

## Presentan el nuevo mapa del genoma humano

Un equipo internacional, con participación española, ha descubierto que lo que hasta ahora se consideraba como 'ADN basura', en realidad es un componente útil y muy importante en el genoma humano. Los resultados se publican a la vez en más de 30 artículos interconectados de las revistas *Nature*, *Genome Research* y *Genome Biology*.

CRG/SINC

5/9/2012 20:00 CEST

Uno de los estudios del proyecto ENCODE. Imagen: Darryl Leja (NHGRI), Ian Dunham (EBI).

Esta semana un equipo de investigadores internacional revela que el denominado 'ADN basura' en el genoma humano es, en realidad, un gran panel de control con millones de 'interruptores' que regulan la actividad de nuestros genes. Sin estos interruptores, los genes no funcionarían y se podrían originar mutaciones que, a su vez, podrían desencadenar enfermedades.

Descubierta por cientos de científicos del proyecto *Encyclopedia of DNA Elements* ([ENCODE](#)), la nueva información es tan exhaustiva y compleja que ha dado lugar a un nuevo modelo de publicación donde los documentos electrónicos y los conjuntos de datos están interconectados. Los resultados aparecen en 30 artículos de acceso abierto conectados entre sí en las revistas *Nature*, *Genome Biology* y *Genome Research*, además de en la revista [Science](#).

El proyecto está liderado por el *National Genome Research Institute* (NHGRI)

en EEUU y el EMBL-*European Bioinformatics Institute* (EMBL-EBI) en Reino Unido. ENCODE presenta un mapa detallado de la función del genoma que identifica 4 millones de interruptores de genes. Esta referencia esencial ayudará a los investigadores a localizar áreas muy específicas de enfermedades humanas.

"Nuestro genoma sólo funciona gracias a los interruptores: millones de lugares que determinan si un gen se enciende o se apaga", explica Ewan Birney del EMBL-EBI, coordinador de análisis del proyecto. "El proyecto Genoma Humano mostró que sólo el 2% de nuestro genoma contiene genes, que son las instrucciones para hacer proteínas. Pero con ENCODE podemos ver que cerca del 80% del genoma está activamente haciendo algo. Hemos encontrado que una gran parte del genoma - de hecho, una cantidad sorprendente - está implicada en controlar cuándo y dónde se producen las proteínas más allá de simplemente fabricarlas".

---

Una cantidad sorprendente del genoma está  
implicada en controlar cuándo y dónde se  
producen las proteínas

Estos descubrimientos ofrecen el conocimiento que se necesita para mirar más allá de la estructura lineal del genoma y ver cómo toda la red está conectada. Tan importante es saber dónde se están ubicados los genes como qué secuencias los controlan. Debido a la compleja estructura tridimensional de nuestro genoma, estos controles a menudo están lejos del gen que regulan si leemos la secuencia linealmente, aunque si se hace de forma tridimensional veremos que se encuentran envueltos alrededor para contactar con ellos.

Si no fuera por ENCODE, probablemente nunca habríamos mirado estas regiones, destacan los investigadores, quienes destacan este paso "enorme" hacia la comprensión del complejo diagrama de cableado del ser humano. ENCODE ayuda a mirar en lo más hondo del circuito de regulación que muestra cómo todas las partes se unen para crear un ser complejo.

El proyecto ha combinado los esfuerzos de 442 científicos en 32 laboratorios

en el Reino Unido, EEUU, Singapur, Japón, Suiza y España. En nuestro país se han involucrado veinte investigadores del Centro de Regulación Genómica (CRG) en Barcelona (aunque algunos están, a fecha de hoy, en otros centros).

### Participación de centros españoles

El científico Roderic Guigó, coordinador del programa de Bioinformática y Genómica del CRG y profesor en la Universidad Pompeu Fabra, ha liderado el grupo de análisis de ARN de ENCODE. También han participado dos investigadores del CNIO y se ha contado con el apoyo del Instituto Nacional de Bioinformática.

Los autores del CRG han colaborado en dos de los artículos publicados en *Nature* (autores principales en uno de ellos), en otros dos en *Genome Biology* y en cuatro de los de *Genome Research* (autores principales en tres de ellos). Investigadores asociados al CRG han diseñado la portada del número especial sobre ENCODE en esta última, inspirándose en el estilo del artista catalán Joan Miró.

Los datos de todo el proyecto han generado cerca de 15 terabytes (15 billones de bytes) de información en bruto, toda ella ahora disponible públicamente. El estudio ha utilizado alrededor de 300 años en tiempo de ordenador para estudiar 147 tipos de tejido y determinar qué enciende o apaga a genes específicos y cómo ese interruptor difiere entre tejidos o tipos celulares.

Los artículos publicados "representan una nueva forma de hacer que los investigadores puedan navegar y acceder a los datos", comenta Magdalena Skipper, editora senior de la revista *Nature*, que ha producido la plataforma de publicaciones gratuita en internet. Todo el contenido de las tres revistas está conectado digitalmente por temas, de tal forma que los lectores pueden seguir su área de interés entre los trabajos y hasta los datos originales.



La portada de Genome Research se inspira en la obra de Miró. Imagen: Genome Research.

### Centro de Regulación Genómica (CRG) y ENCODE

El grupo del CRG coordinado por Roderic Guigó ha participado en el análisis de la actividad transcripcional del genoma, dentro de los subproyectos coordinados por Tom Gingeras de *Cold Spring Harbour Laboratory* (EEUU) y Tim Hubbard del *Wellcome Trust Sanger Institute* (Reino Unido). La extensión de las instrucciones codificadas en el genoma se desencadenan por la transcripción del ADN en ARN. Antes de ENCODE, se consideraba que la actividad transcripcional del genoma iba dirigida a sintetizar moléculas de ARN mensajero para fabricar proteínas.

Pero durante la última década se han desarrollado nuevas tecnologías que permiten monitorizar toda la actividad transcripcional del genoma con una resolución sin precedentes. Utilizando esta tecnología, los investigadores de ENCODE han descubierto una gran actividad transcripcional en el genoma humano que no va dirigido a la fabricación de proteínas. "Las moléculas de ARN son muy abundantes y todavía más diversas en su secuencia, estructura o función de lo que imaginábamos -explica Guigó-, y la biología del ARN será cada vez más crucial para la investigación básica y para las aplicaciones técnicas en biología y medicina en particular".

Participar en el proyecto ENCODE ha sido al mismo tiempo un reto y una recompensa para los investigadores del CRG. "El proyecto ENCODE ha establecido nuevos estándares en relación a la cooperación entre científicos", añade el investigador. "Hemos estado trabajando de forma muy próxima con científicos de todo el mundo. Nuestro grupo de trabajo se reunía semanalmente por teleconferencia y contaba con investigadores de California, la costa este de Estados Unidos, el Reino Unido, Suiza, Singapur y Japón. La reunión era a las 6 de la mañana para los investigadores de California pero a medianoche para los japoneses".

Por su parte, Sarah Djebali, que se ha encargado de la logística de la coordinación en el CRG afirma: "La logística era un reto pero las discusiones entre científicos de todo el mundo, la planificación de los

experimentos, el análisis de resultados y toda esta proximidad y voluntad por compartir han sido una enorme recompensa".

Uno de los principales retos del CRG ha sido el de enfrentarse a la gran cantidad de datos generadas en este proyecto. "Este centro ha sido la base para almacenar los datos sobre ARN y eso, a menudo ha puesto a prueba la capacidad de la infraestructura informática del centro", dice Julien Lagarde, que ha sido el responsable de la informática en este proyecto del CRG.

ENCODE es sólo el primer paso en la larga y compleja tarea de descifrar el significado de la secuencia del genoma. "Esta es, en realidad, la tarea de la biología del siglo XXI. Como investigadores, nos sentimos profundamente privilegiados por poder contribuir en este proyecto" afirma Guigó, que recuerda: "Nuestra participación en este proyecto es, en parte, fruto de unas políticas directas de promoción de la investigación científica. La posibilidad que investigadores de nuestro país participen en proyectos científicos de relevancia global como éste, depende de una apuesta firme y decisiva por apoyar la investigación científica".

Derechos: **Creative Commons**

Creative Commons 4.0

Puedes copiar, difundir y transformar los contenidos de SINC. [Lee las condiciones de nuestra licencia](#)

