

Ciento cincuenta años tras el árbol de la vida

Nuevos retos sobre el origen de las especies

Santi Garcia-Vallvé y Pere Puigbò

El estudio de la evolución y la búsqueda del árbol de la vida (tree of life) están más vigentes que nunca. Los nuevos datos provenientes de las secuenciaciones genómicas no han permitido la construcción de un único árbol donde se representen todas las relaciones evolutivas de las especies, pero han provocado la reconsideración de algunos conceptos y la introducción de nuevos paradigmas. La representación de las relaciones de las especies mediante un árbol quizás no sea la más adecuada y se han sugerido nuevas representaciones como una red (network of life) o un bosque (forest of life).

A partir del momento en que Charles Darwin presentó en 1859 su teoría sobre la «evolución de las especies mediante la selección natural», se inició la búsqueda del árbol de la vida (*tree of life*); un árbol que relacionara todos los seres vivos de la Tierra y que nos mostrase su clasificación y evolución. En la única figura de *El origen de las especies*¹ Darwin ya representó, en forma de árbol, las relaciones entre las especies, mostrando cómo evolucionan (o extinguen) por procesos de cambio constante (véase el recuadro «Darwin y el árbol de la vida»). Uno de los discípulos de Darwin, Ernst Haeckel, representó un árbol de la vida donde, a partir de un tronco común, las ramas divergen en varias direcciones y de cada una de ellas brotan numerosas ramificaciones. Los primeros árboles filogenéticos estaban basados en criterios morfológicos. Posteriormente, cuando se empezaron a obtener las secuencias de proteínas y ácidos nucleicos, se planteó su uso para analizar las relaciones entre las distintas especies. Cuando Emile

Zuckerlandl y Linus Pauling,² en 1965, introducen el concepto de «evolución molecular», se empiezan a construir los primeros árboles filogenéticos a partir de datos moleculares. Para ello se utilizan proteínas o *genes ortólogos* (genes de especies distintas que codifican para la misma proteína y que derivan del proceso de

«Cada árbol filogenético obtenido a partir de una familia de genes o proteínas representa las relaciones evolutivas de esa familia y no tiene por qué representar la historia evolutiva de las especies.»

especiación). Comparando secuencias ortólogas, se puede construir un árbol filogenético utilizando el hecho de que cuando dos secuencias están evolutivamente más relacionadas, menos mutaciones y cambios han sufrido desde su separación a partir de su ancestro común, y más parecidas son.

Uno de los primeros árboles de la vida construido a partir de datos moleculares fue el realizado por Carl Woese y George Fox,³ en 1977, en el que utiliza el RNA de la subunidad pequeña del ribosoma. El rRNA está presente en todas las formas de vida y por ello se consideró como un buen marcador de las relaciones evolutivas entre las especies. En ese árbol se podían diferenciar tres reinos o dominios: eucariotas, procariotas y arqueas. Durante años, éste fue considerado el árbol de referencia para la mayoría de la comunidad científica. La principal limitación para obtener un árbol de la vida completo era el número de secuencias disponibles. Desde entonces, la gran acumulación de secuencias de proteínas y ácidos nucleicos ha permitido hacer más comparaciones e ir refinando el árbol. A partir de 1995, cuando se inició la denominada «era genómica» y la secuenciación de los genomas completos de distintas especies procariotas y eucariotas, el número de secuencias disponibles ya no era la principal limitación. Al ir construyendo árboles filogenéticos a partir de otros genes,

se observó con sorpresa que pocos árboles filogenéticos se correspondían completamente con la topología del árbol obtenido a partir del rRNA. Cada árbol filogenético obtenido a partir de una familia de genes o proteínas (incluido el árbol construido a partir del rRNA) representa, de hecho, las relaciones evolutivas de esa familia y no tiene por qué representar la historia evolutiva de las especies. Hay diferentes razones por las que un árbol filogenético obtenido por la comparación de una sola familia de genes no se corresponda con el árbol de las especies. Entre ellas encontramos las duplicaciones, o *genes parálogos* (genes de una misma especie que derivan de un proceso de duplicación génica) que, a veces, pueden confundir las relaciones de ortología, procesos de pérdida o adquisición de genes, como la transferencia horizontal génica, e incluso las limitaciones de los métodos de reconstrucción de los árboles filogenéticos.

El problema de utilizar un solo árbol filogenético para construir el árbol de la vida, ya fue planteado por el mismo Darwin cuando escribió cómo había que clasificar correctamente las especies. En su libro *El origen de las especies*, Darwin explica que para clasificar correctamente una especie se deben utilizar caracteres con una elevada importancia funcional y prácticamente universales, y que cualquier clasificación elaborada a partir de un solo carácter (como es el caso del árbol del rRNA) puede llevar a una clasificación errónea.

The value indeed of an aggregate of characters is very evident in natural history. Hence, as has often been remarked, a species may depart from its allies in several characters, both of high physiological importance, and of almost universal prevalence, and yet leave us in no doubt where it should be ranked. Hence, also, it has been found that a classification founded on any single character, however important that may be, has always failed; for no part of the organisation is invariably constant.

Darwin, 1859¹

El hecho de que los árboles filogenéticos obtenidos a partir de distintos genes presenten topologías distintas, ha hecho plantear si las relaciones entre las especies se pueden representar mediante un árbol, o si sería mejor una representación en forma de red⁴ (fig. 1a). Si tenemos en cuenta la gran proporción de genes adquiridos

de distintos orígenes que puede tener una especie, parece claro que la historia de la vida no se puede representar formalmente como un árbol. Debido a ello, algunas posturas más extremas llegan incluso a poner en duda el concepto de especie en los procariontes.⁵

Experimentos recientes han demostrado cómo las especies procariontes comparten sus genes fácilmente, y que la transferencia de genes se puede producir tanto por vía vertical como horizontal (véase el recuadro «El intercambio de genes en la naturaleza mediante transferencia horizontal génica»). Este hecho parece incom-

patible con la representación clásica de un solo árbol de la vida. Respecto a este paradigma, todos los árboles obtenidos a partir de las diferentes familias génicas son importantes, sugiriendo que es necesaria una representación más compleja que tenga en cuenta todo el bosque (*forest of life*). El debate acerca del origen de los eucariotes también es complejo. Llevando al extremo la teoría endosimbiótica sobre el origen de mitocondrias y cloroplastos, se ha sugerido que los eucariotes surgieron a partir de la «fusión» de una bacteria con una arquea. Así, la mejor representación sobre el origen de los eucariotes se correspondería con un anillo (*ring of life*)⁶

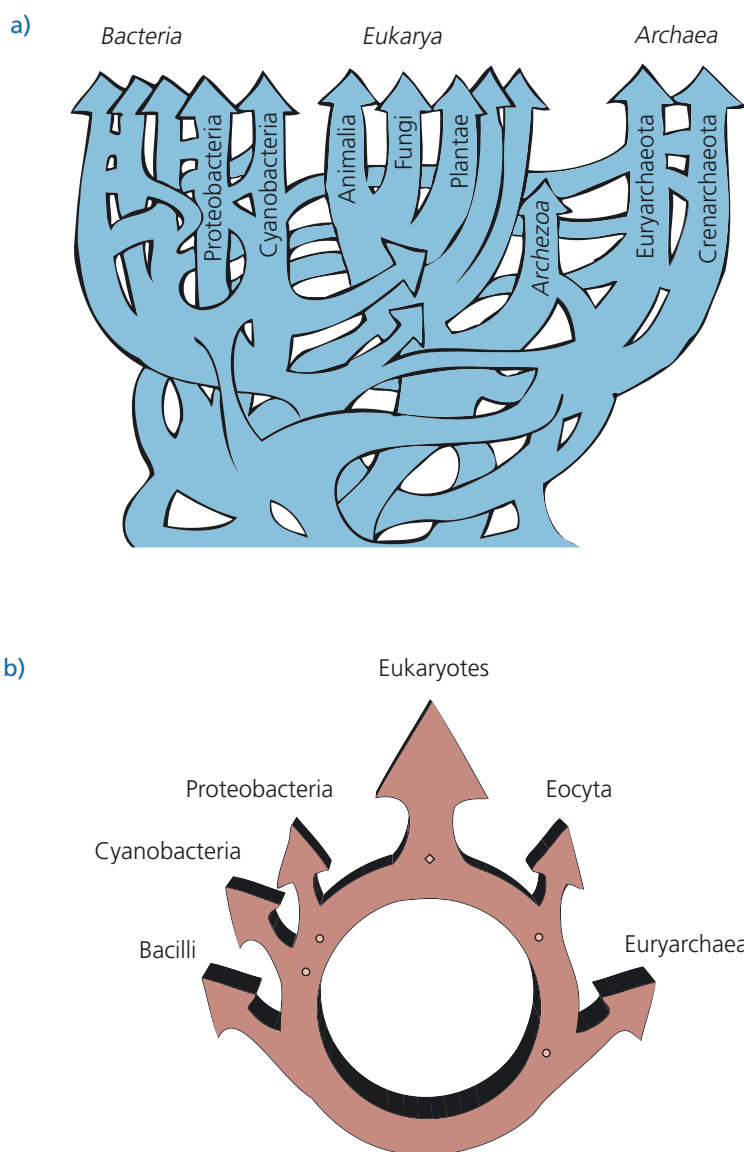


Figura 1 Representación de la relación de las especies mediante a) una red (*network of life*)⁴ o b) un anillo (*ring of life*)⁵.

Darwin y el árbol de la vida

Desde su llegada del viaje en el *Beagle*, en 1836, y con anterioridad a la publicación de *El origen de las especies*,¹ Charles Darwin fue madurando sus ideas sobre el papel de la selección natural y haciendo anotaciones en una serie de cuadernos.² En julio de 1837 empezó el cuaderno «B» de la serie titulada «Transmutación de las especies» donde, en la página 36, esbozó por primera vez las relaciones entre las distintas especies en forma de árbol (Fig. 2a). Junto al diagrama de este árbol se puede leer: *I think.... Case must be that one generation then should be as many living as now. To do this & to have many species in same genus (as is) requires extinction [...]*. Posteriormente, en noviembre de 1859, en la única imagen de la primera edición de *El origen de las especies*,¹ Darwin describe en forma de un árbol más detallado (fig. 2b), la idea sobre cómo evolucionan las especies a partir de la variación y de la selección natural. En esta figura, las pequeñas ramas representadas con líneas de puntos corresponden a

las diferentes variaciones que se van acumulando durante las diferentes generaciones. Sólo las variaciones que sean en algún modo ventajosas, serán preservadas por la acción de la selección natural. Cuando una de estas líneas punteadas supera alguna de las líneas horizontales, las cuales representan una escala de miles o centenares de miles de generaciones, ésta queda marcada y diferenciada de su ancestro, pues Darwin consideraba que ya habría acumulado un número suficiente de variaciones para que fuera registrada en sistemática.

La idea de la existencia de un árbol de la vida (*tree of life*) está implícita en la obra de Charles Darwin, aunque aparece mencionado explícitamente en pocas ocasiones en sus anotaciones y publicaciones. Véanse algunos ejemplos de las apariciones del término **tree of life**: En la *Transmutación de las especies*,² aparece, por ejemplo, en la página 25 del cuaderno «B»: *We need not think that fish & penguins really pass into each other. – The tree of life should*

perhaps be called the coral of life, base of branches dead; so that passages cannot be seen [...], y en la página 263: *At what part of tree of life, can orders like birds & animals separate [...]*.

Sin embargo, en *El origen de las especies*,¹ el término aparece una sola vez en todo el libro, concretamente en el último párrafo del capítulo 4 sobre la selección natural: *The affinities of all the beings of the same class have sometimes been represented by a great tree. I believe this simile largely speaks the truth. ... The green and budding twigs may represent existing species; and those produced during each former year may represent the long succession of extinct species. ... The limbs divided into great branches, and these into lesser and lesser branches, were themselves once, when the tree was small, budding twigs; and this connexion of the former and present buds by ramifying branches may well represent the classification of all extinct and living species in groups subordinate to groups. ... As buds give rise by growth to fresh buds, and these, if vigorous, branch out and overtop on all sides many a feebler branch, so by generation I believe it has been with the great Tree of Life, which fills with its dead and broken branches the crust of the earth, and covers the surface with its ever branching and beautiful ramifications. #*

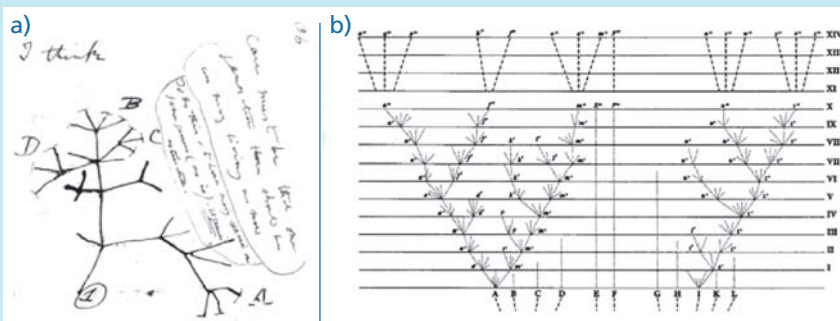


Figura 2 Dos representaciones del árbol de las especies según Darwin: a) diagrama de un árbol evolutivo representado en 1837 en los cuadernos titulados «Transmutación de las especies»; b) única figura de la primera edición (1859) del libro *El origen de las especies* donde representa las relaciones de las especies y el mecanismo evolutivo en forma de árbol.

Notas

¹ Darwin, C. John Murray, Londres, 1959 <<http://tinyurl.com/ylp7ax>>

² Darwin, C. (1837-38) Transmutation of species. Notebook B. <<http://tinyurl.com/cjzfx>>

(fig. 1b) en detrimento, una vez más, del árbol de la vida.

Recientemente también se han hecho intentos para reconstruir un árbol consenso al que poder llamar árbol de la vida. Uno de ellos, elaborado por Ciccarelli y colaboradores⁷ se basó en la concatenación de 31 genes prácticamente universales —presentes casi en todas las especies—, y en su mayoría genes de las proteínas ribosómicas. Sin embargo, otros autores⁸ han sugerido que, lejos de obtener el árbol de la vida, lo que se obtiene en estos casos es el árbol del «uno por ciento», haciendo referencia a que un árbol consenso, semejante a un árbol de las especies, sólo se puede conseguir con el 1 % de las familias génicas, dejando una incerteza del 99 %.

En conclusión, 150 años después de la publicación de *El origen de las especies*,¹ el debate sobre la representación de las especies sigue más vigente que nunca. El fondo de este debate se centra en cómo entendemos los procesos evolutivos. Los nuevos datos genómicos, lejos de resolver la cuestión nos han hecho replantear el concepto de árbol e introducir nuevas hipótesis. Estos cambios de paradigmas parecen inevitables para seguir avanzando en la comprensión de la naturaleza, ya que, como dijo Theodosius Dobzhansky (1900-1975), en biología nada tiene sentido si no es a la luz de la evolución.⁹ #

Santí García-Vallvé

DEPARTAMENTO DE BIOQUÍMICA Y BIOTECNOLOGÍA, UNIVERSITAT ROVIRA I VIRGILI (URV), TARRAGONA

Pere Puigbò

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION, NATIONAL LIBRARY OF MEDICINE, NATIONAL INSTITUTES OF HEALTH, BETHESDA, MD, USA

► Bibliografía

1. Darwin C. John Murray, Londres, 1859. <<http://tinyurl.com/ylp7ax>>
2. Zuckerkandl, E. y Pauling, L. *J Theor Biol* 1965; 8: 357.
3. Woese, C. R. y Fox, G. E. *Proc Natl Acad Sci USA* 1977; 74: 5088.
4. Doolittle, W. F. *Science* 1999; 284: 2124.
5. Bapteste E y Boucher Y. *Trends Microbiol* 2008; 16: 200-207.
6. Rivera, M. C. y Lake, J. C. *Nature* 2004; 431: 152.
7. Ciccarelli, F. D. y col. *Science* 2006; 311: 1283.

El intercambio de genes en la naturaleza mediante transferencia horizontal génica

Hace más de 60 años que conocemos la posibilidad de que los organismos procariotas adquieran DNA de otras especies. Esta adquisición se denomina generalmente, *transferencia horizontal génica*, en oposición a la transmisión vertical, hereditaria de padres a hijos. Existen tres mecanismos mediante los cuales las bacterias pueden adquirir DNA: 1) la *conjugación*, cuando se establece una interacción física entre dos bacterias, generalmente de la misma especie, para la adquisición de un plásmido; 2) la *transducción* o adquisición de DNA mediado por un fago; y 3) la *transformación*, es decir, la adquisición de forma directa de un fragmento de DNA. Aunque experimentos recientes¹ han puesto de manifiesto que casi cualquier gen puede ser transferido, podemos pensar que un gen se impondrá en una población si le da una ventaja selectiva a su huésped. La denominada *hipótesis de la complejidad*, desarrollada por Jain y colaboradores en 1999,² establece que hay menos probabilidad de que los denominados «genes informacionales», aquellos involucrados en la replicación, transcripción y traducción, sean transferidos y se impongan en una población, debido a que estos genes forman parte de redes complejas de interacciones, y la sustitución de uno de ellos sería más difícil. Sin embargo, se ha demostrado que ni estas familias génicas están libres de transferencias horizontales.

El análisis de los primeros genomas completamente secuenciados de bacterias, a finales de los años noventa, puso de manifiesto que la transferencia horizontal génica era un fenómeno más corriente de lo que se creía, estimando que en algunas especies el porcen-

taje de genes adquiridos podría estar entre el 10 y el 20%.³ Existen distintos métodos para predecir qué genes de un genoma han sido adquiridos mediante transferencia horizontal génica.⁴ Entre ellos destacan la comparación de las secuencias mediante métodos heurísticos y la búsqueda de genes composicionalmente anómalos (en el contenido en G+C y uso de codones) respecto a la mayoría de genes de un genoma. La evidencia más clara de un caso de transferencia horizontal es la obtención de un árbol filogenético con una topología distinta a la del árbol de las especies, aunque no es el único proceso que puede generar incompatibilidad entre los árboles obtenidos a partir de las distintas familias génicas.

En eucariotas (pluricelulares) los procesos de transferencia horizontal se ven limitados por el hecho de que sólo se transmitirán a la descendencia aquellas adquisiciones que se produzcan en las células germinales. Sin embargo, la adquisición de genes mediante procesos de simbiogénesis⁵ parece que puede haber jugado un papel importante en la evolución de los eucariotas,⁶ provocando, de nuevo, que tal vez un árbol tampoco sea la mejor forma de representar la evolución de las especies.

Notas

1. Sorek, R. y col. *Science*, 2007; 318: 1449.
2. Jain, R., Rivera, R. C. y Lake, J. A. *Proc Natl Acad Sci USA* 1999; 96: 3801.
3. García-Vallvé, S., Romeu, A. y Palau, J. *Genome Res*, 2000; 10: 1719.
4. Martin, W. *BioEssays* 1999; 21: 99.
5. Margulis, L. y Sagan, D. *Captando genomas*. Kairós, 2003.
6. Keeling, P. J. y Palmer, J. D. *Nat Rev Genet* 2008; 9: 605.

8. Dagan, T. y Martin, W. *Genome Biol* 2006; 7: 118.
9. Dobzhansky, T. *Am Zool* 1964; 4: 443.

► Lecturas recomendadas

- Doolittle, W. F. y Zhaxybayeva, O. *Genome Res* 2009; 19: 744.

- Koonin, E. V. *Nucleic Acids Res* 2009; 37: 1011.
 Lane, C. E. y Archibald, J. M. *Trends Ecol Evol* 2008; 23: 268.
 López, P. y Bapteste, E. *C R Biol* 2009; 332: 171.
 McNerney, J. O., Cotton, J. A. y Pisani, D. *Trends Ecol Evol* 2008; 23: 276.
 Spivak, E. *Ciencia Hoy* 2006; 16: 10.