

# TÉCNICAS DE BIOLOGÍA MOLECULAR PARA LA INVESTIGACIÓN EN ODONTOLOGÍA Y BIOLOGÍA ORAL (2a PARTE)

## MOLECULAR BIOLOGY TECHNIQUES FOR RESEARCH IN DENTISTRY AND ORAL BIOLOGY (PART TWO)

Yarzabal Rodríguez Luis Andrés<sup>1\*</sup>, Buela Salazar Lenys<sup>2</sup>, Sarmiento Ordoñez Jéssica<sup>1</sup>

<sup>1</sup> Docente/Investigador(a), Carrera de Odontología de la Universidad Católica de Cuenca. Cuenca (Ecuador).

<sup>2</sup> Docente/Investigadora, Carrera de Biofarmacia de la Universidad Católica de Cuenca. Cuenca (Ecuador).

\* lyarzabalr@ucacue.edu.ec

### Resumen

La Biología Molecular se dedica al estudio de la estructura y la función de dos importantes biomoléculas –proteínas y ácidos nucleicos- y de las interacciones que se establecen entre ellas. En el primer artículo de esta serie presentamos algunas de las técnicas “moleculares” y sus aplicaciones en el campo odontológico. En este segundo artículo presentamos los métodos de secuenciación de ácidos nucleicos que han permitido conocer en profundidad los detalles acerca de la forma en que se almacena y se descodifica la información genética en los seres vivos. Las secuencias obtenidas a partir del estudio de genes y genomas son útiles, no solo para identificar y clasificar seres vivos (entre los cuales microorganismos causantes de enfermedades) sino para comprender las bases genéticas de muchas enfermedades. Presentamos igualmente las estrategias de secuenciación masiva en paralelo (también llamadas Tecnologías de Secuenciación de Próxima Generación, o NGS por sus siglas en inglés) que han permitido dilucidar la compleja estructura y composición de las comunidades de microorganismos que colonizan la cavidad oral.

**Palabras clave:** Biología molecular; ADN; ARN; Biología oral; Diagnóstico.

### Abstract

*Molecular Biology is dedicated to the study of the structure and function of two important biomolecules - proteins and nucleic acids - and the interactions established between them. In the first article of this series, we presented some of the “molecular” techniques and their applications in the dental field. In this second article, we present the nucleic acid sequencing methods that have allowed us to know in depth the details about the way in which genetic information is stored and decoded in living beings. The sequences obtained from the study of genes and genomes are useful, not only to identify and classify living beings (among which disease-causing microorganisms) but also to understand the genetic basis of many diseases. We also present the massive parallel sequencing strategies (also called Next Generation Sequencing Technologies, or NGS) that have allowed us to elucidate the complex structure and composition of the communities of microorganisms that colonize the oral cavity.*

**Key words:** Molecular Biology; DNA; RNA; Oral Biology; Diagnosis.

## 1 INTRODUCCIÓN

En la primera parte de esta serie de artículos, hemos presentado brevemente las características más importantes de la estructura y función del material genético de todos los seres vivos: el ácido desoxirribonucleico (ADN).<sup>1</sup> Se trata, en efecto, de la molécula biológica que mayor relevancia tiene cuando hacemos referencia a las técnicas “moleculares” que nos han permitido profundizar en la comprensión de diferentes fenómenos biológicos, desde identificar la etiología de una determinada enfermedad, hasta conocer la

magnitud de una infección, ya sea de tipo bacteriano o viral. Hablamos de “biología molecular” cuando nos referimos al estudio de las bases moleculares de la actividad biológica, pero también cuando profundizamos en las interacciones que ocurren entre los ácidos nucleicos y las proteínas.

Las técnicas moleculares se han nutrido, a lo largo del tiempo, de otras técnicas que originalmente se desarrollaron para ser aplicadas en los campos de la bioquímica, la genética y la biología celular, tales como la electroforesis,

la purificación de moléculas biológicas y la digestión enzimática. En décadas recientes, además del desarrollo de técnicas moleculares “húmedas” –así llamadas porque son las que se ponen en práctica en los laboratorios de investigación, empleando generalmente reactivos y soluciones acuosas–, el surgimiento y consolidación de la bioinformática ha significado un impulso inesperado, en el sentido de aportar una nueva dimensión al estudio de los fenómenos antes mencionados.

La genómica y la metagenómica –entre otras tantas ciencias “ómicas”– son algunas de las nuevas disciplinas que han surgido a consecuencia de la cooperación estrecha entre biólogos moleculares e ingenieros de sistemas. Varias de las nuevas técnicas desarrolladas en el contexto de esta interacción han tenido un profundo impacto en diferentes ciencias biomédicas.

En el presente artículo profundizaremos en torno a estos avances, destacando el papel que ha tenido el desarrollo de nuevas técnicas de investigación y estudio en la comprensión de las enfermedades bucodentales. Muchas de estas técnicas están actualmente al alcance de los investigadores que trabajan en nuestras instituciones latinoamericanas, para llevar adelante proyectos de investigación, pese a no contar con el mismo poderío financiero que nuestros colegas de otros países. De allí que hayamos optado por presentarlas en este artículo.

## 2 ¿QUÉ ENTENDEMOS POR BIOINFORMÁTICA?

La aplicación de técnicas de computación y estadística para el análisis de grandes cantidades de datos biológicos se conoce comúnmente con el nombre de “bioinformática”.<sup>2</sup> Aunque sus orígenes se remontan a la década de los 60' del siglo pasado, no será sino hasta bien entrada la década de los 70' que se comenzaría a extender y popularizar su uso entre los investigadores, particularmente entre aquellos interesados en descifrar los secretos del genoma de los seres vivos. Una de las primeras aplicaciones de la incipiente bioinformática consistió en comparar las secuencias de pequeños fragmentos de ADN, cuyas longitudes no superaban los pocos miles de nucleótidos, obtenidas mediante las técnicas de secuenciación desarrolladas por Frederick Sanger y su grupo.<sup>3,4</sup>

A partir de estos trabajos pioneros, se comprendió cabalmente el poder de esta nueva tecnología para, i) alinear entre sí y comparar secuencias de nucleótidos (genes o fragmentos de genes obtenidos a partir de diferentes tipos de seres vivos); ii) predecir la función de genes a partir de secuencias obtenidas al secuenciar genomas o cromosomas completos; iii) ensamblar genomas simples (primero) y mucho más complejos (después) a partir de cortas secuencias nucleotídicas obtenidas aleatoriamente; iv) predecir la

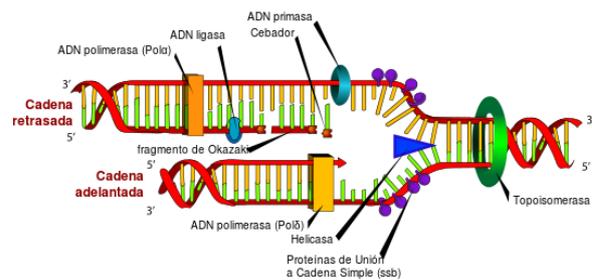
estructura y función de las proteínas codificadas por los mencionados genes; y v) comparar los genomas de distintos seres vivos para establecer sus posibles relaciones evolutivas (lo que denominamos “filogenia molecular”), entre muchas otras aplicaciones.

La bioinformática, que en sus inicios estaba circunscrita a los laboratorios con tecnología avanzada y recursos financieros abundantes, está hoy al alcance de cualquiera que disponga de una computadora personal con buen rendimiento y acceso a Internet de banda ancha. Es gracias a su desarrollo que han podido surgir esas nuevas técnicas “ómicas” a las que hacíamos referencia más arriba.

## 3 SECUENCIACIÓN DE GENES Y GENOMAS

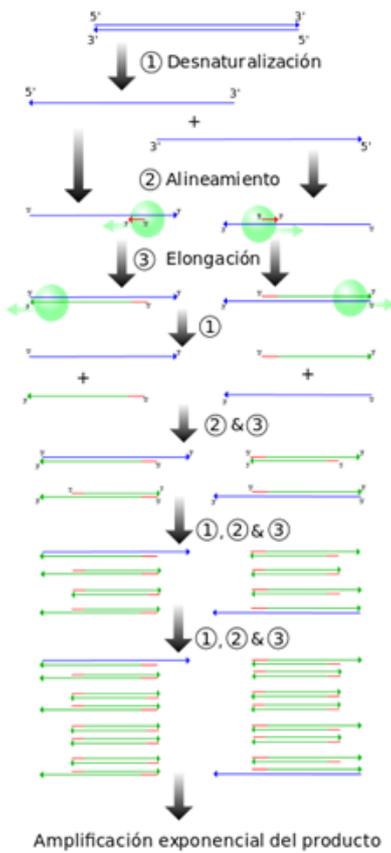
Antes de entrar en los detalles de la genómica y la metagenómica, es necesario hacer un paréntesis para entender cómo se puede obtener la secuencia nucleotídica de un gen o de un genoma completo (lo que llamamos “secuenciar”), y para qué puede ser útil tal información. El método original de secuenciación, desarrollado por Sanger y sus colaboradores<sup>5</sup> –y que por cierto se impuso frente a otros métodos más tediosos y menos eficientes– tenía cierta similitud con la técnica de PCR que describimos en el artículo anterior; esto no es de extrañar, pues ambos métodos se basan en el proceso de replicación de la molécula de ADN, llevada a cabo por una enzima denominada ADN polimerasa.

Esta enzima se encarga de producir una copia exacta de la molécula en doble hélice de ADN, empleando como “molde” a las dos hebras originales (parentales) que la conforman. Para ello utiliza los 4(*desoxi*)nucleótidos (dATP, dCTP, dTTP y dGTP), los cuales une mediante enlaces químicos covalentes (Figura 1).



**Fig. 1. Replicación de ADN.** También llamado síntesis o polimerización, es el proceso mediante el cual cada una de las dos cadenas (hebras) de la molécula parental de ADN son copiadas por una serie de enzimas, entre las cuales destaca la ADN polimerasa. Se destaca también la enzima primasa, encargada de sintetizar el pequeño fragmento de ARN cebador (*primer*) que se menciona en el texto. (Tomado de: *DNA Replication* (esquema bajo licencia de Dominio Público))

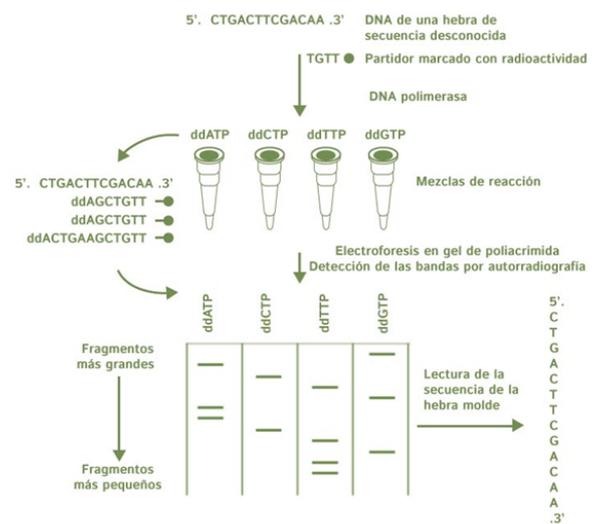
Para que la ADN polimerasa pueda llevar a cabo su función, es necesaria la presencia de un pequeño fragmento de ADN (o de ARN en el caso de células vivas) denominado “cebador” (“primer” en inglés): es a partir de este pequeño fragmento que la ADN polimerasa sintetiza (polimeriza) el resto de la molécula. Precisamente esta particularidad fue la que permitió el desarrollo de las técnicas de secuenciación, primero, y de PCR, más adelante. Teniendo en nuestro poder el “cebador” correcto, es decir, aquel que se asocia de manera específica con su secuencia complementaria en el genoma de una especie, podemos en teoría replicar (o amplificar) cualquier región de un cromosoma o un genoma, de manera perfectamente dirigida y específica (Figura 2).



**Fig. 2. Reacción de PCR.** Amplificación de una región del genoma de cualquier ser vivo mediante Reacción en Cadena de la Polimerasa (PCR). Esta amplificación (o multiplicación) es llevada a cabo por una ADN polimerasa (óvalo verde), a partir de un pequeño cebador (línea roja) en diferentes etapas idénticas (llamados ciclos). Al final de 30 ciclos, el fragmento original flanqueado por los dos cebadores se puede amplificar miles de millones de veces. (Tomado de: De Retama - self-made, Madprime-based. CC BY-SA 4.0.)

Sanger introdujo en esta técnica de secuenciación una modificación genial que determinó el éxito que ha perdurado

a lo largo de más de 45 años: además de los cuatro nucleótidos “normales” introdujo en la mezcla de reacción (llevada a cabo en cada uno de cuatro tubos de ensayo diferentes donde ocurría la polimerización) un nucleótido modificado químicamente. Esto hace que, cuando ese nucleótido es incorporado por la polimerasa, la reacción de polimerización se detiene y no continúa. Sabiendo cual es el nucleótido modificado en cuestión, la genialidad consistió en determinar la longitud de los fragmentos de ADN que se extendían desde el cebador utilizado. Para ello se empleó una técnica ampliamente utilizada en bioquímica: la electroforesis (Figura 3). La visualización de los diferentes fragmentos de ADN en el gel de acrilamida transparente se lograba mediante autorradiografía, pues el cebador empleado había sido marcado radioactivamente.



**Fig. 3. Método de secuenciación de ADN desarrollado por Frederick Sanger.** El fragmento de ADN a secuenciar es sintetizado por una ADN polimerasa a partir de un pequeño cebador o partidor (lo cual determina la especificidad de la región a secuenciar). Este cebador está marcado con radiactividad, lo cual hace que todos los fragmentos sintetizados serán radioactivos. En 4 tubos de ensayo distintos se llevan a cabo 4 reacciones similares de síntesis: en dichos tubos se incluyen los 4 nucleótidos (dATP, dTTP, dCTP y dGTP) y un nucleótido modificado (ddATP, ddTTP, ddCTO ó ddTTP). La reacción de síntesis se detendrá en un punto al azar, cada cuando la polimerasa incorpore un nucleótido modificado en lugar del nucleótido normal. Los productos de esta síntesis se separan por electroforesis, se visualizan por autorradiografía y se lleva a cabo la “lectura” de la secuencia. (Tomado de: Portal médico)

En sus orígenes, la secuencia nucleotídica se obtenía en forma manual; hoy en día, existen aparatos automatizados, que permiten secuenciar ADN de manera rápida, segura (pues ya no se emplea radioactividad sino fluorescencia) y con una enorme precisión (pues la electroforesis se lleva a cabo en el interior de un capilar, lo cual permite una mayor resolución de los fragmentos, que difieren en su longitud en

un solo nucleótido) (Figura 4).

La secuenciación de genomas emplea una aproximación similar, con la salvedad de que se trata de moléculas de ADN mucho más largas, pudiendo alcanzar en ocasiones varias decenas de millones de nucleótidos de longitud. En su momento hablaremos de cómo fue que, partiendo de una estrategia muy tediosa, onerosa y compleja basada en la metodología de Sanger, se ha logrado simplificar y abaratar -de manera exponencial- el costo de esta secuenciación. Por lo demás, existen grandes compañías a nivel mundial que se dedican a prestar este tipo de servicios, lo cual ha reducido aún más el tiempo de procesamiento de las muestras y abaratado los costos del servicio, mejorando significativamente la calidad de dichas secuencias en términos de longitud y confiabilidad.

#### 4 GENÓMICA

El primer genoma viral se secuenció a mediados de los años 70s del siglo pasado: se trataba del genoma de un virus que infecta bacterias, un bacteriófago. Sin embargo, transcurrieron casi 20 años hasta que se lograra secuenciar el primer genoma bacteriano.<sup>6</sup> En efecto, los genomas virales son relativamente pequeños: dependiendo del tipo de virus, los valores extremos oscilan entre los 5,000 y los 500,000 nucleótidos aproximadamente (aunque hay excepciones, como los virus gigantes de los protozoarios). En el caso de las bacterias, sus genomas son mucho más grandes y complejos, pudiendo variar entre los 500mil y los 12millones de nucleótidos. Obviamente, el análisis de tanta información no se puede efectuar por métodos manuales: ¡allí es donde interviene la bioinformática!

A partir de la secuencia completa del genoma de cualquier especie es posible deducir el número de genes que lo componen, pero más importante aún, el "tipo" de genes que están presentes en dicho genoma (haciendo referencia a su función). En el caso de bacterias de interés médico, esto puede ser muy útil, pues se pueden identificar rápidamente genes relacionados con la virulencia de la especie (producción de toxinas o enzimas hidrolíticas, capaces de generar daño tisular) o genes que determinan la resistencia frente a compuestos antimicrobianos. Esto es posible gracias a la eficiencia y rapidez con la que los algoritmos desarrollados por los ingenieros de sistemas pueden comparar unas secuencias nucleotídicas desconocidas (por ejemplo, las que podríamos obtener a través del estudio de un microorganismo de interés) con otras de referencia, almacenadas en un banco de secuencias donde las mismas están claramente identificadas. Por supuesto que construir este tipo de bancos de datos toma mucho tiempo, pero ello ha sido posible gracias al trabajo de decenas de miles de investigadores en el mundo entero a lo largo de varias décadas. Actualmente, la base de datos más utilizada, GenBank, dispone de más de 212millones de secuencias

nucleotídicas<sup>7</sup> y su uso es completamente gratuito.

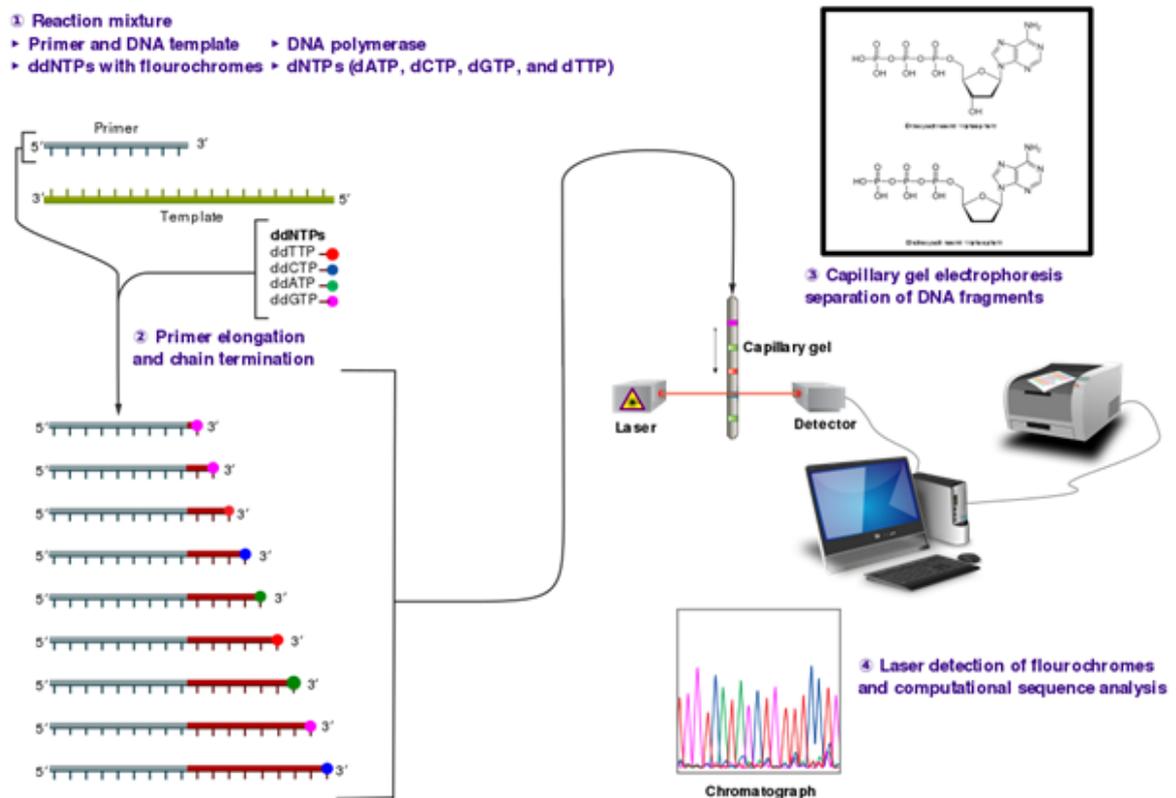
#### 5 TAXONOMÍA Y FILOGENIA

Cuando se dispone de tanta información, convenientemente almacenada en bases de datos accesibles en forma gratuita y universal, se pueden hacer muchas otras cosas. Por ejemplo, si disponemos de la secuencia nucleotídica de un gen que está presente en todas las especies de bacterias conocidas (se le llama "gen universal" o "marcador universal"), es posible que al compararla con las secuencias depositadas en la base de datos obtengamos una lista de las especies que presentan mayor similitud.

Por supuesto, esto solo es posible si se dispone de un algoritmo capaz de hacer el trabajo de búsqueda y comparación lo cual, afortunadamente, tenemos. Uno de los recursos más utilizados del NCBI es la herramienta Blast (Basic Local Alignment Search Tool),<sup>8</sup> la cual permite básicamente alinear secuencias y compararlas. Esta herramienta se utiliza comúnmente para identificar secuencias homólogas (no necesariamente idénticas, pero muy parecidas) que, en principio, codificarían productos (proteínas, en el caso de genes codificantes) con funciones similares en distintas especies de seres vivos. Esto se logra en pocos segundos.

Ya habíamos mencionado, en la entrega anterior, la importancia de la técnica de PCR para el diagnóstico molecular. Sin embargo, muchas veces la PCR es sólo el primer paso en la identificación del microorganismo causante de una determinada patología. Muchas veces no disponemos de marcadores (genes) específicos para tal o cual especie de microorganismo; en ese caso, debemos amplificar primero un gen universal (como el que codifica para el *ADNr16S* en el caso de las bacterias) para luego secuenciarlo y, finalmente, analizar su secuencia.

Mediante el empleo de herramientas como Blast es posible asignar una secuencia a un determinado taxón (grupo de organismos relacionados entre sí). Esto puede hacerse a nivel de especie, género, familia, clase, etc. De esta forma, aún sin haber realizado ningún tipo de estudio adicional, podríamos identificar de manera preliminar la especie a partir de la cual logramos amplificar (o secuenciar) dicha secuencia. A esto se le denomina "identificación molecular". También podríamos identificar, en los microorganismos que hayamos podido aislar a partir muestras de nuestros pacientes, genes relacionados con la manifestación de rasgos de virulencia, o con la resistencia a determinados agentes antimicrobianos (lo que habíamos incluido en la categoría de "epidemiología molecular" en el artículo anterior). Por supuesto, la estrategia de análisis sería la misma que



**Fig. 4. Método de secuenciación automatizada de ADN según el método de Sanger.** El método, muy similar al anterior en lo que se refiere a la síntesis de fragmentos de ADN a partir de cebadores específicos, incluye un último paso de separación de los mismos mediante electroforesis capilar (una variante de la técnica clásica). El secuenciador registra de manera automática el paso de los diferentes fragmentos (marcados esta vez con fluorescencia y no con radioactividad) mediante un rayo láser que incide sobre una fotocelda. La información se registra en una computadora bajo la forma de picos de absorción (absorbancia) que son traducidos en secuencias de nucleótidos. (Tomado de: Estevezj - Obra propia, CC BY-SA 3.0)

describimos más arriba.

Pero incluso podemos ir más lejos, pues haciendo uso de otra serie de algoritmos públicos y gratuitos (como MEGA 6,<sup>9</sup> por ejemplo) es posible establecer la filogenia –esto es, las relaciones evolutivas entre los seres vivos- de los microorganismos que hayamos podido aislar a partir de los tejidos de nuestros pacientes. Si bien es cierto que la taxonomía y filogenia moleculares no son suficientes como para establecer con total seguridad la identidad de un determinado microorganismo, son extremadamente útiles en el proceso de lograrlo. Esto es fundamental, por ejemplo, en el caso de estudios poblacionales, en los que no es tan relevante conocer las particularidades de cada individuo o de cada cepa microbiana, como las características comunes de las poblaciones de microorganismos que circulan en una determinada comunidad.

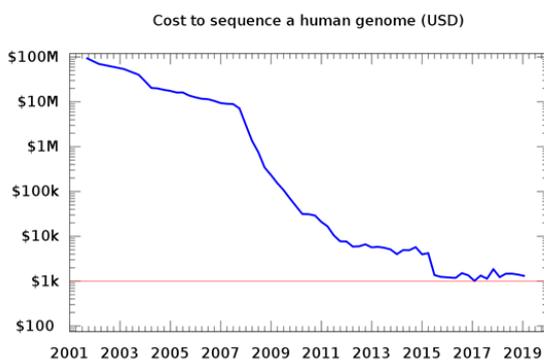
## 6 SECUENCIACIÓN MASIVA EN PARALELO: METAGENÓMICA Y DNA METABARCODING

Mediante el empleo de la técnica de secuenciación de Sanger, automatizada y mejorada, logramos grandes avances en el estudio de genes y genomas. No obstante, los costos inherentes a la secuenciación de moléculas de ADN seguían siendo muy elevados, aún a principios del siglo XXI. De forma tal que pretender secuenciar muchos fragmentos era una meta que estaba al alcance de pocos laboratorios en el mundo. Esto significaba, además, que secuenciar genomas completos o estudiar comunidades complejas de microorganismos, formadas por cientos de especies, era no solamente un reto desde el punto de vista experimental, sino sobre todo una utopía financiera. Esto cambió radicalmente hacia el año 2005, con el desarrollo de lo que se denominó las Técnicas de Secuenciación de Próxima Generación (Next Generation Sequencing Techniques, o simplemente NGS).

Explicar el fundamento de estas técnicas va mucho más allá del objeto del presente artículo. No obstante, conviene

destacar algunos detalles importantes:

- En primer lugar, se trata de técnicas de secuenciación “masiva” o “en paralelo”, en las que simultáneamente se pueden secuenciar millones de fragmentos de ADN, los cuales han sido amplificados previamente mediante variantes de la técnica de PCR;
- Por otro lado, al incorporar secuencias de pocos nucleótidos al extremo de los fragmentos amplificados (lo que se denomina “código de barras” molecular o metabarcoding) es posible distinguir entre grupos de secuencias provenientes de distintas muestras, aunque se secuencien simultánea y masivamente (por ejemplo, secuencias provenientes de distintos pacientes, distintos ambientes, distintos tratamientos, etc);
- Más aún, el tiempo de obtención de resultados es muy corto pues, al secuenciarse en paralelo, se obtienen grandes cantidades de datos (secuencias) en muy poco tiempo (horas); dependiendo de la plataforma empleada, se pueden generar 4 millones de secuencias en 18 horas o 400 millones en 30 horas;
- Finalmente, el costo de secuenciación se ha reducido en forma exponencial, de forma que lo que en 2001 costaba secuenciar 1 millón de nucleótidos (aprox. 8,000USD), hoy en día cuesta 800,000 veces menos (aprox. 0,02USD) (Figura 5).<sup>10</sup>



**Fig. 5. Costo de secuenciación de un genoma humano completo a lo largo de las dos últimas décadas.** Se grafica el costo (en miles de dólares) versus los años transcurridos desde 2001 (año en el que se reportó por primera vez la secuencia completa de un genoma humano) hasta el momento actual. La variación de la pendiente de la curva (línea azul) que se observa a partir del año 2007 es consecuencia directa de la introducción en el mercado de las plataformas de secuenciación de próxima generación (NGS). En la actualidad, algunas compañías ofrecen secuenciar el genoma de una persona por menos de 700 USD. (Tomado de: [Original de Ben Moore, reformateado en Gnuplot por GrendellKhan](#))

El desarrollo de estas nuevas tecnologías (se trata de distintas plataformas, cada una con particularidades que varían sustancialmente) ha permitido impulsar al máximo

una nueva disciplina científica: la metagenómica. Se trata del estudio (a través de su secuenciación) del metagenoma, es decir, del conjunto que resulta al combinar todos los genomas pertenecientes a la totalidad de especies diferentes que colonizan un determinado ambiente.

Pongamos, por ejemplo, la cavidad oral. Hoy en día sabemos que entre 700 y 1000 especies son capaces de residir en ella.<sup>11</sup> Sin embargo, casi un tercio de estas especies no se pueden cultivar ni estudiar en detalle, y solo han sido identificadas mediante métodos moleculares (razón por la cual son llamadas filotipos y no especies).<sup>12</sup> ¿Cómo hacer en caso de querer estudiar los cambios que sufre dicha comunidad a consecuencia de una enfermedad, o de un cambio de hábitos alimenticios, o como consecuencia del uso de una nueva estrategia de higiene oral...? ¿Cómo incluir a muchos sujetos en estudios de este tipo? ¿Cómo comparar la comunidad de microorganismos que caracterizan a un grupo étnico con la de otros grupos? Para todas estas preguntas la respuesta actual es la secuenciación masiva de marcadores moleculares (también denominada “censo taxonómico”).

En la actualidad todas estas técnicas están al alcance de quienes deseen aventurarse en el estudio de las comunidades microbianas que colonizan la cavidad oral de los individuos. Básicamente se requiere de un laboratorio que cuente con el equipamiento necesario para llevar a cabo la extracción y purificación de ADN a partir de muestras de cavidad oral (generalmente realizado mediante el uso de kits comerciales, estandarizados y optimizados), la amplificación por PCR de marcadores universales (de tipo ADNr 16S, si nos interesan las bacterias, y/o de la región ITS, en el caso de comunidades de hongos), y la electroforesis en geles de agarosa (para verificar el estado y la cantidad del material amplificado).

Al igual que sucede en el caso de la secuenciación automatizada, son muchas las empresas que ofrecen el servicio de secuenciación masiva a costos muy accesibles. La posibilidad de contratar este tipo de servicios garantiza, además, que no se incurrirá en gastos adicionales (personal técnico especializado, mantenimiento de equipos, controles periódicos de calidad, vencimiento de reactivos, etc) y que se podrá cambiar rápidamente de tecnología en caso de que la que se encuentra vigente en un determinado momento sea superada por otra menos costosa, más rápida y eficiente (lo cual ocurre cada 2 – 3 años aproximadamente).<sup>13</sup> Después de todo, lo que nos interesa como investigadores es analizar e interpretar adecuadamente la información generada a partir del empleo de estas técnicas, todas ellas altamente automatizadas.

**CONFLICTO DE INTERESES:** Los autores no manifiesta ningún conflicto de interés.

### Referencias

- 1 Yarzabal L, Buella L, Djabayan P. Técnicas de biología molecular para la investigación en odontología y biología oral (1a PARTE). Revista OACTIVA UC Cuenca.2018;3(1):29-36.
- 2 Can T. Introduction to Bioinformatics. En: Malik Yousef and Jens Allmer (eds.), miRNomics: MicroRNA Biology and Computational Analysis, Methods in Molecular Biology, V. 1107. [DOI 10.1007/978-1-62703-748-8\_4], Springer Science+Business Media New York.2014.
- 3 Staden, R. Sequence data handling by computer. Nucleic Acids Research.1977;4(11):4037-51.
- 4 Sanger F, Barrell B, Brown N, Coulson A, Fiddes J, Hutchinson C y col. The nucleotide sequence of bacteriophage  $\phi$ X174. Nature.1977;265:687-95.
- 5 Sanger F, Coulson A. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology.1975;94(3):441-48.
- 6 Fraser C, Gocayne J, White O, Adams M, Clayton R, Fleischmann R y col. The Minimal Gene Complement of Mycoplasma genitalium. Science.1995;270(5235):5235.
- 7 GenBank Statistics. Disponible en:[web](#)
- 8 Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W y col. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research V.1997;25(17):3389-402.
- 9 Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S y col. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution.2013;30(12):2725-29.
- 10 National Human Genome Research Institute. DNA Sequencing Costs: Data. Junio 2019. Disponible en:[web](#)
- 11 Benn A, Heng N, Broadbent J, Thomson W. Studying the human oral microbiome: challenges and the evolution of solutions. Australian Dental Journal.2018;63(1):14-24.
- 12 Kilian M, Chapple I, Hannig M, Marsh P, Meuric V, Pedersen A y col. The oral microbiome an update for oral healthcare professionals. British Dental Journal.2016;221(10):657-66.
- 13 Kulski J. Next-Generation Sequencing — An Overview of the History, Tools, and “Omic” Applications, Next Generation Sequencing - Advances, Applications and Challenges, Jerzy K Kulski, IntechOpen, 2016. [DOI: 10.5772/61964.] Disponible en: [web](#)

**Recibido:** 01 de Agosto del 2019

**Aceptado:** 15 de Agosto del 2019

