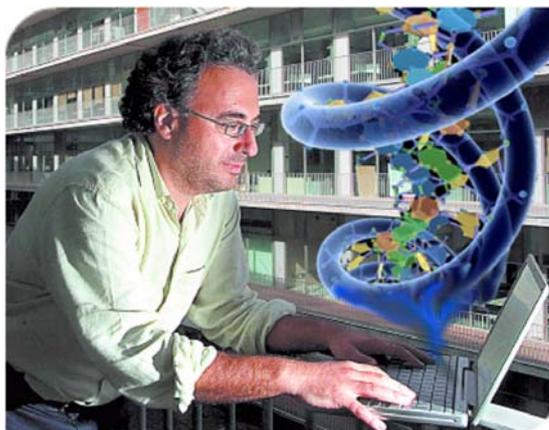


Roderic Guigó, cazador de genes.



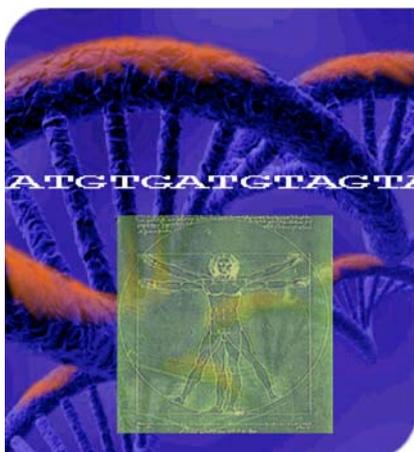
"las secuencias que controlan la actividad de los **genes** están situadas justamente en el DNA basura, así que no es cierto que este no tenga ninguna función"

"Nuestro grupo investiga las señales de la secuencia de DNA que definen dónde se encuentran los genes y, a partir de dichas señales, se desarrollan programas informáticos que, dada una secuencia de DNA, sitúen los genes que codifiquen una proteína"



Entrevista a Roderic Guigó, investigador de el Centre de Regulació **genómica** (CRG) de

Barcelona. Roderic Guigó es un cazador de genes: analiza las secuencias de **ADN** conseguidas con el Proyecto Genoma Humano buscando en qué parte están escondidos los mensajes de la vida.



¿Qué proceso ha seguido la empresa Celera Genomics para **secuenciar el genoma**?

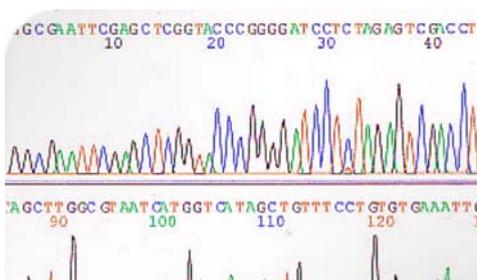
La estrategia que utiliza Celera para secuenciar el genoma de los organismos se denomina *Shot Gun*. Primero el DNA se rompe aleatoriamente en trozos muy pequeños, ya que las máquinas sólo pueden secuenciar fragmentos de 500 o 1000 **nucleótidos**. En el caso del genoma humano, se han generado 50 millones de fragmentos, lo que supone que cada **nucleótido** se halla en promedio en 10 fragmentos distintos. Así, los distintos trozos se pueden superponer por los fragmentos comunes, lo que permite la reconstrucción lineal del genoma. Cada uno de los fragmentos se secuencia en unas máquinas que, básicamente, utilizan el método químico de **secuenciación** de Sanger de los años setenta, sólo que lo hacen automáticamente mediante robots.

exhaustivamente?

Cuando se procede a secuenciar el genoma de un organismo, ¿se hace

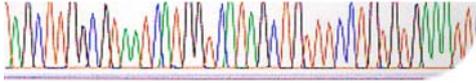
En el caso de la mosca *Drosophila* no ha sido secuenciada una tercera parte de su genoma, la llamada heterocromatina, donde se cree que no hay genes. La heterocromatina de *Drosophila* está muy localizada en unos determinados fragmentos del genoma, como los centrómeros (puntos de unión de los **chromosomas**) y los telómeros (extremos de los cromosomas); y es difícil **clonar** estas regiones por la estructura que el DNA adopta en ellas.

Sólo el 2% del genoma humano codifica para proteínas, lo que implica que un gen está separado de otro por centenares de miles de bases que, en principio, no tienen ninguna funcionalidad. Las regiones repetitivas están entre los genes y, a su vez, las secuencias que controlan la actividad de los genes están situadas justamente en el *DNA basura*, así que no es cierto que este no tenga ninguna función.



En el Proyecto Genoma Humano se utilizan clones de muchos individuos distintos. Celera ha utilizado los genomas de unas personas determinadas. ¿Qué diferencia hay?

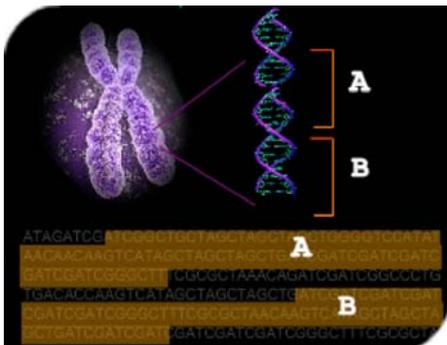
Si se encuentra el cambio de un nucleótido en dos secuenciaci3n diferentes del DNA de un mismo individuo, probablemente se deba a un error de secuenciación. Ahora bien, si se comparan las secuencias de dos individuos diferentes, no se sabe si se trata de un error de secuenciación o de un polimorfismo (es decir, un gen que presenta formas distintas). Secuenciar el DNA de individuos determinados, como hace Celera, permite ver dónde hay



variaciones dentro del DNA.

¿Qué paso viene después del ensamblaje de los fragmentos de DNA secuenciados?

Después del ensamblaje de la secuencia de DNA, se debe determinar dónde se encuentran los genes en dicha secuencia. Nuestro grupo investiga las señales de la secuencia de DNA que definen dónde se encuentran los genes y, a partir de dichas señales, se desarrollan programas informáticos que, dada una secuencia de DNA, sitúan los genes que codifiquen una proteína. Actualmente estamos desarrollando la versión para el genoma humano, llamada *Geneid*, que prevemos hacerla pública en los próximos dos meses. Al margen de esto, hemos hecho un programa llamado *Gff2ps* que permite visualizar estas anotaciones. Cuando se ha predeterminado dónde están los genes, este programa hace un mapa. Este programa informático se hizo público a inicios de año y ya ha sido utilizado para interpretar el genoma de la mosca *Drosophila* publicado recientemente en la revista *Science*.



¿Cómo se puede determinar dónde hay un gen, es decir, un trozo de DNA codificante?

Determinadas secuencias de nucleótidos son señales de la célula que indican dónde hay un gen. Por desgracia, estas señales aparecen con mucha más frecuencia de la que son utilizadas por la célula. Sin embargo, las regiones del DNA donde se encuentran los genes tienen unas características en su composición diferente de aquellas regiones donde no se encuentran los genes. Estas características pueden ser detectadas estadísticamente porque los genes tienden a aparecer más en regiones en las que predominan los nucleótidos G y C que en las que hay más A y T. Por otra parte, se puede comparar una secuencia de DNA con bases de datos de genes conocidos; si coinciden, se determina que la secuencia es codificante.

¿Cuáles son estas bases de datos con genes secuenciados?

Las bases de datos llamadas [EST \(Expression Sequence Tags\)](#) contienen una gran parte de los genes humanos, que han sido obtenidos a partir de la secuenciación de los RNA mensajeros. Es así porque durante la transcripción de DNA en mRNA se eliminan las partes del DNA que no codifican. Estas bases de datos contienen más de 3 millones de secuencias que pertenecen a distintas especies y han sido muy útiles para la secuenciación del genoma humano. No se sabe si estas bases de datos cubren el 50%, el 70% o el 90% de los genes humanos. Seguramente cubren, como mínimo, el 50% de los genes humanos.



Económicamente, ¿es el genoma humano un valor en alza?

La importancia económica de la secuenciación del genoma humano deriva de que hará posible identificar los genes codificados y, en consecuencia, las proteínas que vienen codificadas en esos genes. [...] Muchas enfermedades tienen un componente genético muy importante; conocer el genoma humano permitiría mejorar el tratamiento de algunas enfermedades. Con ello las compañías farmacéuticas podrían obtener muchos beneficios. Además, la secuenciación de especies de interés agrícola o ganadero



permitirá aumentar su rendimiento.

¿Genoma patentado o de libre uso?

En principio, los mecanismos de protección de la propiedad intelectual favorecen que las empresas farmacéuticas inviertan grandes cantidades de recursos en el desarrollo de nuevos medicamentos. Si las patentes se conceden sobre el resultado de investigación esencialmente básica, sin embargo, el resultado puede ser contraproducente, pues se desincentiva la inversión de otras compañías en la investigación de posibles aplicaciones de aquella investigación básica [...] En cualquier caso, no puede ser que las patentes condicionen la investigación básica en los centros de investigación públicos o sin ánimo de lucro; no creo que ello ocurra porque son precisamente las compañías farmacéuticas las primeras interesadas en que esa investigación se produzca libremente.



Esta entrevista es el resultado de extractos de una entrevista realizada por Annia G. Domènech, licenciada en biología, para Biomedica, y extraída de la web de la revista. La entrevista completa está disponible en la web de Biomedica.



[< Anterior](#)

[Següent >](#)