

Rocío Bautista Moreno

Plataforma Andaluza de Bionformática, Universidad de Málaga  
[rociobm@scbi.uma.es](mailto:rociobm@scbi.uma.es)

Cuando Sanger (1), con el método enzimático de terminación de cadenas (método de los didesoxinucleótidos), y Gilbert, con el método de fragmentación química, desarrollaron las primeras aproximaciones a la secuenciación del ADN en los años 70, ninguno de ellos pudo imaginar la velocidad a la cual evolucionaría este proceso treinta años después.

Aunque el método de Maxam y Gilbert permitían secuenciar el ADN original y detectar modificaciones en el mismo, el método de Sanger es el que terminó siendo más popular. En sus orígenes se trataba de un método muy manual, tedioso y peligroso (por el uso de compuestos radioactivos), y con una capacidad de lectura de unas 80 bases. El comienzo de los proyectos de secuenciación en finales de los 90 hizo que esta metodología se adaptara a estrategias menos nocivas gracias a la utilización de didesoxinucleótidos marcados con fluorescencia que se analizaban en una electroforesis capilar y producían un cromatograma o electroferograma, a partir del cual se deducía la secuencia en el ordenador. Esto permitió mejorar, automatizar y aumentar el rendimiento del proceso de secuenciación. Estos nuevos avances posibilitaron el desarrollo de los secuenciadores automáticos de tipo ABI Prism (*Applied Biosystems*) o CEQ-serie (*Beckman Coulter*). Esta tecnología permitía secuenciar a la vez hasta 96 muestras de ADN en unas pocas horas, y la longitud de las secuencias que producía estaba entre 500 y 1000 bases. Estábamos, llegando al límite de lo que se podía conseguir con la primera generación de secuenciadores automáticos. La gran longitud de las lecturas generadas, en comparación con el proceso manual, junto con el desarrollo de las estrategias de secuenciación a gran escala (*Whole Genome Shotgun Sequencing*) facilitaban cada vez más el ensamblaje de las secuencias genómicas.

Empujados por estos avances, en 1990 se firmó el «Proyecto Genoma Humano». Más tarde, en 1995, se publicó el primer genoma secuenciado de un organismo de vida libre, *Haemophilus influenzae* (3). En el 2001, años antes de lo que se tenía previsto, se publicó el primer borrador del genoma humano, que costó casi 3.000 millones de dólares para leer 3.000 millones de nucleótidos (1 \$/nt). Estos costes eran inasumibles para ningún laboratorio, lo que estimuló a los científicos a buscar soluciones más baratas.

En esta búsqueda de abaratar costes se desarrollaron los denominados secuenciadores de segunda generación, capaces de generar cientos de miles de reacciones de secuencias en paralelo (secuenciadores de verdadero alto rendimiento [*high-throughput*]) gracias a la inmovilización de las reacciones en una superficie sólida. De esta forma, la cantidad de reactivos necesarios se minimiza al máximo (podrían llamarse nanorreacciones) y se abarata el coste por base leída. La primera aproximación a la secuenciación masiva en paralelo está basada en la *pirosecuenciación* del ADN (2). Se trata de una técnica no fluorescente que mide la liberación de pirofosfato en una reacción de polimerización mediante una serie de reacciones enzimáticas acopladas que liberan luz cada vez que se incorpora un nucleótido; producen una imagen que se analiza para proporcionar flujogramas que, una vez interpretados por el ordenador, devuelven las secuencias de nucleótidos. Esta tecnología fue desarrollada por la empresa *454 Life Sciences* que fue absorbida poco después por *Roche* ([http://genomics.org/index.php/454\\_GS\\_FLX](http://genomics.org/index.php/454_GS_FLX)). El primer modelo comercial, el GS20, era capaz de secuenciar hasta 20 millones de bases en unas 4 h. Mejorando simplemente la cinética de la reacción química de secuenciación, en el año 2007 apareció el modelo GS-FLX capaz de producir hasta 100 millones de bases en un tiempo similar, en 2009 el GS-FLX-Titanium que genera secuencias de

27

Tabla 1: Comparación de las distintas plataformas de secuenciadores

	Equipo	Compañía	Método de secuenciación	DNA molde	Longitud lecturas (pb)	Tiempo carrera (h)	nt/carrera (Gb)
2ª Generación de secuenciadores	GS-FLX (454)	Roche	Polimerasa (pirosecuenciación)	PCR Emulsión	250-400	10	0,4
	SOLEXA	Illumina	Polimerasa (terminadores reversibles)	PCR Puente	35-75	48	18
	ABI SOLID	Applied Biosystems	Ligasa (octámeros con código de dos bases)	PCR Emulsión	25-75	168	30
3ª Generación de secuenciadores	Helicos tSMS	Helicos BioSciences	Polimerasa	Molécula única	25-45	192	37
	Pacific Biosciences	Pacific Biosciences	Polimerasa	Molécula única	1000	NA	NA
	ZX Genetics	ZX Genetics	Microscopía electrónica	Molécula única	NA	NA	NA

400 bases, lo que duplica la efectividad del modelo anterior, y durante 2010 se espera que lleguen a proporcionar lecturas de más de 600 bases. Como puede inferirse, la longitud de las lecturas generadas se acerca a las obtenidas por el método de Sanger, lo que las convierte en ideales para la secuenciación de genomas *de novo*. Pero su gran inconveniente es que se confunde en las secuencias homopoliméricas.

Al mismo tiempo, otras dos compañías desarrollaron otro tipo de tecnologías para la secuenciación masiva en paralelo del ADN. *Solexa* ([http://www.nature.com/nrg/journal/v11/n1/fig\\_tab/nrg2626\\_F2.html](http://www.nature.com/nrg/journal/v11/n1/fig_tab/nrg2626_F2.html)) utiliza también un método basado en la polimerización del ADN, donde la incorporación de un nucleótido marcado con fluorescencia y protegido en la cadena naciente impide que ésta siga creciendo. Tras detectar la señal fluorescente, se elimina el grupo protector y se puede incorporar otro nucleótido marcado, con lo que se empieza de nuevo el ciclo. La segunda tecnología es la desarrollada por la compañía *SOLiD* (*Sequencing by Oligonucleotide Ligation and Detection*, [http://www3.appliedbiosystems.com/AB\\_Home/applicationtechnologies/SOLiDSystemSequencing/OverviewofSOLiDSequencingChemistry/index.htm](http://www3.appliedbiosystems.com/AB_Home/applicationtechnologies/SOLiDSystemSequencing/OverviewofSOLiDSequencingChemistry/index.htm)) y comercializada por *Applied Biosystems*. Este tecnología secuencia por ligación de octámeros marcados de secuencia conocida a la cadena de ADN, con la posterior detección de la señal fluorescente emitida tras cada ligación. Ambas técnicas, *Solexa* y *SOLiD*, tienen la gran ventaja sobre la pirosecuenciación de resolver de forma fiable las regiones homopoliméricas, y además son más baratas. Sin embargo, su gran desventaja radica en que no son capaces de generar lecturas superiores a las 75 bases, por lo que no se pueden utilizar en las secuenciaciones *de novo*. Estos dos últimos equipos están especialmente diseñados para la resecuenciación de genomas ya conocidos o para estudiar la expresión de

los transcritos. Son capaces de producir gigabases (10<sup>9</sup>) por carrera (Tabla 1) a un coste muy inferior a la pirosecuenciación. Una importante ventaja de todos los secuenciadores de segunda generación es secuencian el ADN sin necesidad de clonarlo primero. Esto es, evita el tedioso trabajo de crear genotecas, replicar los fragmentos en *Escherichia coli* y purificarlos... ¡y conservarlos en el congelador –80! Gracias a la segunda generación de secuenciadores, el coste de cada nucleótido definitivo pasó desde los 10 \$ en 1990 a 0,01 \$ en 2005.

La abaratación por abaratar aún más los costes de secuenciación y aumentar la fiabilidad de las secuencias resultantes ha llevado a los denominados secuenciadores de tercera generación, basados en la secuenciación de una única molécula de ADN (*single molecule real time sequencing*). El primer secuenciador de tercera generación lo vende *Helicos BioSciences* y se basa en la secuenciación a tiempo real de miles de millones de pequeñas moléculas únicas de ADN adheridas a una superficie sólida. Permite generar de forma fiable fragmentos de entre 25 y 45 bases. Dada la pequeñez de las lecturas generadas, esta tecnología está recomendada para la resecuenciación de genomas y no para la secuenciación *de novo*. En un paso más adelante se sitúa la tecnología desarrollada por la empresa *Pacific Biosciences* que promete leer de golpe hasta 1000 nucleótidos, lo que sin duda resolvería todos los problemas asociados a la segunda generación de secuenciadores (esto es: regiones homopoliméricas, repeticiones en tándem, lecturas cortas). Se trata de un enfoque completamente diferente, ya que lo se tiene anclado a una superficie sólida (nanoporo) es la ADN polimerasa. Otra tecnología, encuadrada en los secuenciadores de tercera generación, es la desarrollada por *ZS Genetics*, que utiliza la microscopía electrónica y permite leer la secuencia del ADN directamente sobre una imagen electrónica. La lectura de la secuencia requiere de la replicación previa de una hebra molde de ADN para poder marcarla con bases modificadas con yodo, bromo o triclorometilo antes de analizarlas.

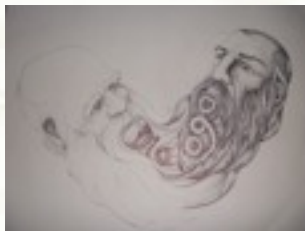
Gracias al avance de la nanotecnología y la microscopía electrónica se ha podido reducir, hasta límites insospechados, la cantidad de reactivos necesarios para lanzar las reacciones de secuenciación, con el consiguiente abaratamiento de los costes. Por tanto, es probable que en un futuro no muy lejano, si la tecnología sigue avanzando al ritmo actual, se haga realidad la utopía de secuenciar de forma fiable el genoma humano en un tiempo récord con un coste inferior a 1000 \$. Además, tampoco podemos olvidar que estas nuevas tecnologías se ya están aplicando a campos como la transcriptómica, el análisis de variabilidad, la identificación de SNP y los análisis de metagenómica.

#### Lecturas recomendadas para saber más:

- Sanger, F., Nicklen S., Coulson, AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl. Acad. Sci. USA*, 74: 5463-5467 (1977)
- Mostafa Ronaghi, Samer Karamohamed, Bertil Pettersson, Mathias Uhlén and Pål Nyrén. Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. *Analytical Biochemistry*. 242: 84-89 (1996)
- Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*. 269:496-512 (1995)

**Lecturas recomendadas para saber más:**

- Tort P. Para leer a Darwin. Alianza, Barcelona, 2001.
- Gould SJ. La estructura de la teoría de la evolución. Tusquets, Barcelona, 2004.
- Darwin C. La variación de los animales y las plantas bajo domesticación. CSIC, Madrid, 2008.
- Darwin C. El origen de las especies : por medio de la selección natural o la conservación de las razas favorecidas en la lucha por la existencia. 6º Ed. Extramuros, Mairena del Aljarafe, 2009.



**Luis Rodríguez Caso**  
Ayudante de Investigación de OPIs  
Estación Experimental La Mayora  
CSIC, Algarrobo Costa (Málaga)  
[caso@eelm.csic.es](mailto:caso@eelm.csic.es)

26



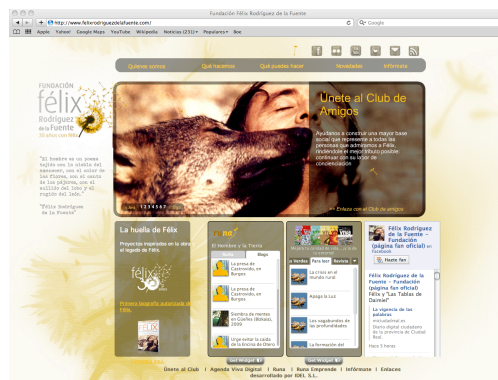
## Foros de la ciencia



### Treinta años con Félix:

Treinta años con Félix es el lema de la Fundación Félix Rodríguez de la Fuente, que mantiene una muy interesante página web que no sólo pretende mantener viva la memoria y la obra de Félix, sino que impulsa proyectos inspirados en la obra y legado de Félix y ofrece información y noticias de interés para todos los identificados con planteamientos conservacionistas. Más información en:

[www.felixrodriguezdelafuente.com](http://www.felixrodriguezdelafuente.com)



con un especial multimedia denominado "Treinta años sin Félix" que incluye la recuperación de *El Hombre y la Tierra*, de todos los audios de la serie de RNE *La Aventura de la Vida*, la realización de la serie documental *La Huella de Félix*, en la que diferentes personajes ilustres recuerdan la figura del naturalista, y más información de interés alojada en el espacio:

[www.rtve.es/treinta-anos-sin-felix](http://www.rtve.es/treinta-anos-sin-felix)

### Un blog sobre Félix en *La Coctelera*:

Para finalizar este Foros de la Ciencia sobre espacios web que homenajean a Félix Rodríguez de la Fuente, cabe mencionar "el blog de Felix Rodríguez de la Fuente" alojado en el espacio web *La Coctelera*:

[www.felixrodriguezdelafuente.lacocctelera.net](http://www.felixrodriguezdelafuente.lacocctelera.net)



### Especial "Treinta años sin Félix":

RTVE esta celebrando el treinta aniversario de la muerte de Félix Rodríguez de la Fuente

Miguel Ángel Medina [medina@uma.es](mailto:medina@uma.es)