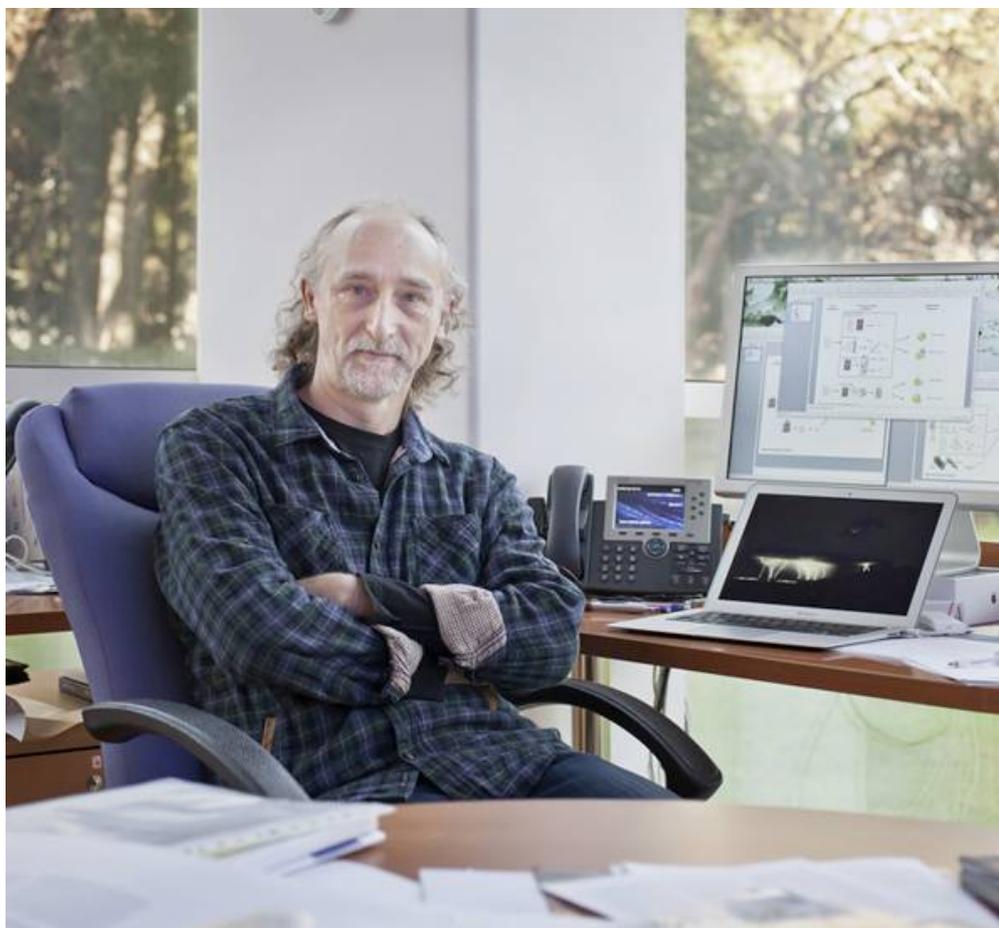


Un equipo del CNIO reduce el tamaño del genoma humano a 19.000 genes

Más del 99% de los genes humanos –aquellos que producen proteínas– tienen un origen anterior a los primates hace más de 50 millones de años. Un nuevo estudio liderado por investigadores del CNIO pone en duda las anotaciones genómicas de distintas especies realizadas hasta la fecha.

CNIO

4/7/2014 12:02 CEST



El investigador Alfonso Valencia. / CNIO

Cómo se metabolizan los nutrientes o se comunican las neuronas en el cerebro son solo algunos de los mensajes contenidos en las 3.000 millones de letras que constituyen el genoma humano.

La detección y caracterización de los genes presentes en esta gran masa de

información es una tarea compleja y ha sido fuente de continuo debate desde los primeros intentos sistemáticos realizados por el Proyecto Genoma Humano hace más de diez años.

Un trabajo liderado por Alfonso Valencia, vicedirector de Investigación Básica y jefe del Grupo de Biología Computacional Estructural del Centro Nacional de Investigaciones Oncológicas (CNIO), y Michael Tress, investigador del grupo, actualiza la cifra de genes humanos –aquellos que generan proteínas– en 19.000, 1.700 menos que los estimados en las últimas anotaciones, y un número muy inferior a los 100.000 que se llegaron a barajar hace años.

El trabajo, que se publica en la revista *Human Molecular Genetics*, concluye que la práctica totalidad de estos genes tienen un origen anterior a la aparición de los primates hace más de 50 millones de años.

La detección y caracterización de los genes
presentes en esta gran masa de información es
una tarea compleja

“El menguante genoma humano”. Así es como describe Valencia las continuas rectificaciones de las anotaciones del genoma humano, que han desembocado en el presente trabajo que contabiliza el número de genes en aproximadamente 19.000.

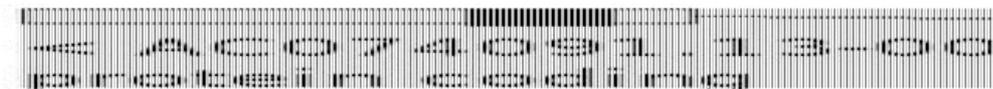
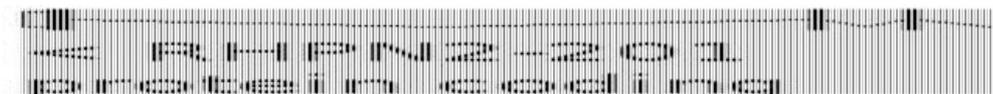
“La parte codificante del genoma –la que produce proteínas– está en constante movimiento”, y añade: “Nadie se podía imaginar hace unos años que con un número tan reducido de genes se podía hacer algo tan complejo”.

Los científicos partieron de análisis proteómicos a gran escala para determinar el mapa de proteínas humanas “a partir del cual identificar aquellos genes que las producen y de esta manera verificar que existen”, explica Valencia.

Para ello, integraron los datos procedentes de siete trabajos anteriores de

espectrometría de masas, la herramienta más potente para detectar moléculas, en este caso proteínas procedentes de más de 50 tejidos humanos.

Los resultados sacaron a la luz poco más de 12.000 proteínas, que casaron con las correspondientes regiones del genoma. Posteriormente analizaron miles de genes que aparecían en otras bases de datos pero no en el estudio.



A diferencia de otros genes (arriba, medio), el gen ACO74091.13 (abajo) no produce proteína y podría ser eliminado junto a otros 1.700 genes de las anotaciones actuales del genoma humano. / Michael Tress (CNIO)

“Un total de 1.700 de los genes que se consideraban como tal no lo son por distintas razones, incluyendo que o bien no generan directamente proteínas o porque su *reading frame* –la lectura de sus letras– no es compatible con la generación de las mismas”, explica Tress.

Menos de diez genes nuevos distinguen a humanos y ratones

Una hipótesis derivada del trabajo es que más del 90% de los genes humanos que producen proteínas tienen su origen en metazoos u organismos pluricelulares del reino animal hace cientos de millones de años; la cifra es superior al 99% para aquellos genes cuyo origen es anterior a la aparición de los primates hace más de 50 millones de años.

“Nuestras cifras indican que las diferencias entre humanos y primates en genes o proteínas son muy pequeñas”, apuntan los investigadores. David de Juan, autor del trabajo e investigador del laboratorio de Valencia sostiene que “el número de genes –originados de nuevo en términos evolutivos– que separan a los humanos de los ratones–[anteriores a los primates en la escala evolutiva– podría ser inferior a diez”.

Esta cifra contrasta con los más de 500 genes humanos originados de nuevo desde primates según datos procedentes de la anotación actual. Los investigadores concluyen: “Las diferencias fisiológicas y de desarrollo entre primates son sobre todo en la regulación y no en las funciones básicas de las proteínas”.

Hacer más con menos

La fuentes de complejidad humana radican más que en el número de genes en cómo estos son utilizados, en las miles de modificaciones químicas que se producen en las proteínas o, según los últimos descubrimientos del proyecto internacional ENCODE, en el control de la producción de estas proteínas por las regiones no codificantes del genoma –incluidas en el catalogado anteriormente como ADN basura, que compone el 90% de la totalidad del genoma–.

Los resultados sacaron a la luz poco más de 12.000
proteínas, que casaron con las correspondientes
regiones del genoma

El trabajo asemeja el número de genes humanos al de otras especies como la de los nemátodos *Caenorhabditis elegans*, unos gusanos de apenas 1 mm aparentemente menos complejos que el ser humano. Pero Valencia prefiere no hacer comparaciones: “Con este estudio ponemos en duda que todo lo que está anotado esté bien, por lo que tendremos que rehacer todos los cálculos, y no solo los del genoma humano”.

Los resultados de la investigación forman parte de GENCODE, consorcio científico integrado en ENCODE en el que participan varios grupos de

investigación de todo el mundo –entre ellos el equipo de Valencia– para introducir las nuevas anotaciones del genoma humano.

“Nuestros datos están siendo discutidos por GENCODE para incorporarlos en las nuevas anotaciones. Cuando esto ocurra habrá que redefinir toda la cartografía del genoma humano, que está siendo utilizado en macroproyectos como los relativos al análisis del genoma del cáncer”, sentencia Valencia.

El estudio es el resultado de una colaboración multidisciplinar que incluye análisis bioinformáticos, genéticos y proteómicos de última generación, en el que han participado investigadores del Centro Nacional de Investigaciones Cardiovasculares (CNIC), además del Wellcome Trust Sanger Institute en Cambridge, Reino Unido, y de la Universidad de California Santa Cruz en EEUU.

Referencia bibliográfica:

Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. Ezkurdia I, Juan D, Rodriguez JM, Frankish A, Diekhans M, Harrow J, Vazquez J, Valencia A, Tress ML. Human Molecular Genetics (2014). doi: 10.1093/hmg/ddu309

Derechos: **Creative Commons**

TAGS

GENOMA | GEN | PROTEÍNA | ADN |

Creative Commons 4.0

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

