

Acerca del origen de la vida: hacia una teoría unificada.

Segunda parte

Erika Martinez y Pablo Argibay

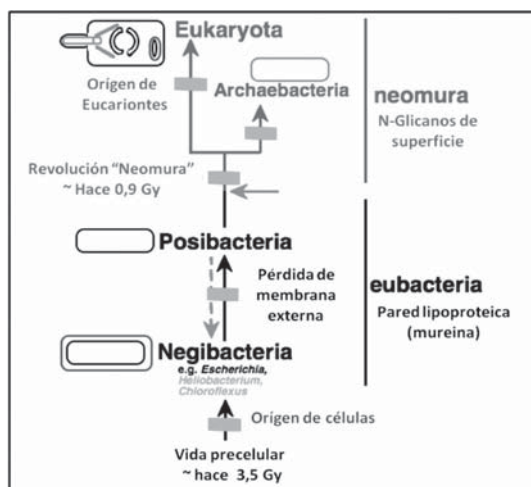
¿QUÉ FUE LUCA¹?

Recordemos de la primera parte que, si bien LUCA es considerado el ancestro común de las actuales formas de vida, no representa los estadios iniciales de la evolución biológica. Antes de formas de vida con sus componentes “empaquetados” en membranas y con una maquinaria de reproducción basada en material genético compuesto principalmente de ADN, se cree que ha existido un período en el cual el dominio de la vida dependía de moléculas de algo parecido al ARN actual. De todas maneras, en algún momento como consecuencia de fenómenos evolutivos de creciente complejidad emergió algo similar a las actuales células. El punto es que cuando en el año 1977 se definieron los tres dominios de la vida: Arquea, Bacteria y Eucariota, cada uno de estos linajes constituyó un candidato potencial para la raíz del árbol de la vida, raíz basada en el concepto “LUCA”. En esta segunda parte desarrollaremos las hipótesis que postulan a LUCA como bacteria, arquea o eucarionte primitivo.

LUCA: BACTERIA

Para Thomas Cavalier-Smith, profesor titular de Biología Evolutiva del Departamento de Zoología de la Universidad de Oxford, LUCA debió haber sido mucho más complejo de lo que Woese supone. Si bien propone una etapa evolutiva relacionada con la hipótesis del progenote, para él, esta etapa tuvo lugar mucho antes de la existencia del último ancestro. Según Cavalier-Smith, la filogenia basada en la reconstrucción de árboles proteicos sugiere que el antecesor celular estaba entre lo que se denomina “neomura”² y eubacterias, mientras que la filogenia basada en el metabolismo y la evolución de enzimas sugiere que el ancestro común estaba más cerca de ser una eubacteria. Por otra parte, los hallazgos paleontológicos sugieren que las eubacterias son mucho más antiguas que los eucariontes. A través de una metodología de análisis de transición en bancos de genes, buscando similitud en secuencias genéticas en eubacterias, Cavalier-Smith llega a la conclusión de que las bacterias son los antepasados comunes.

Figura 1. Relaciones evolutivas entre los cuatro tipos principales de células. La flecha horizontal indica la posición de una supuesta raíz universal entre neomuras y eubacterias. (Tomado y modificado de Cavalier-Smith, 2006).



¹ LUCA es el acrónimo de *Last Universal Common Ancestor*, el antecesor de las actuales formas de vida.

² Neomura (paredes nuevas) es el antepasado de Arquea y Eucariota.

Entre estas, las denominadas clorobacterias son las más antiguas y las arqueobacterias las más modernas con etapas intermedias (Fig. 1).

LUCA: ARQUEA

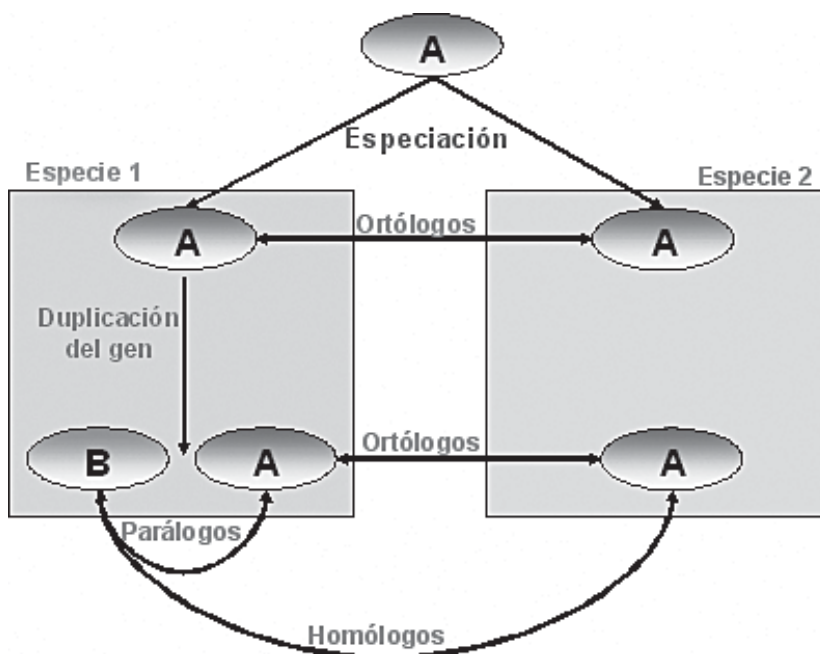
Xue *et al.* proponen que LUCA se encuentra ubicado dentro del dominio arqueo. Se basaron en el hecho de que el análisis mutacional podría ser útil si se enfoca sobre eventos cercanos a los primeros puntos de ramificación del árbol universal. Así midieron la divergencia entre secuencias de genes parálogos (Fig. 2) dentro de un mismo organismo e identificaron aquellos con bajo nivel de separación entre secuencias (especies de evolución extremadamente lenta cerca de la raíz del árbol universal).

El grupo de Xue construyó un árbol filogenético (Fig. 3) a partir de comparaciones entre el conjunto total de los ARNt. Este árbol muestra la clara distinción que existe entre los tres dominios existentes, al igual que los análisis realizados por Woese mediante comparaciones de la subunidad mayor de los ribosomas. Con el objetivo de hallar la raíz para este árbol, Xue llevó a cabo una separación de parálogos, para lo cual primero tuvo que identificar los parálogos de ARNt. Cabe aclarar que ese trabajo ha sido muy criticado, entre otras razones debido a la corta longitud de

las secuencias de ARNt, estas no sirven para la construcción de árboles filogenéticos.

Para la identificación de parálogos de ARNt, Xue *et al.* se basaron en el hecho de que las moléculas de ARNt que comparten la misma estructura de trébol podrían ser parálogos. De este conjunto de moléculas de ARNt, solo fueron consideradas parálogos potenciales dos moléculas de ARNt cuya distancia de separación no exceda 0.2, lo que corresponde a un 18% de diferencia entre secuencias. Entonces, midieron las distancias genéticas entre diferentes ARNt pertenecientes al mismo genoma y encontraron que algunas de esas distancias eran menores de 0.2, y, por lo tanto, consideraron parálogos a estos ARNt. Por ejemplo, encontraron que el ARNt^{Phe} y el ARNt^{Tyr} de algunas especies pertenecientes al dominio arqueo difieren solo en un número pequeño de bases, sumando porcentajes de más del 90% de identidad entre secuencias. Para Xue, tal nivel de identidad significa que esas moléculas de ARNt son parálogos, que nacieron por un evento de duplicación génica que luego sufrió un cambio evolutivo extremadamente bajo. Sin embargo, no sucede lo mismo con las moléculas de ARNt^{Phe} y ARNt^{Tyr} de especies bacteriales y eucariotas. Además, distancias genéticas que no superen el 0.2 se encuentran solo en secuencias de ARNt pertene-

Figura 2. La búsqueda de LUCA y de la raíz del árbol universal se lleva a cabo mediante la comparación entre secuencias génicas de organismos conocidos, lo que permite estudiar el grado de conservación según la distancia evolutiva que separa a estos organismos. Se dice que dos genes son "homólogos" cuando derivan de un ancestro común. A su vez, los genes "homólogos" se clasifican en genes "ortólogos" y genes "parálogos". Los genes ortólogos descienden de un ancestro común de diferentes especies (relación vertical), y por lo tanto se da por sentado que ejercen la misma función en los diversos organismos. Los genes parálogos se encuentran dentro de un mismo organismo y surgieron durante la evolución mediante duplicación génica. Se presume que estos últimos deben tener funciones diferentes o, al menos, cierto grado de especialización.



cientes al dominio arqueo, pero no en bacteria o eucariota. Según Xue, estos resultados describen el carácter primitivo de las arqueas. A partir de los valores de las distancias genéticas que se obtuvieron en este trabajo, Xue propone que LUCA está localizado entre las ramas que conducen a dos especies arqueas (*M. kandleri* y *A. pernix*), debido a que con ambas especies se obtuvieron los valores más bajos de distancias genéticas. El árbol filogenético construido mediante comparaciones de secuencias de ARNt (Fig. 3), muestra la posición de la raíz del árbol y también, sobre una escala térmica de color, las distancias entre especies diferentes, tomando como punto de partida a LUCA. Xue *et al.* mencionan en este punto que, dado que dos especies bacteriales (*A. aeolicus* y *T. maritima*) se ubican más cerca de LUCA que una parte de la especie arquea, seguramente bacteria se separó de arquea poco después de las primeras ramificaciones producidas a partir de LUCA. Para eucariota, este árbol no muestra ninguna especie primitiva (con distancia genética baja), motivo por el cual Xue dice que un punto por determinar es si existe alguna especie eucariota con una distancia genética baja, o bien si jamás existió y este dominio se originó a partir de un procariota, más probablemente de tipo arqueo, que ya poseía un valor de distancia elevado.

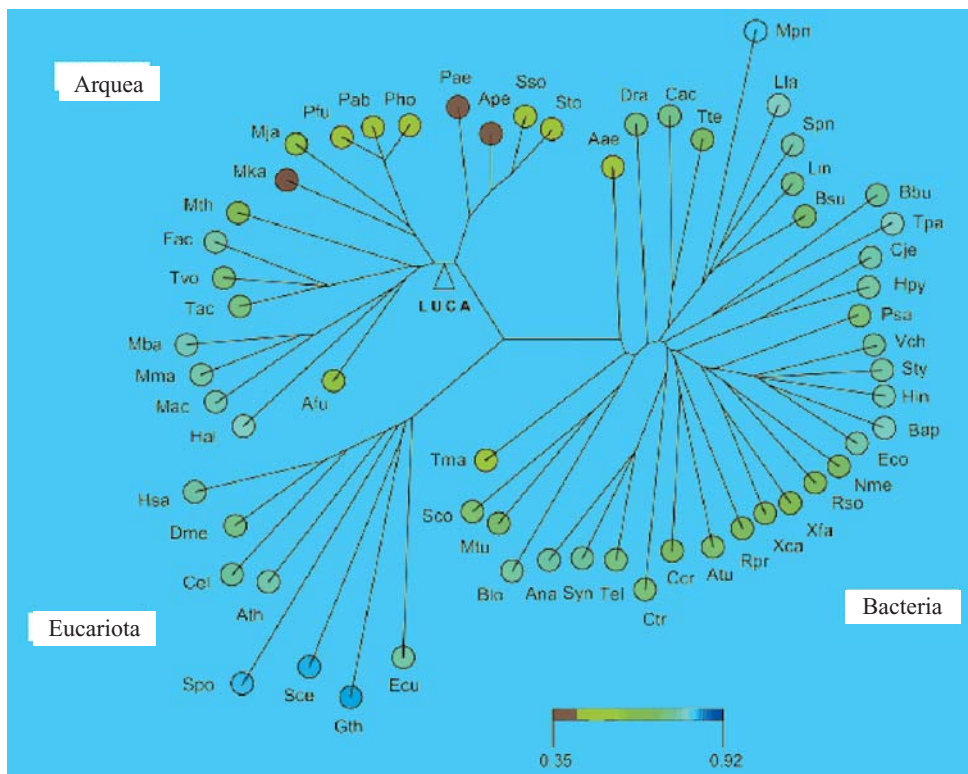
LUCA: EUCARIOTA

La razón por la cual distintos autores proponen diferentes localizaciones para la raíz del árbol universal de la vida se debe al uso de diferentes análisis usando diferentes secuencias o caracteres, lo que finalmente lleva a diferentes probables ubicaciones de la raíz. Como acabamos de ver, Xue *et al.* se basaron en comparaciones entre secuencias parálogas de ARNt para la construcción del árbol de la vida. El grupo de Penny, uno de los que proponen un LUCA de tipo eucariota, se basa en las partes del metabolismo moderno que han persistido desde un estado temprano de la evolución.

Como el resto de las hipótesis vistas en este artículo, Penny *et al.* toman como punto de partida el mundo de ARN. Para ellos este mundo estuvo en un principio poblado por moléculas de ARN que consistían en unos pocos cientos de bases debido a la baja precisión del proceso de replicación. Llamam a estas moléculas de ARN riboorganismos, y es a partir de estos riboorganismos que surge el primer organismo capaz de sintetizar proteínas.

Para poder determinar el tamaño que pudieron tener los genomas de ARN, ellos dicen que el tamaño de los genomas de los virus modernos sería una buena indicación, por lo que proponen que, como máximo, los genomas en el mundo de ARN deben haber tenido 30kb. Este es el número de

Figura 3. El árbol universal construido mediante comparaciones entre moléculas de ARNt. La escala térmica de color expresa los valores de las distancias genéticas para diferentes especies como una medida de cuán agrupados están los ARNt de cada especie.



bases que poseen los corona-virus, los virus que tienen los genomas más largos actualmente. Si bien es el número máximo de bases que podría haber tenido el riboorganismo, este número estuvo limitado por el conocido límite de Eigen. Este límite dice simplemente que la precisión de la información que se copia limita la capacidad codificante del genoma. Si la precisión es alta, pueden mantenerse grandes cantidades de información; si la precisión es pobre, la cantidad de información mantenida es baja.

De acuerdo con esto, el proceso de mutación limita la longitud máxima que una molécula puede tener. Si una molécula excede ese tamaño crítico, el efecto de las mutaciones conducirá finalmente a la destrucción de las generaciones posteriores a esta molécula.

Entonces, el error causado por mutaciones limita el tamaño de las moléculas autorreplicantes tal vez a unos pocos cientos de bases. Sin embargo, la vida como la conocemos hoy requiere moléculas más largas para la codificación de su información genética. En las células modernas existen enzimas que se encargan de reparar los errores producidos por mutaciones, lo que permite que las moléculas

puedan alcanzar tamaños del orden de millones de pares de bases. Pero estas enzimas no estuvieron presentes en el momento en que empezaron a formarse las primeras moléculas autorreplicantes. El problema obvio que salta a la vista es que las moléculas autorreplicantes debían de alguna forma crecer en longitud; sin embargo, la baja fidelidad en el proceso de replicación fuerza a la secuencia a no exceder una longitud máxima.

Lo que Penny propone para resolver este problema se conoce con el nombre de “ciclo de Darwin-Eigen” (Fig. 4). Este ciclo es un mecanismo de realimentación positiva mediante el cual la selección natural de rasgos que reducen la tasa de error permite, en principio, la codificación de mayores cantidades de información. Si mejora la fidelidad de la replicación, entonces aumentará el límite superior impuesto sobre la longitud del genoma, permitiendo la evolución de un genoma de mayor longitud. Esto favorecerá la evolución de una nueva función, que a su vez mejorará aun más la fidelidad replicativa, y así sucesivamente. De esta forma fueron creciendo probablemente las primeras moléculas autorreplicantes.

Figura 4. El ciclo de Darwin-Eigen es un mecanismo de realimentación positivo. Las mejoras sobre la fidelidad de la replicación llevan a un incremento en el límite superior (límite Eigen) impuesto sobre la máxima longitud de un genoma, y permite la evolución de un genoma de mayor longitud, lo que favorece la evolución de una nueva función que podría mejorar aun más la fidelidad de la replicación, y así sucesivamente. De tal modo, la repetición de ciclos sucesivos de este sistema en los primeros riboorganismos pudo, mediante selección natural actuando a favor de incrementos en la fidelidad de la replicación, aumentar el límite superior (límite Eigen) sobre la longitud de los genomas.



Entonces, inicialmente el mundo de ARN estuvo poblado por copias múltiples de genomas. Estos genomas estaban fragmentados (como en los cromosomas) y se emplearon estrategias de recombinación y de redundancia genética contra la pérdida de información debida a la baja fidelidad en la replicación. Dado que los genomas eucariotas exhiben este tipo de organización, mientras que los procariotas usualmente alojan toda la información en un cromosoma, Penny *et al.* sostienen que la arquitectura de los genomas primitivos fue más parecida a la arquitectura de genomas eucariotas.

Según Penny es difícil considerar que la arquitectura del genoma y el procesamiento de ARN procariota precedan a la arquitectura del genoma y al procesamiento de ARN eucariota. La razón es la alta fidelidad que posee la arquitectura del genoma procariota con poco procesamiento de ARN. Esto lleva a la conclusión de que la arquitectura del genoma de LUCA se pareció más a la arquitectura de genomas eucariotas. En el esquema de evolución genómica, los genomas eucariotas no cambian en arquitectura y organización básica en forma significativa, aunque se vuelven mucho más grandes y reclutan proteínas para funciones esenciales.

Ahora bien, sobre la base de esta hipótesis con respecto a la naturaleza de LUCA, es esencial considerar que fuerzas selectivas pudieron dar surgimiento a procariotas.

Una característica que marca diferencias entre genomas procariotas y eucariotas es su organización. Los genomas procariotas son circulares y con un único centro de replicación. Los genomas eucariotas son lineales, fragmentados y poseen múltiples centros de replicación.

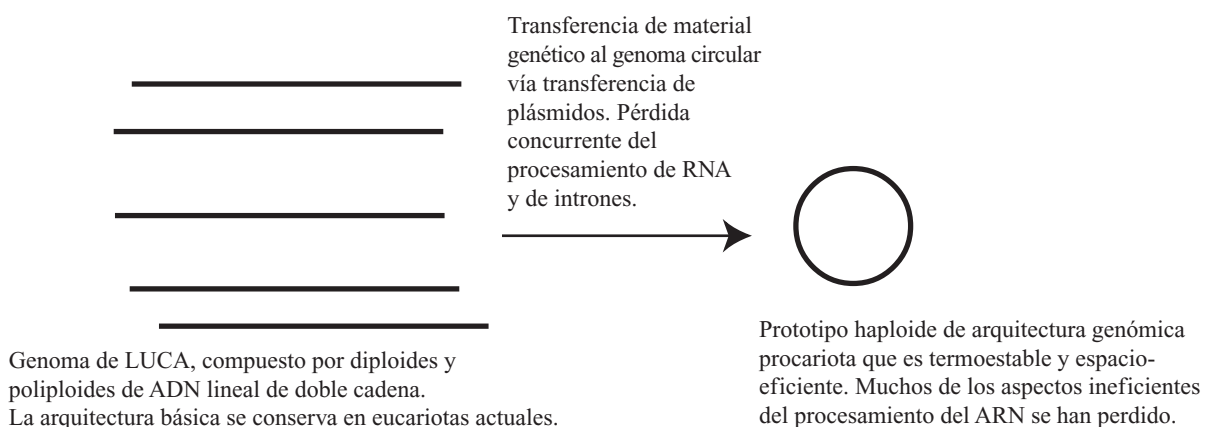
El modelo de transferencia de plásmidos (Fig. 5) propone que, mediante un proceso de transcripción inversa, la información genética alojada en el genoma lineal fragmentado de LUCA fue transferida a una molécula circular (como un plásmido), produciendo el prototipo de la organización genómica procariota. Las presiones de selección que –según Penny– produjeron las ventajas selectivas para el primer genoma circular básicamente fueron dos: termorreducción y selección *r*.

La hipótesis de termorreducción fue propuesta por Forterre y sostiene que los termófilos son derivados a partir de ancestros mesófilos (organismos que crecen a bajas temperaturas), y que buena parte de la organización genómica procariota es el resultado de adaptaciones a la termofilia. Cualquier reducción en el tiempo de procesamiento tanto del ARNm como del ARNr será ventajosa para cualquier organismo viviendo a altas temperaturas. La aceleración del procesamiento de ARN de una hebra (ARNm, ARNr y ARNt) desde horas (eucariotas) a minutos (procariotas) incrementaría la viabilidad de un organismo a altas temperaturas. Una vez que la organización genómica procariota se estableció, tendría otras ventajas que podrían ser explotadas por organismos mesófilos derivados a partir de estos ancestros termófilos.

La selección *r*, junto con la selección *K* son conceptos importantes en ecología y en investigación de la historia de la vida. Los términos vienen de la ecuación para la tasa de crecimiento de una población *r*.

$$r = dN/dt = rN(1 - N/K)$$

Figura 5. El modelo de transferencia de plásmidos para el origen de genomas circulares. Los transcritos de ARN procesado (ARNm, ARNr, y ARNt) son el sustrato para una transcriptasa inversa que produce una copia de ADN de dos hebras sin intrones. Esta copia se integra en un plásmido y, con la acción de ciertas presiones de selección (selección *r* o termorreducción) sobre los organismos, esta nueva arquitectura se ve favorecida, resultando finalmente en la transferencia completa de la información genética al cromosoma circular. Este cambio en la estructura del genoma también dio como resultado la eliminación de muchos aspectos del procesamiento de ARN del metabolismo del linaje procariota y la pérdida de secuencias intrónicas.



donde r es la tasa intrínseca máxima de incremento para la población, N el número de organismos y K la capacidad de transporte. Los organismos r -seleccionados muestran alta tasa de reproducción, tamaño pequeño, ciclos de vida cortos, y se encuentran con frecuencia en medios no predecibles. Compiten por fuentes de nutrientes que fluctúan mucho en cuanto a abundancia. En contraste, los organismos K -seleccionados muestran la inversa de estas propiedades: ciclos de vida largos, gran tamaño, tamaños de población relativamente más estables, y se encuentran en medios más estables.

La selección r generalmente da como resultado un uso de recursos más rápido y eficiente, debido a que la disponibilidad limitada produce fuerte competencia para esto. En general, los procariotas son mucho más r -seleccionados

que los eucariotas ya que el tiempo de generación es más corto, el tamaño de su genoma es menor y el tamaño de la población es más grande.

La selección para un genoma de menor tamaño y para una expresión génica más rápida en procariotas favoreció la pérdida de material genético no esencial (como intrones y spliceosoma) y el uso más eficiente del espacio del genoma. Así, Penny concluye que el genoma procariota es derivado en estructura, habiendo sufrido selección r y/o un estado termofílico, volviéndose más pequeño, más compacto y eficientemente organizado. El genoma ancestro tuvo una organización lineal, fragmentada y contenía intrones. Esto último, sin embargo, no provee nada de información acerca de la organización citológica de LUCA. Tampoco puede decirse si el "fenotipo" de procariotas surgió una vez o dos veces: una para arquea y una para bacteria.

BIBLIOGRAFÍA

- Cavalier-Smith T. Obcells as proto-organisms: membrane heredity, lithophosphorylation, and the origins of the genetic code, the first cells, and photosynthesis. *J Mol Evol* 2001; 53(4-5):555-95.
- Cavalier-Smith T. Rooting the tree of life by transition analyses. *Biol Direct* 2006; 1:19.
- Doolittle WF. Uprooting the tree of life. *Sci Am* 2000; 282(2):90-5.
- Forterre P, Philippe H. The last universal common ancestor (LUCA), simple or complex? *Biol Bull* 1999; 196(3):373-5; discusión 375-7.
- Forterre P. The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie* 2005; 87(9-10):793-803.
- Forterre P. Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. *Proc Natl Acad Sci USA* 2006; 103(10):3669-74.
- Graur D, Li WH. *Fundamentals of molecular evolution*. 2nd ed. Sunderland, Mass.: Sinauer Associates, 2000.
- Jeffares DC, Poole AM, Penny D. Relics from the RNA world. *J Mol Evol* 1998; 46(1):18-36.
- Kurland CG, Collins LJ, Penny D. Genomics and the irreducible nature of eukaryote cells. *Science* 2006; 312(5776):1011-4.
- Moulton V, Gardner PP, Pointon RF, et al. RNA folding argues against a hot-start origin of life. *J Mol Evol* 2000; 51(4):416-21.
- Penny D, Poole A. The nature of the last universal common ancestor. *Curr Opin Genet Dev* 1999; 9(6):672-7.
- Poole A, Jeffares D, Penny D. Early evolution: prokaryotes, the new kids on the block. *Bioessays* 1999; 21(10):880-9.
- Poole A, Penny D, Sjöberg B. Methyl-RNA: an evolutionary bridge between RNA and DNA? *Chem Biol* 2000; 7(12):R207-16.
- Poole AM, Jeffares DC, Penny D. The path from the RNA world. *J Mol Evol* 1998; 46(1):1-17.
- Woese C. The universal ancestor. *Proc Natl Acad Sci USA* 1998; 95(12):6854-9.
- Woese CR. On the evolution of cells. *Proc Natl Acad Sci USA* 2002; 99(13):8742-7.
- Xue H, Tong KL, Marck C, et al. Transfer RNA paralogs: evidence for genetic code-amino acid biosynthesis coevolution and an archaeal root of life. *Gene* 2003; 310:59-66.