

OPINIÓN ENRIQUE LARA PEZZI*

La (re)visión del genoma humano



El autor señala que la magnitud de los resultados 'Encode' es "enorme" para la investigación biomédica y afirma que no hay atajos en la investigación y llevará tiempo asimilarla.

La publicación de la secuencia del genoma humano pasada creó enormes expectativas que, al no se traducirse en avances biomédicos inmediatos, se ha ido diluyendo con los años. Pero sobre todo desveló nuestro gran desconocimiento del genoma, más allá de la secuencia de DNA. Era como si acabáramos de comprar un Lego de 3.000 millones de piezas y descubriéramos con una mezcla de incredulidad y desconcierto que las instrucciones no están en la caja. Además de los más de 20.000 genes codificantes (con información para producir proteínas), se encontró una vasta cantidad de DNA sin función aparente. Este DNA, bautizado algo presuntuosamente como *DNA*

basura, ha sido redescubierto por el proyecto *Encode* (*Encyclopedia of DNA Elements*). Sus resultados, publicados hace dos semanas (ver CF del 10-IX-2012), revelan que prácticamente todo el genoma desempeña algún papel en la célula. Empezamos así a interpretar el genoma, pasando de mirar la secuencia a ver lo que significa.

Los investigadores han mostrado que alrededor del 99 por ciento del genoma participa directamente en algún proceso bioquímico o se encuentra próximo a un elemento que lo hace. Se sabe ahora el tipo de proceso en el que participa el 80 por ciento de las secuencias, aunque aun no la función concreta de todas ellas. La producción de cada una de las proteínas está controlada por varios mecanismos reguladores. *Encode* ha analizado muchos de ellos y ha descubierto cómo se integran entre sí. Gran parte de la información que aportada surge a partir de la secuenciación del RNA y no solo del DNA. Los resultados muestran que aproximadamente el 75 por ciento del genoma se transcribe a RNA. Parte de estos RNAs codifican información que se traducirá a proteínas y el resto tienen actividad reguladora. Además, se han identificado nuevas variantes de proteínas conocidas, generadas por un proceso conocido como *splicing* alternativo del RNA, que permite producir entre 0,5 y 1 millón de proteínas distintas a partir de solo 20.000 genes. El *splicing* alternativo, y en especial su regulación, es en gran medida desconocido y será uno de los focos de interés en la próxima década.

Los resultados de *Encode* tendrán impacto en el desarrollo de nuevas terapias a medio y largo plazo. Una aplicación más inmediata será la reinterpretación de los

estudios de asociación del genoma (GWAS), que intentan establecer una asociación entre la aparición de ciertas enfermedades y la presencia de variaciones genéticas entre individuos. En muchas de las asociaciones identificadas hasta ahora la variación no se encontraba en la parte del gen que se traduce a proteína, sino en regiones alejadas, lo que hacía muy difícil interpretar su impacto en el desarrollo de la enfermedad. *Encode* ha mostrado que muchos de estos nucleótidos variables están en regiones reguladoras, lo que permitirá reinterpretar los datos de GWAS y aclarará la predisposición genética a padecer enfermedades, desde cáncer a cardiovasculares.

'ENCODE' O VISIÓN GLOBAL DEL GENOMA

Aunque parte de esta información había sido descrita a menor escala en los últimos años, *Encode* aporta una visión global sobre el genoma y el papel regulador de muchos de sus elementos. Sin embargo, este sólo es el primer paso. La magnitud de información que aporta es enorme y llevará tiempo asimilarla. Por ejemplo, *Encode* propone una redefinición del concepto de gen y sugiere que sea el transcrito y no el gen la unidad básica hereditaria. La propuesta requiere más debate pero subraya fuertes implicaciones de los nuevos datos. *Encode* muestra que no hay atajos en investigación y el gran impacto se logra mediante grandes consorcios internacionales, mayormente financiados con fondos públicos, donde la vanidad personal se sacrifica en beneficio del interés general.

*Jefe de grupo de investigación del Centro Nacional de Investigaciones Cardiovasculares (CNIC)