

ARTÍCULO

INVARIANZA CRÍTICA DE ESCALA EN LA EVOLUCIÓN DEL CÓDIGO GENÉTICO UNIVERSAL

Marco V. José

Grupo de Biología Teórica, Instituto de Investigaciones Biomédicas
Universidad Nacional Autónoma de México,

Invarianza Crítica de Escala en la Evolución del Código Genético Universal

Resumen

Existen mecanismos evolutivos que tienden a preservar las propiedades inherentes de la vida. El Código Genético Universal (CGU) es el mejor ejemplo de un sistema biológico que se considera como el menos capaz de evolucionar ya que está congelado. En este trabajo se revisan y se discuten recientes trabajos en la modelación de la evolución del CGU partiendo de un código primitivo de RNA. En particular, se ha demostrado que el CGU conserva vestigios de código más antiguos como los llamados códigos Extendidos de RNA tipo I y II, que pudieron ser códigos intermedios entre el primitivo y el CGU. Los genomas que obedecieron al código primitivo, a los Extendidos de RNA y al CGU exhiben la propiedad de invarianza de escala en las series de distancias de sus codones. Esta invarianza de escala a través de la evolución es una típica característica de estados críticos de sistemas complejos. En este sentido la evolución del código genético antes del CGU parece haber obedecido a un proceso Lamarckiano.

Palabras clave: Código Genético Universal; Código Genético Primitivo; Evolución del código genético; Sistemas complejos; Código extendidos de RNA; Evolución no Darwiniana.

Abstract

There are evolutionary mechanisms that tend to preserve inherent properties of life. The Universal Genetic Code (UGC) is the best example of a biological system that is considered to be the less capable of evolving, i.e., is frozen. Herein, recent works on modeling the evolution of the UGC assuming a primeval genetic code are briefly reviewed and discussed. In particular, it has been shown that the UGC conserves vestiges of earlier codes, to wit, Extended RNA codes type I and II which may have been intermediate states between the primeval and the UGC. The primeval, the Extended RNA codes and the UGC show the property of scale invariance in the distance series of any of their codons. This scale invariance is a clear manifestation of a complex system under a critical state. In this sense, the evolution of the genetic code may be an example of a Lamarckian process.

Keywords: Universal Genetic Code; Primeval Genetic Code; Evolution of the genetic code; Complex systems; Extended RNA codes; Non-Darwinian evolution.

Introducción

Uno de los problemas más fascinantes de la biología es determinar el origen y la evolución del Código Genético Universal (CGU). La razón de ello es que todos los seres vivos existentes en la Tierra, con la excepción de algunas pocas variantes, compartimos el mismo código genético. El desciframiento del CGU es una hazaña de la ciencia del Siglo XX. Entre los científicos más importantes que contribuyeron a revelar el CGU están: Marshall Nirenberg, Heinrich J. Matthaei, Philip Leder y Har Gobind Khorana. El código genético consiste en una función en la que a cada triplete de nucleótidos o codón se le asigna un aminoácido. Existen 4 nucleótidos: citosina,

timina (en DNA) o uracilo (en RNA), adenina y guanina denotados, respectivamente por C, T o U, A, y G, y por lo tanto puede haber $4^3 = 64$ tripletes de los cuales 61 codifican por 20 aminoácidos y los otros tres son señales de paro. Típicamente el CGU se representa en dos dimensiones en la llamada Tabla del Código Genético (Crick, 1968). Dado 64 codones y 20 aminoácidos más una señal de terminación puede haber $21^{64} \approx 4 \times 10^{84}$ códigos genéticos posibles. Las pocas variantes del CGU que se han encontrado hasta la fecha (e. g. el código mitocondrial, algunos protozoarios ciliados, *Tetrahymena*) han evolucionado a partir del CGU (Osawa et al., 1992). La mayoría de las variaciones residen en amino ácidos que se incorporaron al código en las últimas etapas de su evolución y que son codificados por un solo codón (metionina y triptófano) y en los codones de paro. Sin embargo, un CGU congelado se refleja en el hecho de que los 20 aminoácidos naturales están en todos los códigos, incluyendo a la mitocondria y al cloroplasto. Las notables excepciones son la selenocisteína y la pirrolisina (Ambrogelly et al., 2007). Así el CGU ha podido evolucionar pero a una tasa glacial.

¿Existe algo especial acerca del único código que gobierna toda la vida en este planeta? F. Crick (1968) propuso que el CGU no tiene nada de especial y que simplemente provenía de un "accidente congelado", esto es, el código no puede cambiar. Este concepto no está lejos de la idea de que alguna vez hubo una era en la que ocurrían milagros. Sorprendentemente la hipótesis del "accidente congelado" no ha podido ser descartada por ninguna de las teorías más avanzadas sobre el origen y la evolución del código. Se ha podido descartar la idea del "accidente" (Ardell & Sella, 2002) pero de una u otra manera todas las teorías incluyen el atributo de congelación del código como veremos más adelante. El carácter universal del código hizo proponer a varios autores (Woese, 2000) que debió haber existido un Último Ancestro Común Universal (UACU en español o LUCA por sus siglas en inglés) a partir del cual surgieron los 3 grandes reinos: Eukarya, Archaea y Eubacteria (Woese, 1968).

En general la evolución se define cómo el cambio en la frecuencia de alelos en una población (Wilson & Bossert, 1971) o como descendencia (vertical) con modificación (Darwin, 1859). No existe una definición de la vida que sea aceptada por todos pero existe un consenso de que la vida no puede entenderse en la ausencia de un código genético y de la Evolución Darwiniana, esto es, material genético capaz de replicarse, cambiar y de heredarse. Sin embargo, si consideramos un CGU congelado entonces la evolución también consiste en mecanismos que preservan o fijan algunas propiedades fundamentales para la vida. Crick (1968) propuso que el código es universal porque en la actualidad cualquier cambio puede ser letal o con una selección muy fuerte en contra de cualquier variación. Es un hecho que la gran diversidad de seres vivos en la historia de la Biósfera no ha sido detenida por un CGU congelado ya que su estructura inherente ha liberado más bien a las formas vivientes del determinismo.

Existe también el consenso de que las primeras formas de vida eran sistemas capaces de estar sujetas a una evolución Darwiniana cuyas propiedades de auto-sustento dependían de la disponibilidad de moléculas orgánicas ya presentes en el medio ambiente primitivo. Una vez que apareció un mecanismo de auto-replicación que permitiera la introducción de la variación y la habilidad de replicar esas variantes, entonces la evolución Darwiniana estaba lista para operar.

La complejidad no es explicada por la evolución Darwiniana la cual se centra en la variabilidad sobre la que opera la selección natural. Sin embargo, existen varios sistemas biológicos auto-organizados que producen estructuras altamente ordenadas que deben operar además de la selección natural. En general, los modelos de complejidad no han sido aceptados debido a que

los principios de auto-organización o propiedades emergentes, típicas de sistemas complejos, son vistos por los biólogos evolutivos como invocaciones disfrazadas de generación espontánea (Lazcano, 2007).

El principal problema de realizar modelos de la evolución del código genético es que no existen rastros fósiles u organismos en la actualidad que tengan un código primitivo o intermedio para compararlos con el CGU. La búsqueda de simetrías en el CGU se ha hecho examinando el tRNA (Eigen et al., 1989; Nicholas & McClain, 1995), o las enzimas aminoacil-tRNA sintetasas (aaRS) (Delarue, 2008; Rodin & Rodin, 2008; de Pouplana & Schimmel, 2001), o por métodos filogenéticos (Nagel & Doolittle, 1995; Woese et al. 2000) o mediante métodos geométricos y algebraicos (Hornos & Hornos, 1993; Sánchez et al. 2005; José et al., 2007).

El código es implementado por moléculas de tRNA (RNA de transferencia) que unen un codón con su anticodón. El anticodón está localizado a la mitad de la estructura del tRNA que tiene forma de clavel, y que coincide con la máxima distancia posible donde se unirá el respectivo aminoácido. Debido a esta separación las moléculas de tRNA no pueden auto-aminoacilarse y por ello existen 20 aaRSs que llevan a cabo esta función. Estas enzimas definen al código genético al relacionar a cada uno de los 20 aminoácidos con sus respectivos tRNAs que presentan sus correspondientes anticodones (Rodin & Rodin, 2008). No solo los tRNAs o los aaRS han sido implicados en la evolución del CGU sino también los ribosomas presentan vestigios de estadios tempranos en la evolución de la vida para establecer la traducción del código genético desde antes del UACU (Johnson & Wang, 2010).

Toda la vida que existe en el planeta Tierra en la actualidad y toda vida de la existen registros geológicos parecen compartir una característica — está basada en genomas de DNA y de proteínas. Sin embargo existen razones convincentes para concluir que la vida basada en DNA y proteínas fue precedida por formas de vida más simples basadas esencialmente en la molécula de RNA. Esta era primitiva se refiere al Mundo de RNA, durante la cual la información genética residía en secuencias de moléculas de RNA y el fenotipo se derivaba de propiedades catalíticas del RNA (Gilbert, 1986; Eigen et al., 1989; Joyce, 1991). La molécula de RNA es ubicua en todos los sistemas biológicos contemporáneos, especialmente con respecto a los procesos celulares más fundamentales y altamente conservados. El RNA actúa como “primer” en la replicación del DNA, como mensajero que lleva la información genética a la maquinaria de traducción, y como catalizador que reside en el corazón de los ribosomas. El RNA instruye el procesamiento de precursores del RNA mensajero (mRNA) durante el “splicing” y edición, y es el mediador de varias transacciones con las proteínas de la célula (Joyce, 1991). El RNA catalítico (ribozimas) asiste en eventos del procesamiento del RNA y en la replicación de genomas virales. Los nucleótidos individuales son importantes moléculas de señalización y sus derivados coenzimáticos participan en la mayoría de las reacciones centrales del metabolismo. El RNA posee características que lo hacen casi ideal para correr un simple sistema Darwiniano: está formado de cuatro diferentes monómeros con propiedades químicas muy parecidas entre sí, sus subunidades se polimerizan fácilmente cuando se una unen a una platilla complementaria, es un poli-anión que es soluble en agua independiente de su secuencia, forma estructuras secundarias simples que son muy tolerantes a variaciones de su secuencia, y puede adoptar estructuras muy diversas con solo una cuantas mutaciones críticas (Fontana & Schuster, 1998; Schultes & Bartel, 2000). Estas mismas características lo hacen menos sofisticado que el DNA y sus sucesores proteicos. La menor reactividad pero mayor estabilidad del DNA lo convierte en una mejor alternativa como material genético, mientras que la gran diversidad química de las subunidades de las proteínas,

los aminoácidos, que incluyen grupos aniónicos, catiónicos e hidrofóbicos, convierten a las proteínas en una mejor elección para las funciones catalíticas. Sin embargo, estas moléculas más sofisticadas no pudieron haber surgido sin las bases que fueron establecidas por el RNA. El reino del Mundo de RNA en la Tierra probablemente empezó alrededor de hace 4.2 mil millones de años y terminó no antes de 3.6 mil millones de años (Joyce, 2002). Eigen y colaboradores (1989) descubrieron relaciones de parentesco en alineamientos de secuencias del RNA de transferencia (tRNA) y concluyeron que el código genético no es más viejo que nuestro planeta, pero es casi tan viejo como éste.

Existen lagunas significativas en el entendimiento científico referente a cómo surgió el Mundo de RNA, el grado de complejidad metabólica que alcanzó, y la manera en la que el RNA produjo genomas de DNA y sus enzimas proteicas.

Modelos matemáticos de la evolución del Código Genético Universal

Sin embargo, ha habido numerosos intentos de “descongelar” el origen y evolución del CGU congelado. Recientemente se han desarrollado 3 modelos sobre del origen y evolución del CGU como un sistema complejo y que son los siguientes:

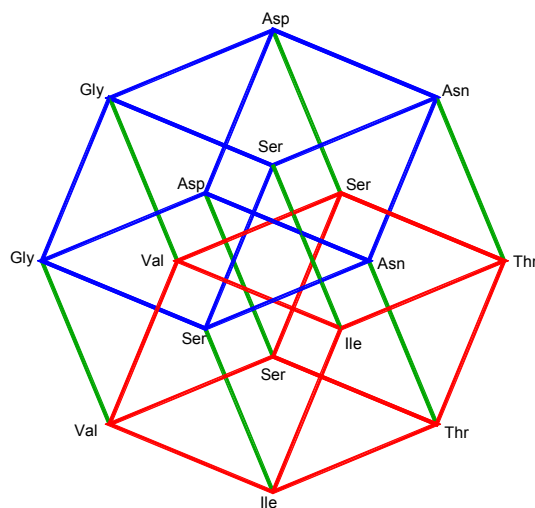
1. La universalidad del CGU ha sido modelada considerando la evolución temprana de una comunidad de entidades tipo células que cooperaban entre sí mediante el fenómeno no genealógico de Transferencia Horizontal de Genes (THG) (Vetsigian et al., 2006). Esta hipótesis sobre la evolución colectiva de replicadores primitivos explica la universalidad del código al través de una combinación de un accidente congelado y de una dinámica no lineal de protocolos evolutivos de innovación compartidos (Vetsigian et al., 2006). Resulta que la universalidad del código es una condición para mantener la THG entre las comunidades de replicadores primitivos, y este flujo de información es una condición para la evolución de cualquier entidad biológica compleja. Debido a que el código genético es universal, esto permite que especies completamente diferentes compartan genes que pueden incorporarse en los genomas de otras especies. El código genético juega un papel doble: no es solamente un protocolo para codificar secuencias de amino ácidos en el genoma sino también es un protocolo de innovación compartida. La dinámica es inherentemente compleja ya que involucra a los elementos de la comunidad y a sus interacciones con el medio ambiente. Los miembros de la comunidad absorben o descartan genes conforme los necesitan en respuesta al medio ambiente. Estos autores concluyeron que: “...the evolution of the genetic code, translation, and cellular organization itself follows a dynamic whose mode is, if anything, Lamarckian” (Vetsigian et al., 2006). La vida en sus orígenes evolucionó de una manera Lamarckiana en donde la descendencia vertical estuvo opacada por las formas más potentes de THG. El refinamiento a través de la THG debió haber disparado una explosión de novedades genéticas hasta alcanzar un nivel de complejidad que generó una transición hacia la era actual de evolución vertical, supuestamente cuando apareció el UACU. Goldenfeld y Woese (2007) consideran como lamentable la concatenación convencional del nombre de Darwin con el de la evolución, ya que otras modalidades deben ser consideradas.

2. Otro modelo interesante no Darwiniano sobre la evolución del código genético es el modelo Auto-Referencial (Modelo AR) (Guimarãez et al., 2008a, 2008b). Este modelo considera al código como una consecuencia de la existencia de pequeñas moléculas prebióticas del RNA de transferencia (tRNA). Estas moléculas del tipo del tRNA poseían la habilidad de formar dímeros mediante la unión de bases nitrogenadas (al través de sus proto-anticodones) y de recibir la adición de otras moléculas. En algún punto de la evolución los aminoácidos fueron capaces de

iniciar el proceso de decodificación que fue una consecuencia de la habilidad de los dímeros de poner a los aminoácidos en contacto cercano para formar péptidos. Cuando los péptidos mostraron la capacidad de unirse a los proto-tRNA y fueron uniones estables, entonces se formó una ribo-nucleoproteína (RNP) y al mismo tiempo apareció un ciclo auto-estimulante de síntesis de proteínas. Es por esto que el modelo se llama AR porque va de los proto-tRNA a sus RNPs. La sencillez y el alto carácter determinista de este proceso sugieren que las poblaciones del UACU estuvieron compuestas en un inicio de líneas con código genéticos similares. El modelo AR es compatible con la propuesta de que al inicio de la historia de la vida en la Tierra, las primitivas proto-bacteria, proto-archaea y proto-eukaryotes pudieron vivir juntos y compartir genes vía la THG: "Eventually this collection of eclectic and changeable cells coalesced into the three basic domains known today" (Woese, 1968). En el modelo AR (Guimarãez et al., 2008a, 2008b), las funciones de las proteínas se retroalimentan positivamente con los ácidos nucleicos formando así un doble puente entre genes y proteínas. Por lo tanto en el modelo AR no tiene sentido hablar de genes o proteínas "egoístas".

3. Usando métodos algebraicos y geométricos se pudieron derivar a partir del código primitivo de RNA, dos códigos genéticos que pudieron generar el CGU (José et al., 2007). El código primitivo de RNA consiste de 16 codones que especifican 8 aminoácidos (Figura 1). El código Extendido de RNA tipo I, consiste de todos los codones del tipo RNY (purina-cualquier base-pirimidina) más codones obtenidos al considerar el código de RNY pero empezando por el segundo (codones tipo NYR) y en el tercer (codones tipo YRN) marcos de lectura. El código Extendido de RNA tipo II, comprende a todos los codones del tipo RNY más los codones que surgen de las transversiones (purina-pirimidina) del código de RNY en la primera (codones tipo YNY) y en la tercera base nucleotídica. Así, en el primer caso, al existir errores en el marco de lectura se generan 18 codones que especifican 17 aminoácidos y tres codones de paro, mientras que en el segundo caso surgen 18 aminoácidos pero ningún codón de paro. El código primitivo de RNA (Fig. 1) así como los códigos Extendidos de RNA fueron representados por hipercubos simétricos en la cuarta dimensión (José et al., 2007), cuya unión disjunta en cada caso generó un hipercubo (Coxeter, 1973), previamente propuesto para representar al CGU en la sexta dimensión (Jiménez-Montaña, *et al.*, 1996).

Figura 1. Código Primitivo de RNA representado en un hipercubo en la dimensión 4 (codones tipo RNY).



Representaciones del CGU basados en novedosos modelos geométrico-algebraicos en 3 dimensiones, denotados Hoteles Genéticos y Hoteles de Aminoácidos han facilitado el probar varias hipótesis sobre su origen y evolución (José et al, sometido). Las representaciones geométricas y algebraicas de los códigos de RNA (primitivo y Extendidos) así como del CGU en 4, 6 o 3 dimensiones han enriquecido su entendimiento al liberarlos de la tortura que vivían al representarlos en 2 dimensiones o Tierra Plana.

La principal evidencia de las dos plausibles rutas evolutivas proviene de la observación de que las propiedades de escalamiento de las series de distancias de algunos codones del código primitivo de RNA y la mayoría de los codones de ámbos códigos Extendidos resultan ser idénticos o muy cercanos a las propiedades de escalamiento de los codones de genomas contemporáneos de procariontes (José et al., 2009). Dicho de otra manera, las magnitudes de los exponentes de escalamiento de las series de distancias de los codones de un procarionte dado NO han cambiado durante el curso de la evolución desde el Mundo de RNA hasta nuestros días (José et al., 2009). De acuerdo a la Teoría de Renormalización de Grupos (Wilson, 1979) se puede decir que estas series de distancias exhiben invarianza crítica de escala (Wilson, 1979; Sornette 2000). Es por eso que varias propiedades estadísticas, de información y de escalamiento son aún detectables en cromosomas actuales de procariontes ya que muestran un estado crítico (con la interesante excepción de algunos codones de algunas Archeas). Por lo tanto, los genomas son sistemas que están constantemente en un estado crítico y que muestran propiedades universales de invarianza de escala. Esta propiedad de ciertas secuencias de RNA/DNA es precisamente una característica de los sistemas complejos.

Así, el CGU conserva vestigios de códigos anteriores, a saber, el código primitivo de RNA y los

código Extendidos de RNA tipo I y II (José et al., 2009). El primero perteneció al Mundo de RNA y el segundo al Mundo de Ribo-Nucleoproteínas. Los códigos Extendidos fueron intermedios entre los ribo-organismos y el UACU de los Reinos de Bacteria, Archea, and Eucarya. Por lo tanto, la mayoría de los procariontes poseen todavía restos de un mundo primitivo de RNA el cual es un palimpsesto sobre el cual ambos códigos Extendidos de RNA fueron adiciones posteriores que formaron el CGU.

Para probar qué tan robustos eran estos resultados se realizaron experimentos de simulación computacional en los que se mutaron aleatoriamente genomas contemporáneos a tasas de 10^{-10} - 10^{-9} por sitio por año durante 3 mil millones de años, y éstas no fueron suficientes para destruir los patrones observados (José et al., 2009).

Conclusión

La mayoría de los procariontes guardan todavía reliquias de su historia desde las primeras etapas de su origen por lo menos desde hace 3 mil millones de años. A pesar de tanto tiempo e independientemente de los detalles de las secuencias de RNA/DNA surge una propiedad emergente típica de un sistema complejo: dada la serie de distancias de un codón de un procarionte determinado, un simple exponente de escalamiento ha permanecido congelado durante las distintas etapas evolutivas del CGU. El uso de las matemáticas nos permite dilucidar procesos que de otra manera sería imposible obtener.

M.V.J. agradece los apoyos de PAPIIT IN205307, y del Macroproyecto: Tecnologías para la Universidad de la Información y la Computación, UNAM, México. Se agradece también las discusiones con Juan R. Bobadilla y Tzipe Govezensky.

Referencias

- * Ambrogelly A, Palioura S, Soll D. (2007). Natural expansion of the genetic code. *Nat Chem Biol.* 3(1):29-35.
- * Ardell, D. H., Sella G. (2002). No accident: genetic codes freeze in error-correcting patterns of the standard genetic code. *Phil. Trans. R Soc. Lond. B*, 357(1427):1625-1642.
- * Coxeter, H.S.M., 1973. *Regular Polytopes*, 3rd edition. Dover Publication Inc., New York.
- * Crick F. H. C. (1968). The origin of the genetic code. *J. Mol. Biol.* 38: 367-379.
- * Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection*. London: John Murray.
- * Delarue M. (2008). An asymmetric underlying rule in the assignment of codons: Possible clue to a quick early evolution of the genetic code via successive binary choices. *RNA* 13: 161–169.
- * de Pouplana L. R., Schimmel P. (2001). Aminoacyl-tRNA synthetases: potential markers of genetic code development. *Trends Biochem. Sci.* 26: 591–596.
- * Eigen, M., Lindemann, B. F., Tietze, M., Winkler-Oswatitsch, R., Dress A., Haeseler A. (1989)

How old is the genetic code? Statistical geometry of tRNA provides an answer. *Science* 244: 673-679.

* Fontana, W., Schuster, P. (1998). Continuity in evolution: on the nature of transitions. *Science* 280, 1451–1455.

* Gilbert, W. (1986). The RNA world. *Nature* 319, 618.

* Guimarães R. C., Moreira C. H. C., Farias S. T. (2008a). A self-referential model for the formation of the genetic code. *Theory Biosci.* 127: 249-270.

* Guimarães R. C., Moreira C. H. C., Farias S. T. (2008b). Self-referential formation of the genetic system. In: *The Codes of Life – the rules of macroevolution*, Ed. Marcello Barbieri, Springer, Dordrecht, The Netherlands- Series Biosemiotics, V. 1, 436 pg, Chapter 5, pg. 69-110, ISBN 978-1-4020-6339-8.

* Goldenfeld N. & Woese C. (2007). Biology's next revolution. *Nature* 445: 369.

* Hornos J. E. M., Hornos Y. M. M. (1993). Algebraic model for the evolution of the genetic code. *Phys. Rev. Lett.* 71: 4401–4404.

* Jiménez-Montaño, M.A., de la Mora-Basañez, C.R., Pöschel, T. (1996). The hypercube structure of the genetic code explains conservative and non-conservative amino acid substitutions in vivo and in vitro. *Biosystems* 39: 117–125.

* Johnson D. B. F., Wang L. (2010). Imprints of the genetic code in the ribosome. *Proc. Natl. Acad. Sci USA.* 55(6): 1546-1552.

* José, M. V., Morgado, E. R., Govezensky, T. (2007). An extended RNA code and its relationship to the standard genetic code: an algebraic and geometrical approach. *Bull. Math. Biol.* 69: 215-243.

* José, M. V., Govezensky, T., García, J. A., Bobadilla, J.R. (2009). *On the evolution of the standard genetic code: vestiges of scale invariance from the RNA World in current prokaryote genomes.* Plos ONE. February 2, vol. 4, e4340.

* José, M. V., Morgado, E. R., Govezensky, T. Genetic Hotels for the Standard Genetic Code: Evolutionary Analysis based upon Novel Three-dimensional Algebraic Models. Submitted: *Bull. Math. Biol.*

* Joyce, G. F. (1991). The rise and fall of the RNA world. *New Biol.* 3, 399–407.

* Joyce G. F. (2002). The antiquity of RNA-based evolution. *Nature* 418: 214-221.

* Lazcano, A. (2007). Prebiotic evolution and the origin of life: is a system-level understanding feasible? En: *Systems Biology. Volume I: Genomics.*, Eds. I. Rigoutsos and G. Stephanopoulos, Oxford University Press, New York, 2007, p. 57–78.

- * Nagel G. M., Doolittle R. F. (1995). Phylogenetic analyses of aminoacyl-tRNA synthetase families. *J. Mol. Evol.* 40: 487–498.
- * Nicholas H. B., McClain W. H. (1995). Searching tRNA sequences for relatedness to aminoacyl-tRNA synthetase families. *J Mol Evol* 40: 482–486.
- * Osawa, S., Jukes, T. H., Watanabe, K., Muto, A. (1992). Recent evidence for evolution of the genetic code. *Microbiol. Rev.* 56: 229-264.
- * Rodin, S. N., Rodin A. S. (2008). On the origin of the genetic code: signatures of its primordial complementarity in tRNAs and aminoacyl-tRNA synthetases. *Heredity* 100: 341-355.
- * Sánchez R., Morgado E. R., Grau R. (2005). A genetic code boolean structure I. The meaning of Boolean deductions. *Bull. Math. Biol.* 67: 1–14.
- * Schultes, E. A., Bartel, D. P. (2000). One sequence, two ribozymes: implications for the emergence of new ribozyme folds. *Science* 289: 448–452.
- * Sornette D (2000) *Critical Phenomena in Natural Sciences*. Chaos, fractals, self-organization and tools. Springer-Verlag.
- * Vetsigian, K., Woese, C., Goldenfeld, N. (2006). Collective evolution and the genetic code. *Proc. Natl. Acad. Sci. USA* 103: 10696-10701.
- * Wilson K. (1979). Problems in physics with many scales of length. *Sci Am* 241: 158–179.
- * Wilson E. O., Bossert, W. H. (1971). *A Primer of Population Biology*. Sinauer Associates, INC. Publishers, Sunderland, Massachusetts.
- * Woese, C. (1968). *The Genetic Code*. Harper & Row.
- * Woese, C. R. 2000. Interpreting the universal phylogenetic tree. *Proc. Nat. Acad. Sci. USA* 97:8392-8396.
- * Woese C. R., Olsen G. J., Ibba M., Söll D. (2000). Aminoacyl-tRNA synthetases, the genetic code and the evolutionary process. *Microbiol. Mol. Biol. Rev.* 64: 202–236.