenia molecular: la estructura real del árbol filogenético es crítica para evaluar las posibilidades de los métodos de secuencias. El árbol filogenético que se suele ver en los libros tiene la forma típica vertical de complejidad creciente mostrado en la figura 6A. En su lugar, Gould [14] ha propuesto una *forma horizontal* que se asemeja más a un arbusto podado que a un árbol (Figura 6B): una gran diversidad al principio, la posterior eliminación de muchos grupos, y la pervivencia de unas pocas ramas que se siguen diversificando más lentamente. Esta idea está apoyada por las dos grandes radiaciones documentadas en el registro fósil: la del Precámbrico de Ediacara (580 Ma) y la del Cámbrico medio de Burgess Shale (530 Ma) [8] donde se ha encontrado mucha más diversidad que la que se ve después. En la fauna fósil de Burgess Shale hay 20 ó 30 grupos diferentes de Artrópodos sin representantes actuales [14, 41].

La tesis de Gould es muy razonable y puede considerarse empíricamente demostrada. Además, nuestro razonamiento contribuye a darle un soporte teórico: la diversificación es mucho más fácil al principio, cuando el material está sin perfeccionar y su optimización puede seguir rumbos diferentes. La mayor diversificación se produjo pues, en *poco tiempo* hace muchos millones de años, y esto puede producir incertidumbre en el reloj en los primeros estados de la evolución. Para un estudio filogenético amplio debe usarse una macromolécula lenta, como el rRNA 18S. Pero esa molécula capaz de medir períodos largos de tiempo no puede tener suficiente precisión para discriminar el gran número de bifurcaciones que ocurrieron muy próximas en el tiempo hace cerca de 600 millones de años. Esto puede explicar bien la incertidumbre observada generalmente en trabajos que tratan de establecer relaciones filogenéticas lejanas basados en el análisis del RNA ribosómico. Una molécula como el citocromo c, sensible a los cambios ambientales, podría quizá haber incrementado temporalmente su velocidad durante ese período manteniendo el registro de tales bifurcaciones lejanas. Quizá el citocromo c contenga esos datos difíciles de encontrar en el rRNA 18S, pero hoy día las modernas técnicas para secuenciar ácidos nucleicos han desplazado a las de secuenciar proteínas; lamentablemente parece que analizar secuencias de citocromo c está hoy día pasado de moda. Nuestro razonamiento sugiere que sería deseable recuperar el análisis de secuencias de citocromo c para terminar este trabajo. No debería haber modas cuando se quiere resolver un problema científico importante.

REVISIONES RECIENTES DE LA FILOGENIA:  
ARTRÓPODOS Y ANÉLIDOS SON GRUPOS ARTIFICIALES

Durante los casi 30 años que se lleva estudiando la filogenia molecular, la mayoría de los trabajos publicados sobre los Metazoos tratan más de detalles de distribución entre órdenes dentro de una clase o familias dentro de un orden, que de una visión global. En general, parece haber más interés en aspectos muy particulares de la filogenia que en problemas más amplios, quizá porque muchos piensan que los problemas grandes ya están resueltos. Es un problema que muchos autores consideren el árbol clásico como un punto fijo para ensayar la validez de nuevos métodos, en lugar de tratar de aplicar esos métodos para probar si el árbol clásico es correcto; muchos autores parecen dispuestos a aceptar sus propios resultados sólo cuando no entran en conflicto con el esquema clásico, como si éste fuese un dogma. Sin embargo, son tantos los datos de filogenia molecular que demuestran que Artrópodos y Anélidos son grupos artificiales que no puede ignorarse la evidencia.

La posible unidad o heterogeneidad de los artrópodos es quizá el problema más importante planteado en la filogenia de los Metazoos. Los trabajos de Manton [26, 27] fueron posiblemente los primeros en señalar una larga serie de características diferenciales entre los distintos grupos de artrópodos y reclamar su falta de homogeneidad. Manton aportó muchos datos que a su juicio demostraban que los artrópodos son un grupo artificial sin consistencia filogenética; a esta idea se unieron Whittington y otros [41]. Puesto que los datos aportados por esos autores eran todos ellos morfológicos, su valor como homologías es difícil de establecer, como hemos visto más arriba. La tesis de Manton ha sido muy criticada, y algunos autores han pretendido quitarle importancia a sus observaciones, pero sus trabajos nunca han dejado de citarse con más datos a favor [4, 5, 42]. Varias revisiones recientes basadas en los caracteres morfológicos [7, 42] vuelven a insistir en el mismo tema, y los datos de Filogenia Molecular y de Bioquímica le están dando la razón, aunque muchas veces a los mismos autores de esos trabajos les parece muy fuerte lo que sus propios resultados están diciendo e intentan buscarle otra explicación: los resultados de Lake [24], Field *et al.* [11], Aguinaldo *et al.* [2] analizando secuencias de rRNA; y Boore *et al.* [6] estudiando la estructura del genoma mitocondrial, confirman la tesis de Manton y las observaciones de otros autores [42]. Igualmente, los Anélidos son también un grupo artificial, pues no hay relación filogenética directa entre Poliquetos y Oligoquetos.

Los últimos resultados de nuestro grupo basados en homologías metabólicas confirman que Artrópodos y Anélidos son grupos artificiales sin relación filogenética próxima: los Poliquetos son parientes próximos de los Quelicerados y Moluscos, mientras que los Oligoquetos e Hirudíneos lo son de los Insectos, Miriápodos y Crustáceos, siendo muy lejana la relación filogenética entre ellos. Los Artrópodos son pues, un grupo polifilético, y las características morfológicas que tienen en común representan sin duda el caso más sorprendente de convergencia adaptativa que ha ocurrido en la evolución de los Metazoos.

**Agradecimiento.** Este trabajo ha sido subvencionado por la Consejería de Educación, Cultura y Deportes del Gobierno de Canarias.

### REFERENCIAS BIBLIOGRÁFICAS

1. Acher, R., 1974, *Biochimie*, **56**, 1-19. Tabla II
2. Aguinaldo, A.M.A., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A. & Lake, J.A. (1997) *Nature*, **387**, 489-493.
3. Ax, P. (1987). *The phylogenetic system*. John Wiley & Sons, Chichester, UK.
4. Bergström, J. (1979) En: *Arthropod Phylogeny* (ed: A.P.Gupta), pp. 33-56. Van Nostrand Reinhold, New York.
5. Bergstrom, J. (1986) *Zool. Scr.* **15**, 189-200
6. Boore, J.L., Collins, T.M., Stanton, D., Daehler, L.L. & Brown, W.M. (1995) *Nature*, **376**, 163-165.
7. Budd, G. E. (1996) *Trends Ecol. Evol.* **11**, 356-358.
8. Conway Morris, S. (1993) *Nature*, **361**, 219-225.
9. Czelusniak, J., Goodman, M., Hewett-Emmett, D., Weiss, M. L., Venta, P. J. & Tashian, R. E. (1982) *Nature*, **298**, 297-300.
10. Dayhoff, M. O. (1971) En: *Molecular evolution*, vol.1 "Chemical evolution and the origin of life. (R. Buvet. & C. Ponnamperna, eds.) Noth-Holland Publ. Co., Amsterdam, pp. 392-419.
11. Field, K., Olsen, G.J., Lane, D.J., Giovannoni, S.J., Ghiselin, M.T., Raff, E. C., Pace, N.R. & Raff, R.A. (1988) *Science*, **239**, 748-753.
12. Goodman, M. (1981) *Progr. Biophys. Mol. Biol.* **38**, 105-164.
13. Goodman, M. Moore, G.W., Matsuda, G. (1975) *Nature*, **253**, 603-608.
14. Gould, S.J. (1989) *Wonderful life. The Burgess Shale and the nature of history*. W.W.Norton & Co., New York.
15. Guyénot, É. (1947) *L'origine des espèces*. Presses Universitaires de France, Paris, p. 36.
16. Heinrich, R. & Hoffman, E. (1991) *J. Theor. Biol.* **151**, 249-283.
17. Heinrich, R., Schuster, S. & Holzhütter, H.-G. (1991) *Eur. J. Biochem.* **201**, 1-21.
18. Heinrich, R., Montero, F., Klipp, E., Waddell, T.G. & Meléndez-Hevia, E. (1997) *Eur. J. Biochem.* **243**, 191-201.
19. Huelsenbeck, J.P. & Rannala, B. (1997) *Science*, **276**, 227-232.

20. Jukes, T.H. & Cantor, C.R. (1969) En: *Mammalian protein metabolism*, (H.N. Munro, de.) Academic Press, New York, pp. 21-132.
21. Kimura, M. (1980) *J. Mol. Evol.* **16**, 111-120.
22. Kimura, M. (1983) *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge, UK.
23. Kohne, D. E. (1970) *Quart. Rev. Biophys.* **33**, 327-375.
24. Lake, J.A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 763-766.
25. Li, W.-H. & Graur, D. (1991) *Fundamentals of molecular evolution*. Sinauer, Sunderland, Mass.
26. Manton, S. M. (1973) *J. Zool. Lond.* **171**, 111-130.
27. Manton, S.M. & Anderson, D.T. (1979) En: *The origin of major invertebrate groups* (M.R. House, ed.) pp. 269-321.
28. Margoliash, E. (1963) *Proc. Natl. Acad. Sci. USA* **50**, 672-679.
29. Meléndez, B. (1982) *Paleontología*, 3rd. edit. Paraninfo, Madrid.
30. Meléndez, R., Meléndez-Hevia, E. & Cascante, M. (1997) *J. Mol. Evol.* **45**, 446-455.
31. Meléndez-Hevia, E. (1985) En: *Historia de la Bioquímica* (A.M. Muncio, coord). Real Academia de Ciencias, Madrid., pp. 109-142.
32. Meléndez-Hevia, E. (1990) *Biomed. Biochim. Acta*, **49**, 903-916.
33. Meléndez-Hevia, E. & Isidoro, A. (1985) *J. Theor. Biol.* **117**, 251-263.
34. Meléndez-Hevia, E. Waddell, T. G. & Montero, F. (1994) *J. Theor. Biol.* **166**, 201-220.
35. Meléndez-Hevia, E., Waddell, T. G. , Heinrich, R. & Montero, F. (1997) *Eur. J. Biochem.* **244**, 527-543.
36. Meléndez-Hevia, E., Waddell, T. G. & Cascante, M. (1996) *J. Mol. Evol.* **43**, 293-303.
37. Meléndez-Hevia, E., Waddell, T. G. & Shelton, E. D. (1993) *Biochem. J.* **295**, 477-483.
38. Meléndez-Hevia, E., Lupiáñez, J. A., García-Salguero, L., Meléndez, B., Peragón, J., Barroso, J. B., Waddell, T. G., Puigjaner, J., Raposo, R. R., Meléndez, R., Navarro, M. F., Cabezas, H. & Canela, E. I. (1997) Enviado para publicación.
39. Ridley, M. (1993) *Evolution*. Blackwell Scientific Publications, Oxford., UK.
40. Sarich, V.M. & Wilson, A.C. (1973) *Science*, **158**, 1200-1203.
41. Whittington, H. B. (1979) En: *The origin of major invertebrate groups* (House, M. R., ed.) Academic Press, London, pp. 253-268.
42. Willmer, P. (1990) *Invertebrate relationships. Patterns in animal evolution*. Cambridge University Press, Cambridge, UK.
43. Wilson, A. C., Carlson, S. S. & White, T. J. (1977) *Ann. Rev. Biochem.* **46**, 573-639.
44. Zuckerkandl, E. & Pauling, L. (1962) En: *Horizons in Biochemistry* (M. Kash & B. Pullman, eds.), Academic Press, New York. Pp. 189-225.
45. Zuckerkandl, E. & Pauling, L. (1965) *J. Theor. Biol.* **8**, 357-366
46. Zuckerkandl, E. (1975) *J. Mol. Evol.* **7**, 1-58.
47. Zuckerkandl, E. (1975) *J. Mol. Evol.* **7**, 167-184.
48. Zuckerkandl, E. (1976) *J. Mol. Evol.* **7**, 269-312.