

Bioinformáticas del CNIO y del CRG investigan una nueva manera de codificar la información del genoma



Barcelona (19/06/2012) - Redacción

El grupo de Alfonso Valencia en el CNIO en Madrid, en colaboración con el de Roderic Guigó del CRG en Barcelona, descubre que en células y tejidos sanos hay proteínas que surgen de combinar varios genes distintos

Secuenciar el genoma humano fue sólo un primer paso. Sigue pendiente un problema de los que hacen historia: decodificar el genoma, entender cómo en una única molécula está empaquetada la información para construir un ser humano. Y lo que queda no son meros cabos sueltos. Un grupo de bioinformáticos del Centro Nacional de Investigaciones Oncológicas (CNIO), en Madrid, en colaboración con científicos del Centro de Regulación Genómica (CRG), en Barcelona, acaban de publicar varios resultados que muestran que podría haber una forma de codificar la información en el genoma aún inexplorada.

Se trata del ARN quimérico, construido no a partir de un gen, sino de fragmentos de varios genes. "Cada vez más datos, algunos muy recientes, nos indican que la codificación de información en el genoma es más compleja de lo que creíamos, y que también pueden producirse RNAs que combinen la información de dos genes distintos", explica Alfonso Valencia, director del Programa de Biología Estructural y Biocomputación del CNIO. "Los hemos llamado ARN quiméricos rememorando los seres mitológicos integrados por partes de dos animales distintos", detalla.

"Aunque nosotros ya habíamos demostrado en 2006 la prevalencia de este fenómeno, el trabajo actual va en la dirección de confirmar su importancia biológica", añade Roderic Guigó, coordinador del programa Bioinformática y Genómica del CRG.

El ADN contiene los genes, que son traducidos a proteínas. El ARN actúa como molécula intermediaria y representa un paso indispensable en el proceso: para traducir un gen en una proteína, primero hay que construir el ARN correspondiente. La visión clásica de cómo está almacenada la información en el genoma dice que la correspondencia es uno-uno, es decir: un gen, un ARN, una proteína. Eso era lo que se esperaba cuando se secuenció el genoma, a principios de la pasada década. Pero enseguida apareció un problema: el genoma humano contiene unos 20.000 genes, mientras que la variedad de proteínas en el cuerpo humano es muy superior.

Hoy se sabe que un único gen puede dar lugar a varias proteínas, algo equivalente a que una palabra tenga significados distintos aunque se escriba igual, como 'solo', 'bonito' o 'polo'. Pero sigue sin estar claro si el fenómeno es muy frecuente -si todos los genes pueden dar lugar a muchas proteínas- o es más bien una rareza. De hecho, el grupo de Valencia también ha hecho avances en esta área. En un trabajo publicado en abril en *Molecular Evolution Biology* demuestran que la traducción de un único gen en varias proteínas, aunque ocurre, es poco habitual.

El ARN quimérico también contribuye a que haya más proteínas distintas que genes. Vendría a ser como si el sistema que lee y traduce los genes pudiera obtener, de dos de ellos, tres o más significados. Así, por ejemplo, 'silla' y 'verano' serían traducciones directas, pero además estarían 'lla-ve'; 'lla-no'; 'si-no'...

ARN quimérico

La existencia del ARN quimérico se conocía ya, y también se sabía que algunos ARN quiméricos son traducidos a proteínas -otros, como ocurre con el ARN normal, no quimérico, se quedan en la fase de ARN-. Pero las proteínas quiméricas se consideraban más bien una rareza propia de procesos patológicos, por ejemplo, el desarrollo de tumores.

Los bioinformáticos del CNIO y del CRG han rebuscado a fondo en las bases de datos de genes, ARN y proteínas, han realizado nuevos experimentos y finalmente han descubierto que hay mucho más ARN quimérico de lo que se creía. También detectan casos de traducción a proteínas en lo que parece ser un proceso normal en tejidos sanos, no sólo tumorales.

Sus resultados se explican en varios trabajos, de los que el último acaba de ser publicado en la revista '*Genome Research*'. Prueba de su gran interés es que otra revista, '*Nature Reviews Genetics*', ya lo ha destacado en una reseña. En concreto, los investigadores del CNIO han identificado 175 transcritos ARN quiméricos presentes en 16 tejidos humanos y doce proteínas quiméricas nuevas.

Este nuevo hallazgo plantea muchas preguntas: ¿Qué peso tiene realmente este proceso en relación a toda la información en el genoma? ¿Hace encajar por fin el número de genes y el de proteínas? ¿Cuántas proteínas quiméricas hay en total? ¿Tienen una función que las caracterice? ¿Por qué existen?

"Acabamos de abrir una vía y esperamos que ahora la sigan muchos grupos", dice Valencia. "En mi opinión, lo más importante de estos trabajos es que demuestran que aún nos queda mucho para aprender a interpretar completamente lo que está escrito en el genoma".

< < Volver