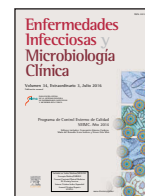




Enfermedades Infecciosas y Microbiología Clínica

www.elsevier.es/eimc



Secuenciación masiva para el diagnóstico y la epidemiología de tuberculosis

Iñaki Comas^{a,b,c,*} y Ana Gil^d

^aUnidad de Genómica de la Tuberculosis, Instituto de Biomedicina de Valencia, Consejo Superior de Investigaciones Científicas, Valencia, España

^bUnidad de Genómica de Tuberculosis, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO), Valencia, España

^cCIBER (Centro de Investigación Biomédica en Red) en Epidemiología y Salud Pública, Instituto de Salud Carlos III, Madrid, España

^dServicio de Microbiología, Hospital Universitario La Fe, Valencia, España

RESUMEN

Palabras clave:

Secuenciación genómica
Mycobacterium tuberculosis
Epidemiología genómica
Genotipado
Resistencia a antibióticos

La tuberculosis (TB) se ha convertido en la enfermedad infecciosa que más personas mata mundialmente por delante del VIH (virus de la inmunodeficiencia humana) y la malaria. La incidencia de TB apenas se reduce en un 2% cada año, por lo que no se podrá eliminar hasta dentro de 200 años. Parte del problema es que las actuales herramientas de control de la TB son antiguas y ya no son suficientes para acelerar la erradicación de la TB. Nuevos métodos son, por tanto, necesarios en los campos de diagnóstico, tratamiento y prevención. La secuenciación genómica masiva puede convertirse en una de esas herramientas. La caracterización genómica de aislados de TB está demostrando su utilidad tanto para entender la epidemiología de la enfermedad como para su diagnóstico, en especial para la detección de resistencias a antibióticos. Sin embargo, las técnicas genómicas y el análisis informático asociado están todavía lejos de usarse de rutina en la mayor parte del mundo. En esta revisión describiremos las actuales tecnologías de secuenciación en el contexto del complejo *Mycobacterium tuberculosis*, las aportaciones de dichas tecnologías al diagnóstico y la epidemiología de la enfermedad, así como los esfuerzos que se están haciendo para dar el salto a su aplicación en el contexto clínico.

© 2016 Elsevier España, S.L.U. Todos los derechos reservados.

Next generation sequencing for the diagnostics and epidemiology of tuberculosis

ABSTRACT

Keywords:

Whole genome sequencing
Mycobacterium tuberculosis
Genomic epidemiology
Genotyping
Antibiotic resistance

Tuberculosis (TB) has overtaken HIV (human immunodeficiency virus) and malaria as the leading cause of death by an infectious disease worldwide. The reduction in the TB incidence is a modest 2% of cases per year, thus we will need 200 years to eradicate the disease. Part of the problem is that TB control tools are decades old and cannot anymore contribute to accelerate eradication of TB. New diagnostics, treatments and vaccines are urgently needed. Next generation sequencing has the potential to become one of these new tools. Genomic characterization of TB isolates is already showing its potential for epidemiology and diagnostics, particularly to identify drug resistance mutations. However, the experimental and bioinformatics skills needed are still far from being standardized and are not easy to incorporate as a routine in clinical laboratories. In this review we will describe current next generation sequencing approaches applied to the *Mycobacterium tuberculosis* complex, their contribution to the diagnostics and epidemiology of the disease and the efforts that are being undertaken to make the technology accessible to public health and clinical microbiology laboratories.

© 2016 Elsevier España, S.L.U. All rights reserved.

Introducción

La identificación de micobacterias por técnicas de biología molecular se inició a finales de los años ochenta del siglo pasado¹. A principios de los noventa, varios grupos de investigación presentaron los resultados obtenidos tras la incorporación de estos métodos para la detección de micobacterias englobadas en el complejo *Mycobacte-*

*Autor para correspondencia.

Correo electrónico: inaki.comas@uv.es (I. Comas).

rium tuberculosis (MTBC)². La conclusión fue que las técnicas moleculares permitían adelantar el diagnóstico de enfermedad tuberculosa frente a los métodos tradicionales de cultivo, aunque nunca sustituirlos como técnica de referencia, debido a que no permitían determinar la infectividad o viabilidad del bacilo, ni las resistencias a los fármacos antituberculosos³. A pesar de que este último aspecto se ha suplido en gran parte con las nuevas tecnologías que permiten detectar genes relacionados con mutaciones asociados a resistencias en un alto porcentaje de cepas, no se ha solucionado por completo¹.

El método de referencia para el diagnóstico de la tuberculosis (TB) es el cultivo, aunque en algunos países la observación directa de bacilos ácido-alcohol resistentes es la única prueba diagnóstica disponible. La microscopía, aunque barata, no es sensible, especialmente en niños y personas que conviven con pacientes positivos para el virus de la inmunodeficiencia humana (VIH); no diferencia entre micobacterias no tuberculosas y especies englobadas en el MTBC ni proporciona información acerca de la sensibilidad. En cultivo en medios habituales de Löwenstein-Jensen y Middlebrook, las micobacterias del MTBC pueden tardar semanas en crecer, retrasándose así los resultados y la determinación de presencia o no de resistencia a los fármacos antituberculosos. Por todo ello, las técnicas moleculares se han ido incorporando al algoritmo diagnóstico de la TB en los diferentes hospitales⁴.

Estas técnicas, además, están teniendo y tendrán un papel clave en la vigilancia y el control epidemiológico de la enfermedad a nivel global. Al ritmo al que se reduce la incidencia global de la enfermedad tardaremos más de 200 años en erradicar la TB. Solo la aparición de nuevas herramientas para el diagnóstico, la prevención y el tratamiento nos hará cumplir los objetivos de erradicación para 2050⁵. En su nueva estrategia, End TB Strategy (2016-2035), la Organización Mundial de la Salud (OMS) incluye como una de sus metas el diagnóstico temprano de la TB, así como un estudio de sensibilidad anti-tuberculosa universal. Los laboratorios tienen, por tanto, un papel crítico en la era posterior a 2015 y en la adaptación de las aproximaciones moleculares. En este último informe mencionan que el nuevo cartucho en desarrollo Xpert Ultra[®] podría incluso reemplazar al cultivo tradicional como primera herramienta en el diagnóstico de la TB. La aparición de tecnologías de secuenciación masiva puede superar las limitaciones de los diagnósticos moleculares actuales y en un futuro ser un método de referencia, como ya lo está empezando a ser para la epidemiología de la enfermedad⁶. En esta revisión se repasará la historia, pero sobre todo el papel presente y futuro de la secuenciación genómica en el diagnóstico y la epidemiología molecular.

Tecnologías de secuenciación masiva aplicadas al bacilo de la tuberculosis

En 1905 una cepa de TB conocida como H37 se aisló asociada a un brote infeccioso en Estados Unidos. Dicha cepa (conocida como H37Rv) y su hermana avirulenta (conocida como H37Ra) se convirtieron en las cepas de referencia más usadas en laboratorios de investigación de todo el mundo. En 1998, H37Rv fue uno de los primeros genomas bacterianos completamente secuenciados y abrió las puertas de la era genómica en el estudio de la TB⁷.

El genoma de H37Rv y su manipulación han permitido un gran número de descubrimientos relevantes para el desarrollo de nuevos antibióticos, vacunas y diagnósticos⁸. La determinación de la secuencia genómica del bacilo de la TB ha permitido estudiar los determinantes genéticos de la resistencia a antibióticos y construir bases de datos asociadas⁹, así como el desarrollo de todo un nuevo conjunto de pruebas moleculares para el diagnóstico de MTBC y para la determinación genotípica de resistencia a los fármacos antituberculosos¹⁴. Por último, el estudio de las diferentes regiones codificantes y no codificantes del genoma ha permitido entender su papel en la biología y la virulencia de la bacteria, siendo el origen de muchas de las vacunas candidatas que están en estudio¹⁰.

Desde el punto de vista de la epidemiología molecular y la microbiología clínica, su impacto ha sido muy importante. Por ejemplo, aunque muchos de los marcadores moleculares que se usan ahora ya existían antes de conocerse la secuencia genómica¹¹, solo a posteriori se han podido caracterizar y poner en el contexto adecuado. Entre los más conocidos se encuentra la región del espoligotipado conocida como CRISPR (*clustered regularly interspaced short palindromic repeats*). Los espoligotipos se usan a veces para definir posibles casos de transmisión, pero sobre todo para clasificar los aislados en diferentes familias génicas¹². Actualmente, el marcador más empleado en epidemiología molecular se conoce como MIRU-VNTR (*mycobacterial interspersed repetitive units-variable number tandem repeats*)¹³. Con este método se analizan diferentes regiones del genoma que se caracterizan por tener un número variable de repeticiones en tándem conocidas como minisatélites. Aunque existe un gran número de minisatélites, por motivos de reproducibilidad se usan 2 grupos estándar conocidos por el número de regiones que interrogan, como MIRU 15 y MIRU 24. Debido a que la frecuencia de cambio genético de dichas zonas (duplicaciones y deleciones) es muy alta, la combinación del número de copias de esos *loci* permite identificar cepas de estrecha relación genética y diferenciarlas de otras que, aunque puedan estar próximas genéticamente, no pertenezcan al mismo grupo de transmisión.

En el año 2005 aparecieron las primeras tecnologías de secuenciación masiva¹⁴. Al contrario que las técnicas usadas hasta ese momento, conocidas como secuenciación por perdigonada (*shotgun sequencing*), estas tecnologías son mucho menos laboriosas al no requerir un paso previo de clonación en *Escherichia coli*. La otra característica fundamental es la cantidad de nucleótidos capaces de determinar en una sola carrera¹⁵. Por poner un ejemplo, actualmente las plataformas de Illumina son las más empleadas en secuenciación bacteriana. El tamaño medio de un genoma de *M. tuberculosis* es de algo más de 4,4 millones de pares de bases (pb) nucleotídicas, o lo que es lo mismo 4,4 megabases. Si solo quisiéramos secuenciar un genoma en una carrera de Illumina MiSeq leeríamos cada base del genoma como media 1.700 veces, por lo que su cobertura sería 1.700×. Con la plataforma de Illumina NextSeq *rapid run* sería 13.600× y en uno de Illumina HiSeq completo 67.700×. Por usar otro ejemplo, en la plataforma de Ion Torrent se podría leer hasta 455 veces cada base del genoma analizado (tabla 1). Por lo tanto, todas ellas generan mucha más cantidad de secuencia que la realmente necesaria para analizar un solo genoma. Lo que realmente se hace es secuenciar decenas o cientos de genomas a la vez, dependiendo de la cantidad de datos generados por cada plataforma. Esta estrategia se denomina *multiplexado* y es la que ha permitido disminuir, en gran medida, los precios de secuenciación genómica. En la tabla 1 se puede consultar el número aproximado de genomas de diferentes cepas del MTBC que se pueden analizar a la vez. Si tuviéramos una colección de 800 cepas la podríamos analizar en una sola carrera con un HiSeq 1500/2500 y el genoma de cada una de ellas se leería como media 80 veces. La cuestión está en que cada cadena de ADN genómico es fragmentada en trozos de entre 500 y 1.000 pb, añadiéndose a cada uno posteriormente una etiqueta nucleotídica que hace que bioinformáticamente se pueda asignar posteriormente cada lectura a la cepa correspondiente. En el laboratorio de FISABIO (Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana) usamos rutinariamente Illumina MiSeq con 20 cepas por carrera. Si solo contamos el coste de las librerías y su secuenciación, sin tener en cuenta la mano de obra, cuesta secuenciar cada cepa entre 80 y 100 euros, aunque va bajando conforme se automatizan.

Otra de las características básicas a la hora de elegir una plataforma de secuenciación es el tamaño de las lecturas que es capaz de generar. Así, se pueden clasificar las plataformas en las que se generan lecturas cortas (entre 100-300 pb) y en las que se generan lecturas largas de hasta varias kilobases (kb) (tabla 1). Las más usadas en el campo de la microbiología clínica y epidemiología de la TB son las

Download English Version:

<https://daneshyari.com/en/article/3400755>

Download Persian Version:

<https://daneshyari.com/article/3400755>

[Daneshyari.com](https://daneshyari.com)