

# INGENIERÍA MOLECULAR.

## Duplicación del ADN y código genético. Parte 2 de 2

Autor:

Carlos J. Alés-Esteban Ing. Ind. por la ETS de Sevilla

### INTRODUCCIÓN

En la primera parte del artículo Ingeniería Molecular se analizó como se duplica el ADN, unos de los procesos biológicos más importantes de la naturaleza, donde se apreciaba las dificultades que la química orgánica ha tenido que solventar para llevarlo a cabo con eficacia. En esta segunda parte se plantea un análisis sobre el Código Genético. Las actuales teorías aceptables sobre el origen de la vida suelen centrarse en cómo en las condiciones tan adversas en las que se encontraba nuestro planeta hace millones de años se pudo producir ese increíble fenómeno que denominamos vida y por el cual las moléculas orgánicas se ensamblan, se desarrollan, se especializan, actúan en cooperación, se adaptan a su entorno y se multiplican. Sin embargo hay muy poca especulación científica sobre cuándo y cómo se pudo elaborar el Código Genético e incluso si en algún momento hubo alguno intermedio al actual más simple. También habrá que considerar que un “código” que sirve para traducir el genoma de cualquier ser vivo, casi por definición, deberá ser poco o nada variable, todo lo más evolutivo en cuanto a la capacidad traductora. En cualquier caso, y puesto que se desconoce, no se va a tratar como se pudo generar, pero si nos podemos preguntar el porqué este y no otro.

### 4. UN PROBLEMA DE CÓDIGOS

La Naturaleza ha desarrollado soluciones que dan pie a plantear cuál hubiese sido la adoptada por un Ingeniero

si hubiese tenido que hacer frente a dicho reto. Evidentemente el problema se va a plantear en un marco teórico, puesto que el Ingeniero desconoce hasta qué punto su solución “técnica” podría ser llevada a cabo por la química orgánica; sería algo así como llegar a la conclusión de que lo “óptimo” para satisfacer la sed es agua, pero estuviésemos en un desierto; sería la solución mejor, pero imposible de llevar a cabo en ese entorno.

En la primera parte de este artículo -duplicación del ADN- se decía que las “letras” del ADN eran purinas/pirimidinas biunívocamente relacionadas, y solo en un apartado se ha nombrado a la Adenina (A) y Timina (T) sin comentar si se usan o no más “letras”, cuando en cualquier libro de Biología esta información se enseña en los primeros capítulos. El motivo de ocultarlo no ha sido otro que poder plantear un problema, al menos para gran parte de los lectores, sin que se supiese “a priori” el resultado real que se da en los organismos vivos.

El enunciado del problema sería: ¿cuál de los posibles modelos de código genético que cubren su objetivo escogería un Ingeniero?

#### 4.1. EL CÓDIGO

Los seres vivos tienen que traspasar de “padres a hijos” su herencia, en concreto necesitan transmitir la composición de las distintas proteínas, que como se comentó, son los componentes biológicos estructurales y funcionales del organismo.

Pero las proteínas son “simplemente” aminoácidos “encadenados” en un determinado orden, recogido en el ADN, por lo que es necesario un Código, de manera que cuando una “máquina-herramienta-biológica”, el Ribosoma, vaya “leyendo” lo escrito con él en el ADN, haga la interpretación y

fabrique una proteína con los aminoácidos en la secuencia que se le indica.

Por tanto, por un lado hay unas secuencias de purina/pirimidina (las letras) que están ordenadas en los cromosomas del ADN y por otro el complejo proteico Ribosoma que va construyendo una cadena en un determinado orden con los 20 aminoácidos básicos.

#### 4.2. UN “INICIO”, UNA “PARADA” y 20 AMINOÁCIDOS BÁSICOS

Las “letras”, tienen que ser capaces de diferenciar los 20 aminoácidos y también dónde se empieza a contabilizar lo que será el primer aminoácido de la proteína, un “inicio”, y cuál será el último una “parada”.

Esto requiere que sea necesario codificar, como mínimo, 22 elementos distintos.

El alfabeto semita poseía 23 signos o letras, el latino primitivo constaba de 22 caracteres y el alfabeto español actual se puede considerar que tiene 27 letras, por tanto, son códigos con magnitudes parecidas.

#### 4.3. CÓDIGOS POSIBLES

Si el ADN tuviese solo 2 bases una purina/una pirimidina, (llamémoslas A y B) un código posible sería aquel que cuando en el ADN se encontrase la A se incorporase el aminoácido 1 y si la B el aminoácido 2 (A=1, B=2). Insuficiente, pues 20 opciones no se podrían identificar. También podría ser que si se encontrasen seguidas, por ejemplo, AA esto se interpretase como que hay que añadir el aminoácido 1, si lo que se encuentra consecutivo es AB se añadiese el aminoácido 2, si fuese BA el 3 y si se encuentra BB el 4; (AA=1, AB=2, BA=3 y BB=4) por lo tanto con dos bases y “leyendo” secuencias agrupadas en dos, se podrían diferenciar 4

aminoácidos. Este código también sería insuficiente para especificar los 22 elementos comentados.

Si se mantienen dos bases pero se “lee” de tres en tres, ya se podría crear una correspondencia con ocho aminoácidos (AAA=1, AAB=2, ABA=3, ABB=4, BAA=5, BAB=6, BBA=7 y BBB=8) es decir  $2^3=8$ . Aún insuficiente. Con dos bases, se tendría que “leer”, como mínimo, en secuencias de cinco, ya que  $2^5=32 > 22$  e incluso quedarían 10 opciones libres.

Por lo tanto:

Primer código posible: 2 bases y “secuencias de lectura de 5 en 5” ya que  $2^5=32 > 22$

Si se amplía el número de bases a 4 (2 purinas y 2 pirimidinas, ABCD) y la agrupación de lectura fuese de dos, este código sería insuficiente pues  $4^2=16 < 22$ , pero si fuese de tres en tres, sería  $4^3=64 > 22$  con 42 opciones libres.

Segundo código: 4 bases y “lectura de 3” ya que  $4^3=64 > 22$

Tercer código: 6 bases y “lectura de 2” ya que  $6^2=36 > 22$

Cuarto código: 8 bases y “lectura de 2” ya que  $8^2=64 > 22$

Quinto código: 10 bases y “lectura de 2” ya que  $10^2=100 > 22$

Así se podría ir formando aún cinco códigos más, con 12, 14, 16, 18 y 20 bases con “secuencia de lectura 2” y finalmente

Undécimo código: 22 bases y “lectura de 1 en 1” ya que  $22^1=22$ ; este último código sería biunívoco y cada purina/pirimidina tendría asociado un aminoácido, un “inicio” o una “parada” y se leería de uno en uno.

#### 4.4. EL CÓDIGO DEL INGENIERO

Antes de averiguar el aplicado, si aún no lo conoce por una buena cultura en Biología, (o incluso si ya lo conoce) intente pensar, como Ingeniero, cuál de los 11 códigos sería el que elegiría, si tuviese en su mano esa decisión,

sabiendo que servirá para sustentar la herencia de todo los seres vivos de la Tierra, y quien sabe, quizás de todo el Universo.

#### 4.5. EXCESO DE POSIBILIDADES DE LOS CÓDIGOS

Excepto el undécimo código, de 22 purinas/pirimidinas y de lectura de uno en uno, todos los demás generan combinaciones extras o sin determinar, ya que superan las 22 necesarias. Como esto ocurre en los 10 primeros, si alguno de estos fuese el usado por la Naturaleza ¿qué cree que se puede hacer con las opciones libres para que siga siendo preciso en la traducción y el complejo Ribosoma no se equivoque de aminoácido?

#### 4.6. SOLUCIÓN

La Naturaleza ha desarrollado el segundo código formado por cuatro bases, dos purinas, la Adenina (A) y la Guanina (G) y dos pirimidinas, la Timina (T) y la Citosina (C) y “lee” secuencialmente de tres en tres, agrupación que se le denomina Codón - nombre fonéticamente poco afortunado-. Por tanto el ADN son unas sucesiones de millones de purinas/pirimidinas (A, G, T, C) que se “leen” de codón en codón, es decir de tres en tres.

La Guanina (G) forma tándem con la Citosina (C) uniendo las dos cadenas del ADN igual que la Adenina (A) con la Timina (T); por tanto los apareamientos de bases son G×C y A×T (nematécnicamente puede ser útil acordarse de Gran Canarias y Alta Tensión).

Por ejemplo, hay pedazos en el ADN humano que se repiten miles y miles de veces, como es el caso de la denominada secuencia Alu de 288 nucleótidos, que está copiada aproximadamente un millón de veces en nuestro genoma, lo que representa un 10% del mismo sin que se le conozca, todavía, claramente su función (a este tipo de secuencias se les denomina también código basura).

En concreto Alu es: (siempre se escribe una sola de las cadenas del ADN,

puesto que la otra es complementaria según G×C y A×T)

(5)→GGC,CGG,GCG,CGG,TGG,CTC,ACG,CCT,GTA,ATC,CCA,GC A,CTT,TGG,GAG,GCC,GAG GCG,GGC,GGC,TCA,CCT,GAG,GTC,AG G,AGT,TCG,AGA,CCA,GCC,TGG,CCA,ACA,TGG,TGA, AAC,CCC,GTC,TCT,ACT,AAA,AAT,ACA,AAA,ATT,AGC,CGG,GCG,TGG,TGG,CGC,GCG,CCT, GTA,ATC,CCA,GCT,ACT,CGG,GAG,GCT,GAG,GCA,GG A,GAA,TCG,CTT,GAA,CCC,GGG,AGG, CGG,AGG,TTG,CAG,TGA,GCC,GAG,ATC,GCG,CCA,CTG,CAC,TCC,AGC,CTG,GGC,GAC,AGA,GCG,AGA,CTC,CGT,CTC,AAA,AA(3)→

La otra cadena (complementaria) sería ←(3)CCG,GCC,CGC,GCC,ACC,GAG, ... ..., TTT TTT ←(5)

Las 288 bases van todas juntas, se ha presentado separadas por comas, en grupos de tres, por hacerlo más didáctico. En definitiva viene a indicar la secuencia de 96 aminoácidos (véase tabla de Código Genético cambiando la T por U).

El complejo proteico Ribosoma “leería” la secuencia Alu, según la tabla de la Fig. 5, de corrido, al igual que un músico una partitura, sintetizando una proteína de 96 aminoácidos concatenados ( $\sum_1^{96} -NC(R_i)C-$ ) en este orden:

Glicina+Arginina+Alanina+Arginina+Triptofano+Leucina+Treonina+Prolina+Valina+Isoleucina+Prolina+Alanina+Leucina+Triptofano+Valina+Alanina+Acido Glutámico+Alanina+... +Lisina.

#### 4.7. REFLEXIÓN SOBRE EL PROBLEMA

Es probable que la decisión de un Ingeniero hubiese sido un código con once purinas y once pirimidinas complementarias y realizar la “lectura” de una en una. Tendría ventajas, hubiese sido biunívoco, sin soluciones en exceso; para la misma información los cromosomas serían 2/3 más cortos, aunque la duplicación del ADN sería muy difícil al “trabajar” con 22 puri-

PRIMERA LETRA	SEGUNDA LETRA				TERCERA LETRA
	U	C	A	G	
U	Fenilalanina	Serina	Tirosina	Cisteína	U
	Fenilalanina	Serina	Tirosina	Cisteína	C
	Leucina	Serina	Parada	Parada	A
	Leucina	Serina	Parada	Triptófano	G
C	Leucina	Prolina	Histidina	Arginina	U
	Leucina	Prolina	Histidina	Arginina	C
	Leucina	Prolina	Glutamina	Arginina	A
	Leucina	Prolina	Glutamina	Arginina	G
A	Isoleucina	Treonina	Asparagina	Serina	U
	Isoleucina	Treonina	Asparagina	Serina	C
	Isoleucina	Treonina	Lisina	Arginina	A
	(Inicio)	Treonina	Lisina	Arginina	G
	Metionina				
G	Valina	Alanina	Ácido aspártico	Glicina	U
	Valina	Alanina	Ácido aspártico	Glicina	C
	Valina	Alanina	Ácido glutámico	Glicina	A
	Valina	Alanina	Ácido glutámico	Glicina	G

Fig. 5: Tabla Código Genético Universal. Codón-Aminoácidos

nas/pirimidinas. Este código hubiese podido ser progresivo, es decir, se irían sumando nuevas purinas/pirimidinas a nuevos aminoácidos y a medida que las proteínas de los organismos vivos hubiesen sido más complejas.

Si eliminamos este código, por la dificultad comentada, los códigos más simples que cumplen su objetivo son el de 2 elementos y lectura de 5, el de 4 elementos y codón de tres (el de la Tierra) y el de 6 elementos y agrupación de 2.

Se podría predecir que son los únicos que se pueden dar, teniendo cada uno ventajas e inconvenientes en la línea de lo ya comentado. Puesto que a “toro pasado”, se sabe que hay soluciones químicas de 4 purinas/pirimidinas, esto desecharía el de codón de 5 quedando solo, si es que hay otro en el Universo, el de 6 purinas/pirimidinas y codón de 2 elementos.

Se desconoce qué pasos o cómo los primeros elementos químicos que tenían “vida” llegaron a desarrollar el Código Universal. Si se ha encontrado que previo a la formación del ADN, hubo una fase intermedia de una Proteína enzimática (Ribozima) equivalente a una sola cadena del ADN -cambiando el azúcar desoxirribosa por ribosa (RNA) y la Timina por Uracilo (otra pirimidina)- y que tuvo simultáneamente propiedades de memorización y de sintetización de proteínas, y que posteriormente daría lugar al ADN, molécula más estable que el RNA.

Pero no se sabe cómo evolucionó ese RNA enzimático y primitivo al ADN que conocemos.

La realidad es que el Código Genético Universal se debió de generar muy en el origen de las primeras formas de vida, ya que una vez que se comienza a memorizar proteínas con una codifi-

cación, los descendientes tienen muy pocas posibilidades de modificar la estructura del código. De hecho toda la Naturaleza que se conoce, bacterias, arqueas, algas, plantas, insectos, peces, aves, dinosaurios, vertebrados, humanos,... utilizan el mismo, tras 3000 millones de años de evolución, y solo la correspondencia entre algún codón y su aminoácido asociado, es decir la traducción, varía en algún hongo, protozoo o en la mitocondria (bacteria fagocitada que se encuentra en las células con núcleo) y en estos casos, las diferencias son de solo cinco o seis codones de los 64 posibles, siendo además algunos de ellos el “inicio” o la “parada”.

Obviamente las “reglas” que han servido para la evolución de las especies y que han generado millones de variantes distintas, algunas tan diferentes como una flor y una ballena, no pueden ser las mismas para la “evolución” de los procesos, o posee otros patrones, ya que en comparación y durante el mismo periodo de tiempo, no se ha modificado nada.

#### 4.8. TRADUCCIÓN

Siguiendo con el problema, y puesto que las posibilidades del Código permiten 42 más que las necesarias ¿Qué se puede hacer con las restantes?

Una solución racional “técnica” sería hacer corresponder a cada aminoácido tres opciones ( $20 \cdot 3 = 60$ ) y con las cuatro restantes especificar “inicio” y “parada”.

De nuevo la química orgánica sorprende ya que no existe un criterio “lógico” en este aspecto, hay aminoácidos, como la Arginina, definidos por seis codones distintos (AGA,AGG,CGA,CGC,CGG,CGU) y otros como el Triptófano por solo uno (UGG); también hay tres codones de “parada” (UAA, UAG y UGA) y uno de “inicio” (AUG) asociado este a la metionina.

Ahora bien todas las combinaciones posibles de la tabla se corresponden, al menos, con un aminoácido, un “inicio” o una “parada”, con lo cual el complejo Ribosoma de ensamblaje

siempre tiene definido su significado.

Si bien se ha analizado cómo está estructurado el código, el lector se habrá planteado cómo se puede pasar de los genes del ADN formado por  $\Sigma_1^n$ -PO(5)azúcar(3)O-purina/pirimidina a una proteína formada por  $\Sigma_1^n$ -NC(R<sub>i</sub>)C-.

Si ya puede parecer compleja la simple duplicación del ADN, la “transcripción y traducción” desde el ADN, doble cadena helicoidal compacta, hasta la síntesis de las proteínas por el Ribosoma, es un proceso que ni el Ingeniero más creativo podría nunca imaginar. Hay que tener en cuenta que con mucha alegría se utiliza el verbo “leer el ADN” (en este artículo también se ha abusado de él) pero la química orgánica ni tiene ojos, ni hay luz, ni tiene manos para ir “leyendo” y “añadiendo” los aminoácidos precisos, por lo que ha tenido que establecer un proceso químico complejo, sorprendente, exacto y fascinante para la traducción.

## 5. CONCLUSIÓN

El ciudadano, en su día a día, se encuentra inmerso en un mundo lleno de “inventos”, tanto humanos que ve, como en “procesos” biológicos internos que no ve, pero en ambos casos no llega a saber muy bien como “funcionan”. Pocas personas, que no sean algo entendidas en la materia, ni siquiera sabrían explicar cómo se produce hielo en su frigorífico a partir de la electricidad y mucho menos aún los procesos que se realizan en sus células, como el ya comentado del ADN.

En la Biología en general, en sus distintas ramas, se puede decir que hay un desfase importante entre los conocimientos científicos alcanzados y el saber popular. También esta afirmación se podría aplicar a aspectos evolutivos, que incluso tienen errores populares muy extendidos. Si se preguntase en una encuesta de donde procede el hombre, la mayoría de la población diría que “*procedemos del mono, según la teoría expuesta en su día por Darwin*”. Gracias a la Filogenia (cien-

cia que estudia la evolución) se sabe que el hombre y el mono solo tuvieron un ancestro común (6) que bien nos puede parecer que tenía algo de mono, pero sin embargo, el linaje humano, ni siquiera en sus orígenes, paso por la etapa de mono. Por tanto, no evolucionó ni del chimpancé, ni del gorila ni del orangután, ni de ningún otro mono. Tampoco el mono evolucionó a partir del hombre, aunque ellos, los monos, también podrían afirmar que el ancestro común tenía algo, en su aspecto, de humano. La verdad es que **Darwin y Wallace** nunca afirmaron que el hombre evolucionó a partir del mono.

También en el pensamiento popular, consciente o inconscientemente, estructurado o no, subyace la idea que tras originarse la vida en la Tierra toda la evolución posterior queda justificada en base a la teoría evolutiva de Darwin (actualmente enmarcada en el Neodarwinismo) y la selección natural. Si bien esto es cierto a partir de la formación de los sistemas orgánicos, siendo la célula su unidad más elemental, no se puede aplicar a como se formaron, ni organizaron los elementos y procesos que la constituyen.

Desde que se nace, estamos tan acostumbrados a ver cómo crecen las plantas, se reproducen los animales o funciona nuestro organismo, que todo parece muy “natural” y por tanto fácil. Cuando nos sumergimos en ese mundo de la biología molecular y celular, nos descubre muchas sorpresas, ya que creyendo encontrar procesos relativamente simples, puesto que sin “conocimiento” no deberían ser muy complejos, asombra ver las dificultades a las que se tiene que enfrentar, y más aún, como están resueltos.

Si bien se acepta que se ha llegado a estos procesos de una manera evolutiva y aunque esta evolución molecular no esté basada en los principios expuestos por Darwin y Wallace para las especies, estos procesos poseen soluciones tan complejas y acertadas, que justo sería asumir que la química orgánica es inteligente, al contrario que la química inorgánica, y si no, al menos, reconocerle la cantidad suficiente de

inteligencia como la que es necesaria para resolver cualquier problema de Ingeniería.

## PARA SABER MÁS:

- (1) Francisco J. Ayala. “Darwin y el Diseño Inteligente”. Madrid. Alianza Editorial SA, 2010. 231p. ISBN: 978-84-206-4822-4
- (2) James D. Watson. “ADN - El secreto de la vida”. Madrid. Santillana Ediciones Generales, SL. 2006. 429p. ISBN: 84-306-0514-2
- (3) E. López Azpitarte y I. Nuñez de Castro. “Cruzando el Puente. Bioética”. Madrid. Editorial San Pablo. 2011. 603p. ISBN: 978-84-285-3855-8.
- (4) Alberts, Bray, Hopkin y varios. “Introducción a La Biología Celular”. 3ª Edición. México. Editorial Médica Panamericana. 2011. 732p. ISBN: 978-607-7743-18-7
- (5) Watson, Baker, Bell y varios. “Biología Molecular del Gen”. 5ª Edición. Buenos Aires. Editorial Médica. Panamericana. 2004. 776p. ISBN: 978-84-7903-505-1
- (6) Milton H. Gallardo. “Evolución. El Curso de la vida”. Buenos Aires. Editorial Médica Panamericana. 2011. 503p. ISBN: 978-950-06-0278-5
- (7) Jeremy M. Berg, John L. Tymoczko y Lubert Stryer. “Bioquímica” 6ª Edición. Editorial Reverté, SA. 2008 - 1.026 p. ISBN 978-84-291-7600-1
- (8) Wikipedia. Gráficos y artículos.