



15 JUN 07 | Genética

## Tal vez debamos modificar nuestra idea del genoma humano

Los primeros datos del proyecto ENCODE obligan a replantear los conocimientos sobre genética.

(Jano Online)

Los resultados ponen en duda la clásica visión de que el genoma humano consiste en un conjunto relativamente pequeño de genes junto con una gran cantidad de "ADN basura"



El proyecto ENCODE se puso en marcha en el momento en que se completó el genoma humano. Como bien saben los lectores, en aquel momento se dispuso de la secuencia formada por 30 millones de pares de bases de ADN, pero a partir de ahí había que comenzar a leer "el libro de la vida" y a comprenderlo.

Con ese fin hace cuatro años nació el proyecto ENCODE (ENCyclopedia Of DNA Elements), coordinado por el Instituto de Investigación Nacional del Genoma Humano (NHGRI).

Los resultados de la fase piloto se acaban de publicar en un conjunto de 29 artículos –uno en "**Nature**" y 28 en "**Genome Research**"-, y la gran sorpresa es que la idea que tenían los científicos de lo que es un gen y sobre cómo funciona nuestro genoma puede cambiar de forma sustancial, pues estos primeros resultados ponen en entredicho la visión tradicional sobre nuestro programa genético como una colección ordenada de genes independientes. Al contrario, la idea que se desprenden de los datos del proyecto ENCODE es la de una **gran red** en la que los genes, los elementos reguladores y otros tipos de secuencias de ADN interactúan de formas complejas y superpuestas, no descifradas por completo todavía.

La primera fase de esta gran investigación ha consistido en elaborar una "lista de componentes" de todos los elementos funcionales que hay en el 1% del genoma humano. Se trataba de saber si era viable hacerlo tanto técnica como económicamente para pasar a abordar el genoma al completo. Esta fase piloto ha sido desarrollada por 35 grupos de 80 organizaciones de todo el mundo.

Según declara el director del NHGRI, **Francis S. Collins**, "este impresionante esfuerzo ha desvelado muchas y excitantes sorpresas e ilumina el camino a futuros esfuerzos para explorar el contenido funcional del genoma humano completo. Gracias a ello y a ideas innovadoras del consorcio ENCODE, la comunidad científica necesitará replantearse algunas de las asunciones largamente sostenidas sobre qué son los genes y qué hacen, además de cómo han evolucionado los elementos funcionales del genoma. Esto podría tener implicaciones notables en los esfuerzos para identificar las secuencias de ADN implicadas en muchas enfermedades humanas". Tal como lo explica el NHGRI, "la finalización del Proyecto Genoma Humano en abril de 2003 fue un enorme logro, pero la secuenciación del genoma marcó sólo el primer paso hacia el objetivo de utilizar este tipo de información para el diagnóstico, tratamiento y prevención de enfermedades. Disponer de la secuencia del genoma humano es similar a disponer

## FOROS

[¿Cómo será la Medicina en 2020?](#)

[IntraMed cumple diez años](#)

[Burnout](#)

[Internado Rotatorio](#)

[Psiquiatría](#)

[Tocoginecología](#)

[Clínica Médica](#)

[Cirugía](#)

[Pediatria](#)

[Cardiología](#)

[Búsqueda Bibliográfica](#)

[más >](#)



Estudio  
Internacional  
Síndrome de  
Burnout en  
Profesionales  
de la Salud

Participe ahora

Encuestas

IntraMed

de todas las páginas de un manual de instrucciones necesario para hacer un cuerpo humano. Los investigadores todavía deben aprender cómo leer el lenguaje del manual para poder identificar cada parte y comprender cómo los componentes trabajan juntos para contribuir a la salud y la enfermedad".

En los últimos años, los investigadores han realizado grandes progresos al utilizar datos de la secuencia de ADN para identificar genes, que se definen tradicionalmente como los componentes del genoma que codifican proteínas. Sin embargo, los genes sólo representan una pequeña fracción del genoma humano – entre el 1,5 y el 2%, y existen evidencias que demuestran que otras partes del genoma también desempeñan funciones importantes. A pesar de ello, hasta ahora, la mayoría de los estudios se han concentrado en elementos funcionales asociados con genes específicos y no han proporcionado perspectivas sobre elementos funcionales a lo largo del genoma. El proyecto ENCODE representa el primer esfuerzo sistemático para determinar dónde están ubicados todos los elementos funcionales y cómo están organizados.

En la fase piloto, los investigadores han concebido y evaluado estrategias de alta capacidad de procesamiento para identificar elementos funcionales en el genoma. Esos elementos incluyen genes que codifican para proteínas, genes que no codifican para proteínas, elementos reguladores que controlan la transcripción de genes y elementos que mantienen la estructura de los cromosomas y actúan como mediadores de la dinámica de su replicación.

El proyecto ENCODE se centró en 44 regiones, que cubren alrededor del 1% de la secuencia del genoma humano, o unos 30 millones de pares de bases de ADN. Las regiones fueron seleccionadas estratégicamente para proporcionar una sección transversal representativa del genoma humano completo. Así pues, el consorcio ENCODE ha generado más de 200 conjuntos de datos y ha analizado más de 600 millones de datos.

"Nuestros resultados revelan importantes principios sobre la organización de elementos funcionales en el genoma humano, proporcionando nuevas perspectivas sobre todos los aspectos, desde la transcripción del ADN a la evolución de los mamíferos. Concretamente, hemos ganado un notable conocimiento sobre las secuencias de ADN que no codifican proteínas, y sobre las cuales antes sabíamos muy poco", señala Ewan Birney, jefe de la anotación del genoma en el Instituto Europeo de Bioinformática (EBI), del Laboratorio Europeo de Biología Molecular, en Hinxton, Inglaterra, y quien dirigió la integración masiva de datos y el trabajo de análisis.

Los resultados más importantes del consorcio ENCODE incluyen el descubrimiento que apunta que la mayor parte del ADN del genoma humano se transcribe a ARN, y que las transcripciones resultantes se superponen unas a otras de forma extensiva. Este extenso patrón de transcripción pone en duda la visión largamente sostenida que el genoma humano consiste en un relativamente pequeño conjunto de genes, junto con una gran cantidad de "ADN basura", que no es biológicamente activa.

Los nuevos datos indican que el genoma contiene muy pocas secuencias que no se usen y que, en realidad, se parece más a una red compleja. En esta red, los genes son sólo uno de los muchos tipos de secuencias de ADN con un impacto funcional. "Nuestra perspectiva de transcripción y genes deberá evolucionar," declaran los investigadores en el artículo de "Nature", destacando que el modelo de red del genoma "plantea algunas interesantes cuestiones mecanísticas" a las que todavía hay que dar respuesta.

Otras sorpresas en los datos de ENCODE tienen mayores implicaciones para nuestra comprensión de la evolución de los genomas, en especial de los genomas de mamíferos. Hasta hace poco, los investigadores pensaban que la mayoría de secuencias del ADN importantes para la función biológica se encontraban en áreas del genoma sujetas a restricciones evolutivas –es decir, en áreas en que era más probable que la secuencia se conservara durante el proceso evolutivo. Sin embargo, el trabajo de ENCODE ha revelado que alrededor de la mitad de los elementos funcionales en el genoma humano no parecen haber sido sometidos a restricciones evolutivas, al menos cuando son examinadas mediante los métodos más actuales empleados por biólogos computacionales en el proyecto ENCODE.

Según los investigadores, esta falta de restricción evolutiva podría indicar que los

genomas de muchas especies contienen una reserva de elementos funcionales, incluyendo transcritos de ARN, que no proporcionan beneficios específicos en términos de supervivencia o reproducción. A medida que esta reserva avanza en la evolución, los investigadores especulan que puede servir como "almacén para la selección natural", al actuar como fuente única de elementos funcionales para cada especie y de elementos que realizan funciones similares entre las especies a pesar de tener secuencias aparentemente distintas.

Otros aspectos destacables del estudio ENCODE incluyen:

Identificación de numerosos puntos de inicio de la transcripción de ADN previamente no conocidos.

Pruebas de que, contrariamente a las perspectivas tradicionales, es igual de probable que las secuencias reguladoras estén situadas antes que después del inicio de la transcripción.

Identificación de firmas específicas en histonas, que son las proteínas que organizan el ADN, y correlación de estas firmas con distintas funciones genómicas.

Comprensión más profunda sobre cómo la replicación de ADN se coordina mediante modificaciones en histonas.

"La magnitud de la complejidad de la estructura del genoma, que ha puesto de manifiesto el proyecto ENCODE, era ciertamente inesperada y obligará a replantear alguna de las estrategias que se utilizan actualmente para establecer relaciones entre las variaciones que observamos en la secuencia del genoma y manifestaciones patológicas. El objetivo –comprender cómo la secuencia del genoma determina nuestras características biológicas– está quizás más lejos de lo que esperábamos hace cuatro años; pero estamos desde luego, más cerca de él de lo que estábamos" dice el Dr. Roderic Guigó, del Centro de Regulación Genómica en Barcelona, uno de los nueve científicos que ha dirigido el trabajo publicado en "Nature".

El conjunto de autores de los artículos de ENCODE incluye investigadores de organizaciones académicas, gubernamentales e industriales ubicadas en Australia, Austria, Canadá, Alemania, Japón, Singapur, España, Suecia, Suiza, el Reino Unido y los Estados Unidos. El proyecto ENCODE ha estado abierto a todos los investigadores interesados dispuestos a obrar de acuerdo con las directrices del consorcio.

Concretamente, en España, un grupo de científicos pertenecientes a diversos centros ubicados en el Parque de Investigación Biomédica de Barcelona (PRBB) han contribuido sustancialmente al proyecto ENCODE. El NHGRI aportó 2 millones de dólares a un proyecto en el Instituto Municipal de Investigación Médica (IMIM), liderado por Roderic Guigó, quien actualmente es coordinador del Programa de Bioinformática y Genómica en el Centro de Regulación Genómica (CRG), y profesor de la Universitat Pompeu Fabra (UPF). Este proyecto, en el que también participaron investigadores del Reino Unido, Suiza y Estados Unidos ha producido la anotación génica de las regiones ENCODE. Esta anotación ha sido empleada como modelo de referencia por muchos de los análisis de ENCODE. Junto a Tom Gingeras, de Affymetrix, una empresa de California especializada en biochips, Roderic Guigó también ha coordinado el llamado grupo de Genes y Transcripción, uno de los cinco grupos en que el proyecto ENCODE está organizado. La investigación en este grupo ha producido algunos de los resultados más inesperados en el seno del proyecto. En especial, la aplicación combinada de enfoques experimentales y computacionales mostró que, contrariamente a asunciones previas, la mayor parte de la secuencia del genoma humano puede transcribirse y que los genes no pueden considerarse como pequeñas entidades del genoma, sino mejor como redes interconectadas de transcritos.

De forma independiente, el Dr. Xavier Estivill, coordinador del Programa Genes y Enfermedad en el Centro de Regulación Genómica (CRG), también ha participado en la sección de "Variación" del proyecto. Su grupo ha investigado la variación en el número de copias en las regiones ENCODE. El estudio computacional de los elementos funcionales del genoma en regiones donde el número de copias es variable ha revelado un exceso notable de exones codificadores y secuencias transcritas así como depleción en pseudogenes, arrojando nueva luz sobre la relevancia funcional de las regiones donde el número de copias es variable y reforzando su importancia como mediadoras de diferencias entre humanos y la susceptibilidad a las enfermedades.

"El análisis de variación que se ha llevado a cabo en el marco del proyecto ENCODE es el más exhaustivo conseguido hasta el momento y debería servir de guía a futuras investigaciones de variabilidad del genoma humano. Este hecho debería tener un impacto en estudios de fármaco-genómica y en el análisis de susceptibilidad a las enfermedades" comenta Xavier Estivill, coordinador del Programa Genes y Enfermedad, del Centro de Regulación Genómica, en Barcelona.

Además de las contribuciones al artículo de "Nature", los investigadores del PRBB son autores en cuatro de los artículos de ENCODE publicados en "Genome Research". Estos artículos tratan de RNAs no codificadores, pseudogenes, y nuevos transcritos. En concreto, la contribución de los investigadores del PRBB ha sido relevante en el artículo que describe un nuevo enfoque tecnológico para estudiar de manera exhaustiva la actividad transcripcional de los genes que codifican para proteínas. Los resultados de este artículo apoyan con firmeza la naturaleza del genoma como una red de transcritos superpuestos. La primera autora de este artículo, en aquellos momentos investigadora del IMIM, es France Denoeud, que actualmente se encuentra en GENOSCOPE, Francia.

Cabe destacar también que el artículo en "Nature" reconoce explícitamente la utilización exhaustiva que los investigadores del PRBB han realizado del ordenador Mare Nostrum del Barcelona Supercomputing Center-Centro Nacional de Supercomputación para realizar alguno de los análisis computacionales.



[Términos y condiciones de uso](#)

Todos los derechos reservados | Copyright 1997-2007

