

La medicina tras la secuenciación del genoma humano

por **José M.^a Mato**

*Conferencia pronunciada
el 26 de abril de 2005*

Forum Deusto

La medicina tras la secuenciación del genoma humano

José M.^a Mato*

En 1905, su *annus mirabilis*, Albert Einstein publicó cinco trabajos en la revista *Annalen der Physik* que cubrían tres áreas distintas de la física: la fotoelectricidad, el movimiento browniano, y la teoría de la relatividad especial. Estos cinco trabajos constituyen el principal testamento científico del genio de Einstein. A los cien años de su publicación, su figura y su obra siguen despertando un enorme interés. Así, en 1999, la revista estadounidense *Time* eligió como personaje del siglo a Einstein y, para honrar su memoria, el 2005 ha sido elegido el año de la física.

La obra de Einstein representa el momento estelar de un legado universal en física que tuvo su comienzo en el siglo xvii con Galileo, explorando la gravedad y el sistema solar, y con Newton, descubriendo las leyes del movimiento y el cálculo, y que continuaron, además de Albert Einstein, otros muchos físicos geniales entre ellos Maxwell, Lorentz, los Curie, y Planck.

¿Qué ocurría, entre tanto, con la investigación en ciencias de la vida? En 1905, Robert Koch, uno de los fundadores de la bacteriología y descubridor del bacilo de la tuberculosis y del cólera, recibía el Premio Nobel de Fisiología y Medicina por sus estudios sobre la etiología de la tuberculosis y la teoría microbiana de la enfermedad; y un año después, en 1906, Santiago Ramón y Cajal recibía el Premio Nobel por descubrir la neurona, o célula nerviosa, como la unidad funcional del sistema nervioso. Cajal, como es bien conocido, había mejorado la técnica de tinción de nitrato de plata desarrollada por

* Director General de CIC bioGUNE (Centro de Investigación Cooperativa en Biociencias). Derio, Bizkaia, España.

Camilo Golgi, con quien compartió el Premio Nobel, lo cual le permitió estudiar la estructura fina del tejido nervioso del cerebro y de la médula espinal en embriones y animales jóvenes. Esta tinción de plata, específica para las neuronas, no sólo le sirvió para determinar la estructura detallada de la retina, y trazar la estructura y conexiones de las neuronas de la sustancia gris y de la médula espinal, sino que tuvo, asimismo, un gran valor para el diagnóstico de los tumores cerebrales.

En 1900, tan sólo unos años antes de que apareciesen publicados los famosos trabajos de Einstein, la comunidad científica, gracias a los trabajos que habían realizado de forma independiente el holandés Hugo de Vries, el francés Carl Correns, y el austriaco Erich von Tschermak, había redescubierto las ideas que el gran Gregor Mendel había publicado en 1866 sobre la herencia. ¿Cómo explicar este olvido? Dice Marcel Proust que «la causa de que una obra de talento sea pocas veces admirada en seguida se debe a que quien la ha escrito es extraordinario, a que poca gente se le parece. Es su obra misma la que, fecundando los raros espíritus capaces de comprenderla, los hará crecer y multiplicarse». Esta parece ser la explicación para el caso de la obra de Mendel. El fraile agustino de Brünn había enviado una copia de su trabajo a Charles Darwin, que había publicado su famosísimo tratado sobre *El Origen de las Especies* tan sólo unos años antes, en 1859. Darwin, al igual que Alfred Wallace, un interesante naturalista Victoriano injustamente olvidado, había llegado a la conclusión de que cuando se producen cambios heredables en un organismo, los que ofrecen alguna ventaja selectiva persisten y se van haciendo cada vez más comunes con el paso de las generaciones. Pero carecían de un mecanismo que explicase cómo se heredaban estos cambios. Aquí es donde el trabajo de Mendel les podría haber ayudado, pero Darwin y Wallace desconocían el trabajo de Mendel —parece que Darwin nunca llegó a leer el artículo que sobre su obra Mendel le había enviado, las páginas del manuscrito no fueron separadas. Y es que la metodología estadística empleada por Mendel en su trabajo era compleja y ajena a la forma de experimentar y pensar de los biólogos en esa época. La biología de finales del siglo XIX era una disciplina fundamentalmente descriptiva y no estaba preparada para comprender a Mendel.

Así que a comienzos del siglo XX, cuando la física alcanzaba uno de sus momentos estelares, la moderna biología se encontraba, aún, en sus comienzos. Se acababa de descubrir que el cerebro estaba hecho

de neuronas, pero se desconocía totalmente cómo se comunicaban entre ellas. Mendel había sido redescubierto, pero se desconocía cuál era la sustancia de los genes —el término «gen» lo acuñó Wilhelm Johannsen en 1909. La teoría microbiana de la enfermedad estaba bien asentada, pero hubo que esperar hasta 1940, ya en plena guerra mundial, para que Howard Florey y Ernst Chain aislaran la penicilina pura y publicaran un histórico trabajo en el que demostraban la acción antibacteriana de la penicilina. En 1945, Fleming, Florey y Chain recibieron el Premio Nobel de Fisiología y Medicina por el descubrimiento de la penicilina y su uso terapéutico.

La historia de la penicilina, en la que trabajaron conjuntamente microbiólogos, químicos, bioquímicos, y clínicos, y de la que se beneficiaron en la segunda mitad del siglo xx cientos de millones de personas, muchos de los cuales seguramente habrían muerto sin la ayuda que recibieron de los antibióticos, ejemplifica, mejor que ninguna otra, el enorme impacto que la investigación en biología puede llegar a tener sobre la sociedad, y justifica el enorme interés que actualmente suscitan los temas relacionados con la investigación biomédica. Otro ejemplo evidente de los beneficios de la investigación biomédica sobre la salud es la historia de las vitaminas. También a principios del siglo xx, en 1906, Frederick Hopkins demostró que los alimentos contenían «*factores accesorios*» que eran necesarios para mantener la vida, además de las proteínas, carbohidratos, grasas, minerales y agua. En 1911, Casimir Funk, un colaborador de Hopkins, identificó que la naturaleza del componente de los alimentos que prevenía el beriberi —una enfermedad neurodegenerativa cuyo nombre deriva de la palabra senegalesa que significa «extrema debilidad»— era una amina (un tipo de sustancia que contiene nitrógeno), de modo que llamó a este compuesto vitamina, término que procede de amina vital. Hopkins recibió el Premio Nobel de Fisiología y Medicina en 1929, que compartió con Christiaan Eijkman por su descubrimiento de que el arroz descascarillado producía en los pollos una enfermedad parecida al beriberi y que sanaban cuando se les suministraba arroz natural y que llevó al descubrimiento de la vitamina B₁, la tiamina. El descubrimiento de las vitaminas, que ocupó el centro del escenario de la investigación bioquímica durante los años veinte y treinta, acabó con epidemias como el beriberi y la pelagra —una enfermedad crónica, con manifestaciones cutáneas y perturbaciones digestivas y nerviosas— que anualmente mataban a decenas de miles de personas en el mundo. Como en el caso de los antibióticos, hubo que esperar varias décadas entre el descubrimiento de las vitaminas y su uso terapéutico. Y aunque la investigación sobre vitaminas

vivió sus mejores momentos hace ya muchas décadas, aún nos sigue proporcionando importantes sorpresas, como el relativamente reciente descubrimiento de que el suplemento de las mujeres en edad fértil con ácido fólico, otra vitamina, previene la mayoría de los casos de espina bífida —una terrible enfermedad en donde la deficiencia materna en ácido fólico impide que se cierre el tubo neural durante el desarrollo fetal en el embarazo.

Si durante las primeras décadas del siglo xx la investigación bioquímica estuvo dominada, principalmente, por los buscadores de vitaminas, a partir de la década de los treinta y hasta los cincuenta el dominio fue, sin duda alguna, de los investigadores del metabolismo y de los buscadores de enzimas —las enzimas son proteínas que inician y proporcionan el sitio para que se produzcan reacciones específicas entre sustancias sin verse ellas modificadas durante este proceso. La Enciclopedia Britannica define el metabolismo como «la suma de todas las reacciones químicas que tienen lugar dentro de cada célula de un organismo vivo, proporcionando energía para los procesos vitales y sintetizando nuevas sustancias orgánicas».

Las reacciones metabólicas son de dos tipos. Unas son reacciones anabólicas, es decir usan energía para sintetizar moléculas complejas a partir de sustancias sencillas. Y las otras son reacciones catabólicas, es decir liberan energía al romper moléculas complejas en otras más sencillas. La primera etapa en el proceso de liberación de energía a partir de los alimentos tiene lugar, en los animales, en el tubo digestivo; las moléculas complejas presentes en los alimentos, como por ejemplo las proteínas, se rompen en sus componentes más pequeños, los denominados aminoácidos. Del mismo modo, los carbohidratos complejos presentes en los alimentos se rompen en azúcares más sencillos, como la glucosa; y los lípidos se descomponen en ácidos grasos y glicerol. Los productos de esta primera etapa del metabolismo entran a continuación en las células, donde son transformados, más correctamente oxidados, liberando así la energía necesaria para los procesos vitales que toda célula tiene que llevar a cabo, así como para la síntesis de nuevas moléculas. Los detalles de esta última fase de liberación de energía durante el catabolismo de los alimentos fueron dilucidados por Hans Krebs en 1937, quien recibió, por estos trabajos, el Premio Nobel de Medicina y Fisiología en 1953. Entre 1947 y 1970 el comité Nobel reconoció en numerosas ocasiones el impresionante trabajo realizado por estos investigadores del metabolismo y buscadores de enzimas. Así, en 1947 recibieron el Premio Nobel el argentino Bernardo Houssay y el

matrimonio Carl y Gerty Cori por sus estudios sobre el metabolismo de la glucosa; en 1955 Vincent du Vigneaud, por su trabajo sobre el metabolismo de los aminoácidos; en 1964 fueron premiados Konrad Bloch y Feodor Lynen, por sus trabajos sobre el metabolismo del colesterol y de los ácidos grasos; y en 1970 el también argentino Luis Federico Leloir, por sus estudios sobre la síntesis del glucógeno, y Julius Axelrod por dilucidar la síntesis de las catecolaminas.

Si hasta la década de los cincuenta la investigación biomédica estuvo dominada por estos buscadores de enzimas e investigadores del metabolismo, a partir de la década de los sesenta el dominio ha sido, incuestionablemente, de los buscadores de genes. Llama la atención esta falta de interés, durante las últimas décadas, por el estudio del metabolismo cuando las enfermedades metabólicas —como la diabetes tipo 2, la obesidad, o el síndrome metabólico— afectan en la actualidad, aproximadamente, al 20% de la población de los países desarrollados, lo cual las convierte en uno de los principales problemas de salud pública en Occidente así como en una auténtica epidemia.

¿Cómo y por qué se produjo este cambio de interés de los bioquímicos a mediados del siglo xx desde el estudio del metabolismo y los enzimas a la genética y los genes? El redescubrimiento en 1900 del trabajo de Mendel sobre la herencia y la teoría cromosómica —que fue desarrollada a finales del siglo xix y según la cual en todas las células que proceden de la división de un huevo fecundado, la mitad de los cromosomas son de origen paterno y la mitad de origen materno— impulsaron el desarrollo de la genética en el siglo xx. El término «gen», como he comentado anteriormente, fue acuñado por el danés Wilhelm Johannsen en 1909 para denominar unas partículas hipotéticas que, suponía, portaban los cromosomas y determinaban la herencia. Unos años antes, en 1902, Archibald Edward Garrod descubrió la primera enfermedad congénita del metabolismo, la alkaptonuria, un defecto en el metabolismo de la tirosina y la fenilalanina. En su artículo de 1902 Garrod decía: «No existe ninguna razón para pensar que la consanguinidad entre padre y madre pueda producir en su descendencia una condición como la alkaptonuria, y debemos más bien buscar una explicación de ella en alguna peculiaridad de los progenitores, pero que tiene las mayores probabilidades de manifestarse en la descendencia de la unión de dos miembros de una familia en la que es transmitida», y añadía, «la ley de la herencia descubierta por Mendel ofrece una razonable explicación a la alkaptonuria». Y en 1926, Thomas Morgan, que recibiría el Premio Nobel de Medicina y Fisiología en 1933 por sus es-

tudios de genética utilizando la mosca de la fruta, *Drosophila melanogaster*, escribía: «Con el mismo criterio con el que los químicos dan por reales los átomos invisibles y los físicos los electrones, el estudioso de la herencia apela a unos elementos invisibles llamados genes». Morgan había demostrado que los genes se encontraban ordenados linealmente en los cromosomas y que podían experimentar cambios repentinos permanentes, o mutaciones, que producían un cambio concreto determinado por el gen como el cambio de color rojo al blanco en la mosca de la fruta. ¿Cuál era la naturaleza química de los genes? ¿Cómo puede ser, que tratándose de moléculas orgánicas, se conserven invariables con plena estabilidad? Morgan previó la estructura de los genes cuando escribió: «Es difícil, con todo, evitar la sugestión fascinante de que el gen sea constante precisamente porque venga constituyendo una entidad química organizada». La respuesta a la pregunta que se hacía Morgan sobre que era esa entidad química organizada de la que estaban hechos los genes la dieron Oswald Avery y sus colaboradores Colin MacLeod y Maclyn McCarty en 1944.

En 1928, el médico británico Fred Griffith había descubierto el extraño fenómeno de la transformación. Cuando se inyectaba un ratón simultáneamente con un neumococo no patogénico (tipo II) y un neumococo infeccioso (tipo III), pero cuya actividad infecciosa había sido destruida mediante tratamiento con calor, el animal rápidamente sucumbía a la infección. La única explicación posible era que el neumococo tipo III (virulento pero muerto) había transformado al neumococo tipo II (no patogénico pero vivo). En 1935 Oswald Avery comenzó a purificar el agente responsable de la transformación en neumococos. Avery continuó este trabajo durante casi una década y en 1944, en colaboración con Colin MacLeod y Maclyn Mc Carty, publicó un artículo en la revista *The Journal of Experimental Medicine* en el que describía como se podía llevar a cabo la transformación de neumococos *in vitro*, así como los métodos que habían usado para purificar y caracterizar el agente infeccioso. Sus resultados indicaban, fuera de toda duda razonable, que este factor no era una proteína sino DNA. El factor resistía temperaturas a las que las proteínas se desnaturalizaban. Estudios espectrofotométricos indicaban que el factor contenía sólo DNA, sin trazas de proteína o RNA. El análisis químico demostraba que la contaminación por proteínas era menor del 1%. Estudios enzimáticos demostraban que la actividad transformante que habían purificado era resistente a la acción de enzimas que atacan proteínas y a la de fosfatasa que degradan RNA, pero era degradado por una enzima que rompía el DNA. Asimismo, sus resultados indicaban que

tan sólo 0.3×10^{-9} gramos de DNA eran suficientes para inducir la transformación de neumococos en un tubo de ensayo.

Hoy nos sorprende que Avery, MacLeod y McCarty, los protagonistas de uno de los más grandes descubrimientos del siglo xx, nunca recibieran el Premio Nobel por identificar el DNA como la sustancia de los genes. También nos sorprende que, a pesar de la impresionante evidencia acumulada por Avery y sus colaboradores indicando que los genes estaban hechos de DNA, la comunidad científica viese con escepticismo esta nueva teoría durante casi una década. Así, por ejemplo, en 1951, siete años después de que apareciera la publicación de Avery demostrando que los genes estaban hechos de DNA, Hermann Muller —un gran genetista colaborador de Morgan— decía que nada se sabía sobre la sustancia de los genes. Y no es que el trabajo de Avery hubiese pasado desapercibido, es que no se supo valorar su importancia. Para entender esta situación, no sólo hay que recordar que entonces la mayoría de los bioquímicos y genetistas creían que los genes estaban hechos de proteínas y los ácidos nucleicos eran todavía pobremente comprendidos —la estructura de los nucleótidos y la naturaleza de los enlaces químicos que los unían para formar largos polímeros era aún sujeto de debate— sino que la idea de un código genético no era una condición previamente aceptada para explicar la teoría de la herencia. Esta idea fue introducida gradualmente, después de que Francis Crick y James Watson descubrieran en 1953 la estructura en doble hélice del DNA. En realidad, la idea de un código genético no fue totalmente aceptada hasta 1963, cuando, de manera independiente, Har Gobind Khorana, Marshall Nirenberg, y Robert Holley descifraron el código genético, trabajo por el que recibieron el Premio Nobel de Fisiología y Medicina en 1968. En 2003, con motivo de la celebración del cincuentenario del descubrimiento de la estructura del DNA, la comunidad científica ha honrado a Avery, MacLeod y McCarty reconociendo la enorme importancia de su trabajo.

En 1952 Alfred Hershey y Martha Chase, utilizando fagos en lugar de neumococos, llegaron a la misma conclusión que Avery había llegado 8 años antes —es decir, a que el DNA era la sustancia de los genes. Hershey y Chase demostraron que cuando un fago infecta una bacteria y se reproduce en ella, solamente necesita transferir el DNA. Los fagos son virus bacterianos compuestos únicamente de DNA y proteína, que habían sido introducidos como un nuevo modelo experimental unos años antes por Max Delbrück, Salvador Luria y el propio Alfred Hershey

para descubrir, de una vez por todas, el misterio de los genes (estos tres investigadores recibieron el Premio Nobel de Medicina y Fisiología en 1969 por sus trabajos sobre fagos).

El descubrimiento más importante realizado en biología en la segunda mitad del siglo xx fue, sin duda, la determinación de la estructura en doble hélice del DNA. Y, aunque los protagonistas indiscutibles de esa historia fueron el británico Crick, fallecido a finales de 2004, y el estadounidense Watson, es importante recordar que la determinación de la estructura del DNA sólo fue posible gracias al desarrollo de las técnicas de rayos X para analizar la estructura molecular de todo tipo de moléculas, que habían llevado a cabo a principios de 1910 William Lawrence Bragg y su padre, William Henry Bragg, en el Laboratorio de Cavendish de Cambridge, y por los que ambos recibieron el Premio Nobel de Física en 1915. En este laboratorio de biología estructural —como lo llamaríamos actualmente— se había establecido, antes de iniciarse la Segunda Guerra Mundial, Max Perutz, un bioquímico vienés que, en compañía de su colaborador John Kendrew, estudiaba la estructura de la hemoglobina y de la mioglobina. Por este extraordinario trabajo, que les llevó alrededor de 40 años, Perutz y Kendrew recibieron el Premio Nobel de Química en 1962. Francis Crick, un físico que había decidido abandonar la física y dirigir sus investigaciones hacia la biología después de trabajar para el gobierno británico en la fabricación de minas magnéticas durante la guerra, se había incorporado al laboratorio de Perutz al finalizar ésta y, aunque su proyecto inicial era estudiar la estructura de la mioglobina, con la llegada de Watson al laboratorio de Perutz, —Watson se había trasladado a Europa después de finalizar la tesis doctoral siguiendo el consejo de su mentor, Salvador Luria—, ambos se pusieron a estudiar la estructura del DNA.

De manera simultánea, al finalizar la guerra, John Randall, un físico que había desempeñado un papel destacado en la construcción de radares que avisaran de las incursiones de los bombarderos de Hitler, había establecido otro laboratorio de difracción de rayos X para el estudio de macromoléculas en el University College de Londres. Este laboratorio estaba dirigido por Maurice Wilkins quien, con la colaboración de Rosalind Franklin, estudiaba la estructura de fibras de DNA. Franklin demostró la existencia de dos formas de DNA, la A y la B, y determinó, asimismo, la densidad, las dimensiones celulares exactas y la simetría de la forma A, evidencia que sugería muy firmemente que la estructura del DNA tenía dos cadenas, y no sólo una, como había propuesto el bioquímico Linus Pauling, cuya hélice proteica α inspiró a Watson

y Crick su modelo de doble hélice del DNA (Pauling recibió el Premio Nobel de Química en 1954 por sus trabajos sobre el enlace peptídico y la estructura de proteínas, y el de la Paz, por su oposición a las armas nucleares, en 1962). Wilkins, a su vez, demostró que las imágenes de difracción de rayos X de las fibras de DNA se ajustaban al modelo de doble hélice.

De la parte de la química, Watson y Crick apoyaron su modelo en doble hélice en la información sobre la composición química del DNA obtenida por Edwin Chargaff. El DNA está compuesto de cuatro compuestos químicos, denominados nucleótidos o bases: adenina (abreviada A), citosina (abreviada C), guanina (abreviada G), y timina (abreviada T). Chargaff había demostrado que en cualquier muestra de DNA la cantidad de A es igual a la de T, y la de C igual a la de G, lo que se conoce como la regla de Chargaff. Asimismo, el modelo de la doble hélice se sirvió de los datos suministrados por John Mason Gulland cuyos resultados sugerían que, en solución, el DNA se mantenía unido mediante enlaces de hidrógeno. Basados en esta información, el 25 de abril de 1953 Watson y Crick publicaron su famoso artículo en la revista *Nature* titulado «Estructura molecular de los ácidos nucleicos» al que acompañaban otros dos artículos que confirmaban y validaban el modelo en doble hélice del DNA, uno firmado por Maurice Wilkins y sus colaboradores, y el otro por Rosalind Franklin y sus colaboradores. En 1963 Crick, Watson, y Wilkins recibieron el Premio Nobel de Medicina. Rosalind Franklin había fallecido varios años antes de cáncer de mama. Franklin fue especialmente maltratada por James Watson en su famoso libro, *La doble hélice*, publicado en 1968, en el que narra la historia del descubrimiento de la estructura del DNA. El machismo y prejuicio con que se refiere a Rosalind Franklin serían, afortunadamente, inaceptables en la actualidad. En 2003, con motivo de la celebración del cincuentenario del descubrimiento de la estructura del DNA, la comunidad científica ha honrado también a Rosalind Franklin reconociendo la enorme importancia de su trabajo.

El descubrimiento de la doble hélice abrió la puerta a resolver el problema del papel de los genes en la síntesis de proteínas. Así, durante los siguientes años se descifró el código genético, y en los años sesenta la relación entre DNA, RNA y las proteínas quedó definitivamente establecida.

La herencia tiene un doble aspecto: la transmisión de caracteres de una generación a la siguiente y la expresión de estos caracteres durante el proceso de diferenciación y desarrollo mediante el que

un organismo se va construyendo a sí mismo. La transmisión de caracteres y su expresión fueron elegantemente unificadas al describir Watson y Crick la estructura tridimensional del DNA. Forma y función en el DNA son una misma cosa. Las cuatro bases que constituyen el DNA son la única parte variable de la estructura de esta molécula. Entre ambas hebras de DNA, una vez la secuencia de una de ellas está fijada, el apareamiento de bases determina la secuencia complementaria de la otra hebra. Como consecuencia de esta sencilla regla, si ambas hebras se separan, cada una puede ensamblar sobre ella una copia exacta de su previo compañero, produciendo dos dobles hélices idénticas durante la división celular y garantizando así la transmisión de caracteres. La libertad de la secuencia de las bases a lo largo del DNA permite codificar, como mediante un alfabeto de cuatro letras, las especificaciones que constituyen las características de un individuo.

Descifrar la secuencia de los 3.000 millones de pares de bases que constituyen el genoma humano, es una misión que se encontraba implícita en el descubrimiento de la estructura del DNA. En 1953 este proyecto era, sin embargo, imposible de abordar. Aún no se habían producido los avances tecnológicos necesarios que lo harían posible. El conjunto de técnicas que iban a permitir manejar grandes bloques de DNA y que harían factible la secuenciación de genomas enteros, y que se conocen actualmente con el nombre de «DNA recombinante» o «ingeniería genética», se desarrollaron durante los años sesenta y setenta. El comité Nobel ha reconocido en numerosas ocasiones el trabajo realizado por estos investigadores que hicieron posible la ingeniería genética. Así, en 1959 recibió el Premio Nobel de Fisiología y Medicina Arthur Kornberg por el descubrimiento de la DNA polimerasa, la enzima que sintetiza el DNA. Kornberg compartió el Premio Nobel con Severo Ochoa, de cuyo nacimiento se cumplen ahora cien años. En 1978 fueron galardonados con el Nobel de Fisiología y Medicina Werner Arber, Daniel Nathans y Oliver Smith, por el descubrimiento de las enzimas de restricción, un grupo de enzimas que permiten cortar el DNA en lugares específicos. Y en 1980 recibió el Premio Nobel de Química, Paul Berg, por llevar a cabo el primer experimento de ingeniería genética. Berg compartió el Premio Nobel con Fred Sanger y Walter Gilbert, que lo recibieron por desarrollar las técnicas que han permitido secuenciar el DNA —para Sanger era su segundo Premio Nobel de Química, el primero le fue galardonado en 1958 por desarrollar las técnicas que permitieron la secuenciación de proteínas.

En 1972, Berg y sus colaboradores describieron cómo habían obtenido *in vivo* una molécula híbrida que contenía el DNA del oncogén SV₄₀ y el DNA de un fago. Berg y sus colaboradores demostraron que la molécula de DNA resultante podía ser integrada en los cromosomas de una célula de mamíferos. Pero como contenía también secuencias de un fago, podía a su vez replicarse autónomamente en bacterias. Con este trabajo se abrió por primera vez la posibilidad de introducir genes en un cierto organismo al mismo tiempo que se podía usar a las bacterias para amplificar las moléculas de DNA híbridas que se habían creado *in vitro*. Del mismo modo que en 1953 el trabajo de Watson y Crick abrió las puertas de la biología molecular, el trabajo de Berg y colaboradores en 1972 supuso el comienzo de la ingeniería genética.

Conviene hacer aquí hincapié en que la ingeniería genética no se desarrolló porque hubiese una industria biotecnológica que la demandase, ni tampoco por el afán de curar enfermedades, sino que tenía como principal finalidad encontrar genes y comprender su función biológica. La tecnología para manipular DNA se desarrolló fomentando la investigación básica y creativa, es decir: delegando en el científico la responsabilidad de determinar lo que considera una pregunta interesante. Una vez más, la historia demuestra que el proceso de la investigación no puede estar organizado, tiene que aparecer de forma espontánea a través del talento individual. La industria se dio cuenta de las enormes posibilidades que ofrecía la ingeniería genética después de que Berg publicase su famoso trabajo en 1972. Para entonces, la biología molecular era ya, en el sentido que expresa Thomas Kuhn en el libro *La estructura de las revoluciones científicas*, una disciplina bien establecida y algunos científicos habían comenzado ya a mirar en otras direcciones, aún inexploradas, como el ciclo celular, la biología del desarrollo y de la evolución, la bioinformática, etc.

En 2003, coincidiendo con el cincuenta aniversario del descubrimiento de la estructura en doble hélice del DNA se hizo disponible una versión «refinada» del genoma humano. Previamente, el 26 de junio de 2000, en un acto histórico presidido por Tony Blair desde el Reino Unido y por Bill Clinton desde los Estados Unidos, se había hecho público el primer borrador de la secuencia del genoma humano un poco antes de que apareciese publicado en las revistas *Nature* y *Science*.

La secuenciación del genoma humano ha sido el proyecto de colaboración internacional en biología más caro y ambicioso que jamás se ha llevado a cabo. Por el interés con el que ha sido seguido por los medios de comunicación y la sociedad sólo es comparable al proyecto

de llevar al hombre a la Luna. El proyecto genoma humano emergió de varias iniciativas independientes llevadas a cabo a mediados de los ochenta por Robert Sinsheimer, un biólogo molecular de prestigio que había sido nombrado presidente del campus de Santa Cruz de la Universidad de California en 1977, y por Charles de Lisi, director de la Oficina de Salud y Medio Ambiente del poderoso Departamento de Energía (DOE) estadounidense. El proyecto genoma humano apareció como una aplicación lógica de las tecnologías del DNA (métodos para secuenciar DNA, clonar y manipular genes) que se habían desarrollado en los años sesenta y setenta. Para llevar a cabo este proyecto se creó un consorcio internacional en el que desde el principio los Estados Unidos y el Reino Unido jugaron un papel destacado. En una reunión internacional que se celebró a finales de febrero de 1996 en Bermudas, los principales responsables del proyecto acordaron que todos los datos de la secuencia del genoma humano serían de dominio público para así optimizar su uso e impulsar los beneficios de este proyecto para la sociedad. A mediados de 1996 la comunidad científica estaba ya en disposición de iniciar de forma sistemática el proyecto de secuenciar el genoma humano (se disponía de financiación y se había llegado a un acuerdo internacional sobre el acceso a la información que se fuese generando) y se calculó que podría estar finalizado a mediados del 2003.

Pero en 1998, cuando el proyecto público llevaba ya tiempo haciendo accesibles sus datos del genoma humano, una empresa comercial privada creada por Craig Venter, un investigador que había desarrollado su carrera científica en los prestigiosos Institutos Nacionales de la Salud (NIH) estadounidenses, irrumpió ruidosamente en este escenario con el propósito de secuenciar el genoma humano en dos años con objetivos comerciales. Sus argumentos eran, que con una gestión privada y una estrategia diferente su empresa podría secuenciar el genoma humano mucho más rápido que el consorcio público. De esta manera se inició lo que los medios dieron por denominar la carrera por la secuenciación del genoma humano entre el sector público y el privado. Sin embargo, un mínimo análisis objetivo de la situación real que se creó desbarata esa simplista explicación. En primer lugar, el objetivo de Celera, la empresa presidida por Venter, no era secuenciar el genoma humano con la precisión que se estaba haciendo en el consorcio público (menos de un error por cada diez mil pares de bases) sino producir lo que se conoce como un «borrador», una secuencia no totalmente finalizada. En segundo lugar, cuando Celera comenzó su proyecto de secuenciar el genoma humano no lo hacía desde cero, sino que partía con toda la información que había hecho accesible durante años el consorcio

público. Además, mientras que el consorcio público para el proyecto genoma humano diariamente hacía accesibles sus datos, Celera hacía uso de estos datos pero mantenía en secreto los suyos. Como se puso de manifiesto cuando en el 2000 aparecieron publicados los análisis del genoma humano del consorcio público y el de Celera, sin los datos del proyecto público la compañía de Venter no habría sido capaz de generar un borrador de la secuencia del genoma humano. En el 2002, al no poder tener derechos exclusivos sobre el genoma humano, Celera abandonó su ambición de vender información genética para dedicarse a la búsqueda de nuevas dianas terapéuticas y Craig Venter dimitió como presidente de la compañía. Mientras tanto, las acciones de Celera cayeron desde 247 dólares en marzo de 2000 a 15 dólares en mayo de 2002.

¿Para qué nos sirve disponer de la secuencia del genoma humano? ¿Nos ayudará a comprender cómo un organismo tan sencillo como una bacteria pudo evolucionar hasta dar lugar a algo tan complejo como un ser humano? ¿Nos servirá para decir qué nos hace ser seres humanos en lugar de, por ejemplo, chimpancés, cuando existe tan sólo una diferencia del uno por ciento entre el genoma de estos otros primates y el nuestro? ¿Nos ayudará a comprender como en tan sólo unas semanas un óvulo fertilizado se desarrolla hasta formar un feto reconocible como humano en el interior del útero de la madre, y a entender en detalle la complejidad, integración y reproducibilidad que caracterizan este proceso? ¿Nos ayudará a comprender por qué somos sensibles a ciertos virus como el SIDA, la hepatitis B y C, o la influenza, o a contraer la malaria? ¿Nos servirá para saber por qué algunos individuos son más sensibles que otros a desarrollar enfermedades complejas como la obesidad, cirrosis, cáncer, o Parkinson? ¿Nos ayudará a desarrollar nuevas y mejores medicinas?

El conocimiento de la secuencia del genoma humano no nos llevará de manera inmediata a contestar estas u otras preguntas similares, pero sin duda nos ayudará a encontrar su contestación. Así, por ejemplo, la finalización de la secuenciación del genoma de distintos organismos eucarióticos (que poseen núcleo) como levadura, arroz, mosca, ratón y humano, ha permitido compararlos y confirmar la existencia de alrededor de un centenar de genes específicos de los organismos eucarióticos que no se encuentran presentes en los organismos procarióticos (que no poseen núcleo, como por ejemplo las bacterias). Es decir, para conseguir la complejidad estructural que supuso pasar de las células sin núcleo, como las bacterias, a las células con núcleo fueron tan sólo ne-

cesarios alrededor de cien genes nuevos. Entre los genes identificados los hay que codifican proteínas del citoesqueleto, otros que están implicados en funciones de la membrana celular (como las denominadas proteínas G) y también los hay que están implicados en la regulación de la división celular, procesamiento de RNA y proteólisis, además de ciertas enzimas (proteína quinasas y fosfatasas) específicas.

La utilización de una estrategia similar para determinar cuántos genes nuevos fueron necesarios para hacer posible la aparición de los primeros organismos eucarióticos multicelulares dio como resultado únicamente tres genes: un factor de transcripción (los factores de transcripción regulan la expresión de otros genes), una proteína que une RNA y una proteína que une selenio. Aunque este resultado esté en parte influenciado por el hecho de que los dos organismos que se compararon divergían «tan sólo» en 300 millones de años —es decir procedían de un ancestro eucariótico común presente quizá hace 1.000 millones de años— lo cierto es que estos resultados claramente indican que es la manera de usar los genes ya existentes y no la aparición de nuevos genes lo que hizo posible la evolución de formas más complejas de vida. En otras palabras, procesos morfológicamente tan distintos como el desarrollo de un gusano, una mosca, un ratón o un ser humano, utilizan básicamente la misma «caja de herramientas moleculares» es decir, los mismos genes. La secuenciación y comparación del genoma de otros organismos eucarióticos unicelulares y multicelulares aclarará esta teoría en el futuro y nos permitirá identificar qué genes específicos hay en esa «caja de herramientas moleculares».

Del mismo modo, la comparación de la secuencia del genoma de organismos como la mosca de la fruta *Drosophila melanogaster* y el gusano *Caenorhabditis elegans*, ha demostrado que la mayoría de los genes reguladores (factores de transcripción, componentes de rutas de señalización y factores de crecimiento) se han conservado en el reino animal. La comparación de estos genomas con el de dos especies de arroz indica que hay numerosos genes reguladores que son específicos de las plantas y que no se encuentran en el reino animal. De otra parte, en las plantas como en los animales los principales genes reguladores (genes implicados en precisar el momento de la floración, la morfología de las flores, hormonas y factores de crecimiento) se encuentran, en general, bien conservados.

Finalmente, la disponibilidad de la secuencia del genoma de ratón y del humano ha permitido concluir que la mayoría de los genes de ambos genomas están bien conservados y, aunque aún no se dispone del

borrador de la secuencia del genoma de otros primates, la información disponible permite concluir que el genoma humano y el del chimpancé deben de ser casi idénticos. Para muchos, la búsqueda de diferencias entre el genoma humano y el del chimpancé puede ser de gran valor para identificar qué genes están implicados en el lenguaje o en las capacidades cognitivas. Por ejemplo, recientemente se ha descubierto que las mutaciones de un gen denominado *FOXP2* producen severos desórdenes en el lenguaje. Este gen produce una proteína que difiere en tan sólo dos aminoácidos respecto de la del chimpancé. De qué forma participa esta proteína en el lenguaje es, sin embargo, desconocida. Pero la búsqueda de diferencias entre el genoma humano y el del chimpancé también puede ayudar a explicar la diferente susceptibilidad de ambas especies a desarrollar ciertas condiciones o enfermedades. Por ejemplo, en las mujeres la menopausia es universal mientras que esta condición es poco frecuente en las hembras de chimpancés. De manera similar, los humanos son mucho más susceptibles que los chimpancés a infectarse con ciertos virus como el del SIDA, hepatitis B y C, influenza, o a contraer malaria; y asimismo, mientras que la cirrosis, el cáncer, Alzheimer y la aterosclerosis son frecuentes en humanos, estas enfermedades son muy raras en chimpancés. ¿Se deben estas diferencias entre el hombre y el chimpancé en la capacidad de desarrollar un lenguaje o en la susceptibilidad a padecer ciertas condiciones o enfermedades a la existencia de diferencias en sus genomas, o por el contrario se debe a que ambas especies usan los mismos genes de forma distinta?

Dada la universalidad de los llamados genes reguladores y la gran conservación de la secuencia del genoma entre especies separadas por más de mil millones de años de evolución como el ratón y el hombre, ¿cómo explicar la gran variedad de formas que han adoptado los organismos durante la evolución y la organización de los diferentes órganos y tejidos? Consideremos, por ejemplo, la forma del pico de las aves, el cual ha experimentado grandes cambios morfológicos durante la evolución. Estas diferencias están íntimamente ligadas a la adaptación de las aves a nuevos ambientes. De hecho, Charles Darwin argumentaba ya en *El Origen de la Especies* que la selección natural gobernaba el tamaño y la forma del pico de las aves. Las aves usan el pico para alimentarse y alimentar a sus crías, fabricar nidos, defenderse, atacar y limpiarse. Consecuentemente, la forma y el tamaño del pico de las aves se ha adaptado al tipo de comida con el que se alimentan. Así, mientras que el pico del pato es largo, ancho y plano, para facilitarle coger peces y plantas acuáticas, el de la codorniz es corto, fino y cóncavo, lo que le facilita alimentarse de pequeños granos de semillas. Los huesos y cartí-

lagos que forman el pico de las aves derivan todos de un único grupo de células embrionarias que reciben el nombre de células de la cresta neural. Recientemente unos investigadores han intercambiado las células de la cresta neural entre embriones de pato y de codorniz. Mientras que el embrión de pato trasplantado con células de la cresta neural de codorniz desarrolló pico de codorniz (corto, fino y cóncavo), el embrión de codorniz trasplantado con células de la cresta neural de pato tenía pico de pato (largo, ancho y plano). El pato y la codorniz utilizan los mismos genes para regular el tamaño y la forma del pico, pero ambas especies difieren en el momento del desarrollo en el que expresan estos genes reguladores y, consecuentemente, en la red de interacciones en la que participan. Es decir, las diferencias morfológicas entre el pico del pato y de la codorniz se deben más a variaciones espacio-temporales en la expresión de ciertos genes «clave» que a diferencias en el tipo de genes que ambas especies utilizan para construir su pico.

¿Por qué durante años se ha favorecido la hipótesis de que la complejidad en biología se obtenía incrementando el número de genes, en lugar de cambiando la manera de usar los genes ya existentes? La biología molecular, a través del estudio directo de los genes, ha desarrollado un concepto determinista de la biología que descansa sobre el principio de la especificidad. Según este concepto, que ha sido fundamental para entender la función de los genes y las proteínas, a un conjunto de moléculas le corresponde una única estructura tridimensional es decir un único genotipo. ¿Cómo puede, entonces, producirse diversidad de fenotipos a partir de un único genotipo? Lo que ocurre es que la especificidad es un término cualitativo y abstracto. En biología, las interacciones entre moléculas (DNA, RNA, proteínas, metabolitos) se caracterizan no por la exclusividad, sino, al contrario, por la multiplicidad de posibles interacciones entre unas moléculas y otras. Las funciones biológicas, consecuentemente, están constituidas no por procesos lineales sino por redes de interacciones entre los componentes de la célula. Como cada una de estas posibles interacciones está sujeta a fluctuaciones ambientales, tiene una cierta probabilidad de realizarse y una estabilidad diferente, las interacciones en biología se describen mejor como un proceso estocástico. La cuestión es ahora conocer cuáles de estas posibles interacciones van a producirse en un determinado momento y, por consiguiente, como se produce un organismo, un tejido o una función biológica única de entre todas las posibles.

El azar desempeña un papel fundamental en biología no sólo generando errores durante la replicación del DNA, lo que da lugar a

mutaciones y a la evolución, sino también determinando el destino de cada célula durante la diferenciación celular y morfogénesis. Sin embargo, el comportamiento de las células en un sistema biológico se caracteriza por ser ordenado: cada célula desempeña una función específica dentro de unos estrictos parámetros y, en general, este comportamiento es robusto. Es decir, una vez que una célula ha escogido un destino, este proceso es resistente a cambios internos y externos. ¿Cómo consigue una célula funcionar de manera ordenada en un medio bioquímico desordenado? La contestación posiblemente se encuentra en la estructura de las redes de interacciones entre moléculas de DNA, RNA, proteínas y metabolitos que sustentan la fisiología celular.

De acuerdo con el dogma básico de la biología molecular, la complejidad biológica se encuentra almacenada en el DNA. Así, es generalmente aceptado que el almacenamiento de información, el procesamiento de información, y la ejecución de los diversos programas celulares reside en distintos niveles de organización: el genoma de la célula, el transcriptoma (el conjunto de los factores de transcripción de un organismo), el proteoma (el conjunto de proteínas de un organismo) y el metaboloma (el conjunto de metabolitos de un organismo). Sin embargo, la estanqueidad de estos niveles de organización está siendo actualmente cuestionada. Así, aunque la información de larga duración se almacena casi exclusivamente en el genoma, el proteoma es crucial para almacenar la información de corta duración; y el proceso de ir obteniendo información del genoma, que está controlado por los factores de transcripción, se encuentra fuertemente influenciado por el estado del metaboloma. Esta integración entre los diversos niveles de organización nos obliga a mirar a las funciones biológicas como distribuidas en grupos de componentes heterogéneos que interactúan unos con otros dentro de grandes redes.

La aspiración de esta nueva biología, que ha recibido el nombre de biología de sistemas, no es sólo la de proporcionar una teoría global que nos sirva para entender los procesos biológicos complejos, como el desarrollo embrionario o la morfogénesis, sino que aspira también a conocer la estructura de las redes de interacciones entre moléculas de DNA, RNA, proteínas y metabolitos que sustentan la fisiología celular para así facilitar la identificación de nuevas dianas terapéuticas, el desarrollo de nuevos fármacos, y la optimización de los tratamientos médicos.

Consideremos, por ejemplo, un gen cuyas variaciones inducen una cierta enfermedad. Esta susceptibilidad no será aparente si exis-

ten circuitos que compensen los efectos de esta variabilidad y sólo se hará evidente si dichos circuitos compensatorios se interrumpen por alguna razón. El conocimiento de estos circuitos compensatorios posibilitará el desarrollo de nuevos medicamentos que permitan restaurar su función cuando se encuentre interrumpida previniendo así el desarrollo de la enfermedad. El estudio de la estructura de las redes biológicas que regulan las funciones celulares (proliferación y muerte celular, inflamación, etc.) facilitará, asimismo, el desarrollo de nuevos tratamientos contra el cáncer y otras enfermedades complejas, así como el diseño de circuitos genéticos sintéticos que detecten los niveles de expresión de ciertos genes (supresores de tumores, citoquinas, oncogenes, etc.) y los active o inactive cuando sea necesario, como se hace en la actualidad en ciertos diseños de ingeniería. Comprender las leyes que gobiernan la estructura de estas redes biológicas será uno de los principales retos de la biología durante las próximas décadas. Para desentrañar los principios básicos que regulan la estructura, función y evolución de estas redes biológicas será necesario una aproximación multidisciplinar que combine métodos clásicos con las nuevas técnicas de genómica, proteómica, metabolómica y biología computacional.

Para finalizar, quisiera recordar que la historia de la ciencia está llena de ejemplos que confirman el carácter impredecible de la misma. ¿Podría acaso alguien haber anticipado que los trabajos de Avery sobre neumonía le llevarían a descubrir que el DNA era la sustancia de los genes? Pero no sólo son impredecibles los resultados de la investigación científica, sino también sus aplicaciones. ¿Podría acaso alguien haber anticipado hace 50 años, cuando se descubrió la estructura del DNA, que las tecnologías basadas en los ácidos nucleicos se convertirían en el principal exponente de la economía basada en el conocimiento en el siglo XXI? De otra parte, es importante recordar que, con frecuencia, las promesas que algunos políticos y científicos hacen sobre las aplicaciones de la ciencia no se suelen cumplir. Recordemos, por ejemplo, la del presidente estadounidense Richard Nixon, prometiendo que en diez años se curaría el cáncer; o la de algunos científicos que en los años ochenta nos prometían que el xenotrasplante (es decir, el trasplante a humanos de órganos de animales, principalmente del cerdo) se encontraba a la vuelta de la esquina y acabaría con las listas de espera para el trasplante de órganos; o la de aquellos otros que hace ya cerca de veinte años nos prometieron que la terapia génica serviría para curar numerosas enfermedades, entre ellas, por supuesto, de nuevo el cáncer. Ninguna de estas tecnologías

ha encontrado aún su aplicación terapéutica y cuando lo hagan, lo más probable es que inicialmente tengan un impacto discreto sobre la salud, ya sea mejorando un tratamiento existente, o la calidad de vida de ciertos pacientes, o aumentando en unos cuantos años la supervivencia a alguna enfermedad incurable. Ahora le ha llegado el turno a la terapia celular y en particular a las células madre, en donde desde 1999, cuando dos laboratorios estadounidenses consiguieron cultivar por primera vez células madre embrionarias humanas, se han producido más testimonios que sustancia. Comprender las interacciones entre genes, proteínas y metabolitos que mantienen a las células madre embrionarias en estado indiferenciado, así como los cambios que se producen cuando diferencian, establemente, en otros tipos celulares, como neuronas, células beta pancreáticas, o hepatocitos, es una pregunta biológica de enorme interés, y no es necesario justificar estas investigaciones con promesas sobre futuras aplicaciones más o menos especulativas. Es como si Watson, Crick, Franklin, o Wilkins, hubiesen tenido que justificar sus investigaciones sobre la estructura del DNA prometiendo que iban a curar el cáncer.

Es fundamental comprender que la ciencia y sus aplicaciones son impredecibles y que sólo progresan fomentando un entorno intelectual libre, tolerante y multidisciplinar. La historia del SIDA, en la que trabajaron y siguen trabajando conjuntamente virólogos, biólogos moleculares, bioquímicos, químicos, y clínicos, ejemplifica, mejor que ninguna otra, no sólo lo impredecible que son las aplicaciones de la ciencia, sino también el enorme impacto que la cadena de descubrimientos que a lo largo de más de 100 años unen a Mendel, Avery, Watson, Crick, Sanger, Arber, Berg, etc. y que culminaron con la secuenciación del genoma humano ha tenido, y seguirá teniendo, sobre nuestra sociedad. Los primeros casos de SIDA se diagnosticaron en 1981, y en menos de una década, gracias a las modernas técnicas de biología celular y molecular, se aisló el virus del HIV, se aprendió a cultivarlo, se secuenció su genoma, se desarrolló un test diagnóstico eficaz, se estudió su ciclo de vida, y se desarrollaron una batería de fármacos que transformaron esta enfermedad de mortal en crónica. De no haber sido por la moderna biología, la epidemia del SIDA se habría extendido por Estados Unidos y Europa como lo ha hecho por África, la expectativa de vida en Occidente no estaría hoy en los 80 años, y el gasto médico en Occidente, a causa de esta imparable epidemia, se habría disparado de tal manera que el imponente crecimiento económico que hemos experimentado en los 20 últimos años se habría reducido drásticamente.

