

EL CÓDIGO GENÉTICO CUMPLE 40 AÑOS

LUIS FRANCO VERA *

* Real Academia de Ciencias Exactas, Físicas y Naturales. Universidad de Valencia. Departamento de Bioquímica y Biología Molecular. 46100 Burjassot, Valencia. luis.franco@uv.es

INTRODUCCIÓN

En 1799, durante la campaña de Napoleón en Egipto, las tropas francesas descubrieron una estela de poco más de 1 m de alta, con inscripciones en dos sistemas egipcios de escritura, —caracteres jeroglíficos y demótico— y en griego clásico. Debido al lugar en que se encontró, la estela, que se conserva actualmente en el *British Museum*, se conoce hoy día con el nombre de piedra de Rosetta. La comparación del texto —un decreto de Ptolomeo V— en los diferentes idiomas fue determinante para que los especialistas, entre los que destacó el egiptólogo francés Jean-François Champollion, consiguieran descifrar los caracteres jeroglíficos en 1822.

Si el desciframiento de la piedra de Rosetta fue decisivo para la Egiptología, en la Biología Molecular se dio un acontecimiento de similar importancia. Había que descifrar el *código genético*, algo necesario para conocer como las células son capaces de *traducir* un lenguaje de ácidos nucleicos a otro de proteínas. En las líneas que siguen se intenta poner de manifiesto cómo el desciframiento del código genético resultó ser tan apasionante como el de los caracteres jeroglíficos, con la diferencia de que en vez de durar 23 años, el trabajo se terminó en poco más de 5. La competencia entre los diversos laboratorios implicados en la ingente tarea fue, posiblemente, uno de los factores que permitieron culminar la investigación en un tiempo récord, ya que todos trabajaron contra reloj en su afán de alcanzar la solución asegurando la prioridad de sus descubrimientos.

EL PLANTEAMIENTO DEL PROBLEMA

En 1866, Gregor Mendel empleó por primera vez el término *gen*. Aunque había de pasar mucho tiempo hasta que se perfilara su concepto —se puede decir que aún no se ha terminado de perfilar— y se conociera que el DNA es el componente de los genes, las investigaciones del fraile agustino echaron los cimientos de la Genética moderna. Casi un siglo más tarde, en 1958, Francis Crick, que 5 años antes había propuesto con James Watson la estructura de doble hélice para el DNA, postuló el *dogma central*, según el cual, el flujo de información genética desde el DNA hasta las proteínas —las destinatarias de esa información— pasa por el RNA (1). Aunque la publicación formal del *dogma central* tuvo lugar en 1958, la idea de que el RNA servía de intermediario estaba en la mente de Crick y en la de otros investigadores desde unos años antes, como veremos enseguida. En líneas generales, los procesos implicados en ese flujo se han revisado en otro lugar (2) y allí se han comentado las razones por las que se denomina *transcripción* al paso de la información desde el DNA hasta el RNA y *traducción* al tránsito de esa información desde el RNA hasta las proteínas. Baste ahora con recordar que, aunque hay varios tipos de RNA, sólo uno, el RNA mensajero, abreviado como mRNA, desempeña ese papel de intermediario informativo entre el DNA y las proteínas.

En 1955 Paul Zamecnik había descubierto que la traducción, o biosíntesis de proteínas, tiene lugar en los ribosomas (3), unos pequeños orgánulos presentes en todas las células. Los ribosomas de las células pro-

carióticas tienen un coeficiente de sedimentación¹ de 70S y están formados por dos subunidades de 50 y 30S. Los ribosomas eucarióticos son algo mayores, de 80S, y también poseen dos subunidades, en este caso de 60 y 40S. En todos los casos, las subunidades ribosomales están constituidas por proteínas y RNA. Este RNA ribosomal, del que existen varias moléculas de diverso tamaño, se abrevia como rRNA. Aunque en un principio se pensó que el rRNA desempeñaba el papel de mensajero, Jacob y Monod, en 1961, postularon la existencia del mRNA como molécula independiente (4) y ese mismo año se demostró su presencia tanto en procariotas como en eucariotas.

Dejando aparte los problemas mecánicos, tan complejos que siguen siendo objeto de una activa investigación, la transcripción es conceptualmente fácil. Es bien sabido que la información contenida en una molécula de DNA depende de su secuencia, del orden en que se encuentran encadenados los nucleótidos que lo componen, de la misma manera que el sentido de una frase escrita depende del orden en que se hayan colocado las letras que la forman. Los nucleótidos que constituyen el DNA difieren en sus bases, que pueden ser 4: adenina, guanina, timina y citosina, que en lo sucesivo designaremos de ordinario sólo con sus iniciales, A, G, T y C. El RNA es también un polímero lineal formado por 4 nucleótidos, pero en este caso T está reemplazada por el uracilo, U, una base de rasgos estructurales muy similares. Teniendo en cuenta estos datos y la conocida propiedad de las bases de aparearse específicamente, G con C y A con T (o U), es fácil comprender que si una cadena de DNA tiene, por ejemplo, la secuencia ...ATTGGCTACGAT-TACGGTA... cuando se transcriba su hebra complementaria, el mRNA resultante se leerá: ...AUUGGCU-ACGAUUACGGUA... Pero las cadenas polipeptídicas que forman las proteínas están constituidas por 20 aminoácidos, por lo que la traducción presenta un problema conceptual más serio: ¿cómo pasar de un alfabeto de 4 nucleótidos a otro de 20 aminoácidos? Otro problema añadido, y no pequeño, es la falta de complementariedad entre bases y aminoácidos. Si bien existe un apareamiento específico entre las bases que forman los nucleótidos, examinando las estructuras de

bases y de aminoácidos, no se ve ninguna relación estructural entre ambos grupos de compuestos.

Pero vayamos por partes. En primer lugar, ¿cuántos nucleótidos del RNA mensajero son necesarios para determinar la colocación de un aminoácido? Es evidente que, habiendo sólo 4 nucleótidos, la respuesta no puede ser *uno*. Si así fuera, sólo se podrían determinar 4 aminoácidos. ¿Y si fueran 2? Entonces, la combinatoria nos dice que cabrían $4^2=16$ posibilidades: AA, AG, AU, AC, etc. Siguen siendo insuficientes para los 20 aminoácidos. Si fueran 3, es decir, si cada triplete de nucleótidos (o de bases, que es la parte variable de un nucleótido a otro) determinara la colocación de un aminoácido, las posibilidades serían $4^3=64$. Ahora sobran posibilidades, pero esto no constituye un problema si hay algunos tripletes que no corresponden a ningún aminoácido o si cada aminoácido viene determinado por más de un triplete. Es esencial comprender que en el flujo direccional de información, que va del mRNA a las proteínas, lo problemático sería que una combinación de bases pudiera corresponder a más de un aminoácido. Esa ambigüedad daría al traste con la fidelidad con que debe producirse la traducción, ya que la maquinaria traduccional se encontraría con indeterminaciones que no se podrían resolver fácilmente. Por el contrario si, como veremos que sucede, los tripletes UUA y CUG corresponden al aminoácido leucina, no hay problemas. Tanto cuando aparezca uno como otro en el mRNA, la maquinaria de traducción colocará leucina en el lugar correspondiente de la cadena polipeptídica y, como nunca hay un flujo de información desde proteínas hacia RNA, las células no se encontrarán con situaciones ambiguas.

El primer investigador que llegó a la conclusión de que cada unidad de codificación, que pronto comenzó a designarse *codón*, estaba constituida por un triplete de nucleótidos fue el físico ruso Georgii Gamov (5). Lo hizo por un simple razonamiento matemático, como se ha expuesto más arriba, ya que la formación de Gamov no era biológica ni química. Pero ello no fue obstáculo para que, tras el descubrimiento de la doble hélice, se interesara profundamente por los problemas planteados por la transmisión del mensaje

¹ El coeficiente de sedimentación es un parámetro relacionado con la velocidad a la que sedimenta una partícula en un campo centrífugo. Se mide en unidades Svedberg (S), denominadas así en honor de Theodor Svedberg, pionero de la ultracentrifugación aplicada a la bioquímica.

genético. Comprendió que la solución llegaría a través de un esfuerzo multidisciplinar y en 1954, con un innegable toque de buen humor, fundó el “RNA Tie Club”, llamado así porque sus 24 miembros —20 por el número de los aminoácidos y 4 por el de los nucleótidos— llevaban una corbata con una hélice bordada, que diseñó el propio Gamov. La finalidad del club era “resolver el enigma de la estructura del RNA y comprender cómo construye las proteínas”. Francis Crick era, por derecho propio, uno de los miembros del club; le correspondía el apelativo de “Tyr”, por el aminoácido tirosina.

Centremos ahora la atención en el segundo de los problemas enunciados anteriormente, es decir, en el modo por el que puede cada aminoácido reconocer su triplete específico. Las actividades del “RNA Tie Club” fueron decisivas al respecto. Francis Crick propuso en 1955 la *hipótesis del adaptador* según la cual “alguna estructura, hasta ahora desconocida, transporta los aminoácidos y los coloca en el orden correspondiente a la secuencia sobre una hebra de ácido nucleico”. La hipótesis no llegó a publicarse en una auténtica revista científica, aunque como diría Crick más tarde, fue su artículo no publicado más influyente. Pero dejemos que explique él mismo la hipótesis:

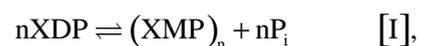
«La idea fundamental era que resulta muy difícil considerar cómo el DNA o el RNA en cualquier posible estructura pueden constituir un molde inmediato para las cadenas laterales de los veinte aminoácidos. Lo que se esperaría de cualquier estructura posible es una distribución específica de grupos atómicos que puedan formar enlaces de hidrógeno. Por lo tanto, propuse una teoría según la cual existirían veinte adaptadores (uno para cada aminoácido), junto con veinte enzimas especiales. Cada enzima sería capaz de unir un aminoácido concreto a su propio adaptador específico. Estos complejos se trasladarían por difusión al molde de RNA. Una molécula de adaptador podría encajar sólo en aquellos lugares del molde de ácido nucleico en los que pudiera formar los enlaces de hidrógeno necesarios para mantenerlo en su lugar. Al hacer eso, el adaptador habría llevado a su aminoácido precisamente al lugar correcto en que hacía falta» (6)

No hubo que esperar mucho para confirmar la hipótesis del adaptador. En 1958 el grupo de Zamecnik, el descubridor del papel del ribosoma en la síntesis de proteínas, describió la existencia de unas moléculas pequeñas de RNA que desempeñaban el papel sugerido por Crick (7). Habida cuenta de su papel, esa nueva especie de RNA pasó a denominarse RNA de transferencia, tRNA. Cada aminoácido reconoce y se une a un tRNA específico, gracias a la acción de la correspondiente aminoacil-tRNA sintetasa. Y la presencia en el tRNA de un triplete complementario al codón, que llamaremos a partir de ahora el *anticodón*, permite que encajen los diversos aminoacil-tRNAs sobre el mRNA en un orden correcto. Todo este conjunto de interacciones tiene lugar sobre el ribosoma, que se encarga de separar de sus tRNAs los aminoácidos, ya ordenados, y de encastrarlos entre sí mediante la formación de los enlaces peptídicos².

HACIA EL DESCIFRAMIENTO DEL CÓDIGO

En este momento puede proponerse ya el tema central de este artículo. La existencia de 64 posibles tripletes de nucleótidos plantea inmediatamente la cuestión: ¿qué aminoácido codifica cada uno de los 64 codones? Descubrir este código, el *código genético* como se comenzó a llamar desde el principio, atrajo la atención de varios laboratorios en los inicios de la década de 1960, cuando ya se estaba en posesión de los conocimientos que se han recogido más arriba.

Pero los antecedentes de la historia hay que buscarlos en 1955. En ese año, en el laboratorio de Severo Ochoa se realizó un descubrimiento de crucial importancia. Consiguieron caracterizar la polinucleótido fosforilasa, enzima que cataliza la reacción reversible:



en la que XMP es uno cualquiera de los 4 nucleótidos que componen el RNA. Hay que caer en la cuenta de que el producto, que en el esquema se ha representado

² El lector interesado puede encontrar los detalles de este proceso, que caen fuera del alcance de este capítulo, en cualquier texto actual de Bioquímica o Biología Molecular.

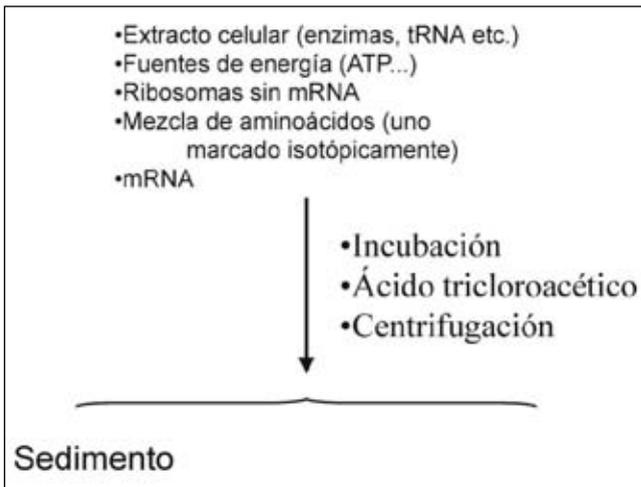


Figura 1. Esquema de un experimento de traducción *in vitro*. Tras la incubación de una mezcla de todos los componentes indicados, se añade ácido tricloroacético, que precipita las proteínas y ácidos nucleicos. Evidentemente, si el aminoácido marcado isotópicamente se ha incorporado a una cadena polipeptídica naciente, aparecerá radiactividad en el sedimento.

como $(XMP)_n$, es un polímero de ribonucleótidos, es decir un RNA. Lo que habían hecho Ochoa y sus colaboradores era sintetizar *in vitro*, por primera vez, un RNA, un ácido nucleico (8). Puede comprenderse la emoción del Profesor Severo Ochoa cuando realizó este descubrimiento. Muchos años más tarde, en 1989, lo recordaba en una conferencia pronunciada en Valencia, que comenzó con las palabras:

«Quisiera hoy discurrir sobre la satisfacción que puede producir el dedicar la vida a la investigación científica. Pocas veces he sentido emoción más intensa que cuando creí haber hecho descubrimientos de alguna trascendencia» (9).

La polinucleótido fosforilasa no tiene *in vivo* la función de catalizar la síntesis de RNA. A la vista del dogma central, la transcripción traslada al RNA la información contenida en el DNA. Es necesaria la presencia de un DNA que actúe como *molde* para que la transcripción tenga lugar y la polinucleótido fosforilasa ni requiere DNA, ni une los nucleótidos en un orden determinado, sino que lo hace al azar. De hecho, *in vivo* la enzima funciona según el esquema [I], pero de derecha a izquierda, es decir, destruyendo el RNA. El que la enzima no se encargara *in vivo* de sintetizar el RNA no quitó valor al hecho de haber conseguido sintetizar un RNA, al contrario, la investigación supuso el premio Nobel para Ochoa. Pero, para el fin que per-

sigue el presente artículo, el interés del descubrimiento radica en que hizo que Ochoa se interesara por el desciframiento del código genético.

Gracias a los trabajos de los grupos de Zamecnik y de Nirenberg, se conocían los requisitos necesarios para llevar a cabo la biosíntesis de proteínas en sistemas acelulares, es decir, *in vitro*. Esencialmente, son necesarios, además de los ribosomas y del mRNA, un extracto que contenga todas las enzimas, tRNAs y factores precisos, una fuente de energía que proporcione la necesaria para la formación del enlace peptídico y, por descontado, los aminoácidos que se polimerizarán al formar la cadena polipeptídica. La fig. 1 esquematiza un experimento de traducción *in vitro* y en él se aprecia cómo el sedimento final, que contiene los polipéptidos sintetizados, será radiactivo siempre que el aminoácido marcado isotópicamente se haya incorporado. El grupo de Ochoa comenzó a utilizar la reacción catalizada por la polinucleótido fosforilasa para sintetizar polinucleótidos artificiales, con la idea de utilizarlos en una reacción como la esquematizada en la fig. 1 en sustitución del mRNA. Pero cuando estaban comenzando ese trabajo, en el V Congreso Internacional de Bioquímica celebrado en Moscú se enteraron de que otro grupo, el de Marshall Nirenberg se les había adelantado. Así lo narra Carlos Basilio, un colaborador de Ochoa, testigo del histórico momento:

«Cuando acabábamos de iniciar nuestro trabajo, Marshall Nirenberg comunicó en el Congreso Internacional de Bioquímica de Moscú que el ácido poliuridílico promovía la síntesis de polifenilalanina.

»Algunos científicos asistentes al Congreso contaron que los oyentes quedaron electrizados por la noticia. Cuando volvió de Moscú un miembro de nuestro Departamento y nos comunicó la novedad, quedamos literalmente petrificados» (10).

El experimento de Nirenberg, que se publicó poco más tarde (11), había seguido el protocolo esquematizado en la fig. 1. Y para preparar un polinucleótido que hiciera las veces de mensajero habían empleado la reacción de la polinucleótido fosforilasa, partiendo exclusivamente de UDP. Evidentemente, en estas condiciones sólo se podía obtener un homopolímero, el ácido poliuridílico, poli(U). Usando éste como mensajero, únicamente se conseguía que apareciera radiactividad en el sedimento cuando el aminoácido radiactivo era la fenilalanina. Y entonces, en el sedi-

mento aparecía un homopolipéptido, la polifenilalanina, poli(Phe). El experimento fue especialmente relevante por varias razones. La primera de ellas, era que se demostraba sin lugar a dudas el papel del mRNA en la síntesis de proteínas: en el poli(U) sólo existen los codones UUU y, por tanto, sólo puede esperarse la síntesis de un polipéptido repetitivo. Pero en segundo lugar, el experimento significó el inicio del desciframiento del código genético: el codón UUU codifica la fenilalanina.

Es fácil imaginar la decepción de los miembros del laboratorio de Ochoa; otro grupo se había adelantado a su investigación utilizando sus mismas armas. Pero lejos de desanimarse, la reacción de Ochoa y sus colaboradores fue lanzarse a una frenética carrera para conseguir descifrar el significado de los 63 codones restantes. Por supuesto el equipo de Nirenberg no abandonó el trabajo y en los dos años siguientes el grupo de Ochoa publicó 14 artículos sobre el código genético y el de Nirenberg 12. Con homopolinucleótidos se consiguió demostrar que el codón AAA corresponde a la lisina y el CCC a la prolina. Los resultados con poli(G) no fueron concluyentes y entonces comenzó el trabajo con polinucleótidos mixtos. El equipo de Ochoa consiguió una ligera ventaja sobre el de Nirenberg, al ser el primero en comprobar que los polinucleótidos que contenían dos bases diferentes eran capaces de dirigir la síntesis de polipéptidos que incluían diferentes aminoácidos. Por ejemplo, el producto formado por la reacción de la polinucleótido fosforilasa usando como sustratos UDP y CDP a una relación molar 5:1 promovía la incorporación mayoritaria de fenilalanina, pero también se incorporaban serina y leucina y, en menor proporción, prolina (12). Años después, Ochoa mantenía vivo el recuerdo del momento en que obtuvieron los primeros resultados de su experimento:

«Cuando, en nuestro primer experimento, Lengyel, Speyer y yo observamos ansiosamente el contador de radiactividad y vimos que los polinucleótidos que contenían dos bases diferentes promovían la incorporación de diversos aminoácidos, nuestro entusiasmo no tuvo límites. (...) Siempre recordaré aquel momento como uno de los más emocionantes de mi vida» (9).

Como la polinucleótido fosforilasa produce polinucleótidos de secuencia al azar, el producto obtenido por polimerización de U y C, puede contener codones

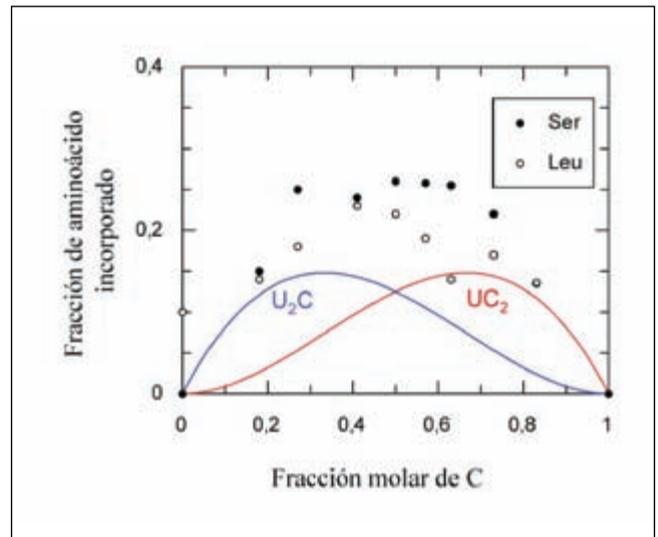


Figura 2. Resultados experimentales de la incorporación de serina y leucina dirigida por polinucleótidos formados por U y C en función de la fracción molar de este último nucleótido. Las curvas representan la probabilidad de encontrar tripletes de composición U₂C o UC₂.

de composición U₃, U₂C, UC₂ o C₃. A su vez, los codones U₂C pueden tener la secuencia UUC, UCU o CUU, mientras que los UC₂ pueden ser UCC, CUC o CCU. Con la salvedad de los codones UUU y CCC, que ya estaban asignados por los experimentos con homopolinucleótidos, quedaba la ingente tarea de comprobar cuál o cuáles de los otros 6 codones correspondían a los aminoácidos incorporados. Para abordar esta cuestión, tanto el grupo de Ochoa como el de Nirenberg utilizaron una estrategia similar: analizar la proporción de los aminoácidos incorporados en función de la proporción de U y C en la mezcla de polimerización. Los resultados experimentales para serina y leucina se recogen en la figura 2. Hay que tener en cuenta que, estadísticamente, la frecuencia con la que aparecerán en el producto de reacción los codones que contengan más U disminuirá a medida que disminuya la relación U:C. Por ejemplo, se puede calcular esta frecuencia para los codones de composición UC₂ y U₂C, con los resultados que se recogen en las curvas de la figura 2. A primera vista puede parecer sorprendente que los puntos experimentales no se ajusten a ninguna de las curvas, pero hay que tener en cuenta que, como el número de tripletes posibles es superior al de aminoácidos, el código genético puede estar *degenerado*, es decir, a cada aminoácido pueden corresponder varios tripletes, una posibilidad que se ha apuntado antes y que hizo notar por primera vez

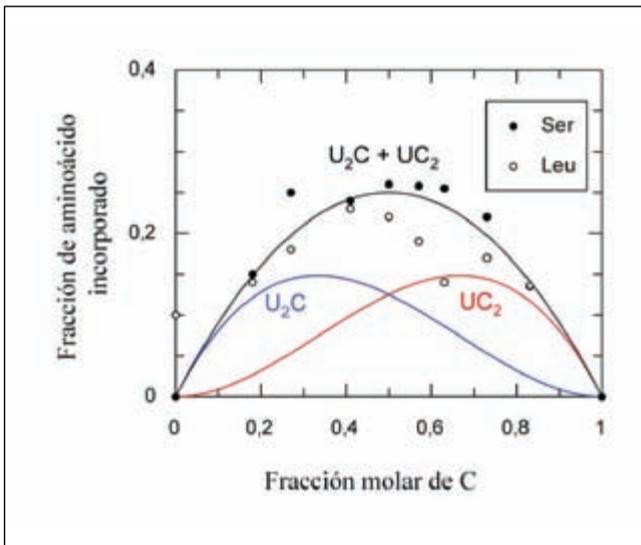


Figura 3. Se muestran los mismos resultados experimentales de la figura 2, pero se ha añadido la suma de las probabilidades de encontrar tripletes de composición U₂C o UC₂. Los resultados muestran que a los aminoácidos serina y leucina les corresponden al menos dos de los tripletes UUC, UCU, CUU, UCC, CUC o CCU.

Gamov (5). Por eso, en la figura 3 se añade la suma de las dos curvas, es decir la frecuencia de UC₂+U₂C en función de la proporción de C. Puede verse que todos los puntos experimentales se ajustan ahora razonablemente bien a la curva resultante, lo que indica que a los aminoácidos serina y leucina les corresponden más de uno de los tripletes UUC, UCU, CUU, UCC, CUC o CCU. Pero, en buena lógica, teniendo en cuenta la degeneración del código, alguno o algunos de estos codones podría también corresponder a fenilalanina (además del ya conocido UUU) o a prolina (además del CCC). Para aclarar esta cuestión, se analizó, por un procedimiento similar, la frecuencia de aparición de estos dos últimos aminoácidos, con los resultados de la figura 4. Analizando estos resultados con los criterios mencionados antes, se puede concluir que para la fenilalanina, además del codón UUU, existe uno U₂C y para la prolina, además del CCC existe otro UC₂.

De esta forma se fueron dando pasos entre 1962 y 1964. Por supuesto, los grupos de Ochoa y de Nirenberg analizaron todas las posibles combinaciones de nucleótidos en la mezcla de polimerización, con lo que obtuvieron resultados con poliUA, poliUG, etc.

En el transcurso de estos experimentos se obtuvo, además, la comprobación experimental de algo que se venía dando por supuesto, pero que no estaba aún demostrado de modo fehaciente: que los codones eran realmente tripletes (13). Pero el rompecabezas que implicaba el desciframiento total del código genético era demasiado complejo para que se pudiera resolver sólo con esa aproximación y hubo que esperar a que aparecieran nuevas técnicas.

LA SOLUCIÓN

El apartado anterior terminaba anunciando la necesidad de una metodología nueva para despejar las numerosas incertidumbres que se presentaban en el desciframiento del código genético. En este momento hizo su aparición en escena Gobind Khorana, que describió un procedimiento para la síntesis de polinucleótidos de secuencia alternante. El método daba un pequeño rodeo: se sintetizaba primero un DNA artificial de secuencia alternante y, mediante transcripción *in vitro* se lograba después el polirribonucleótido alternante (14). Así, el polirribonucleótido que se designa como poli(UC) tiene la secuencia ...UCUCUCUCUCUC... y, evidentemente, contiene una sucesión de dos tripletes: UCU y CUC. El producto obtenido en un sistema de traducción *in vitro* como el de la fig. 1, es un polipéptido de secuencia alternante, que puede ser Leu-Ser-Leu-Ser... o bien Ser-Leu-Ser-Leu..., ya que la traducción puede empezar por cualquier punto³. La conclusión de este experimento es que UCU y CUC corresponden a leucina o serina. Aún no se puede asignar cada triplete a uno de esos dos aminoácidos, pero por exclusión se obtiene otra conclusión. Si antes se había concluido que un triplete de composición U₂C pertenecía a la fenilalanina, ahora se puede descartar que sea UCU. De la misma manera se puede descartar que CUC corresponda a prolina. El método de Khorana permite también la síntesis de polinucleótidos con repetición de tres bases.

Con todas esas aproximaciones experimentales, se dio un paso de gigante en la dilucidación del código genético, pero aún faltaban otras aportaciones decisivas. La primera de ellas vino otra vez de la mano del

³ Es evidente que, *in vivo*, la traducción no puede empezar por cualquier lugar del mRNA. Pero los sistemas acelulares que se emplean *in vitro*, aunque son válidos para estudios relacionados con el código genético, no garantizan un inicio correcto de la traducción.

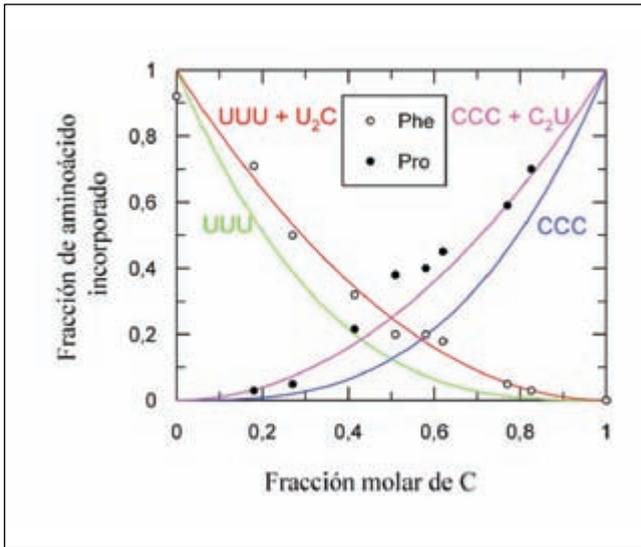


Figura 4. Resultados experimentales de la incorporación de fenilalanina y prolina dirigida por polinucleótidos formados por U y C en función de la fracción molar de este último nucleótido. Las curvas representan la probabilidad de encontrar tripletes formados homogéneamente por U o por C, así como la suma de estas probabilidades y las de encontrar, respectivamente, tripletes de composición U_2C o UC_2 . Los resultados muestran que a la fenilalanina, además del codón UUU le corresponde otro de la serie UUC, UCU, CUU, UCC, CUC o CCU y la prolina, además de CCC, posee algún codón de esa serie.

grupo de Nirenberg. En 1964 comprobaron que los ribosomas son capaces de unir un tRNA específico en presencia de un trinucleótido cuya secuencia sea la del codón correspondiente. La unión puede medirse fácilmente, porque los ribosomas se retienen en filtros de nitrocelulosa (15). Así pues, se incubaba una mezcla de ribosomas, un aminoacil-tRNA marcado isotópicamente y un trinucleótido, se filtra la mezcla y luego se mide la radiactividad retenida en el filtro. Si no aparece radiactividad en el filtro, es porque no se ha retenido el aminoacil-tRNA, o lo que es lo mismo, porque el trinucleótido *no es* el codón correspondiente al aminoácido cargado en el tRNA. Por el contrario, si el filtro aparece radiactivo, la conclusión es que el trinucleótido empleado *es* un codón perteneciente al aminoácido. El primer codón que se asignó de este modo fue GUU, como correspondiente a valina (16). A esta asignación siguieron otras llevadas a cabo por el mismo método.

La segunda aportación fue la de Holley. Este investigador había perfeccionado los métodos de secuenciación de RNA y consiguió llevar a cabo la primera determinación de la estructura primaria de un tRNA de

levadura específico de alanina, abreviadamente, tRNA^{Ala}. Evidentemente, la secuencia del anticodón permite, por complementariedad, deducir la del codón. Con este método indirecto se llegaron a descifrar tres codones.

En las líneas precedentes se han resumido las apasionantes investigaciones que, comenzadas con el experimento fundacional de Nirenberg con la asignación del codón UUU a la fenilalanina en 1961, permitieron en poco más de 5 años llegar a asignar con certeza 59 de los 63 tripletes restantes. A dos de ellos, UAA y UAG, se les adjudicó el papel de terminadores de la cadena polipeptídica, basándose en datos genéticos (para una revisión contemporánea a la investigación, ver la ref. 18). Como se ha apuntado, los logros fueron fundamentalmente de los laboratorios de Nirenberg, Ochoa, Khorana y Holley, aunque participaron otros muchos. La figura 5 recoge las aportaciones de los diferentes grupos al desciframiento del código. Como resultado de estas brillantes contribuciones a la Biología Molecular, Nirenberg, Khorana y Holley recibieron en 1968 el premio Nobel de Fisiología o Medicina “por su interpretación del código genético y su función en la síntesis de proteínas”. Ochoa, como se ha comentado, lo había recibido en 1959 por la síntesis enzimática de polirribonucleótidos.

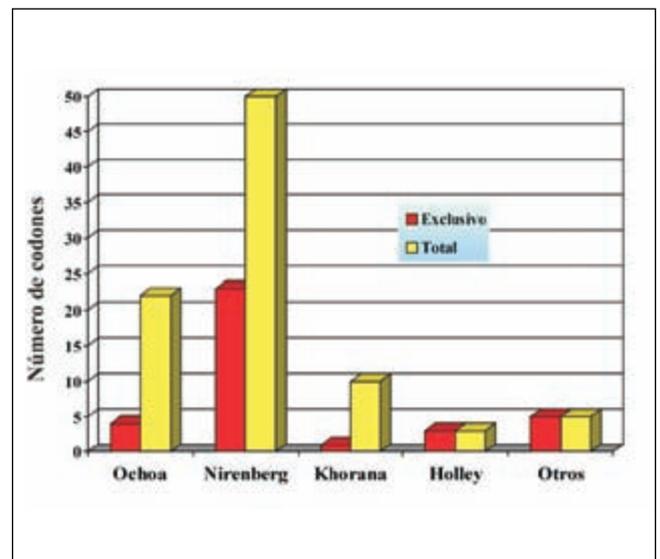


Figura 5. Número de codones asignados por los diversos grupos de trabajo. En amarillo se representa el total de codones asignados en cada laboratorio, mientras que las barras rojas indican los descubiertos exclusivamente por el grupo indicado.

1ª base (5')	2ª base				3ª base (3')
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Figura 6. El código genético. La tabla es de triple entrada: en la primera columna se indica la primera base del codón (extremo 5'); la segunda base se lee en las columnas centrales (coloreadas en amarillo) y la tercera (extremo 3') aparece en la última columna.

Quedaban, pues, 4 tripletes dudosos, UGA, AGG, AGC y CUG. Para el primero se dudaba si era también un codón de terminación o si pertenecía a la cisteína. Poco después se demostró que la primera opción era la válida. Los otros tres se asignaron provisionalmente a arginina, serina y leucina, respectivamente. El tiempo demostró que esas asignaciones provisionales eran válidas. De esta manera, se llegó a construir la tabla que resume el código genético (Fig. 6) y que resulta tan familiar a cualquier estudiante de Biología.

En 1966 todavía quedaban algunos *flecos sueltos* en el estudio del código genético. Como se ha comentado antes (véase la nota 3), es evidente que la traducción del mRNA no puede empezar por cualquier lugar si ha de producir una proteína específica. Téngase en cuenta que, al estar constituido el código genético por tripletes, el desplazamiento de uno o dos nucleótidos en la pauta de traducción significaría cons-

truir un polipéptido de estructura primaria totalmente diferente⁴. En 1966 se obtuvieron los primeros datos experimentales sobre la iniciación de la traducción en procariontes y en los siguientes años el problema quedó centrado. El triplete AUG sirve no sólo para codificar la metionina, sino también como señal de iniciación. De hecho, las cadenas polipeptídicas nacientes, tanto en procariontes como en eucariotes, comienzan con el aminoácido metionina⁵. En procariontes existen dos tRNAs específicos de metionina, el tRNA_f^{Met} y el tRNA_m^{Met}. El primero se emplea para la iniciación de la traducción y la metionina que transporta se encuentra N-formilada. Así pues, las cadenas polipeptídicas nacientes, estrictamente hablando, en procariontes comienzan por N-formil-metionina (fMet). Para que se inicie la traducción, se ha de formar un complejo de preiniciación entre fMet-tRNA_f^{Met}, la subunidad 30 S del ribosoma y el mRNA. Se requiere para ello la participación de dos proteínas no presentes en el

⁴ En este sentido, se llama cambio de fase a ese desplazamiento de la pauta de lectura que ocurre en la naturaleza, por ejemplo, cuando se produce una delección o inserción de uno o dos nucleótidos.

⁵ El que la traducción comience por metionina no significa estrictamente que todas las cadenas polipeptídicas maduras posean este aminoácido como residuo N-terminal. La mayoría de las proteínas se sintetizan en forma de precursores y, muy frecuentemente, en la maduración se pierden aminoácidos del extremo N-terminal, incluida la metionina con la que comienza la traducción.

ribosoma, los denominados factores de iniciación IF-1 e IF-3, así como de un tercer factor de iniciación, IF-2, que liga GTP. Cerca del codón de iniciación, AUG, se encuentra en todos los mRNAs una secuencia de bases, conocida como secuencia Shine-Dalgarno en honor de sus descubridores, que interacciona con la subunidad pequeña del ribosoma y ayuda a que el fMet-tRNA_f^{Met} se sitúe con su anticodón interactuando con el codón AUG. En este momento, se ensambla la subunidad 50S del ribosoma para formar el complejo de iniciación, con el aporte energético procedente de la hidrólisis del GTP, efectuada por el propio IF-2, que tiene actividad de GTPasa.

El segundo tRNA específico de metionina, tRNA_m^{Met}, se ocupa de posicionar al aminoácido en los lugares determinados por los codones AUG internos durante la fase de elongación de la cadena polipeptídica. La subunidad mayor del ribosoma tiene tres localizaciones, denominadas E, P y A. El complejo de iniciación tiene situado el fMet-tRNA_f^{Met} sobre el sitio P. A continuación, un nuevo aminoacil-tRNA, aquél cuyo anticodón sea complementario del codón siguiente al de iniciación, se coloca sobre la localización A. Entonces, la peptidil transferasa, una actividad enzimática que reside en una de las moléculas de rRNA de la subunidad 50S —concretamente en el rRNA 23S⁶— cataliza la formación del enlace peptídico entre el grupo amino libre del aminoácido recién ensamblado y el carboxilo de la formil-metionina, que estaba unido al extremo 3' de su tRNA. Tras la formación del enlace peptídico, tiene lugar la fase de translocación, en la que el peptidil-tRNA resultante se sitúa sobre la localización P, el tRNA_f^{Met} vacío pasa a la localización E antes de abandonar el ribosoma y el sitio A queda libre y enfrentado al siguiente codón, listo para recibir al siguiente aminoacil-tRNA. Para esta fase de elongación, se requiere nuevamente GTP como suministrador de energía (tanto para el ensamblamiento del aminoacil-tRNA sobre el sitio A, como para el cambio de conformación del ribosoma requerido para la translocación), así como de otros tres factores proteicos de elongación, EF-Tu, EF-Ts y EF-G.

La elongación de la cadena polipeptídica continúa hasta que en el mRNA se encuentra un codón de terminación. Cuando éste se enfrenta a la localización A, ésta permanece libre, ya que no existe ningún tRNA que pueda complementar los codones de terminación. Esto sirve de señal para que la peptidil transferasa se limite a catalizar la hidrólisis del enlace éster que une el polipéptido formado al último tRNA, con lo que la cadena se libera, el ribosoma se desensambla y el mRNA se libera. También para esta fase de terminación se requiere energía aportada por el GTP y la presencia de factores de terminación o de liberación (RF1, RF2 y RF3).

En eucariotas los procesos de iniciación, elongación y terminación son muy similares a los que se acaban de apuntar. También existen dos tRNAs para la metionina, pero en este caso, la metionina iniciadora no está formilada. Otra diferencia se encuentra en los factores de iniciación, elongación y liberación.

El examen de la Fig. 6 permite concluir que, como era de esperar, no hay ambigüedades en el código genético. De hecho, el que el triplete AUG signifique metionina tanto para la iniciación de la cadena como para las posiciones internas del aminoácido no constituye ambigüedad alguna, pues la existencia de dos tRNA^{Met}, como se acaba de mencionar, salva las posibles dificultades. Lo que sí existe es degeneración, una degeneración que se da en grado variable. Así, mientras que para algunos aminoácidos, como el triptófano, sólo existe un triplete, para otros, como la leucina, existen 6. El número de tripletes no está relacionado con la abundancia de cada aminoácido en las proteínas. Por ejemplo, por término medio, la lisina es mucho más abundante que la histidina en las proteínas y, sin embargo, ambos poseen el mismo número de codones.

Además de las cuestiones relativas a la iniciación, elongación y terminación de la traducción, que se han mencionado más arriba⁷, en 1966 quedaba aún otro *fleco suelto* en el estudio del código genético. Cuando

⁶ Se trata, pues, de una *ribozima*, como se denominan las enzimas que, en vez de ser de naturaleza proteica, están constituidas por RNA. Las ribozimas son muy raras en los organismos actuales, aunque parece que tuvieron un protagonismo importante en los primeros tiempos de la vida en nuestro planeta.

⁷ Por caer fuera del alcance de este artículo, estas cuestiones se han tratado aquí de una forma sucinta. El lector interesado puede encontrar más detalles en cualquier texto actual de Bioquímica y Biología Molecular.

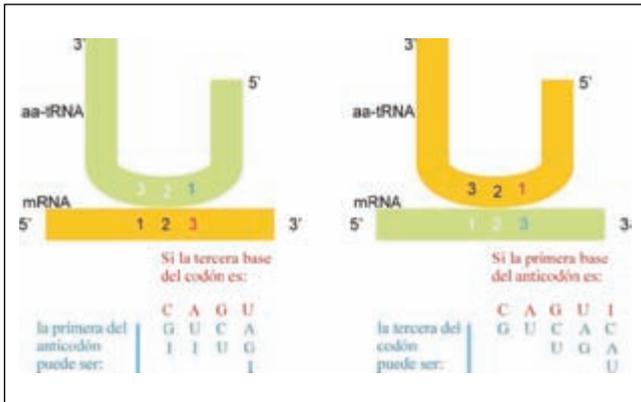


Figura 7. Hipótesis del balanceo. La figura recoge de un modo esquemático la hipótesis formulada por Crick. Se representa la interacción codón-anticodón, indicando las diversas posibilidades de apareamiento que existen para la tercera base del codón (izquierda) o para la primera del anticodón (derecha).

se comenzó a fraccionar la mezcla de los tRNAs, se observó que, como era de esperar por la degeneración del código, se podían encontrar varios tRNAs para un mismo aminoácido, pero el número total de especies distintas de tRNA no llega a ser 61, el número de codones que codifican aminoácidos⁸. Eso quiere decir que un anticodón puede interactuar con más de un codón. Este sorprendente hallazgo, junto con el no menos sorprendente de que al secuenciar el tRNA^{Ala} se encontró como anticodón una secuencia, IGC, que contenía una base anómala, la inosina (17). Un año más tarde, al secuenciar el tRNA^{Ser} de levadura, se encontró el anticodón IGA (19). Teniendo en cuenta que, como en todos los casos de apareamiento de bases en ácidos nucleicos, el apareamiento codón-anticodón implica que cada cadena tiene una orientación distinta, la base rara inosina debe aparearse con la tercera base del codón, precisamente la variable entre los 4 codones de la alanina y entre 4 de los 6 de la serina (Fig. 6). A partir de estos datos, y tras un profundo examen del código genético, Crick elaboró la *hipótesis del balanceo*, según la cual el apareamiento de las dos primeras bases del codón con las correspondientes del anticodón es estricto, pero el de la tercera base presenta una mayor flexibilidad (20). Crick llegó incluso a proponer unas reglas para el apareamiento de esa tercera base del codón que las investigaciones posteriores demostraron ser ciertas. La figura 7 muestra de modo

esquemático la hipótesis del balanceo. Más recientemente se han formulado algunas precisiones a la hipótesis elaborada por Crick (21).

La historia del desciframiento del código genético comenzó, pues, en 1961 con el trascendental descubrimiento de que el triplete UUU codifica a la fenilalanina y prácticamente culminó cinco años más tarde con el establecimiento de la hipótesis del balanceo. Es cierto que algunos detalles particulares tardaron un par de años en aclararse totalmente, pero en esencia, el ingente trabajo de descifrar los 64 codones se realizó en tan sólo 5 años, constituyendo una de las páginas más brillantes de la investigación biológica.

¿CUÁL ES LA SITUACIÓN 40 AÑOS DESPUÉS?

Han transcurrido 40 años desde que se terminó el desciframiento del código genético. ¿Qué nuevas perspectivas ha introducido la investigación en este periodo? La primera concierne a la universalidad del código genético. El problema, una vez más, fue ya planteado por Crick en su discurso al recibir el premio Nobel en 1962. En esa memorable ocasión, decía:

«Hay una cuestión general adicional acerca del código genético que nos podemos formular en este momento. El código, ¿es universal, es decir, el mismo en todos los organismos? Los datos preliminares sugieren que muy bien podría ser» (22).

Efectivamente, en líneas generales se puede decir que el código genético es universal. Pero algunas excepciones a esa regla general se han ido acumulando en los años posteriores. La primera se observó en 1979, cuando se descifró el código genético mitocondrial⁹. En las mitocondrias de vertebrados, AGA y AGG son codones de terminación, en vez de codificar arginina, mientras que UGA codifica triptófano en vez de ser un codón de terminación. Además, AUA codifica a la metionina. El código genético mitocondrial de levadura presenta alguna semejanza con el de mitocondrias de vertebrados, pero tiene también otras variantes

⁸ Aunque el número de especies de tRNA varía de unos organismos a otros, por término medio se puede estimar en 40.

⁹ Las mitocondrias son unos orgánulos subcelulares presentes en células eucarióticas, que poseen su propio DNA y contienen las maquinarias necesarias para realizar la transcripción y la traducción. Gozan, por tanto, de una cierta autonomía genética.

con respecto al código estándar recogido en la figura 6. También se han observado algunas pequeñas variaciones en bacterias, en las que es frecuente encontrar GUG y UUG como codones de iniciación¹⁰.

Un caso aparte lo constituyen la selenocisteína y la pirrolisina. Se trata de aminoácidos naturales recientemente descubiertos. Aunque su presencia se encuentra limitada a algunos casos concretos, su incorporación a las proteínas exige codones propios. En arqueobacterias metanogénicas la pirrolisina está codificada por UAG, que actúa de ordinario como un codón de parada (Fig. 6), mientras que la selenocisteína está codificada por UGA, otro codón normal de parada, tanto en procariotas como en eucariotas. Pero el mecanismo de codificación es en este caso muy particular. La maquinaria de biosíntesis de proteínas inserta selenocisteína al encontrar el codón UGA sólo cuando se encuentra presente en el mRNA una secuencia auxiliar, denominada elemento SECIS (de *selenocysteine insertion sequence*), que se caracteriza por poseer estructuras primaria y secundaria características (23). Los elementos SECIS se pueden localizar en diversas posiciones relativas al codón UGA dependiendo de si el organismo es procariótico o eucariótico.

La degeneración del código conlleva, como se ha indicado, que para muchos aminoácidos exista más de un codón. En estos casos se plantea una cuestión interesante: ¿se usan indistintamente todos los codones posibles o hay alguna preferencia entre ellos? La ingente acumulación de datos de secuencia de genomas que se ha producido en los últimos años permite asegurar que no hay una utilización indistinta de los diversos codones y que cada organismo tiene unas determinadas preferencias, que están relacionadas con los niveles de cada uno de los tRNAs (24). El conocimiento de la utilización de los diversos codones es no sólo de interés académico, sino que se trata, por ejemplo, de un factor a tener en cuenta al diseñar estrategias experimentales para la producción de proteínas recombinantes (25).

En resumen, aunque se han observado algunas excepciones al código genético descrito en la década de 1960, sorprende que en la inmensa mayoría de los casos, todos los organismos, tanto procariotas como

eucariotas, utilicen el mismo sistema de codificación. La conservación de los mecanismos moleculares entre organismos muy distantes en la escala evolutiva es así, también en este caso, una característica constante de los seres vivos.

CONCLUSIONES

Como se ha comentado antes, todo estudiante de Biología e, incluso, toda persona culta ha visto alguna vez una representación del código genético similar a la recogida en la figura 6. Pero al observarla, se corre el riesgo de pasar por alto el inmenso trabajo que costó el desciframiento del código y contemplar la figura como si hubiera *llovido del cielo* sin esfuerzo alguno por parte de los investigadores. Esta breve revisión tenía por objeto principal hacer patente ese trabajo a cuya culminación concurren muchos factores. De una parte, la innegable valía científica de los protagonistas de la aventura que, con las forzosas omisiones, han quedado reseñados en las páginas precedentes. Pero no menos importancia tuvo el clima de competitividad que surgió entre los diferentes laboratorios implicados. Esa competitividad y el propio afán de superación han catalizado numerosos avances de la ciencia, de la misma manera que favorecen la obtención de buenas marcas en una competición deportiva.

Estamos acostumbrados a ver cómo son muchos los atletas que empiezan una competición, cómo los más van cayendo eliminados en las rondas de clasificación, cómo en la final van quedando algunos descolgados y cómo, al terminar, sólo uno se ha alzado con la victoria. Y también estamos habituados a contemplar cómo numerosos grupos de investigación se embarcan en un determinado tema, cómo algunos sólo consiguen resultados triviales o poco significativos, cómo otros muchos, aún con buenos resultados, no llegan a dar un salto cualitativo en el avance científico.

Pero la analogía puede ir más allá de considerar la investigación como una excitante aventura o del simple afrontarla con ánimo deportivo. También las célebres palabras del barón de Coubertin valen para quienes se aplican al *deporte* de la investigación. Hay

¹⁰ Estas y otras variantes del código genético estándar se pueden encontrar navegando en la página web del National Center for Biotechnology Information (National Institutes of Health, U.S.A.) (<http://www.ncbi.nlm.nih.gov/Taxonomy/>) para buscar el apartado *Genetic codes*.

que tener en cuenta que, ciertamente, científicos que hayan protagonizado grandes revoluciones ha habido pocos, pero que los buenos científicos, los que realizan su trabajo con honestidad, preparación, dedicación y competencia, son muy numerosos. Y que, posiblemente, sin su esfuerzo y sin sus resultados, muchas veces pequeños, no se habrían podido dar los descubrimientos definitivos. Pretender que sólo los grandes genios sean capaces de hacer ciencia equivale a firmar la partida de defunción del desarrollo científico, cuyos grandes avances suelen ir precedidos por numerosos hallazgos más modestos que, además, contribuyen a crear una atmósfera adecuada para la aparición de los genios.

Pero, sigamos más allá con la analogía deportiva. Poner el acento en la participación más que en la victoria es compatible con el lema olímpico: *citius, altius, fortior*. Dicho de otro modo, hacer énfasis en la importancia de la participación no significa que haya que erradicar la competencia. Antes al contrario, sin una sana competitividad seguramente no sería posible llegar más arriba o hacerlo más rápidamente. Y en la investigación científica, el deseo de profundizar en el conocimiento, y aún el de quemar etapas en ese empeño, ha de presidir el afán del investigador.

Si se admiten las premisas anteriores, se pueden establecer varias conclusiones, pero dos parecen las fundamentales. En primer lugar, que no vencer en esa competencia no significa conformarse con la mediocridad. El investigador debe tender a que su trabajo sea siempre de excelencia. Por otro lado, es evidente que la competencia que favorece el desarrollo y avance de la investigación debe ser respetuosa con las normas éticas: no puede significar, de ninguna manera, un deseo de triunfar a costa de no se sabe qué concesiones. Es esta última la actitud que, por mantener la analogía que se viene empleando, bien podríamos llamar *deportividad*, que se resume en el empleo exclusivo de medios lícitos, en el *saber ganar* y también, ¿por qué no?, en el *saber perder*. Excelencia y deportividad son, pues, dos de los pilares en que se debería apoyar toda investigación.

BIBLIOGRAFÍA

1. CRICK, F. H. C. (1958) «On protein synthesis» *Symp. Soc. Exptl. Biol.* **12**, 138–163.

2. FRANCO, L. (2003) «Doble hélice, genes y cromosomas» *Rev. Real Acad. Cienc. Exact. Fis. Nat. (Esp)* **97**, 203–222.
3. LITTLEFIELD, J. W., KELLER, E. B., GROS, J. Y ZAMECNIK, P. C. (1955) «Studies on cytoplasmic ribonucleoprotein particles from the liver of the rat» *J. Biol. Chem.* **217**, 111–123.
4. JACOB, F. Y MONOD, J. (1961) «Genetic regulatory mechanisms in the synthesis of proteins» *J. Mol. Biol.* **3**, 318–356.
5. GAMOW, G. (1954) «Possible relation between deoxyribonucleic acid and protein structures» *Nature* **173**, 318.
6. CRICK, F. (1988) «What Mad Pursuit: A Personal View of Scientific Discovery», 182 pp. Basic Books, New York.
7. HOAGLAND, M. B., STEPHENSON, M. L., SCOTT, J. F., HECHT, L. I. Y ZAMECNIK, P. C. (1958) «A soluble ribonucleic acid intermediate in protein synthesis» *J. Biol. Chem.* **231**, 241–257.
8. GRUNBERG-MANAGO, M., ORTIZ, P. J. Y OCHOA, S. (1955) «Enzymatic synthesis of nucleic acid-like polynucleotides» *Science* **122**, 907–910.
9. OCHOA, S. (1989) «La emoción de descubrir» Conferencia pronunciada en Valencia el 5 de octubre, con motivo de la inauguración de las nuevas instalaciones del Museo Ochoa. Publicada por el Instituto de Investigaciones Citológicas. Caja de Ahorros de Valencia.
10. BASILIO, C. (1961), citado en: GANDÍA BALAGUER, A. (1999) «El pensamiento científico de Severo Ochoa: desde la fosforilación oxidativa hasta el código genético y correspondencia científica». Fundación Ramón Areces, Madrid.
11. NIRENBERG, M. W. Y MATTHAEI, J. H. (1961) «The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides» *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1588–1602.
12. LENGYEL, P., SPEYER, J. F. Y OCHOA, S. (1961) «Synthetic polynucleotides and the amino acid code» *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1936–1942.
13. MATTHAEI, J. H., JONES, O. W., MARTIN, R. G. Y NIRENBERG, M. W. (1962) «Characteristics and composition of RNA coding units» *Proc. Natl. Acad. Sci. U. S. A.* **48**, 666–677.
14. KHORANA, H. G., JACOB, T. M., MOON, M. W., NARANG, S. A., Y OHTSUKA, E. (1965) «Studies on polynucleotides. XLII. The synthesis of deoxyribo-polynucleotides containing repeating nucleotide sequences. Introduction and general considerations» *J. Am. Chem. Soc.* **87**, 2954–2956.
15. LEDER, P. Y NIRENBERG, M. W. (1964) «RNA code-

- words and protein synthesis. The effect of trinucleotides upon the binding of sRNA to ribosomes» *Science* **145**, 1399–1407.
16. LEDER, P. Y NIRENBERG, M. W. (1964) «RNA code-words and protein synthesis, II. Nucleotide sequence of a valine RNA codeword» *Proc. Natl. Acad. Sci. U. S. A.* **52**, 420–427.
 17. HOLLEY, R. W., APGAR, J., EVERETT, G. E., MADISON, J. T., MARQUISEE, M., MERRILL, S. H., PENSWICK, J. R. Y ZAMIR, A. (1965) «Structure of a ribonucleic acid» *Science* **147**, 1462–1465.
 18. BRETSCHER, M. S. Y JONES, O. W. (1967) «The Biochemistry of the Genetic Code» pp. 217–263, en *Techniques in Protein Biosynthesis* (Campbell, P. N. y Sargent, J. R., eds.), vol. I. Academic Press, London.
 19. ZACHAU, H. G., DÜTTING, D Y FELDMANN, H. (1966) «The structures of two serine transfer ribonucleic acids» *Hoppe Seylers Z. Physiol. Chem.* **347**, 212–235.
 20. CRICK, F. H. C. (1966) «Codon-anticodon pairing: The wobble hypothesis» *J. Mol. Biol.* **19**, 548–555.
 21. S. MATSUYAMA, T. UEDA, P. F. CRAIN, J. A. McCLOSKEY Y K. WATANABE (1998) «A novel wobble rule found in starfish mitochondria. Presence of 7-methylguanosine at the anticodon wobble position expands decoding capability of tRNA» *J. Biol. Chem.* **273**, 3363–3368.
 22. CRICK, F. H. C. (1964) «On the genetic code» en *Nobel Lectures. Physiology or Medicine, 1942-1962*. Elsevier, Amsterdam.
 23. F. ZINONI, A. BIRKMANN, W. LEINFELDER Y A. BOCK (1987) «Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon» *Proc. Natl. Acad. Sci. U. S. A.* **84**, 3156–3160.
 24. M. A. SANTOS, G. MOURA, S. E. MASSEY Y M. F. TUIITE (2004) «Driving change: the evolution of alternative genetic codes» *Trends Genet.* **20**, 95–102.
 25. S. JANA Y J. K. DEB (2005) «Strategies for efficient production of heterologous proteins in *Escherichia coli*» *Appl. Microbiol. Biotechnol.* **67**, 289–298.