



EL CENTRO DE REGULACIÓN GENÓMICA

Tras el genoma humano

En busca de la partitura de la vida



© CRG, 2006

© Luis Ángel Fernández Hermana, 2006.

Julio-agosto 2006

Autor:

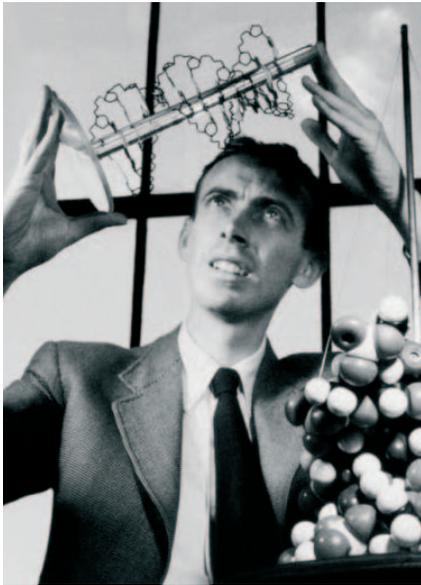
Luis Ángel Fernández Hermana
Centro de Regulación Genómica de Barcelona

Diseño y producción:

Estudi Freixes

Tras el genoma humano

En busca de la partitura de la vida



James Watson, descubridor, junto con Francis Crick, de la estructura en doble hélice del DNA. Ambos recibieron el Premio Nobel de Medicina en 1962 "por sus descubrimientos de la estructura molecular de los ácidos nucleicos y su significación para la transferencia de información en los materiales vivos".

En 1986, el premio Nobel Renato Dulbecco publicaba un artículo en la revista Science en el que postu-

laba la necesidad de secuenciar todo el genoma humano si se quería luchar seriamente contra el cáncer. El científico sostenía que la estrategia de aislamiento y secuenciación de genes individuales que se practicaba en aquel entonces resultaba demasiado lenta y ardua, y era de dudosa efectividad. Se necesitaba un salto adelante que mostrara con mayor claridad las relaciones entre algunos genes, en particular los asociados a algunas enfermedades hereditarias.

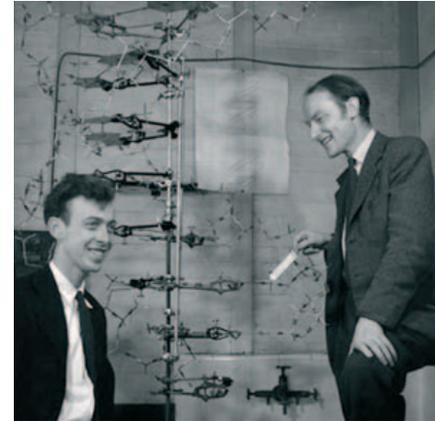
Esta propuesta se inscribía en una corriente de opinión entre los biólogos más destacados y reconocidos del momento, quienes promovieron diversos encuentros auspiciados por el Departamento de Energía de EEUU, como la cumbre de Alta (Utah, EEUU) de diciembre de 1984, que muchos señalan como el primer acto de la gran obra de la secuenciación del genoma humano. De todas maneras, éstas eran palabras mayores en aquella época (y hablamos de poco más de 20 años). La gran preocupación, por supuesto, era cómo abordar una tarea tan enorme –en realidad, tan desconocida en cuanto a sus posibles dimensiones- y qué exigencias tecnológicas plantearía para poder acometerla y obtener resultados en un tiempo que todos adjetivaban como "razonable", aunque sin atreverse a marcarlo en el calendario. Estábamos a punto de doblar hacia un nuevo siglo y, en esas circunstancias, las décadas se nos caían de la boca con una incontenente alegría.

Finalmente, estos encuentros desembocaron en la articulación de uno de los proyectos científicos más singulares del siglo XX: en 1990, el Proyecto Genoma Humano se puso en funcionamiento en EEUU bajo la dirección de James Watson, uno de los dos científicos que había desentrañado la estructura en doble hélice del DNA en 1953. Cuando Watson asistió a una de las reuniones sobre el genoma humano que se celebraron en Valencia a principios de los 90, tuve la oportunidad de entrevistarle para El Periódico de Cataluña y preguntarle sobre su estimación acerca de cuándo tendríamos secuenciado todo el genoma

humano. “Después del 2010 y probablemente cerca del 2025” soltó casi mecánicamente, una respuesta que sin duda estaba cansado de repetir. “Esta es una empresa gigantesca, avanzamos rápido, pero despacio en relación con la dimensión y complejidad del código genético que debemos descifrar”.

A su lado estaba Francis Crick, su colega de aventura científica en el descubrimiento de la estructura del DNA y con quien compartió el premio Nobel, que movía la cabeza de derecha a izquierda en un evidente gesto de escepticismo. “Para comienzos del próximo siglo, es decir, dentro de unos 10 años a más tardar, ya tendremos el mapa completo del genoma humano. Créame, haga caso a mi titular, es el bueno”. Y lo fue. Vaya si lo fue. El 26 de junio del 2000, en una conferencia de prensa mundial simultánea en cinco países y con el estrado central en la Casa Blanca, se presentó públicamente el primer borrador de la secuenciación del genoma humano, aproximadamente un 90% del total. En el acto estaban presentes, además del presidente de EEUU Bill Clinton, dos de los protagonistas de uno de los saltos más espectaculares en la historia del género humano, en general, y de la ciencia, en particular: Francis Collins, sucesor de Watson en el Proyecto Genoma Humano de carácter público, y Craig Venter, presidente de Celera Genomics, una empresa que había mantenido una dura pugna contra todos los laboratorios públicos del mundo al tratar de unir la secuenciación del genoma humano con la opción de rentabilizar económicamente las aplicaciones científicas y farmacéuticas que se derivaran de esta iniciativa.

El conocimiento sobre nuestro código hereditario experimentó un nuevo impulso cuando, en abril de 2003, el Proyecto Genoma Humano publicó lo que denominó “una secuenciación global de alta calidad del genoma humano”. Como señaló entonces Francis Collins en un artículo en Nature (Vol. 422; 24/4/03), el nuevo mapa veía la luz a los cincuenta años del descubrimiento de la estructura de la doble hélice del DNA. “La era de la genómica es ahora una realidad”, dijo.

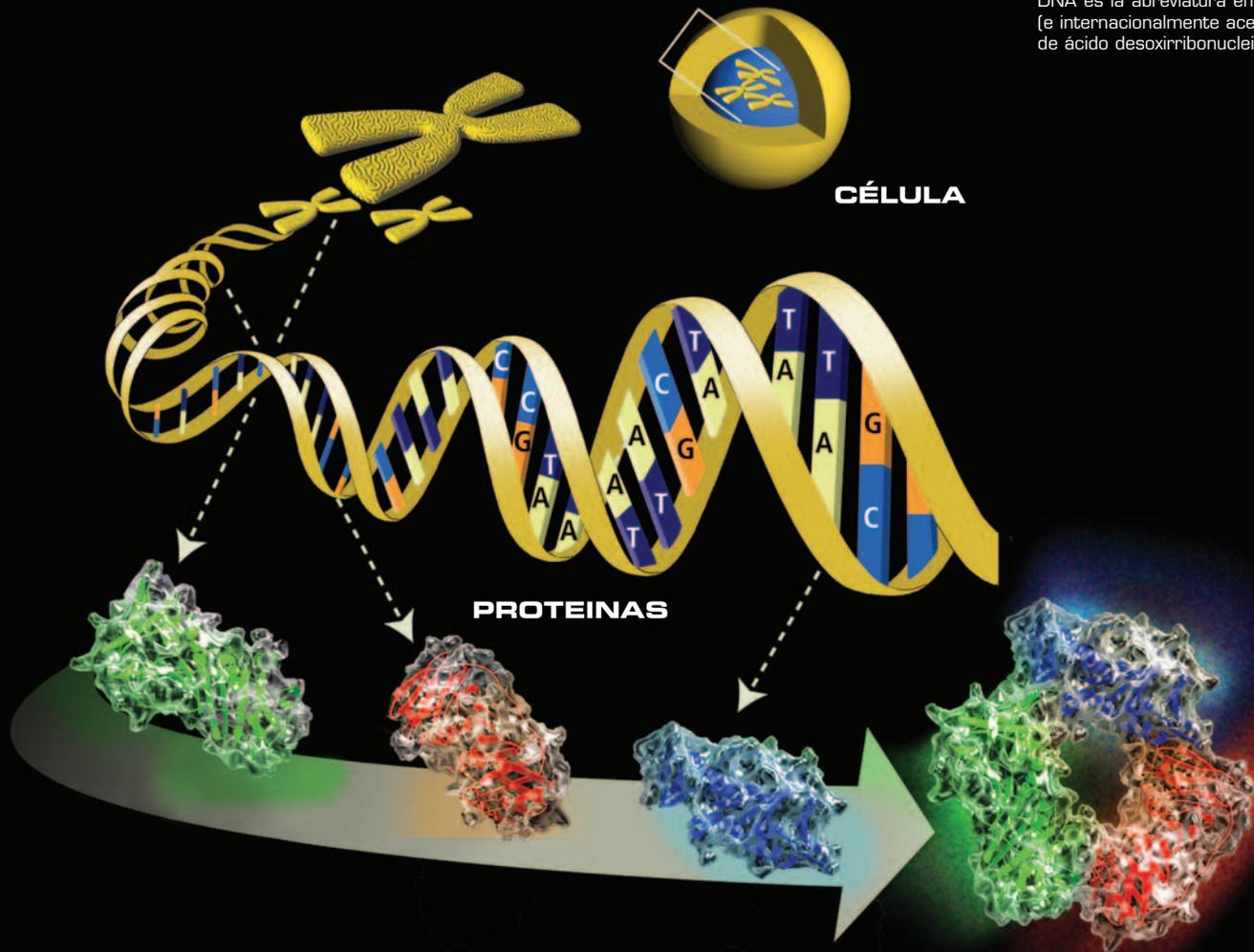


Watson y Crick.

“Para comienzos del próximo siglo, es decir, dentro de unos 10 años a más tardar, ya tendremos el mapa completo del genoma humano. Créame, haga caso a mi titular, es el bueno”

(Francis Crick, a principios de los años 90).

DNA es la abreviatura en inglés (e internacionalmente aceptada) de ácido desoxirribonucleico.



En la mayoría de los organismos vivos complejos, el DNA reside en el núcleo celular en forma de cromatina y organizado en cromosomas. El DNA está formado por dos hebras antiparalelas (orientadas en sentido opuesto) entrelazadas en forma de espiral, la famosa doble hélice. Cada hebra está formada por cuatro nucleótidos que se diferencian sólo en sus bases: dos purinas, denominadas adenina (A) y guanina (G), y dos pirimidínicas, denominadas citosina (C) y timina (T). Cada base de una hebra se empareja con una base de la otra, siempre la A con la T y la G con la C. Se estima que en el genoma humano tiene unos 3.000 millones de pares de bases.

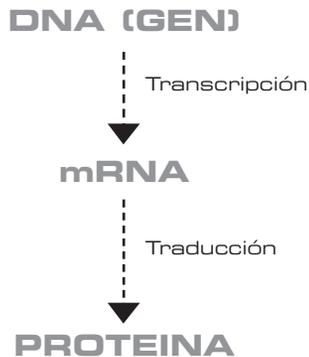
Y lo fue también para Barcelona y España. Porque en diciembre del año 2000, siguiendo la estela de la avalancha de nuevos descubrimientos relacionados con el código que regula la vida, se creaba en Barcelona el Centro de Regulación Genómica (CRG). A la discreción con que apareció este instituto en el paisaje científico catalán y español no fue ajeno el estentóreo eco del trueno de la secuenciación del genoma humano. Pero se trataba de un centro de nuevo cuño en el panorama científico español, cuya matriz era el formidable cambio experimentado en la comprensión y el desentrañamiento del código genético humano y la potencial aplicación de estos conocimientos a un amplísimo espectro de campos y disciplinas, no sólo el de la salud, aunque ese fuera el objetivo más conocido por el público.

Este breve ensayo trata de mostrar cómo la era de la genómica se encarna en un centro de investigación como el CRG y, viceversa, cómo dicho centro se convierte a su vez, en el contexto particular de Barcelona y con una organización que desafía a las más asentadas tradiciones de nuestra ciencia pública, en impulsor de una de las aventuras más prometedoras del conocimiento: el estudio de la regulación de los genes que contienen las instrucciones de la vida. El CRG forma parte del selecto grupo de centros de investigación en el mundo dedicados a desentrañar cómo el DNA que recibimos de nuestros progenitores instruye al huevo fecundado para iniciar un proceso de desarrollo y diferenciación embrionaria que produce un organismo con una forma y una función característica, así como que, con prácticamente los mismos genes, se generen individuos tan distintos como, por ejemplo, un gusano, un ratón o un ser humano.

Historias de un paisaje secuenciado

¿Qué es un gen?

Visión canónica del gen



Un gen en el sentido tradicional es una secuencia lineal de nucleótidos en la molécula de DNA que contiene la información para sintetizar moléculas con funciones celulares específicas, como proteínas, la mayoría enzimas, o distintos tipos de RNA. El DNA de un gen primero se transcribe en un RNA y este RNA mensajero (mRNA) se traduce en una proteína. Los genes ocupan lugares específicos a lo largo de cada uno de los cromosomas. De ahí que la secuenciación del DNA sea el método para desentrañar la localización y las funciones de los genes. El conjunto de los cromosomas de una especie se denomina genoma.

En periodismo se suele decir que, a lo largo de la producción de un diario, las oportunidades

para que un titular finalmente no exprese correctamente el contenido del artículo que encabeza son numerosas. A veces, incluso, lo milagroso es que suceda lo que supuestamente todo el mundo supone: que titular y contenido se correspondan sin que el primero diga más de lo necesario o diga cosas que ni siquiera son necesarias. Algo semejante podríamos decir del genoma. La delicada maquinaria hereditaria compuesta por unos de 22.000 genes en el caso de los seres humanos, la cascada de productos que estos generan con su actividad y sus interrelaciones, ofrece múltiples oportunidades para que el mensaje original se distorsione a lo largo del camino y acabe produciendo titulares o artículos que no tienen mucha relación con lo que dijeron las fuentes originales. Cuando sucede este tipo de aberración, el resultado suele ser un problema que, dependiendo de las oportunidades y de una multitud de circunstancias, oscila desde lo inocuo hasta la muerte, pasando por diferentes alteraciones que solemos denominar enfermedades.

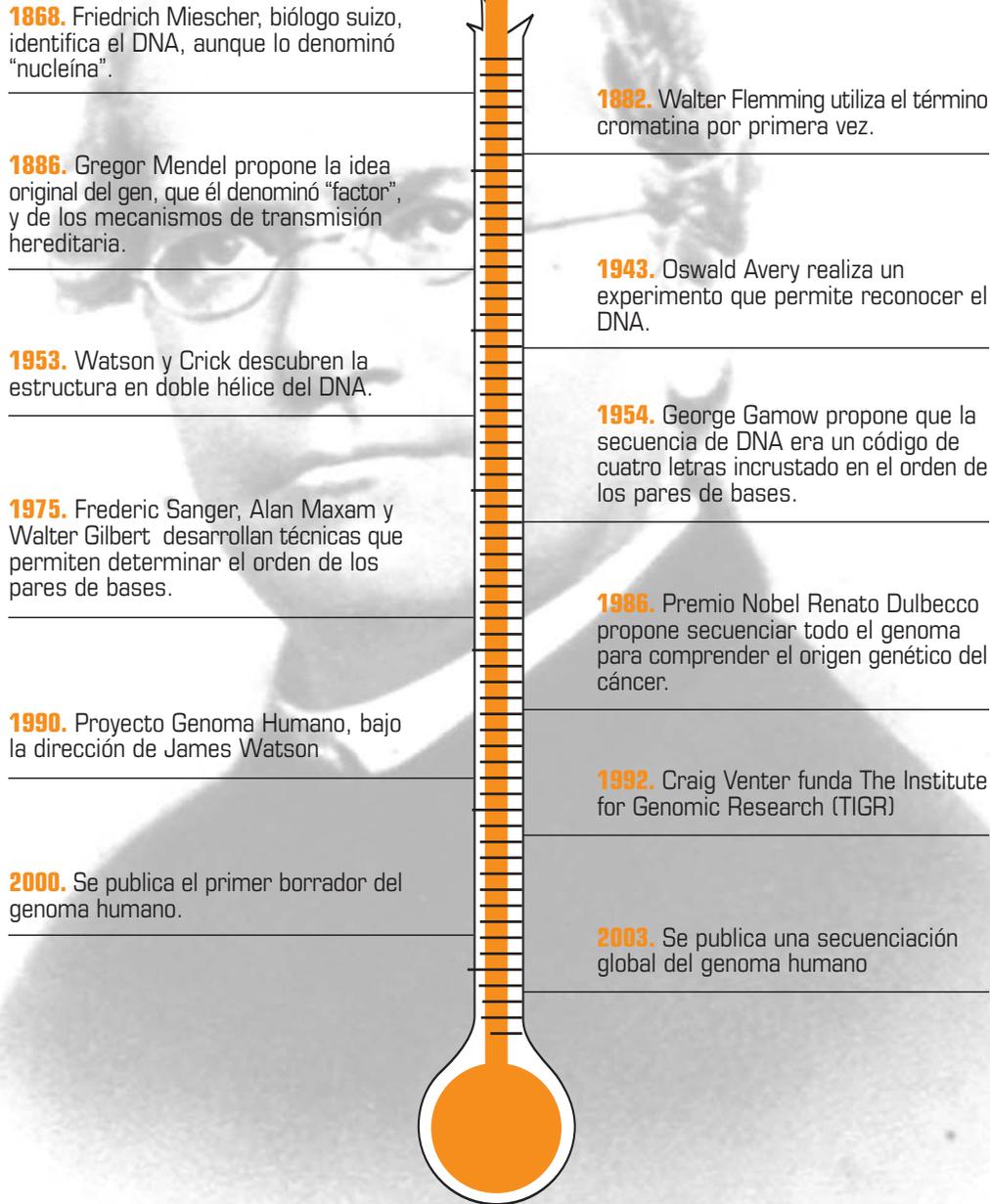
Ahora bien, cuando decimos genes, ¿de qué estamos hablando exactamente? Hasta hace poco, la respuesta entre los biólogos no dejaba muchas rendijas por donde pudieran escaparse destellos de heterodoxia. Hasta que la secuenciación del genoma humano no alcanzó velocidad de cruce, la definición canónica nos decía que el DNA, la cadena de nucleótidos que contiene la información de un ser vivo y que éste transmite a su descendencia, estaba integrado por un número determinado de genes. Cada gen tenía la información para fabricar una proteína. El proceso consistía, invariablemente, en copiar esas instrucciones mediante una molécula (el RNA mensajero, o mRNA) que las trasladaba a la “fábrica” para elaborar la proteína en cuestión. Un gen, un mensajero y una proteína o una enzima. A veces pasaban cosas “raras”, pero no las suficientes como para alterar esta mirada estática sobre el paisaje genético.

Esta visión todavía predomina en buena parte de la biología molecular actual, a pesar de que la secuenciación del genoma la ha puesto en estado de sitio. Hoy día, si alguien quiere boicotear un simposio sobre biología, sólo le basta lanzar la explosiva pregunta “¿Y qué es un gen?”. Garantizado: no habrá fumata blanca ni mintiendo, ni en el bar, y más de uno le calificará de “terrorista celular”. Pero sí habrá una mayoría considerable que se inclinará por las definiciones hechas a retazos o en sentido negativo: “Lo que no es un gen es que sólo sea el conjunto de instrucciones de una proteína”. En apenas unos años, informaciones fragmentadas, líneas de investigación aparentemente desconectadas, descubrimientos que se pensaban con efectos solamente allí donde se habían localizado, han ido convergiendo hasta conformar una imagen, todavía movida, borrosa, con vastas zonas donde predominan los oscuros sobre los claros, pero que comienza a transformar radicalmente las ideas que se consideraban mejor asentadas de la biología molecular.

Los cimientos de un cambio

La secuenciación del DNA, pues, ha puesto sobre la mesa no sólo un cúmulo casi inmanejable de información y un robusto arsenal de potentes tecnologías, sino, sobre todo, los pilares básicos para mirar al genoma de otra manera. La que muchos denominan nueva investigación genómica se sustenta en unos pocos cimientos, pero de suficiente entidad como para prometer un cambio considerable en nuestra manera de ver –y actuar sobre- la arquitectura biológica de la vida. Estos cimientos, aún a riesgo de dejarnos partes importantes en el cajón, podríamos agruparlos así:

- El empalme alternativo de exones
- La epigenética
- La bioinformática y las nuevas tecnologías aplicadas a la genómica
- La biología celular y las técnicas de imagen aplicadas a procesos celulares y de desarrollo



1868. Friedrich Miescher, biólogo suizo, identifica el DNA, aunque lo denominó "nucleína".

1886. Gregor Mendel propone la idea original del gen, que él denominó "factor", y de los mecanismos de transmisión hereditaria.

1953. Watson y Crick descubren la estructura en doble hélice del DNA.

1975. Frederic Sanger, Alan Maxam y Walter Gilbert desarrollan técnicas que permiten determinar el orden de los pares de bases.

1990. Proyecto Genoma Humano, bajo la dirección de James Watson

2000. Se publica el primer borrador del genoma humano.

1882. Walter Flemming utiliza el término cromatina por primera vez.

1943. Oswald Avery realiza un experimento que permite reconocer el DNA.

1954. George Gamow propone que la secuencia de DNA era un código de cuatro letras incrustado en el orden de los pares de bases.

1986. Premio Nobel Renato Dulbecco propone secuenciar todo el genoma para comprender el origen genético del cáncer.

1992. Craig Venter funda The Institute for Genomic Research (TIGR)

2003. Se publica una secuenciación global del genoma humano

- La biología de sistemas
- Las terapias basadas en la comprensión y la reparación de los sistemas de regulación genómica.

Estas seis columnas conforman, a su vez, como no podía ser de otra manera, las líneas troncales de investigación del Centro de Regulación Genómica de Barcelona, cuyo objetivo primordial es tratar de entender qué está escrito en el texto genético, cómo está escrito y cómo se interpreta. Y, acompañados de los principales responsables de desarrollarlas, vamos a tratar de comprender algunas de ellas para, a su vez, obtener un acercamiento más íntimo a las particularidades del CRG y el lugar que ocupa esta institución en el mapa de la investigación mundial sobre la regulación genómica y el estudio de las enfermedades causadas por el funcionamiento anómalo del genoma.

Pocos genes, pero muy trabajadores

El empalme alternativo de exones

El primer sobresalto que sufrió la visión clásica de la biología molecular ocurrió discretamente en

1977, cuando los grupos de Richard Roberts y Phillip Sharp descubrieron que los genes de algunos virus no contenían la información para fabricar proteínas dispuesta de una manera continua, sino en fragmentos que debían empalmarse correctamente para generar un mensaje traducible por la maquinaria que usa la célula para fabricar proteínas. Poco después se verificó que esta disposición truncada de los mensajes genéticos era la norma en los organismos multicelulares y, más sorprendente aún, que había formas alternativas de empalmar los fragmentos informativos de un gen y generar así no una, sino más de una proteína a partir de un mismo gen. En un derroche de generosidad, a veces podían generarse miles de proteínas a partir de un único gen mediante este proceso de empalme alternativo.

El empalme alternativo supuso un giro copernicano en la forma de contemplar las interacciones génicas, aunque al principio nadie se atrevió a considerarlo como un recurso frecuente del arsenal de la regulación genómica. ¿En qué consistía en realidad el empalme alternativo? En el DNA, los genes se configuran como una cadena de perlas unidos por un hilo constituido de “exones” e “intrones” [véase recuadro página 14]. Los exones son los que se copian y se utilizan como una “plantilla” original para fabricar proteínas. Durante mucho tiempo, a los intrones se les consideraba como código basura, que los genes eliminaban en el proceso de transcripción. No se entendía muy bien qué hacían allí, entre los genes y mezclándose con los exones. Algunos los consideraban reminiscencias de genomas primitivos.

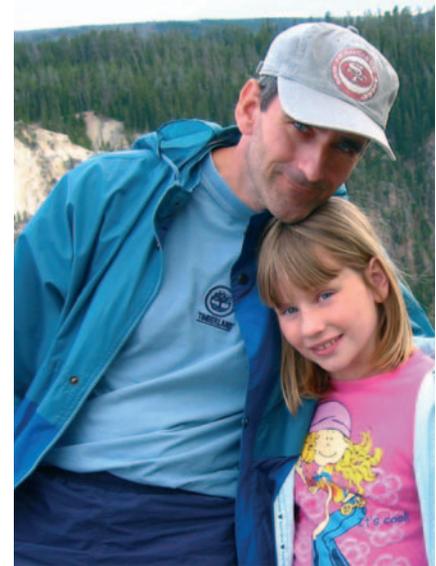
“Lo que nosotros investigamos es lo que sucede entre el DNA, donde tenemos la información genética, y los actores de la célula, que son las proteínas. En ese espacio tenemos a un intermediario, el RNA mensajero (mRNA), que es la copia del DNA que se va a convertir en proteínas. El dogma central de la genética es que un gen produce un RNA que produce una proteína. Sobre esta base se ha desarrollado lo que

sabemos de genética. Pero este dogma a veces es excesivamente complicado y a veces excesivamente sencillo. Si bien al gen siempre le consideramos el protagonista, estamos descubriendo que a veces son los RNAs, bien en su versión de mensajeros multipotentes o en su versión de moléculas reguladoras, los que juegan un papel preponderante en el funcionamiento del genoma”, explica Juan Valcárcel, responsable en el CRG del “Grupo de empalme alternativo de pre-mRNA durante la diferenciación, desarrollo y enfermedad”.

Juan Valcárcel considera que el nuevo conocimiento que se está adquiriendo sobre el funcionamiento del genoma está brindando una respuesta a uno de los grandes misterios de la secuenciación del DNA humano: el relativo escaso número de genes que compone nuestra dotación genética, apenas unos 22.000, una cantidad muy parecida a la del genoma del ratón y no muy superior a la de gusanos e insectos. Cuando se reveló esta similitud, la gente no sabía si incrementar su admiración por los animales más simples o sucumbir a un ataque de subestima. Pero muchos de los genes de nuestro genoma son capaces de utilizar los exones para, barajándolos, fabricar una cantidad considerable de productos diferentes. Ahí podríamos tener la explicación de la complejidad del ser humano y, en general, de los mamíferos.

Mensajeros de muchos reinos

La sorpresa de 1977 fue el sordo pistoletazo de salida de una carrera que aceleraba su marcha a medida que aumentaba la información y ésta iba encajando en el rompecabezas genómico. En 1993 se descubrió que los intrones eran capaces de fabricar microRNAs, unas diminutas moléculas que no sólo transcribían mensajes, sino que portaban otras misiones que, a final de cuentas, propiciaban ciertas actividades de los genes. En otras palabras, actuaban como factores de regulación génica: ahora enciendo este gen, ahora lo apago, ahora permito que pase esta sustancia química, ahora no, o dejo que se corte este fragmento o que se pegue con aquel otro; todo lo cual concluía quizá con una proteína nueva o con una actividad del genoma que



“Lo que nosotros investigamos es lo que sucede entre el DNA, donde tenemos la información genética, y los actores de la célula, que son las proteínas.”

(El científico Juan Valcárcel paseando con su hija Flora).



se desarrollaba a bastante distancia del gen que había iniciado todo el proceso. Esto acababa con el paradigma de un gen, un RNA, una proteína, una función. A partir de entonces, se podía cortocircuitar este proceso y pasar del gen a la función sin detenerse en la proteína.

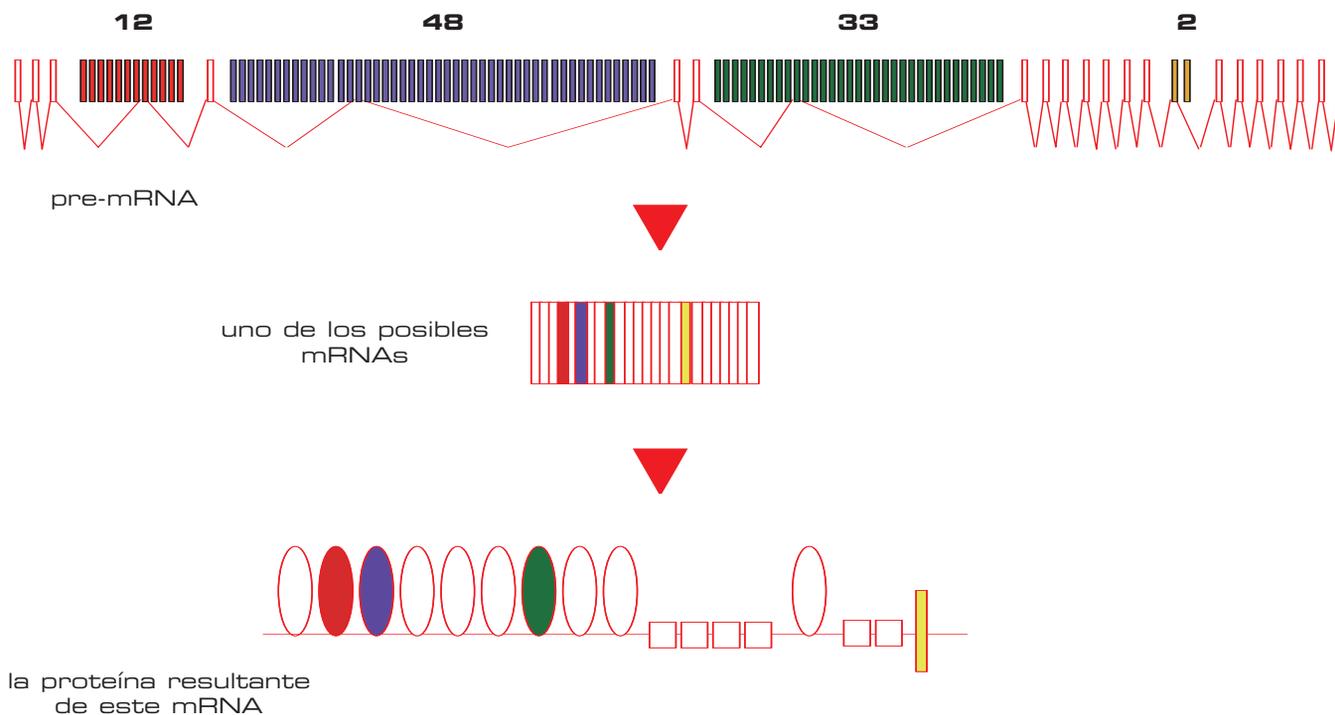
Las investigaciones de los últimos 4 años han deparado galaxias enteras de nuevos tipos de RNA a los que se les está tratando de encontrar trabajo, es decir, saber a qué se dedican. Están desde los RNA no codificantes (nc-RNA, no transmiten instrucciones para fabricar proteínas o enzimas), hasta RNA muy pequeños (microRNA) cuyas tareas en muchos casos no se conocen todavía muy bien. Se sabe que no fabrican proteínas, pero no se están quietos: regulan funciones que podríamos denominar de "asistencia". Por ejemplo, quizá estén involucrados en la actividad de algunos mRNA que salen de un gen y producen una proteína completamente diferente de la que producen otros mRNA de ese gen.

Por este camino, se han ido descubriendo nuevos aspectos de la actividad del genoma que abren campos de enormes posibilidades. El mecanismo de transcripción de las proteínas codificadas en el DNA parece no reconocer fronteras, literalmente: con frecuencia, lo importante no es la contigüidad, que cada secuencia trabaje con la siguiente, como se pensaba hasta ahora, sino que también pueden alcanzar a secuencias que se encuentran a cientos de miles de bases de distancia. La implicación es que la regulación de muchos genes (si se expresan o no, qué fabrican o dejan de fabricar, qué fragmentos instruyen que se corten y se peguen, etc.) puede estar controlada por regiones que pertenecen a otros cromosomas. Valcárcel dice que la confirmación de estos procesos apunta a un orden muy diferente del que se pensaba que regía hasta ahora: en vez de genes que se encargan de controlar su parcela y fabricar su proteína, el código genético es una especie de maquinaria que funciona como un continuo, sin un comienzo ni un fin nítido.

El "empalme alternativo" es el mecanismo por el que los genes fabrican productos en función de la célula, según que ésta se halle en el hígado, la sangre, el ojo, la vejiga, un músculo o el corazón. Como se sabe, en todos los casos, esas células poseen siempre el mismo genoma, pero éste actúa en consecuencia según el lugar donde se encuentre: en unos casos, el mismo gen fabricará las proteínas necesarias para construir el hígado, o los compuestos inmunológicos de la sangre o el tejido de órganos radicalmente diferentes en su conformación física y en sus funciones. ¿Cómo lo hace? ¿Cómo sabe qué parte activar de la maquinaria en cada circunstancia? Al parecer, aquí es donde interviene una multitud de diferentes micro mRNA para transmitir mensajes que permitirán transcribir correctamente las órdenes en cada caso y en cada lugar. Roderic Guigó, responsable del grupo de Bioinformática del CRG, publicó un artículo en 2006 donde confirmaba que la transcripción de los mensajes podía comenzar en una secuencia de DNA asociada con una proteína y seguir a través del gen para concluir en una proteína completamente diferente, produciendo una transcripción "fusionada". Este proceso le permitiría a la célula generar una gran variedad de proteínas a partir de un número limitado de exones.

Ejemplo de empalme alternativo de exones

Un solo gen puede dar hasta más de 38,000 proteínas



Los exones están indicados por barritas verticales. Las barritas blancas son exones comunes, mientras que los coloreados son exones alternativos. La línea indica el modo cómo se empalman los exones en este caso particular.

Actividad galáctica

En otras palabras, a lo mejor el concepto de gen con el que hemos operado hasta ahora simplemente era una quimera (lo cual tiene sus bemoles), o sólo arroja luz sobre algunos aspectos esenciales de sus funciones. El conocimiento que está proporcionando la investigación apunta de hecho a la interacción de intrincadas redes génicas y de vasto alcance, con capacidad para desplegar las instrucciones que encienden, apagan, aceleran o frenan la actividad de los genes en diferentes regiones, por más alejadas que éstas se encuentren en el genoma. Esto es lo que se denomina regulación genómica. Y sus implicaciones, como puede uno fácilmente imaginarse, son fenomenales, sobre todo para la forma como hemos entendido muchas enfermedades hasta ahora y las nuevas posibilidades que se abren para combatirlas.

Valcárcel explica esta conexión aportando el punto de vista desde el empalme alternativo de exones: “Durante el proceso de fabricación de proteínas, la maquinaria de procesamiento de los mensajes genéticos tomará los exones que son apropiados para la función deseada y creará con ellos el producto solicitado. Ahora bien, si esta maquinaria olvida algún ingrediente o si, por el contrario, se incluye alguno que, en realidad, no corresponde al producto que se desea fabricar, aquí pueden empezar los problemas”. Si la proteína no satisface la demanda específica de la célula –por decirlo de alguna manera- entonces ésta puede comportarse erráticamente. Y cuando se habla del genoma humano y las células que fabrica, por “comportamiento errático” podemos entender un amplísimo abanico de opciones, que van desde células con mensajes que las induzcan al suicidio (apoptosis) o, por el contrario, en el otro extremo, a su reproducción acelerada y descontrolada sin tiempo para diferenciarse de acuerdo a las necesidades del órgano en el que se encuentran. Este suele ser el inicio de un proceso canceroso, como nos explicará más tarde Luciano Di Croce, investigador de la leucemia en el CRG.

“Lo que a nosotros nos interesa es este proceso de diversificación por el que un gen produce diferentes proteínas. El DNA va a producir un RNA. Pero éste no tiene la información escrita de manera continua. Es como si tuvieras un libro con las primeras páginas perfectamente escritas, después hay 40 páginas en blanco y después continúa el texto. Es una especie de Rayuela de Julio Cortázar en más de un sentido. Lo que significa es que el mensaje de un gen está repartido en trozos. Hay muchos fragmentos del libro que no significan nada, pero se hace RNA a partir de ahí”, explica Valcárcel.

El equipo de Valcárcel trabaja sobre tres ejes básicos: el reconocimiento del lugar donde se produce el empalme alternativo; los mecanismos que regulan este empalme y los programas celulares de la regulación del empalme alternativo. “Nuestros trabajos demuestran que diferentes mutaciones afectan a los pasos necesarios para que ocurra el empalme alternativo, y esto está relacionado con distintas enfermedades y síndromes, como la atrofia muscular espinal o la fibrosis quística, entre otras”, señala Valcárcel. Por ejemplo, muchas mutaciones en genes se pensaban que no tenían ninguna significación porque no iban a cambiar el mensaje de ese gen, se llamaban mutaciones silenciosas. Resulta que esas mutaciones no son silenciosas en absoluto, sino que pueden producir mensajeros diferentes que alteran la interpretación de la información y generan productos que pueden ocasionar enfermedades.

Otro campo de trabajo es el estudio de genes que pueden producir una proteína que induce al suicidio celular, o lo evita. “Intentamos encontrar los factores que modulan esa actividad. Esto tiene mucha importancia en el funcionamiento del sistema inmunitario. Por ejemplo, cuando los linfocitos T ven un antígeno comienzan a proliferar para defender el organismo contra la infección. Ahora bien, cuando consiguen su propósito, esas células tienen que morir; porque si no comienzan a combatir contra el propio organismo. Y las personas que no son capaces de cambiar el empalme de este gen podrían desarrollar enfermedades autoinmunes. Por eso tratamos de ver qué pedazos se unen con qué pedazos y qué es lo que hacen a partir de las consiguientes transferencias de información”.

¿Orden vs. Caos?

En el fondo, se trata de entender las reglas que determinan los cambios en la actividad del genoma en función de la expresión de los factores reguladores. Ahora bien ¿existen esas reglas o son tan sólo el reflejo de nuestra propensión a establecer un orden en lo que nos rodea que nos permita clasificarlo y explicarlo? La respuesta de Valcárcel hace justicia a su origen gallego: “Existen, pero no se las ve tan claramente como esperábamos”.

En algunos casos, las reglas ahora son evidentes. Como sucede en este título digno de Agatha Christie: “La molécula que produce muerte celular, o no”. En el caso mencionado de la producción de linfocitos T para combatir una infección, si los RNA transmiten la información para que se unan determinados fragmentos del gen, la proteína resultante promueve entonces la muerte celular y se evita un ataque contra el propio sistema inmunitario. Ahora bien, si las instrucciones de los RNA concluyen con la unión de fragmentos diferentes, entonces previenen la muerte celular y la célula prolifera sin control.

“Nuestro objetivo final es tomar una secuencia genómica y poder predecir qué va a hacer. Si se conociera su patrón de lectura, en caso de enfermedad se podría incluso guiar a los reguladores genómicos para que produjeran las proteínas necesarias para devolver la normalidad al organismo. Pero no entendemos aún cómo está escrito el empalme alternativo en el genoma hasta ese punto”, nos explica Valcárcel, quien en 2006 publicó un artículo en la revista científica de la Organización Europea de Biología Molecular (EMBO en sus siglas en inglés), donde se apunta a que el conjunto de RNAs que produce un genoma está en expansión y, por lectura inversa, no está nada claro, o cada vez está menos claro, cuál es el papel que juega lo que antes se denominaba genoma basura, que muchos consideraban que podía llegar a ser hasta el 90% del total. Como sostiene Valcárcel en dicho artículo, la investigación en la regulación genómica está revelando mucha materia oscura que, como en el caso del universo, se sospecha su existencia

y es la que le da sentido al conjunto, pero cuesta mucho encontrarla y desentrañar su composición y funcionamiento.

Los trabajos de Valcárcel y su grupo están contribuyendo a redefinir qué es un gen, aunque a veces sea sólo por exclusión y sin que les horrorice el borrón y cuenta nueva. “El concepto de gen se desdibuja un poco y ya nos podemos despedir del concepto del DNA que produce un RNA que produce una proteína. Ahora sabemos que hay muchos RNA y muchas proteínas. Y que un mismo fragmento de DNA puede contener más de un gen. Hasta los intrones u otros RNA pueden cumplir alguna función a pesar de pertenecer a genes de otras regiones. No obstante, el concepto de gen como entidad física, que puede interpretarse de muchas maneras para cumplir distintas funciones, sigue teniendo mucha validez. Lo que pasa es que el gen es muchísimo más complejo de lo que nos pensábamos hasta ahora”, sostiene el científico del CRG.

Cuando entrevisté a Watson y Crick en Valencia, predominaba la idea de los mapas físicos y funcionales del genoma humano. El esfuerzo se centraba en localizar el gen ya fuera por su posición física en el DNA o por el tipo de proteína que expresaba, o las funciones que desencadenaba. Por eso, cuando se suponía que el número de genes estaba arriba de los 80.000, no era de extrañar la largueza de miras de los científicos cuando se les pedía una fecha para concluir los susodichos mapas. Secuenciar tamaño genoma, con las tecnologías disponibles en aquel entonces, suponía una agenda de varias décadas. Curiosamente, la secuenciación final del genoma humano demuestra que, con muchísimos menos genes de los pensados, se trata sin embargo de una maquinaria mucho más complicada, rica y diversa, a la vez que menos dependiente del lugar físico que ocupan los genes. Su arquitectura laberíntica nos proporciona una comprensión diferente del organismo humano, de su evolución e incluso de los efectos de su interacción de social, una visión mucho más holística e integral de todo el complejo hereditario. Y todo esto tiene que ver con el conocimiento que va surgiendo de la mano de la regulación genómica.

El texto y su contexto

La secuenciación ha añadido una dimensión a la lectura del genoma de consecuencias inesperadas: la globalidad. El disponer de un paisaje integral, donde la tecnología permite contemplar de golpe el estado en que se encuentran todos los genes, los RNA, los microRNA e, incluso, los ncRNA de un genoma, ha transformado la tarea de los científicos desde el clásico “síndrome del pionero”, el avanzado que va relevando un territorio desconocido colina a colina, valle a valle, gen a gen, al “síndrome de la infantería de despliegue rápido”, que con pocos movimientos puede ocupar una vastísima proporción de dicho territorio, sino todo, y repasarlo cuantas veces quiera o necesite. La tarea, sin embargo, es más fácil enunciarla que ejecutarla.

El artículo publicado en la revista científica del EMBL el 2 de marzo de 2006, Juan Valcárcel y Luis Mendes Soares, miembro de su equipo en el CRG, lo titulan: “El transcriptoma en expansión: el genoma como el Libro de Arena”. La referencia a la absorbente obra de Jorge Luis Borges no es sólo un recurso literario, sino una guía para mantener los pies en tierra y orientarse en un territorio multidimensional, repleto de recovecos y de, estirando el símil, universos paralelos. “Imagine que le dan un libro y un diccionario”, ofrece Valcárcel. “Puede ir traduciendo poco a poco, página a página, pero de ahí a entender toda las interacciones entre los personajes y su encaje histórico y cultural, eso no se lo da el diccionario, sino el poseer una visión global del texto y del contexto. Y de eso va la biología hoy”.

Es muy importante ser capaz de leer un libro, pero, como Gutenberg comprendió rápidamente, es mucho más importante ser capaz de leer muchos libros diferentes, o, en el caso que nos ocupa, muchos genomas distintos. Si se lee desde Góngora a Borges, por ejemplo, no sólo se pueden extraer las ideas fundamentales de todo un período de la literatura, sino que además se puede ver lo que se ha conservado, lo que ha ido desapareciendo, la riqueza que subyace en combinaciones nuevas... Esa es la paradoja en la que estamos sumergidos: sabemos



que una parte sustancial de lo que somos está en el libro de los genomas. Pero hay que leerlo muchas veces y de muchas maneras para comenzar a comprenderlo. “Eso es lo que tratamos de hacer desde el CRG”, rubrica Valcárcel

Juan Valcárcel hizo su tesis en el Centro de Biología Molecular en Madrid. Después trabajó cinco años con Michael Green en EEUU, uno de los científicos más sobresalientes en las investigaciones sobre el empalme alternativo desde el Programa en Medicina molecular de la Facultad de Medicina de la Universidad de Massachusetts. En 1996, Valcárcel se trasladó al Laboratorio Europeo de Biología Molecular (EMBL en sus siglas en inglés) en Heidelberg, Alemania, para iniciar su propio grupo de investigación, que se trasladó al CRG de Barcelona en el 2002. Y, seña de identidad del CRG, los investigadores de su equipo proceden de EEUU, Alemania, Suecia, Argentina, Francia, Portugal, España...

De la partitura a los músicos y los instrumentos

Epigenética

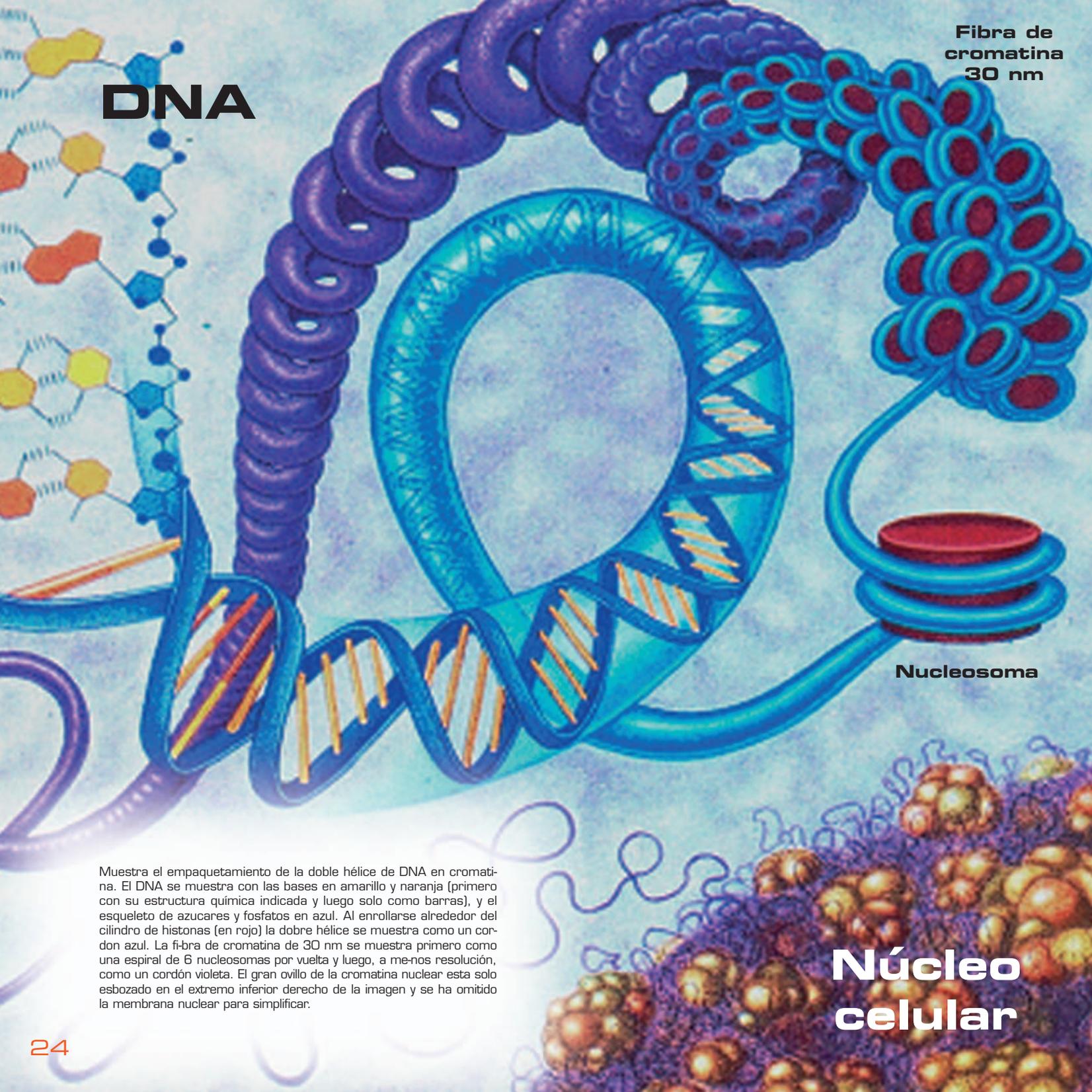
Si la duda sobre lo que en realidad es un gen recorre hoy la biología molecular, no es menor otra

que surge pareja con la anterior: ¿está todo escrito en los genes? Y no nos referimos a la evidente influencia de los factores medioambientales o sociales en las pautas de comportamiento de los seres vivos. Nos referimos a la propia maquinaria hereditaria: ¿basta el DNA y su libro de mensajes para cumplir con la función de crear individuos de la diversidad y heterogeneidad que contemplamos a nuestro alrededor o que poblaron alguna vez la Tierra? La respuesta parecer ser que es no. Cada vez se concretan más los indicios, que se habían venido acumulando sobre todo desde finales de los años 80 y principios de los 90 del siglo pasado, de que en la modulación de la regulación genómica también intervienen factores “externos” que interaccionan con el DNA. Aunque no se sabe mucho al respecto, lo cual obliga a una cierta cautela en su valoración actual por parte de la comunidad científica, lo cierto es que para un sector considerable de la biología molecular, estos factores apuntan a un cambio de paradigma con nombre propio: epigenética, es decir, aquellos sucesos que ocurren fuera del DNA pero que afectan a su funcionamiento y a la manera como los genes cumplen sus funciones.

El DNA está contenido en la célula dentro de un paquete de cromatina [véase imagen]. Y este paquete no es un simple “papel de envolver”. Todo lo contrario, no sólo organiza el empaquetamiento de la doble hélice, sino que ejecuta una serie de decisiones trascendentales, como regular el acceso a diferentes partes del código genético, o permitir el encendido o apagado de ciertos genes, entre otras interacciones. La cromatina juega el papel de director de orquesta que no sólo comprende la partitura que debe interpretarse en cada momento, sino qué músicos y con qué instrumentos hay que interpretarla. De ahí su creciente importancia, porque, cuando funciona bien, funciona estupendamente. Pero si comete errores, sus alteraciones pueden causar numerosas enfermedades, en particular el cáncer.

DNA

Fibra de
cromatina
30 nm

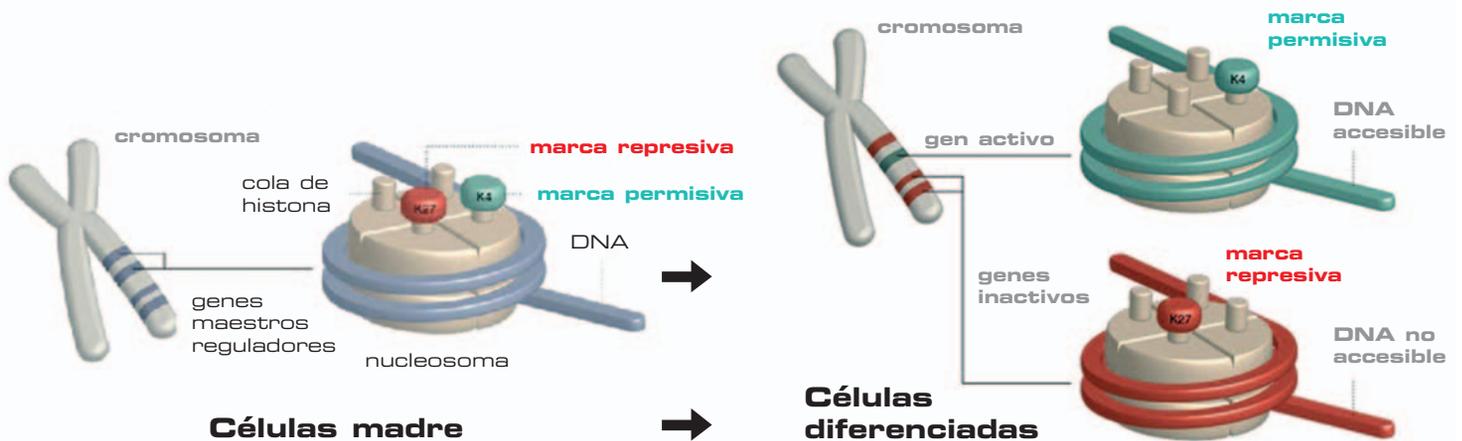
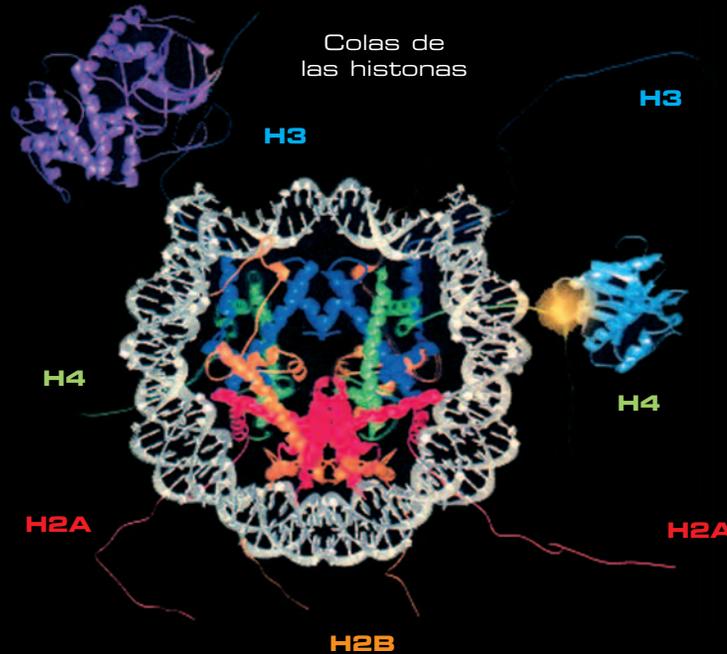


Nucleosoma

Muestra el empaquetamiento de la doble hélice de DNA en cromatina. El DNA se muestra con las bases en amarillo y naranja (primero con su estructura química indicada y luego solo como barras), y el esqueleto de azúcares y fosfatos en azul. Al enrollarse alrededor del cilindro de histonas (en rojo) la doble hélice se muestra como un cordón azul. La fibra de cromatina de 30 nm se muestra primero como una espiral de 6 nucleosomas por vuelta y luego, a menor resolución, como un cordón violeta. El gran ovillo de la cromatina nuclear está solo esbozado en el extremo inferior derecho de la imagen y se ha omitido la membrana nuclear para simplificar.

Núcleo
celular

Muestra un nucleosoma en cierto detalle. La parte globular del octámero de histonas ocupa el interior del círculo formado por las casi dos vueltas de DNA (la doble hélice de DNA se muestra en gris claro) y las colas de las histonas se extiende como antenas fuera del nucleosoma. Cada histona está coloreada de modo distinto para facilitar su reconocimiento. Hay dos copias de cada una de las cuatro histonas, de ahí que se hable de un octámero. El eje de simetría del octámero lo marcan dos histonas H3 (en azul) que interaccionan con dos histonas H4 (en verde). A cada lado del tetrámero de H3/H4, por encima y por debajo del plano de la imagen, hay un dímero H2A (en rojo)/ H2B (en marrón). Una de las colas de H3 es reconocida por una proteína violeta y una de las colas de H4 por una proteína azul.



Las marcas en las colas de las histonas definen la accesibilidad del DNA en nucleosomas. La imagen muestra un cromosoma humano condensado (gris claro) con la posición de algunos genes maestros en gris oscuro (genes que regulan la expresión de otros genes). En las células madre (mitad izquierda de la imagen), las colas de los nucleosomas sobre estos genes maestros contienen marcas represivas (en rojo) y marcas permisivas (en verde). Esto hace que el DNA (en violeta claro) sea parcialmente accesible. Cuando las células se diferencian, la mayoría de los genes maestros que no serán expresados (indicados en rojo) están empaquetados en nucleosomas que sólo tienen marcas represivas y el DNA de estos genes no es accesible. Algunos de los genes maestros, que serán expresados en este tipo especial de células (indicados en verde) están empaquetados en nucleosomas que sólo tienen marcas permisivas y el DNA de estos genes es accesible. Las marcas epigenéticas son distintas en cada tipo de célula, lo que explica que sean capaces de expresar distintos grupos de genes maestros y por tanto distintas partes del genoma.

La epigenética, pues, es la parte de la biología molecular que se ocupa de la implementación del programa genético, tanto durante el desarrollo embrionario y la diferenciación celular, como durante el funcionamiento normal del organismo adulto. Por esta razón, la epigenética y la cromatina se han convertido en áreas prioritarias de la investigación biológica. Que es lo mismo que decir que centros de EEUU, Japón, Inglaterra, países emergentes de Asia y organizaciones científicas de la UE, compiten con proyectos por valor de hasta cientos de millones de euros con el fin de establecer las modificaciones epigenéticas que acompañan a la diferenciación de las células madre y al mantenimiento de la identidad celular.

Esta nueva área de investigación concentra una buena parte de los estudios más avanzados sobre los mecanismos que regulan cómo los genes son expresados y que no son obvios si sólo se estudia la secuencia de nucleótidos del DNA. Todas las células de nuestro organismo tienen el mismo genoma y, sin embargo, cumplen funciones muy diferentes porque expresan distintos fragmentos del texto genético. Esta diferenciación en distintos tipos de células se “cocina” en las modificaciones epigenéticas de distintas regiones del genoma por factores que determinan los patrones espaciales y temporales de expresión génica, los silencios o los hitos bioquímicos que señalizan la ejecución y progre-



Miguel Beato, director del CRG y experto en epigenética.

Cómo se conforma el empaquetamiento del DNA en la cromatina: En un primer nivel, o nucleosoma, el material genético, la doble hélice del DNA, se compacta unas 7 veces enrollándose con casi dos vueltas alrededor de un cilindro de histonas. A este nivel los nucleosomas forman una estructura parecida a la de un “collar de cuentas”. En un segundo nivel, los nucleosomas se compactan formando una fibra de 30 nanómetros (1 nanómetro = 1 milmillonésima de metro) donde los nucleosomas están empaquetados los unos sobre los otros en disposiciones regulares gracias a interacciones entre las histonas. Es como si se tratara de un parking compacto donde todo el espacio está aprovechado al máximo. En el tercer nivel, el DNA alcanza una compactación todavía mayor formando lazadas más o menos densas que se hace visible en los cromosomas durante la metafase de la división celular. Esta estructura de la cromatina no es estable y fija sino sumamente dinámica y cambiante en respuesta a los programas que ya porta la célula y también a señales provenientes de otras células o del mundo exterior. La modificación de la estructura y la compactación de la cromatina se lleva a cabo mediante marcas químicas introducidas en las colas de las histonas que, como antenas, sobresalen en la superficie de los nucleosomas. Distintas modificaciones de las colas de las histonas (principalmente acetilación, metilación) forman una especie de código que determina qué regiones del genoma van a ser accesibles para la maquinaria que lee los genes y, con ello, la identidad de cada célula. Además estas modificaciones de la cromatina pueden extenderse al mismo DNA (metilación) y se heredan durante la división celular garantizando que, por ejemplo, cuando una célula del hígado se divide genere dos células del hígado.

“Ahora tenemos la partitura genómica ante nuestros ojos, pero sólo entendemos -con limitaciones- una parte mínima de su contenido, la escrita en una tonalidad familiar, los genes clásicos. El resto sigue siendo un misterio. Conocemos la música final, somos nosotros, pero no entendemos como están escritas las notas en la infinidad de líneas entrelazadas, anotaciones paralelas y variaciones que contiene el genoma.”

[Miguel Beato]

sión de la partitura musical en que se basa el funcionamiento de todo el conglomerado genético.

En un arreglo orquestal de la complejidad del genoma, resulta fundamental que “alguien” controle las pausas, los momentos de activación o desactivación de los genes, los procesos que van a desencadenar en función de la localización espacial de cada uno de ellos, la modulación de las respuestas. Esto supone disponer de una información “global” que, casi por definición, no puede estar contenida en cada gen, ni en el DNA. Aquí es donde intervienen los factores epigenéticos, en particular, la última estrella emergente, la cromatina.

La cromatina, o material coloreado, recibe su nombre precisamente de la facilidad con que cambia de color cuando se la tiñe con ciertos colorantes usados por los histólogos, los especialistas en tejidos orgánicos. Pero a este rasgo de su forma, se une otro capital, de corte bioquímico, al organizar la estructura que el DNA adopta en el núcleo de las células. De hecho, la explicación de la evolución de organismos complejos con grandes genomas, como el de los seres humanos, se basa en este “matrimonio” por ahora insoluble: por un lado, la aparición de la membrana nuclear, con la separación de la lectura del mensaje genético en dos compartimentos, la síntesis del RNA en el núcleo y la síntesis de proteínas en el citoplasma, y, por el otro, la organización del DNA en cromatina.

La tarea de la cromatina es un milagro y la forma de ejecutarla un dechado de elegancia. Consiste en empaquetar la molécula del DNA, que en nuestras células mide dos metros, para acomodarla en una esfera de unas siete micras de diámetro (1 micra = 1 millonésima de metro), el núcleo de la célula. Y ello sin que, por una parte, pierda ninguna de sus funciones o propiedades y, por la otra, la información relevante del DNA continúe siendo accesible para el buen funcionamiento de la célula. La cromatina ejecuta este verdadero acto circense en tres instancias sirviéndose de las histonas, unas proteínas de bajo peso molecular, que los seres eucariotas, como los humanos, hemos logrado conservar muy bien a lo largo de la evolución (véase recuadro). Se

podrían utilizar muchas metáforas para explicar este verdadero abracadabra, pero estamos hablando de algo así como meter un millón de euros en billetes de 5, 10, 20, 40, 100, 200 euros en una cajita de cerillas, de la que podemos sacar las cantidades que necesitamos en cada ocasión con dos dedos y sin apenas abrirla. Ahí queda eso.

En menos de un lustro, la cromatina se ha convertido en una especie de primera bailarina del ballet genético. Su intervención decisiva en la regulación de los genes que empaqueta ha capturado la atención de los científicos y, de paso, ha pavimentado el camino de la investigación epigenética en los centros más avanzados del mundo que trabajan sobre la regulación genómica. En el caso del CRG, esta tarea le corresponde al “Grupo de la Cromatina y la Expresión Génica”, liderado por Miguel Beato, también director del centro.

Miguel Beato comenzó a prestarle atención a las funciones de regulación genómica de la cromatina apenas se publicaron los primeros trabajos que apuntaban en tal sentido en 1988. Pero, en aquel entonces, la información sobre el genoma no había alcanzado la masa crítica necesaria como para obtener una visión avanzada de lo que estaba sucediendo en esta envoltura del DNA, ni mucho menos sospechar que los cambios en la cromatina podían afectar al conjunto de los genes. De todas maneras, el término epigenética ya existía para referirse a fenómenos que ocurrían en el núcleo de la célula, los cuales habían sido descritos tan temprano como 1942 por Conrad Waddington (1905-1975). De hecho, como en casi todas las cosas relacionadas con la ciencia, hay quien encontró los antecedentes en Aristóteles y su concepto de la epigénesis o el desarrollo del individuo a partir de materia amorfa.

El punto de partida, sin embargo, podemos situarlo cuando la secuenciación del genoma mostró páginas de la partitura hereditaria que no se correspondían con la melodía que sonaba hasta entonces en los centros de investigación. Por una parte, parecía que el código genético, como todos los buenos códigos dignos del nombre, no contenía toda la explica-

ción de lo que hacía. Por la otra, los genes que daban lugar a proteínas, y a los que se concedía un lugar preponderante en esta orquesta, constituían una pequeña parte de nuestro genoma (menos del 3%). La indagación comenzó a sacar a la luz un segundo código, una especie de “generador electrónico de música” capaz de interpretar partituras y conformar orquestas. Ese segundo código está encriptado en la cromatina. La investigación de la estructura y la dinámica de la cromatina, que directa o indirectamente ha merecido hasta ahora cuatro premios Nobel, ha mostrado, además, algo fundamental: sus alteraciones suceden por oscilaciones en los procesos de metilación y acetilación del DNA y las proteínas. Cuando éstas se perturban, la cromatina, o bien no acierta a silenciar genes que determinan la identidad celular o promueven la proliferación celular, o, inversamente, reduce la actividad de los genes relacionados con la supresión de tumores o con la reparación del DNA, en ambos casos dando lugar al cáncer.

La cromatina también decide cuándo una célula madre abandona su nicho estacionario y se pone a trabajar, es decir, comienza a diferenciarse. “Las células madre son como un menú abierto en cromatina, tienen marcas epigenéticas permisivas y represivas en los genes que determinan la identidad celular, por eso son pluripotentes. Cuando se diferencian, silencian desde la cromatina –o epigenéticamente– todos los genes que no van a necesitar y mantienen activos sólo aquellos que van a utilizar. Se trata de identificar las señales que hacen que las células madre se diferencien y de entender cómo actúan para fabricar una extremidad, o glóbulos rojos, mientras que, al mismo tiempo, permanecen como células madre”, explica Miguel Beato. [Véase recuadro].

Los trabajos del grupo de Miguel Beato alimentan –y se alimentan también de– las otras líneas de investigación que se desarrollan en el CRG, en particular de los grupos que trabajan sobre el cáncer u otras enfermedades asociadas al funcionamiento de los genes. Todos, en el fondo, buscan esas señales que silencian o activan partes del genoma en una célula, ya sea cuando la melodía suena bien, o, sobre todo, cuando el coro de genes empieza a desafinar. En estos casos, para reparar

estas disfunciones y regresar al “tono sano”, las terapias que están surgiendo tocan la cromatina, como sucede, por ejemplo, con los inhibidores de las acetilasas de histonas. “Todavía nos falta una visión integrada global, que nos permita comprender cómo funciona este control. Los cambios en la cromatina afectan a grupos de genes. Ahí se juega la funcionalidad y la identidad de la célula. La necesidad de una visión global nos obliga a regresar a las matemáticas y la estadística, porque hay que manejar vastísimas cantidades de datos que nos den una idea de lo que está sucediendo”, señala Beato.

Si se le pregunta a Beato cómo llega uno a la cromatina cuando, a lo mejor, se ha pasado la vida investigando otras cosas, su respuesta procede de una especialidad donde descifrar es el principio de todas las cosas: “Yo me he dedicado a la pintura durante una larga época, y me parece que hacer investigación es como escuchar un cuadro, ver qué te dice, que falta, qué sobra. Estudiabas un sistema de regulación génica y sentías que faltaba algo. En nuestro caso era la cromatina. En 1988 estábamos estancados con nuestro sistema y no podíamos interpretar nuestros resultados considerando que el DNA estaba “desnudo”, porque no conocíamos bien su empaquetamiento. Entonces, unos investigadores de los Institutos Nacionales de la Salud de EEUU apuntaron a que la organización del DNA en cromatina no era aleatoria, sino que adoptaba posiciones muy precisas, las cuales podían ser muy importantes para entender cómo se lleva a cabo la lectura de los genes. El problema es que resulta muy fácil diseñar experimentos incorporando a la cromatina, a los nucleosomas, pero es extremadamente complicado hacerlos. Aun así, hemos avanzado en la mejor comprensión de cómo la célula utiliza la estructura de la cromatina para afinar la regulación de sus genes”.

El grupo de Beato trata de incorporar las lecciones de esta experiencia con la cromatina en el estudio del cáncer de mama. “Actualmente tenemos 80 tumores analizados, lo cual ya nos permite intuir criterios útiles para elaborar el diagnóstico genético”, afirma el científico.

Las redes: nadie sabe para quien trabaja

La Era de la bioinformática

La secuenciación del genoma, y prácticamente toda la ciencia dedicada a desentrañar la maqui-

naria hereditaria de los seres vivos, es deudora de la tecnología. Sobre todo de las tecnologías de la información: los ordenadores, las aplicaciones informáticas, la supercomputación y, finalmente y por encima de todo, la red de redes, Internet. Y viceversa. Sería un ejercicio inútil intentar dilucidar quien fue primero en esta ocasión, si el huevo o la gallina, si el genoma o la tecnología de la información. Sus caminos se entrecruzan de tal manera que, unas veces la ciencia y otras la tecnología, actuaron como promotores de desarrollos y avances que después se esparcieron como manchas por la comunidad científica y el resto del tejido social.

Sin embargo, hemos llegado a un punto donde la discusión ya no es posible: sin las tecnologías de la información, sobre todo sin la creciente extensión, densidad y porosidad de Internet, el proyecto de desciframiento y secuenciación del genoma (de los seres vivos de hoy y de hace cientos de millones de años), así como el desarrollo de aplicaciones, productos de simulación y terapias basadas en el apoyo prestado por estas tecnologías, simplemente sería una tarea imposible. “Hemos llegado a un punto donde las plataformas de genotipación para detectar SNPs¹, los chips o microarrays que permiten analizar un genoma completo de golpe, las herramientas informáticas para comparar millones de bases de nucleótidos y detectar dónde están las diferencias, etc., son imprescindibles para estudiar el genoma”, nos cuenta Roderic Guigó, responsable del “Grupo de Bioinformática y Genómica” del CRG, quien, junto con Josep F. Abril, participó en la aventura de Celera para secuenciar el genoma humano.

¹ Hay más de 10 millones de polimorfismos de nucleótido único (single nucleotide polymorphisms, o SNPs, en inglés) en el genoma humano. Los SNPs son sitios específicos dentro del genoma en los cuales algunas personas tienen un nucleótido presente mientras que otras tienen uno diferente. Los SNPs empiezan su existencia como cambios de secuencia y eventualmente se establecen en una población. Esta sustitución debe ocurrir en una proporción significativa (más del 1 por ciento) de una población grande para ser considerado un SNP. Cuando los SNPs ocurren dentro de un gen, la proteína puede tener una función distinta.

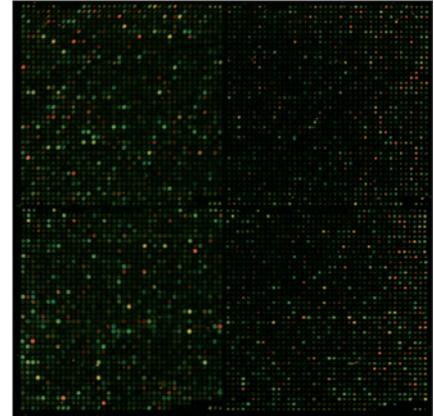
En apenas 10 años, vastas parcelas de la actividad científica han pasado a depender de la automatización de los experimentos, de la producción, recogida y análisis computacional de datos, de la cuantificación de procesos de todo tipo. En biología sucede además un fenómeno particular. Los datos de la secuencia del DNA son de tipo computacional,

casi una metáfora de un programa informático, un conjunto de instrucciones que no es binario sino cuaternario: adenina (A), timina (T), guanina (G) y citosina (C), las cuatro bases que componen la doble hélice. “Los mecanismos fundamentales de procesamiento de estas instrucciones son computacionales. Por ejemplo, en el caso de la traducción de RNA mensajeros a proteínas, se trata de convertir una secuencia en el alfabeto cuaternario en una secuencia en un alfabeto de 20 símbolos, los cuales corresponden a los aminoácidos con los que se construyen las proteínas. Las instrucciones para la conversión entre secuencias (mensajes) están especificadas en el llamado código genético”.

El grupo de Bioinformática colabora con los otros grupos del CRG para analizar los datos de sus experimentos sobre la cromatina, el empalme alternativo o la detección de oncogenes. Pero, al mismo tiempo, aprovecha su destreza en el uso de los ordenadores y la gestión del creciente volumen de información sobre el genoma, para abrir campos propios de investigación y experimentación. Además, en biología molecular la computación no sólo es una de las formas de abordar ciertos problemas, a veces es la única. No se trata de una subordinación de la ciencia a la tecnología, sino una relación dialéctica, dos formas de investigar inextricablemente relacionadas.

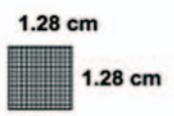
En concreto, con el equipo de Miguel Beato, Guigó está estudiando cómo se forma el paquete de la cromatina y cómo varía la regulación de los genes según las regiones del organismo donde actúa. ¿Cómo es capaz de reconocer sólo las regiones donde se transcriben a RNA? En el caso del grupo de Juan Valcárcel, los bioinformáticos están tratando de entender cómo se produce el empalme alternativo, por qué la célula es capaz de reconocer ciertas fronteras, cortar el “hilo” allí, empalmar y fabricar una molécula mRNA con instrucciones para fabricar una proteína diferente desde el mismo gen.

Con Raúl Méndez, un jefe de grupo junior del CRG, se estudia cómo dependiendo de unas señales unas veces se produce la proteína y

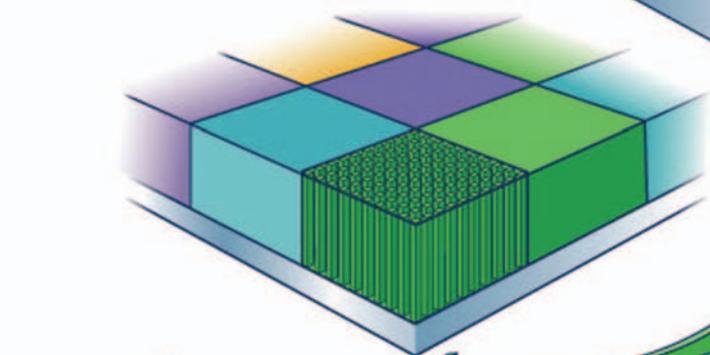
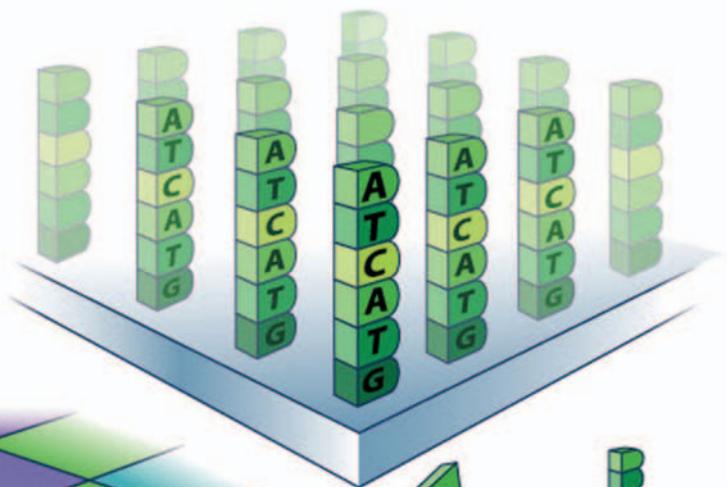


Uno de los cambios tecnológicos más radicales en la secuenciación del genoma lo ha aportado el microarray—también conocido como chip del genoma o biochip—, que permite observar todo un genoma simultáneamente y comprobar el estado en que se encuentran los genes, los RNA o los otros productos e interacciones generados por la actividad del DNA. El microarray es una colección de oligonucleótidos de DNA complementario dispuestos en hileras fijadas. En estos chips se pueden ver de golpe qué genes están encendidos, apagados o funcionando “a medio gas”.

Ejemplo de un tipo de microarray



Tamaño real de GeneChip



500,000 celdas en cada GeneChip array



Millones de cadenas de DNA por celda

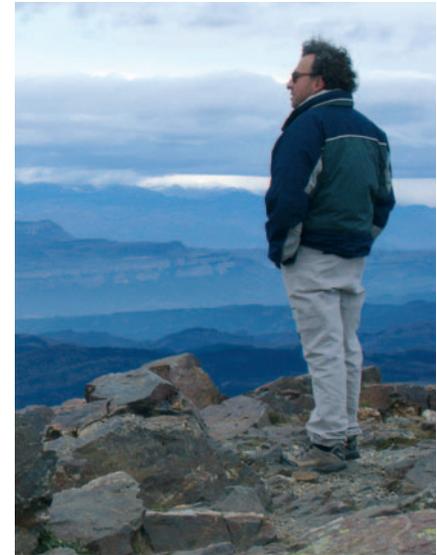


Una cadena de 25 pares de bases

otras se degrada. Guigó toma un marcador y dibuja en la pizarra: “A lo mejor cuando estamos hablando de una célula de los ojos, quizás sólo funciona este gen, pero no estos otros dos. Pero en una célula del estómago funcionan estos dos, no el tercero. Y al hablar del espermatozoide, funcionan los tres al mismo tiempo. O en el cerebro, ninguno de ellos. Esta combinatoria es la que produce la diversidad y heterogeneidad de nuestro organismo y de lo que somos como humanos. Esto quiere decir que no todos los genes acaban en proteínas en todos los tejidos. Y por eso hay individuos distintos. Puede ser que una persona que tiene cáncer es porque éste gen, por ejemplo, por razones que todavía desconocemos, nunca se expresa”. Hasta ahora, los esfuerzos se habían concentrado en las perlas, los genes, pero, como dice Guigó con una metáfora expresiva, “a lo mejor el hilo –los exones y los intrones- es de oro”. Lo que tiende un puente hacia las investigaciones de Juan Valcárcel.

Para sacar conclusiones significativas es fundamental saber moverse entre los datos, extraerles el jugo mediante comparaciones. Por eso, el equipo de Guigó no trabaja con una especie en particular, o con un gen o un DNA. Ahora está analizando con Valcárcel el paramecio, cuyo genoma muestra algunas peculiaridades que no ocurren en humanos. Y aprovecha también los trabajos que se hacen en el CRG con otros modelos de genoma, como la levadura, la mosca del vinagre, el pez cebra, el ratón o el ser humano. El análisis comparativo es fundamental porque permite entender cuáles son los cambios relevantes que se registran en los genomas y, sobre todo, en la regulación genómica.

“El desarrollo actual de las tecnologías, sobre todo de los chips de DNA, apunta a que, en un plazo aproximado de entre 10 años y 15 años, los hospitales serán capaces de secuenciar los genomas de los pacientes y detectarán rápidamente muchas anomalías génicas.” Disponer del genoma de todos los individuos planteará problemas de privacidad muy serios, como reconoce Guigó. Ahora se trabaja con un genoma anónimo, una quimera de individuos distintos, de tejidos o líneas unicelulares diferentes. “En un chip como éste, parecido al de las



Roderic Guigó, responsable de Bioinformática en el CRG.

“El desarrollo actual de las tecnologías, sobre todo de los chips de DNA, apunta a que, en un plazo aproximado de entre 10 años y 15 años, los hospitales serán capaces de secuenciar los genomas de los pacientes y detectarán rápidamente muchas anomalías génicas.”

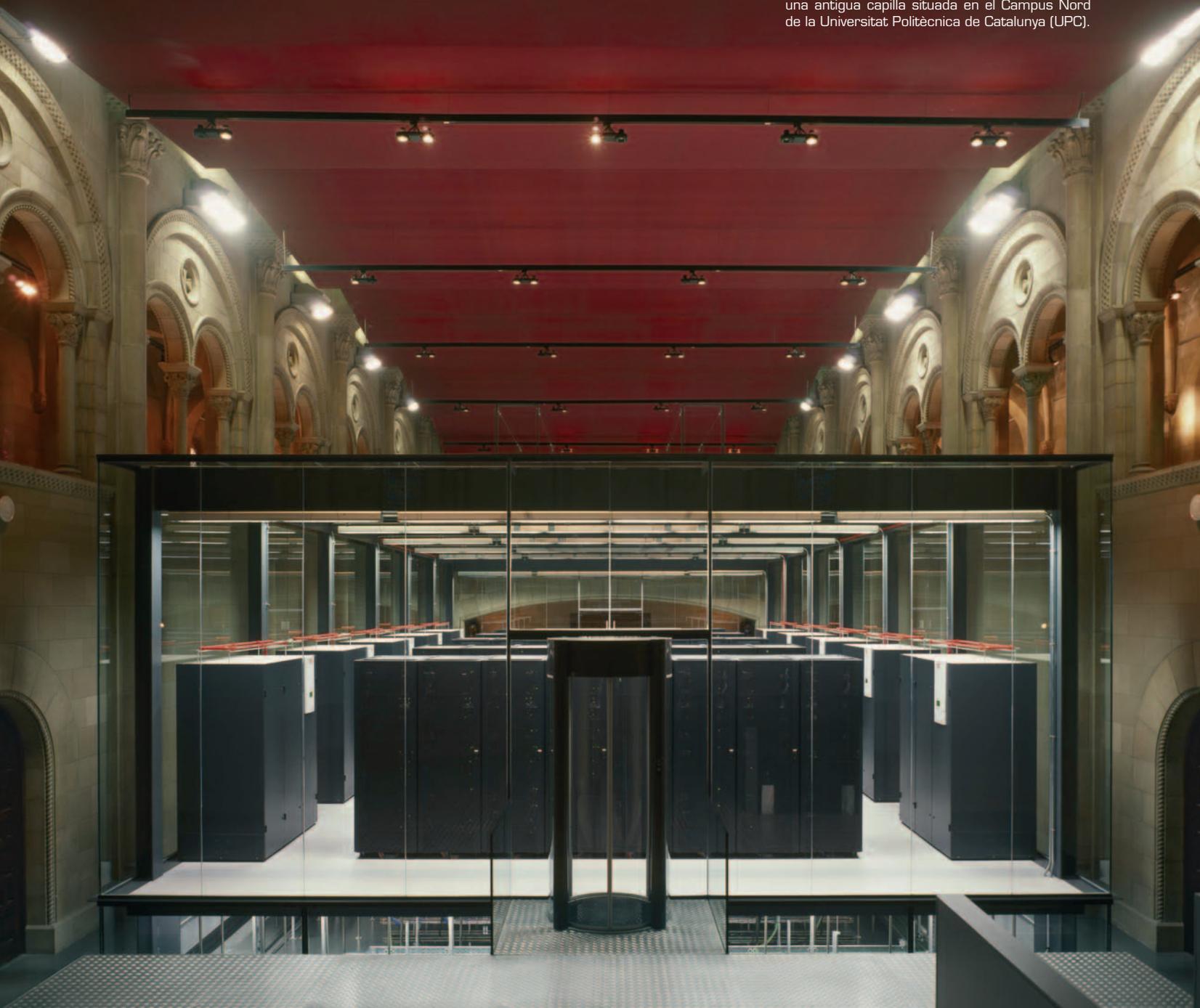
[Roderic Guigó]

tarjetas de crédito, hay un 1% de todo el genoma humano. Si lo comparo con el mío, puedo ver cómo está el mío en esas regiones, qué genes están activos. En cuatro años todo el genoma estará en un chip así o más pequeño”.

Las técnicas de hibridización, que apenas aparecieron a principios de los años 90 del siglo XX, han propinado un empujón fenomenal a los trabajos del genoma. “Hace cinco años”, señala Guigó, “la obtención de cada uno de estos puntos podía representar la tesis doctoral de una persona, casi cuatro años de trabajo. Ahora lo haces en una tarde”. Cada uno de esos puntos es lo que se ve en el chip –un microarray- al pasarlo por un lector: una página con luces de colores ordenadas como semáforos en fila. “Aquí vemos los mensajeros, según su color sabemos si están encendidos, apagados o funcionando sólo a medio gas. En este caso, todos los mensajeros de un 1% del genoma. Pero podemos hacer microarrays de todos los mensajeros, o de todo un genoma, o de tejido de cáncer de mama, y ves genes que se encienden en un caso y en otros no”.

El grupo de Guigó, uno de los pocos bioinformáticos en España que contribuyó al matrimonio entre el genoma e Internet durante los años 90, es una pieza clave del CRG, si es que cabe hacer una distinción de este tipo en una minigalaxia de científicos escogidos únicamente por la excelencia de sus trabajos y sus destacadas aportaciones a la biología molecular. Por una parte, ordena el paisaje de las investigaciones de los otros grupos del centro y lo integra en el marco de un análisis global que permite puntualizar procesos, anomalías, mutaciones oncogénicas, tendencias y líneas de investigación, todo ello con un alto valor predictivo. Por la otra, elabora su propia agenda de trabajo y convierte a la computación en el laboratorio virtual donde desarrolla sus investigaciones. Hacerlo desde Barcelona, además, supone algunas ventajas notables, como es la posibilidad de utilizar la capacidad de cálculo de Mare Nostrum, uno de los ordenadores más potentes del mundo instalado en la Universitat Politècnica de Catalunya.

El Mare Nostrum es uno de los ordenadores más potentes del mundo. Está emplazado en una antigua capilla situada en el Campus Nord de la Universitat Politècnica de Catalunya (UPC).



Refabricamos

Biología de sistemas

En los últimos años, algunos centros de investigación de biología molecular y empresas de biotecnología han abierto un despacho nuevo. El rótulo de la puerta dice: Biología de sistemas. La tendencia es tan reciente, que a veces no se sabe con seguridad cuál es el perfil idóneo para este nuevo puesto. No se estudia en la universidad, la literatura que lo fundamenta es escasa y dispersa (aunque considerable cuando se la logra armar como un rompecabezas siempre sorprendente) y el encaje con los otros departamentos, culturalmente asentados en la mejor tradición de la investigación en biología molecular, no resulta fácil. Pero la biología de sistemas ha venido para quedarse. No sólo eso: a pesar de su juventud, reclama un puesto de avanzada, como se puede observar en los mejores y más prestigiosos centros de biología del mundo, donde hacen girar sobre esta nueva área la toma de decisiones estratégicas relacionadas con lo que ya muchos denominan “nueva biología”.

Y bien, ¿en qué consiste esta nueva disciplina? Aquí surge la primera sorpresa: “En realidad, más que una disciplina nueva es una forma de ver las cosas”. Así se explica Luis Serrano, uno de los científicos europeos más prestigiosos en el campo de la biología de sistemas, que dirige un equipo en el EMBL en Alemania y que ahora hace lo propio desde el CRG. Serrano es el puente entre el instituto de Barcelona y el EMBL para coordinar trabajos e intercambiar científicos e informaciones entre las dos instituciones. Por tanto, esta “forma de ver las cosas” es la que, en el fondo, le está dando forma y proyección a las cosas que se hacen.

¿Cómo? “Hasta ahora, los grupos de investigación trabajaban (bueno, y trabajan) en aspectos muy concretos: un gen, una proteína, una célula. Lo que hacemos desde la biología de sistemas es tratar de adquirir una visión global, en vez de centrarnos en el funcionamiento de una célula pancreática, tratamos de ver cómo funciona esa célula en el páncreas dentro del cuerpo humano”, explica Serrano.

¿Por qué ahora tiene tanto predicamento la biología de sistemas, por qué no se buscaba antes esta visión integral, holística, del genoma, de

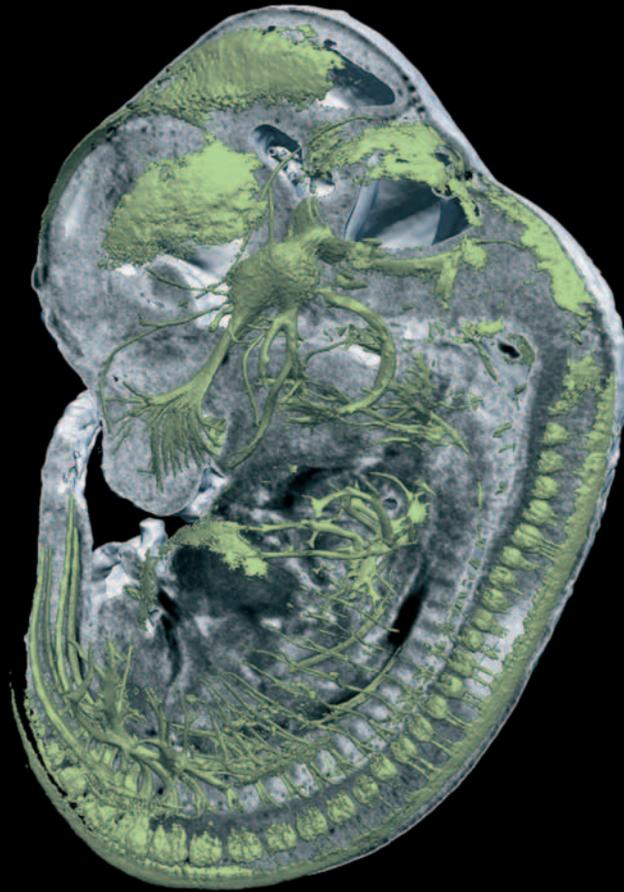
las relaciones entre el todo y sus partes? La respuesta la hemos ido explicando a lo largo de todo este texto de distintas maneras: densidad de información. En un abrir y cerrar de ojos se ha pasado de la escasez de información, o demasiado concentrada como referencia a problemas muy locales, a disponer de una masa abundante de datos que proceden desde todos los lugares y zonas de interés. Esta avalancha de información, que plantea incluso colosales retos desde el punto de vista de su gestión operativa, es la costurera que ha hilvanado las distintas áreas del conocimiento que iban despuntando a medida que cuajaba el proyecto del genoma, como la proteómica o la genómica. “No sólo tenemos una gran cantidad de datos, sino que son datos globales. Podemos ver en un chip cuántos genes se están expresando en una célula humana o qué proteínas se están fabricando. La bioinformática nos permite trabajar con una extraordinaria precisión sobre lo que está sucediendo en el organismo. Esa información hay que digerirla y entenderla de forma cuantitativa para que nos permita hacer predicciones, que es el objetivo esencial de la biología de sistemas”, explica Luis Serrano.

O sea, que sobre la capa de la biocomputación, ahora se extiende la de la biología de sistemas con tres patas de apoyo: 1) los datos, todo lo que se está cosechando hoy día en laboratorios de todo el mundo y que nutre gigantescos bancos de datos, 2) los modelos matemáticos de análisis y simulación y 3) la cuantificación. Con estos tres apoyos se puede dar el salto de verificar un hecho puntual (qué pasa si cambio este gen o esta proteína) a entender cómo este cambio se integra dentro del funcionamiento de un organismo y ser capaces de predecir el efecto de otros cambios. La biología de sistemas es la denominación que asume la globalización en el campo de la biología molecular.

Si la bioinformática está pidiendo a gritos nuevos perfiles profesionales que sepan aunar biología, computación y gestión de la información, la biología de sistemas clama por la conformación de equipos integrados por especialistas en diversas disciplinas capaces de analizar y modelizar el genoma desde una visión global. Equipos donde convivan biólogos, físicos, químicos, informáticos, matemáticos... Visto desde la experiencia y los

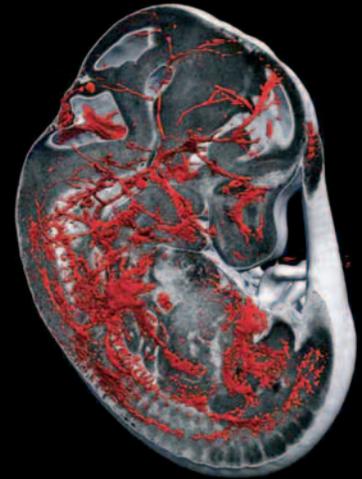


La nueva unidad de Biología de Sistemas del CRG está dirigida por Luis Serrano.



Izquierda: "La nueva tecnología de imaging denominada OPT (siglas en inglés de Tomografía Óptica de Proyección, desarrollada por el Dr. James Sharpe, del CRG) nos permite visualizar la actividad de las proteínas en 3D en el contexto de un embrión de ratón completo. En este caso, la distribución de la proteína del neurofilamento destaca el sistema nervioso en desarrollo.

Abajo: "Una imagen OPT del embrión de un ratón destaca el sistema sanguíneo en desarrollo en 3D".



hábitos de la comunidad científica en España, suena como si se estuviera incubando otro de esos choques culturales que tanto han caracterizado nuestra historia. Serrano lo expresa con claridad: “Es cierto que hay un problema psicológico de entrada que se debe superar. Todo estriba en una cuestión de enfoque: debo dejar de pensar en mi proteína, en mi gen, en mi mutante, y comenzar a ver lo que hago de manera global”.

En otras palabras, si se trabaja sobre el desarrollo del ojo en el pez, se necesitará disponer de información sobre la formación del ojo en todo tipo de seres vivos, qué genes están comprometidos, qué proteínas, qué mutantes se han producido en otros organismos. El mensaje de inicio es: “No se puede seguir haciendo biología concentrado cada uno en su pequeña parcela”. El siguiente paso es disponer de modelos y simulaciones significativas de las piezas que van conformando el gran rompecabezas de los seres vivos. Y pasar incluso de la simulación al producto, es decir, a la posibilidad de reproducir mecanismos biológicos completos que permitan reparar anomalías, lo que el resto de los mortales llamamos enfermedades. “El ordenador puede representar una célula entera o más. El desafío sería poder fabricarla”, asegura Serrano.

El científico ha dejado su plaza de director de programa en el EMBL, donde inició los primeros trabajos en este campo hace 6 años, para dirigir el grupo de biología de sistemas en el CRG, aunque la conexión con el laboratorio europeo ha quedado más reforzada que antes. De hecho, todos los miembros de su equipo son europeos, algunos con experiencia de trabajo en el EMBL. El CRG tendrá finalmente cinco grupos de biología de sistemas, integrado cada uno por unas 8-10 personas, que desarrollarán software para simular redes génicas complejas, prepararán experimentos para ver cómo funcionan ciertos procesos biológicos, diseñarán genes y proteínas para estudiar cómo interaccionan y qué funciones producen y estudiarán cómo los fármacos interrumpen o abren nuevas rutas con el fin de desarrollar nuevas terapias más eficaces e individualizadas.

Con estas cartas en la mano, Serrano no tiene dudas: “Vamos hacia una medicina personalizada”. Sobre esto, más algo más adelante

“Hasta ahora, los biólogos trabajaban en un gen, una proteína, una célula. La biología de sistemas trata de adquirir una visión global de todos los genes. En vez de centrarnos, por ejemplo, en el funcionamiento de una célula pancreática, tratamos de ver cómo funciona esa célula en el páncreas dentro del cuerpo humano”.

[Luis Serrano]

El bucle

Terapias humanas

A pesar de su juventud, el CRG cubre los flancos más importantes y las entretelas más intrigantes y preocupantes de la actividad del genoma: el amplio abanico cubre desde descifrar códigos locales de los genes y los mRNA, así como aspectos particulares de las secuencias, de su funcionamiento y de los factores que regulan los genes que expresan proteínas, hasta la reconstrucción del argumento de un texto troceado, inconcluso, mediante el análisis y la comparación computacional, la experimentación en el laboratorio y la aplicación de una estrategia global de análisis.

El paso siguiente es relacionar este conocimiento con posibles patologías y proponer estrategias que permitan reparar daños o desarrollar terapias para recomponer la actividad anómala del genoma. No se trata de cerrar el círculo, sino de elevar el bucle para retroalimentar todo el proceso de generación de conocimiento. Porque la consecuencia directa de todo el trabajo de los diferentes grupos científicos se expresa sobre todo en el momento en que el proceso se altera: cuando aparecen mutaciones, la célula se comporta erráticamente y se desencadenan las enfermedades asociadas a errores genéticos.

El equipo de Luciano Di Croce, por ejemplo, que lidera el “Grupo de eventos de epigenética en el cáncer”, aprovecha el conocimiento que se está adquiriendo sobre cómo la organización de la cromatina modula la transcripción de los genes y cómo estos mecanismos epigenéticos pueden jugar un importante papel en la iniciación y la progresión del cáncer. Di Croce, tras tres años en el Instituto Europeo de Oncología de Milán, se trasladó al CRG atraído tanto por la oferta científica, como por la posibilidad de trabajar en Barcelona (y al lado del mar, aunque no lo dice en voz alta, pero se le nota). Actualmente trabaja en los errores en la regulación genómica que causan la leucemia promielocítica aguda, LPA (o APL en sus iniciales inglesa), que afecta a un 15% de los adultos con leucemia mieloide aguda.

Esta leucemia ocurre por una traslocación de cromosomas, que se descubrió por primera vez en 1992. Dos cromosomas separados, el 15 y el 17, se unen debido a un error genético. En realidad se parten en dos y una parte del 15 se fusiona con el 17 y lo mismo sucede con las otras dos partes. O sea, un gen que está en el punto de fractura se une con otro gen del otro punto de fractura. Eso quiere decir que se forman dos genes nuevos, que son recíprocos el uno del otro. Los dos genes de cada uno de los extremos, al unirse expresan una nueva proteína que no existía, ni nadie la esperaba, la PML-RAR (o proteína de leucemia promielocítica). Por si fuera poco, dicha proteína se comporta como un factor de regulación genómica: se pega al DNA y se convierte en un potente represor génico.

Al expresar esta proteína en la célula, ésta se vuelve leucémica. ¿Qué quiere decir? Que no puede diferenciarse, no alcanza la madurez y no es funcional. Es la que se encuentra en la sangre del paciente leucémico, incapaz de cumplir con sus funciones. Además, como no ha logrado diferenciarse, su mecanismo de autorreproducción se dispara y prolifera descontroladamente. En síntesis, el paciente leucémico tiene muchas células en la sangre que no sirven para nada.

“Estamos investigando cómo actúa esta proteína, que no es hereditaria. Si lo averiguáramos, podríamos revertir lo que hace. Nuestro objetivo es la proteína, no el gen nuevo, porque eso no habrá forma de arreglarlo por ahora. Hemos descubierto que la proteína añade un grupo metil al DNA. Si le damos al paciente un compuesto que elimina ese grupo metilo, la célula vuelve a funcionar hasta la diferenciación completa. Es un ejemplo de terapia basada en el estudio de la regulación genómica. Ahora estamos en los ensayos clínicos con este compuesto químico –la decitabina- que neutraliza el metilo”, explica Di Croce.

Este fue el primer ejemplo donde se demostró que una lesión genética se puede convertir en un daño epigenético, gracias a lo cual ahora se sabe que si se liquida la fuente del error, el gen traslocado, no se soluciona nada, porque la célula ya está “tocada” y sigue reproduciendo



El científico italiano Luciano Di Croce de paseo con su hija Lara.

“Estamos investigando cómo actúa esta proteína que vuelve leucémica a la célula. Hemos descubierto que añade un grupo metilo al DNA. Si le damos al paciente un compuesto que elimina ese grupo metilo, la célula vuelve a funcionar normalmente. Es un ejemplo de terapia basada en el estudio de la regulación genómica”.

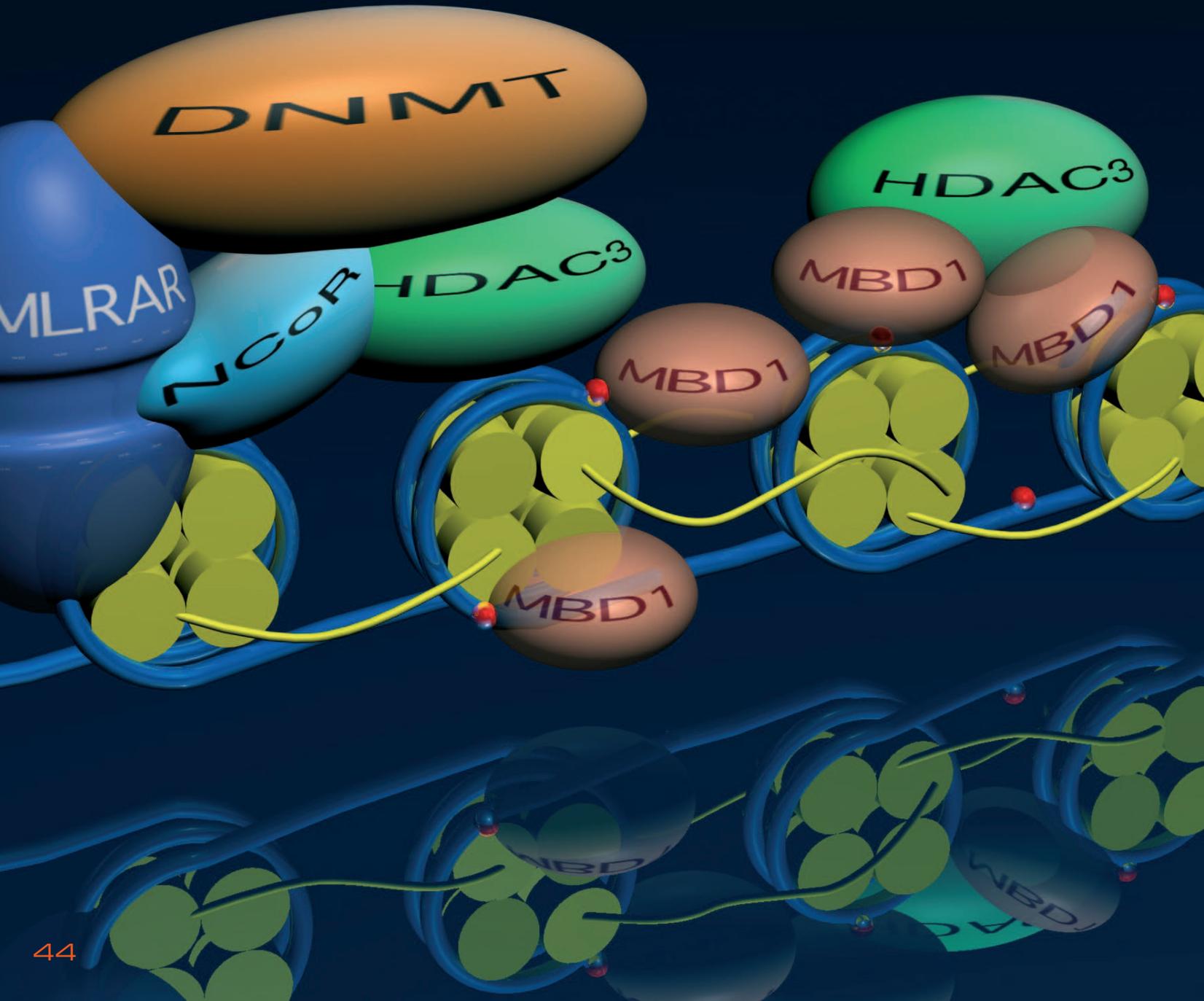
[Luciano Di Croce]

células inmaduras. Por tanto, hay que ir a por ella cortando de raíz lo “que la vuelve loca”, en palabras de Di Croce. El estudio del científico italiano se ha convertido en un modelo para entender mejor la epigenética. Y también para comprender otros tipos de leucemia que funcionan a partir de una proteína que no debería estar y que se sabe qué genes apaga o suprime. Estas traslocaciones, asegura Di Croce, son bastante frecuentes y se deben a factores diversos como los rayos UV o el tabaco. Ahora bien, estas proteínas nuevas suelen ser tan tóxicas para la célula que la célula muere y el sistema se limpia él solo. En otros casos, sin embargo, la proteína, como la producida por la traslocación de los cromosomas 15 y 17, induce la proliferación celular descontrolada, la primera característica de la leucemia.

Cuando viene un paciente con leucemia, el grupo de Di Croce ya no tiene que buscar gen a gen, como antes. Ahora ve todos los genes al mismo tiempo, de golpe, gracias a los microarrays. Esta misma tecnología sirve a los otros grupos del CRG que trabajan sobre diferentes genes. Xavier Estivill dirige el “Grupo de las Causas Genéticas de las Enfermedades”. Estivill, ex-director del Centro de Genética Médica y Molecular del Instituto de Recerca Oncològica (IRO) y del Servicio de Genética del Hospital Clinic, se vino con parte de su equipo al CRG en 2002 aportando una de las experiencias más extensas y densas en España en la investigación de enfermedades genéticas.

Uno de los objetivos de los trabajos de este grupo es detectar y clasificar las mutaciones puntuales y genómicas más frecuentes en enfermedades neurológicas y psiquiátricas. Estivill ya posee en estos momentos importantes colecciones procedentes de pacientes con diferentes enfermedades psiquiátricas y su grupo ha identificado variantes que predisponen al desarrollo de la anorexia, la bulimia y los trastornos adictivos. Una de las técnicas más potentes de la que se sirve es la genotipación a gran escala para estudiar SNPs o regiones del genoma que varían en su número de copias. “Analizamos en paralelo decenas de miles de variantes del genoma y regiones que varían de forma notable entre sujetos. Entre dos individuos hay más de 100 regiones

Imagen de proteínas que modifican el DNA y las histonas





El científico Xavier Estivill ante un reto diferente al habitual: pasear a su hija Marta a bordo de una bala de paja.

que son distintas, lo que supone unos 20 millones de nucleótidos. La información que hemos obtenido permite concentrarnos en los genes y regiones de mayor importancia funcional”, explica el científico. Pero aclara de inmediato: “La importancia de nuestro trabajo y de nuestros resultados no depende de tener detectados más y más genes mutados, sino en obtener información sobre cómo y en qué circunstancias se expresan, lo cual permite entonces diseñar estrategias terapéuticas para la corrección del efecto de su disfunción.”

El equipo de Estivill posee una notable experiencia en el estudio de cómo el retraso mental en un niño, o los trastornos psiquiátricos de una persona, pueden originarse en una mutación que genera pequeños cambios en una página del código genético, pero de enormes consecuencias porque impide que no se puedan comprender capítulos completos del libro de la vida.

“La importancia de nuestro trabajo y de nuestros resultados no depende de tener detectados más y más genes mutados, sino en obtener información sobre cómo y en qué circunstancias se expresan, lo cual permite entonces diseñar estrategias terapéuticas para la corrección del efecto de su disfunción”.

[Xavier Estivill]



El CRG por dentro... y por fuera

El entorno

El CRG está sustentado por el compromiso de encontrar nuevas vías para generar conocimiento sobre el genoma y atacar las enfermedades asociadas a los errores genéticos desde la perspectiva única y singular de cada ser humano. La promesa de una nueva forma de medicina regresa al centro del escenario, ocupado hace años por las “balas mágicas” y otras terapias que no estuvieron a la altura de los sueños de sus protagonistas. Ahora cabe preguntarnos: ¿estamos realmente ante una nueva forma de medicina, o es el mismo modelo de siempre con diferente y más lustroso ropaje?

Nadie mejor que el director del CRG, Miguel Beato, para contárnoslo, quien además nos acompañará en esta parte final:

MB.: Sin ningún género de dudas estamos ante una nueva forma de medicina, y cualquiera puede comprenderlo. Por ejemplo, todos sabemos que un mismo medicamento tiene efectos muy diversos en diferentes pacientes. ¿Cuál es la base de esa variación? Si somos capa-

Vista desde las instalaciones del CMIMA, centro donde el CRG ha estado ubicado desde el año 2002



ces de ver todo el genoma de un individuo de una manera rápida (mediante análisis de SNPs), podremos clasificar esa variación y entender por qué en un grupo de población ese medicamento tiene un alto poder curativo y en otros o no tiene efecto o produce reacciones negativas. Esto, de partida, ya ofrece una ventaja evidente en el diagnóstico. Cuanto mejor entendamos la identidad genética global de un individuo, más cerca estaremos de diseñar terapias específicas y más ajustadas a las características de cada uno. Y en cuanto a las enfermedades genéticas clásicas, lo que estamos obteniendo con los nuevos métodos genómicos es una información muy precisa sobre cómo una mutación provoca una enfermedad. Claro, lo que no se puede decir es que esta nueva medicina va a curarlo todo y lo hará mañana.

Tan importante como el dinero para investigar, que frecuentemente ocupa una atención excesiva, es el contexto de la investigación, tanto en lo referente a recursos humanos y materiales, como a los aspectos jurídicos, las políticas fiscales, o las relaciones con las empresas de capital riesgo, las biotecnológicas o las farmacéuticas. Pero, sobre





todo, como saben todos los grandes centros de investigación del mundo, la localización física, el acceso a los recursos pertinentes y la calidad del paisaje circundante suelen ser muy valoradas por quienes están en la élite de la comunidad científica, en particular si ese paisaje incluye también a otros grupos científicos de reconocida calidad. En todos estos aspectos, el CRG es un instituto privilegiado y, por qué no decirlo así, un bicho raro dentro del panorama científico español.

El centro se creó en diciembre del 2000 por iniciativa del Departament de Universitats, Recerca i Societat de la Informació (DURSI) de la Generalitat de Catalunya [ahora Departament de Educació i Universitats], como una fundación sin ánimo de lucro en la que participan, entre otros, el Departament de Educació i Universitats y el Departament de Salut de la Generalitat y la Universitat Pompeu Fabra (UPF). Se trata de lo que Beato denomina un “centro de investigación de última generación”, basado en un modelo no funcionarial de organización de la investigación, algo no muy frecuente, por no decir bastante inaudito, entre nosotros.

El edificio del CRG está integrado en el Parc de Recerca Biomèdica de Barcelona (PRBB), donde conviven instituciones tales como el Departamento de Ciencias Experimentales y de la Salud (CEXS) de la UPF, el Instituto Municipal de Investigación Médica (IMIM), o el Centre de Medicina Regenerativa de Barcelona. El magnífico edificio en el frente marítimo de la Barceloneta, al costado del Hospital del Mar (un centro de investigación clínica en expansión) y el Puerto Olímpico, ya supone un fuerte atractivo para muchos investigadores que suman el disfrute de este atractivo paisaje a la calidad de los centros de investigación reunidos en el PRBB. El PRBB es, al mismo tiempo, un eslabón esencial de la Bioregión, una iniciativa que trata de casar un entorno biomédico de avanzada con un apoyo empresarial potente, algo de lo que siempre ha adolecido este tipo de investigación en nuestro país.

LAFH. ¿Cómo empezó el CRG? ¿Por qué se planteó un centro de regulación genómica en Barcelona?

MB.: Antes de empezar el CRG se hizo la Facultad de Ciencias Experimentales y de la Salud de la UPF, que fue por lo que vine a Barcelona. Donde yo estaba en Alemania teníamos una licenciatura de biología humana y la UPF pensaba hacer una facultad experimental de biología con una orientación humana. Esa experiencia fue muy positiva, sobre todo, entre otras cosas, porque ya adoptamos los criterios para reclutar científicos como se hace en EEUU: se ponían anuncios en Nature, seleccionábamos a los candidatos con criterio exclusivo de calidad y adecuación, los invitábamos a venir a Barcelona para una entrevista y un seminario... Finalmente, recibíamos la evaluación de expertos externos y escogíamos al mejor. Esta experiencia me convenció de que en Barcelona se empezaban a hacer las cosas de otra manera. A pesar de todo, aquella Facultad era muy pequeñita y no nos permitía crear una masa crítica de investigación para competir a nivel internacional.

LAFH.: Además, la idea original de crear en paralelo una facultad de medicina en la UPF no fue bien recibida por el resto de universidades de Cataluña.

MB.: Efectivamente. Pero, lo que son las cosas, aquella oposición al proyecto se saldó con el nombramiento de Andreu Mas-Colell como consejero del DURSI. El proyecto de la facultad de medicina quedó aparcado de momento, pero se concibieron entonces otros nuevos. Así surgió la idea del CRG que, en principio, debería haber dirigido Manolo Perucho, pero no quiso venir porque todavía estaba muy ocupado en EEUU y me encargué yo. Mas-Colell me pidió un plan, se lo hice y le gustó.

LAFH.: ¿Cuándo se dijo usted: esto empieza a funcionar?

MB.: Empezamos a reclutar personal con el criterio que ya habíamos ensayado: anuncios en Nature y entrevistas en profundidad. En el 2001 incorporamos a un par de científicos jóvenes de muy alto nivel: Raúl Méndez, de Worcester (Massachusetts, EEUU), y Josep Vilardel, del Albert Einstein de Nueva York. Comenzó a correrse la voz de lo que hacíamos. El primer "fichaje" sonado fue el de Juan Valcárcel, que dirigía un grupo en EMBL y se podía haber quedado allí perfectamente.

Además, Juan y su mujer, Fátima Gebauer, que era líder de un equipo en EMBL, tenían una excelente oferta de EEUU. Cuando en el 2002 les convencimos de que se vinieran, eso despertó una cierta expectación. Juan es reconocido como uno de los científicos más prometedores de Europa y su llegada mejoró la calidad de los candidatos. Ahora tenemos candidatos que nunca pensé que se iban a interesar por venir aquí. En 2004 pusimos un límite a nuestro crecimiento y hubo que empezar a decir que no a gente muy buena.

LAFH.: Dos años antes se viene Xavier Estivill con parte de su equipo del IRO, un pelotón importante.

MB.: Si, aquello fue también en 2002, cuando nos vinimos transitoriamente al Centro Mediterráneo de Investigaciones Marinas y Ambientales (CMIMA) hasta que nuestro espacio en el nuevo edificio del PRBB estuviera listo. Fue una operación especial que nos reforzó mucho en el terreno de la investigación con implicaciones clínicas. Por la misma época también aceptaron venir Pura Muñoz, una experta en desarrollo muscular, también procedente del IRO, Thomas Graf, alemán de origen austríaco, que había coordinado un programa en el EMBL y que aporta contribuciones esenciales sobre las células madre de la sangre, y Luciano Di Croce, del Instituto Europeo de Oncología de Milán. Finalmente también conseguimos incorporar a Roderic Guigó, un gran experto en bioinformática genómica, así como a Isabelle Vernos y Luis Serrano, ambos procedentes del EMBL y con un consolidado prestigio en sus respectivos campos, la biología celular y la biología de sistemas. Después hemos tenido que lidiar con el aluvión de candidatos que acaban de llegar o están por llegar, la mayor parte procedente del extranjero. Varios de estos científicos coordinan alguno de los seis programas que actualmente desarrollamos en el CRG, o de los grupos dentro de los programas:

- Programa de Bioinformática y Genómica (Roderic Guigó)
- Program de Biología Celular y Desarrollo (Isabelle Vernos, en funciones)

“Sabemos que tenemos un problema -que en el fondo hay que convertir en nuestra ventaja- y es que lo que hacemos es demasiado nuevo y que en nuestro país no hay una cultura de apoyo privado a la investigación como ocurre en el Reino Unido o EEUU, y debemos promoverla con la calidad de nuestro trabajo”.

- Programa de Diferenciación Celular y Cáncer
(Thomas Graf)
- Programa de Genes y Enfermedad
(Xavier Estivill)
- Programa de Regulación Génica
(Miguel Beato)
- Programa de Biología de Sistemas
(Luis Serrano)

LAFH.: Ustedes seleccionan en base a la experiencia, al curriculum, a la entrevista, al seminario... Pero este es un campo muy nuevo, muy exigente y puede ser que, por las razones que sea, el candidato seleccionado no dé la talla. ¿Qué pasa entonces?

MB.: Tenemos un Consejo Científico Asesor (CCA) integrado por 10 científicos muy destacados y reconocidos en el sector, miembros relevantes de prestigiosos centros de investigación de Europa y EEUU. El consejo nos evalúa periódicamente a todos, incluido a mí. Y si la investigación que hacemos no está a la altura de la mejor a nivel internacional, pues hay que buscarse otro lugar para trabajar. Lo que dice el Consejo va a misa. Hemos elaborado una carrera científica que aceptan todos los científicos del CRG y basada en evaluaciones por expertos externos. Así se funciona en otros lugares del mundo. La existencia del CCA ha tenido un doble efecto: convenció a mucha gente para venir aquí y disuadió a alguno que prefirió la cobertura funcionarial. Un sistema externo de control, sin dar garantías a nadie de permanencia, es duro, pero es dinero público y no podemos garantizar puestos vitalicios a nadie. Aquí no hay un solo funcionario.

LAFH.: En este contexto, el proceso de toma de decisiones por los diferentes grupos, e incluso por usted en cuanto líder de uno de los programas y director del centro, aparece como un factor clave. Las cosas deben estar muy claras como para que el CCA sepa qué va a evaluar.

MB.: Tenemos una escala de reuniones y encuentros que forman parte de una liturgia inexcusable. Los jefes de grupo nos encontramos una vez al año para un retiro de dos días, en el que discutimos y cruzamos

ideas libremente para ir perfilando estrategias. Una vez al mes nos encontramos para tomar las decisiones que afectan al CRG y repasar los problemas de un grupo. Cada semana, cada programa tiene un “data seminar” en el que se discute uno de los proyectos de cada grupo e intervienen estudiantes, postdocs y jefes de grupo. Igualmente, el comité de dirección se reúne mensualmente y se comunica con el CCA con una cierta regularidad. En el CRG, los jefes de grupo tienen un amplio margen de independencia, en un marco colectivo de discusión, para escoger y supervisar las líneas científicas sobre las que vamos a trabajar.

LAFH.: Usted ha conseguido finalmente un acuerdo con el Laboratorio Europeo de Biología Molecular. ¿Cuál es la importancia estratégica de esta colaboración?

MB.: El European Molecular Biology Laboratory, más conocido entre nosotros como EMBL, es el centro de referencia de la investigación biológica en Europa. El objetivo de este acuerdo es que el CRG sea parte del EMBL en España, un centro avanzado de investigación biomédica cofinanciado por dicho organismo. El acuerdo actual crea la Unidad EMBL/CRG de Investigación en Biología de Sistemas, dirigida por Luis Serrano.

En Europa hay otros centros financiados por EMBL, en Hinxton (Inglaterra), en Hamburgo (Alemania), en Grenoble (Francia) y en Monterotondo (Italia). Un acuerdo con el EMBL supone la elevación de la calidad del CRG a unos estándares que no son habituales, por no decir desconocidos, en España, porque obliga a tener la calidad del EMBL. El acuerdo actual de cooperación EMBL/CRG durará 9 años, el mismo tiempo que los contratos de los jefes de grupo. El presupuesto inicial aportado por el Ministerio de Educación y Ciencia es de 12,7 millones de euros. Gracias a este acuerdo tendremos relaciones con el programa de biología computacional de Heidelberg y asistencia para desarrollar una unidad puntera de microscopía, lo que tendrá ramificaciones en el desarrollo de instrumentos ópticos avanzados que se podrán utilizar en las instalaciones de Barcelona.

LAFH.: “Elevar la calidad al nivel del EMBL”. En los últimos 20 años hemos escuchado muchas veces en España declaraciones parecidas sobre la pretendida equiparación de nuestra ciencia con la de centros de referencia en Europa y, sobre todo, EEUU. Sí, hay científicos de alto nivel y que cada vez publican más en las mejores revistas científicas, pero institucionalmente el trecho que queda por recorrer es notable. ¿Qué será lo diferente en esta ocasión?

MB.: Para llegar a ser un centro EMBL hay dos requisitos importantes que debemos cambiar con respecto a nuestra estructura científica habitual:

1) Movilidad, la gente joven no puede tener contratos indefinidos, sino de cinco años con un máximo de nueve, tras evaluación externa, y después deben irse; 2) Internacionalidad, es decir, al menos el 50% de los jefes de grupo debe ser extranjero. Cuando cumplamos estos criterios, entonces EMBL podría pensar en co-financiar el CRG. Para ello tenemos que convencer a los 19 países miembros del EMBL de que el CRG es un centro europeo, que integra gente de diferentes países en competición abierta y que ofrecemos algo que merece la pena.

LAFH.: Otra piedra de toque en la que hemos tropezado con frecuencia desde la Ley de la Ciencia de 1986 es la relación ciencia – industria. La transferencia de innovaciones, productos, tecnologías o aplicaciones desarrolladas por los centros públicos de I+D parece convertirse en el famoso camello que quiere pasar por el ojo de la aguja.

MB.: Somos muy conscientes de esta dificultad. Nosotros hemos creado un Consejo Empresarial que empezó en 2004 y ya ha dado algunos frutos, como la creación de una posición de “product manager” que sirve de conexión entre el CRG y las empresas. Tenemos la ventaja de que nos movemos en el entorno de los centros y empresas más importantes al nivel nacional y en parte internacional, pero eso no es ninguna garantía de éxito. Necesitamos que confluyan muchos factores, como la industria farmacéutica, las cajas de ahorro, las fundaciones... cosa que todavía no está sucediendo. Sabemos que tenemos un problema –que en el fondo hay que convertir en nuestra ventaja- y es que lo que hacemos es demasiado nuevo y que en nuestro país no hay una cultura de apoyo privado a

la investigación como ocurre en el Reino Unido o EEUU, y debemos promoverla con la calidad de nuestro trabajo.

LAFH.: Si hiciéramos un mapa de la investigación de la regulación genómica, ¿qué lugar ocuparía el CRG?

MB.: En EEUU hay muchos centros muchísimo más importantes que el nuestro. Nos llevan una considerable ventaja de años y recursos, no somos entidades comparables. De todas maneras, hace sólo dos años se creó en Boston un centro de regulación genómica parecido al nuestro financiado por un millonario. Entre los centros europeos con los que quisiéramos medirnos, aparte del EMBL, mencionaría al Instituto de Patología Molecular (Austria), fundado por la farmacéutica Boeringher Ingelheim, parecido al CRG pero más rico, el Instituto de Biología Celular y Genética del Max Plank de Dresde (Alemania), o el Instituto John Gurdon de Cambridge (Inglaterra) del Wellcome Trust. Una de nuestras ventajas es que tenemos un gran hospital en la puerta de al lado y centros académicos y de investigación dedicados a cuestiones de salud en el mismo edificio. Además, el PRBB se encuadra dentro de una estrategia muy ambiciosa de desarrollo de la investigación biomédica en Cataluña que se está propagando a través de otros focos de investigación de muy alto nivel, como el IRBB del Parc Científic de la UB, el IDYBAPS del Hospital Clinic, los grupos de investigación de Bellvitge, Sant Pau o la Vall d'Hebron. Todo esto, englobado en la iniciativa de la Bioregión, que debería activar las empresas biotecnológicas, crea una posición privilegiada porque nos convierte en una potencia biomédica considerable. Lógicamente, muchos otros centros, incluso en España, hacen cosas similares a nosotros, a fin de cuentas estamos hablando de biología molecular y del genoma, que es la biología de frontera, pero éste contexto marca frecuentemente la diferencia respecto a la calidad y la importancia estratégica. Este contexto y la visión global que anima a lo que hacemos en el CRG. Lo que Barcelona ha conseguido crear al borde del mar, al lado del Puerto Olímpico, no me cabe la menor duda de que tendrá un impacto notable sobre la ciencia y la salud de los ciudadanos, a nivel local y global.

