# El CÃ³digo GenÃ©tico

Alumbac - 24/03/2010

El cÃ³digo genÃ©tico es el conjunto de normas por las que la informaciÃ³n codificada en el material genÃ©tico (secuencias de ADN o ARN) se traduce en proteÃ-nas (secuencias de aminoÃ¡cidos) en las cÃ©lulas vivas. El cÃ³digo define la relaciÃ³n entre secuencias de tres nucleÃ³tidos, llamadas codones, y aminoÃ¡cidos. Un codÃ³n se corresponde con un aminoÃ¡cido especÃ-fico.

La secuencia del material genÃ©tico se compone de cuatro bases nitrogenadas distintas, que tienen una funciÃ³n equivalente a letras en el cÃ³digo genÃ©tico: adenina (A), timina (T), guanina (G) y citosina (C) en el ADN y adenina (A), uracilo (U), guanina (G) y citosina (C) en el ARN.

Debido a esto, el nÃºmero de codones posibles es 64, de los cuales 61 codifican aminoÃ¡cidos (siendo ademÃ¡s uno de ellos el codÃ³n de inicio, AUG) y los tres restantes son sitios de parada. La secuencia de codones determina la secuencia aminoacÃ-dica de una proteÃ-na en concreto, que tendrÃ¡ una estructura y una funciÃ³n especÃ-ficas.

La transferencia de laÂ informaciÃ³n consiste en que el genoma de un organismo se encuentra en el ADN o, en el caso de algunos virus, en el ARN la porciÃ³n de genoma que codifica una proteÃ-na o un ARN se conoce como gen. Esos genes que codifican proteÃ-nas estÃ¡n compuestos por unidades de trinucleÃ³tidos llamadas codones, cada una de los cuales codifica un aminoÃ¡cido. Cada subunidad nucleotÃ-dica estÃ¡ formada por un fosfato, una desoxirribosa y una de las cuatro posibles bases nitrogenadas en la configuraciÃ³n en doble hÃ©lice ademÃ¡s, estos puentes siempre se forman entre una adenina de una cadena y una timina de la otra y entre una citosina de una cadena y una guanina de la otra. Esto quiere decir que el nÃºmero de residuos A y T serÃ¡ el mismo en una doble hÃ©lice y lo mismo pasarÃ¡ con el nÃºmero de residuos de G y C. En el ARN, la timina (T) se sustituye por uracilo (U), y la desoxirribosa por una ribosa.

Cada gen codificante de proteÃ-na se transcribe en una molÃ©cula plantilla, que se conoce como ARN mensajero o ARNm. Ã‰ste, a su vez, se traduce en el ribosoma, en una cadena aminoacÃ-dica o polipeptÃ-dica. En el proceso de traducciÃ³n se necesita un ARN de transferencia especÃ-fico para cada aminoÃ¡cido con el aminoÃ¡cido unido a Ã©l covalentemente, guanosina trifosfato como fuente de energÃ-a y ciertos factores de traducciÃ³n. Los ARNt tienen anticodones complementarios a los codones del ARNm y se pueden "cargar" covalentemente en su extremo 3' terminal CCA con

aminoÃ¡cidos. Los ARNt individuales se cargan con aminoÃ¡cidos especÃ-ficos por las enzimas llamadas aminoacil ARNt sintetasas, que tienen alta especificidad tanto por aminoÃ¡cidos como por ARNt. La alta especificidad de estas enzimas es motivo fundamental del mantenimiento de la fidelidad de la traducciÃ³n de proteÃ-nas.

Hay 4Â³ = 64 combinaciones diferentes de codones que sean posibles con tripletes de tres nucleÃ³tidos: los 64 codones estÃ¡n asignados a aminoÃ¡cido o a seÃ±ales de parada en la traducciÃ³n. Si, por ejemplo, tenemos una secuencia de ARN, UUUAAACCC, y la lectura del fragmento empieza en la primera U (convenio 5' a 3'), habrÃ-a tres codones que serÃ-an UUU, AAA y CCC, cada uno de los cuales especifica un aminoÃ¡cido. Esta secuencia de ARN se traducirÃ¡ en una secuencia aminoacÃ-dica de tres aminoÃ¡cidos de longitud.

El cÃ³digo genÃ©tico es compartido por todos los organismos conocidos, incluyendo virus y organulos, aunque pueden aparecer pequeÃ±as diferencias. AsÃ-, por ejemplo, el codÃ³n UUU codifica para el animoÃ¡cido fenilalanina tanto en bacterias, como en arqueas, como en eucariontes. Este hecho indica que el cÃ³digo genÃ©tico ha tenido un origen Ãºnico en todos los seres vivos conocidos.

Gracias a la genÃ©tica molecular, se han distinguido 22 cÃ³digos genÃ©ticos, que se diferencian del llamado cÃ³digo genÃ©tico estÃ¡ndar por el significado de uno o mÃ¡s codones. La mayor diversidad se presenta en las mitocondrias, orgÃ¡nulos de las cÃ©lulas eucariotas que se originaron evolutivamente a partir de miembros del dominio Bacteria a travÃ©s de un proceso de endosimbiosis. El genoma nuclear de los eucariotas sÃ³lo suele diferenciarse del cÃ³digo estÃ¡ndar en los codones de iniciaciÃ³n y terminaciÃ³n.

## Especificidad y continuidad

NingÃºn codÃ³n codifica mÃ¡s de un aminoÃ¡cido, ya que, de no ser asÃ-, conllevarÃ-a problemas considerables para la sÃ-ntesis de proteÃ-nas especÃ-ficas para cada gen. Tampoco presenta solapamiento: los tripletes se hallan dispuesto de manera lineal y continua, de manera que entre ellos no existan comas ni espacios y sin compartir ninguna base nitrogenada. Su lectura se hace en un solo sentido (5' - 3'), desde el codÃ³n de iniciaciÃ³n hasta el codÃ³n de parada. Sin embargo, en un mismo ARNm pueden existir varios codones de inicio, lo que conduce a la sÃ-ntesis de varios polipÃ©ptidos diferentes a partir del mismo transcrito.

## DegeneraciÃ³n

El cÃ³digo genÃ©tico tiene redundancia pero no ambigÃ¼edad (ver las tablas de codones). Por ejemplo, aunque los codones GAA y GAG especifican los dos el Ã¡cido glutÃ¡mico (redundancia), ninguno especifica otro aminoÃ¡cido (no ambigÃ¼edad). Los

codones que codifican un aminoÃ¡cido pueden diferir en alguna de sus tres posiciones.

De una posiciÃ³n de un codÃ³n se dice que es cuatro veces degenerada si con cualquier nucleÃ³tido en esta posiciÃ³n se especifica el mismo aminoÃ¡cido. Por ejemplo, la tercera posiciÃ³n de los codones de la glicina (GGA, GGG, GGC, GGU) es cuatro veces degenerada, porque todas las sustituciones de nucleÃ³tidos en este lugar son sinÃ³nimos; es decir, no varÃ-an el aminoÃ¡cido. SÃ³lo la tercera posiciÃ³n de algunos codones puede ser cuatro veces degenerada. Se dice que una posiciÃ³n de un codÃ³n es dos veces degenerada si sÃ³lo dos de las cuatro posibles sustituciones de nucleÃ³tidos especifican el mismo aminoÃ¡cido. Por ejemplo, la tercera posiciÃ³n de los codones del Ã¡cido glutÃ¡mico (GAA, GAG) es doble degenerada. En los lugares dos veces degenerados, los nucleÃ³tidos equivalentes son siempre dos purinas (A/G) o dos pirimidinas (C/U), asÃ- que sÃ³lo sustituciones transversionales (purina a pirimidina o pirimidina a purina) en dobles degenerados son antÃ³nimas. Se dice que una posiciÃ³n de un codÃ³n es no degenerada si una mutaciÃ³n en esta posiciÃ³n tiene como resultado la sustituciÃ³n de un aminoÃ¡cido. SÃ³lo hay un sitio triple degenerado en el que cambiando tres de cuatro nucleÃ³tidos no hay efecto en el aminoÃ¡cido, mientras que cambiando los cuatro posibles nucleÃ³tidos aparece una sustituciÃ³n del aminoÃ¡cido.Â Que el cÃ³digo genÃ©tico sea degenerado es lo que determina la posibilidad de mutaciones sinÃ³nimas.

La degeneraciÃ³n aparece porque el cÃ³digo genÃ©tico designa 20 aminoÃ¡cidos y la seÃ±al parada. Debido a que hay cuatro bases, los codones en triplete se necesitan para producir al menos 21 cÃ³digos diferentes. Por ejemplo, si hubiera dos bases por codÃ³n, entonces sÃ³lo podrÃ-an codificarse 16 aminoÃ¡cidos (4Â²=16). Y dado que al menos se necesitan 21 cÃ³digos, 4Â³ da 64 sodones posibles, indicando que debe haber degeneraciÃ³n.

Esta propiedad del cÃ³digo genÃ©tico lo hacen mÃ¡s tolerante a los fallos en mutaciones puntuales. Por ejemplo, en teorÃ-a, los codones cuatro veces degenerados pueden tolerar cualquier mutaciÃ³n puntual en la tercera posiciÃ³n, aunque el codÃ³n de uso sesgado restringe esto en la prÃ¡ctica en muchos organismos; los dos veces degenerados pueden tolerar una de las tres posibles mutaciones puntuales en la tercera posiciÃ³n. Debido a que las mutaciones de transiciÃ³n (purina a purina o pirimidina a pirimidina) son mÃ¡s probables que las de transversiÃ³n (purina a pirimidina o viceversa), la equivalencia de purinas o de pirimidinas en los lugares dobles degenerados aÃ±ade una tolerancia a los fallos complementaria.

Agrupamiento de codones por residuos aminoacÃ-dicos, volumen molar e hidropatÃ-a
Una consecuencia prÃ¡ctica de la redundancia es que algunos errores del cÃ³digo genÃ©tico sÃ³lo causen una mutaciÃ³n silenciosa o un error que no afectarÃ¡ a la proteÃ-na porque la hidrofilidad o hidrofobidad se mantiene por una sustituciÃ³n

equivalente de aminoÃ¡cidos; por ejemplo, un codÃ³n de NUN (N =cualquier nucleÃ³tido) tiende a codificar un aminoÃ¡cido hidrofÃ³bico. NCN codifica residuos aminoacÃ-dicos que son pequeÃ±os en cuanto a tamaÃ±o y moderados en cuanto a hidropatÃ-a; NAN codifica un tamaÃ±o promedio de residuos hidrofÃ-licos; UNN codifica residuos que no son hidrofÃ-licos. Estas tendencias pueden ser resultado de una relaciÃ³n de las aminoacil ARNt sintetasas con los codones heredada un ancestro comÃºn de los seres vivos conocidos.

La relaciÃ³n entre el ARNm y el ARNt a nivel de la tercera base se puede producir por bases modificadas en la primera base del anticodÃ³n del ARNt, y los pares de bases formados se llaman "pares de bases wobble" (tambaleantes). Las bases modificadas incluyen inosina y los pares de bases que no son del tipo Watson-Crick U-G.

El origen del cÃ³digo genÃ©tico

A pesar de las variaciones que existen, los cÃ³digos genÃ©ticos utilizados por todas las formas conocidas de vida son muy similares. Esto sugiere que el cÃ³digo genÃ©tico se estableciÃ³ muy temprano en la historia de la vida y que tiene un origen comÃºn en las formas de vida actuales. AnÃ¡lisis filogenÃ©tico sugiere que las molÃ©culas ARNt evolucionaron antes que el actual conjunto de aminoacil-ARNt sintetasas.

El cÃ³digo genÃ©tico no es una asignaciÃ³n aleatoria de los codones a aminoÃ¡cidos.Por ejemplo, los aminoÃ¡cidos que comparten la misma vÃ-a biosintÃ©tica tienden a tener la primera base igual en sus codones y aminoÃ¡cidos con propiedades fÃ-sicas similares tienden a tener similares a codones.[

Experimentos recientes demuestran que algunos aminoÃ¡cidos tienen afinidad quÃ-mica selectiva por sus codones. Esto sugiere que el complejo mecanismo actual de traducciÃ³n del ARNm que implica la acciÃ³n ARNt y enzimas asociadas, puede ser un desarrollo posterior y que, en un principio, las proteÃ-nas se sintetizaran directamente sobre la secuencia de ARN, actuando Ã©ste como ribozima y catalizando la formaciÃ³n de enlaces peptÃ-dicos (tal como ocurre con el ARNr 23S del ribosoma).

Se ha planteado la hipÃ³tesis de que el cÃ³digo genÃ©tico estÃ¡ndar actual surgiera por expansiÃ³n biosintÃ©tica de un cÃ³digo simple anterior. La vida primordial pudo adicionar nuevos aminoÃ¡cidos (por ejemplo, subproductos del metabolismo), algunos de los cuales se incorporaron mÃ¡s tarde a la maquinaria de codificaciÃ³n genÃ©tica. Se tienen pruebas, aunque circunstanciales, de que formas de vida primitivas empleaban un menor nÃºmero de aminoÃ¡cidos diferentes,[12] aunque no se sabe con exactitud que aminoÃ¡cidos y en que orden entraron en el cÃ³digo genÃ©tico.

Otro factor interesante a tener en cuenta es que la selecciÃ³n natural ha favorecido la degeneraciÃ³n del cÃ³digo para minimizar los efectos de las mutaciones[13] . Esto ha llevado a pensar que el cÃ³digo genÃ©tico primitivo podrÃ-a haber constado de codones de dos nucleÃ³tidos, lo que resulta bastante coherente con la hipÃ³tesis del balanceo del ARNt durante su acoplamiento (la tercera base no establece puentes de hidrÃ³geno de Watson y Crick).

DNA to Amino Acid Conversion
DNA Sequence -> Protein Sequence converter
DNA to protein translation (6 frames/17 genetic codes)

Tablas del cÃ³digo genÃ©tico

The Genetic Codes â†' Genetic Code Tables