



Enfermedades Infecciosas y Microbiología Clínica

www.elsevier.es/eimc



Formación Médica Continuada: Infecciones por micobacterias

Epidemiología molecular de la tuberculosis[☆]

Pere Coll^{a,b,c,*} y Darío García de Viedma^{d,e,f,*}

^a Servicio Microbiología, Hospital de Sant Pau, Barcelona, España

^b Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, España

^c Institut de Recerca, Hospital de Sant Pau, Barcelona, España

^d Servicio Microbiología Clínica y Enfermedades Infecciosas, Hospital General Universitario Gregorio Marañón, Madrid, España

^e Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

^f CIBER Enfermedades respiratorias, CIBERES, Madrid, España



INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 27 de diciembre de 2017

Aceptado el 13 de enero de 2018

On-line el 17 de febrero de 2018

Palabras clave:

Tuberculosis

Epidemiología molecular

Genotipificación

R E S U M E N

La aplicación de técnicas de genotipificación ha permitido discriminar entre los aislados de *Mycobacterium tuberculosis* obtenidos en el laboratorio. Esta singularización a nivel de cepa abrió las puertas a estudios de epidemiología molecular que han permitido progresar en nuestro conocimiento de la transmisión de este patógeno en entornos socio-epidemiológicos cada vez más complejos. La estabilidad genética de este microorganismo ha llevado al desarrollo de metodologías específicas, que son revisadas en detalle en este capítulo. Además de las aplicaciones epidemiológicas, repasamos cómo dan respuesta asimismo a diversos interrogantes diagnósticos y clínicos. Por último, nos ocupamos de describir la nueva revolución genómica en el estudio de la tuberculosis, los métodos en los que descansa y las posibilidades inéditas que ofrece, incluyendo nuevas vías de integración de los lenguajes moleculares y genómicos en propuestas innovadoras de trabajo posgenómico, más adaptadas a la realidad de nuestro entorno.

© 2018 Elsevier España, S.L.U.

y Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica. Todos los derechos reservados.

Molecular epidemiology of tuberculosis

A B S T R A C T

The application of genotyping tools allowed us to discriminate between the *Mycobacterium tuberculosis* isolates obtained in the laboratory. The differentiation between single strains opened the door to molecular epidemiology studies, which had helped us to progress in our knowledge of how this pathogen is transmitted in the progressively more complex socio-epidemiological scenario. The genetic stability of this microorganism led to develop specific methodologies, which are thoroughly revised in this chapter. In addition to their application in epidemiology, we review, how they can offer a response to different diagnostic and clinical challenges. Finally, we focus on describing the novel genomic revolution we are experiencing in the analysis of tuberculosis, the methodology in which it is based and the novel possibilities it offers, including new routes of integrating both the molecular and genomic languages in innovative post-genomic proposals, better suited to our real-life context.

© 2018 Elsevier España, S.L.U. and Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica. All rights reserved.

Keywords:

Tuberculosis

Molecular epidemiology

Genotyping

Introducción

La epidemiología molecular comprende una gran variedad de técnicas que tienen como objeto comparar secuencias de ácidos nucleicos de 2 o más aislamientos. El término clon o grupo clonal en epidemiología hace referencia al grupo de aislamientos relacionados por el hecho de descender de un ancestro común más o menos lejano. Las cepas relacionadas provienen, pues, de

[☆] “Nota: sección acreditada por el Consell Català de Formació Continuada de les Professions Sanitàries. Consultar preguntas de cada artículo en: <http://www.elsevier.es/eimc/formacion>”.

* Autores para correspondencia.

Correos electrónicos: pcoll@santpau.cat (P. Coll), dgvedma2@gmail.com (D. García de Viedma).

la expansión clonal de un precursor único y poseen un nivel de similitud entre sus genotipos y fenotipos significativamente superior al que se encontraría entre aislamientos no relacionados de la misma especie seleccionados arbitrariamente. La definición de clon es probabilística y el nivel de similitud necesario para su definición debe tener en cuenta el taxón estudiado, el marcador utilizado y el tiempo que dure la investigación epidemiológica. Por ello, antes de escoger un marcador molecular debe formularse de forma clara y precisa aquella pregunta en términos epidemiológicos cuya respuesta es preciso conocer, definir el nivel de relación genética que es necesario determinar para responder a la pregunta anterior, escoger los marcadores que son capaces de discriminar a este nivel de relación y verificar la eficacia de los métodos seleccionados.

Estudios de secuenciación comparativa en *Mycobacterium tuberculosis* (*M. tuberculosis*) han detectado poca variación genética entre cepas, mínimo intercambio horizontal de genes y una estructura poblacional clonal. A pesar de ello, existe una variación genética en la población de *M. tuberculosis* que puede ser debida a polimorfismos de un único nucleótido (SNP), polimorfismos de secuencia larga (LSP) o polimorfismos de secuencias repetidas. Esta variación puede detectarse mediante marcadores. Algunos de estos marcadores visualizan variaciones que tuvieron lugar en un tiempo más o menos lejano y sirven para clasificar las cepas en «linajes» de interés filogenético. Otros marcadores detectan variaciones mucho más recientes y permiten tipificar las cepas para reconocer cadenas de transmisión.

Marcadores para estudios filogenéticos de *M. tuberculosis*

Polimorfismo de un único nucleótido (single nucleotide polymorphism)

Los SNP pueden ser sinónimos o no sinónimos. En los sinónimos, el cambio nucleotídico no comporta un cambio de aminoácido y, por tanto, no está sometido a presión selectiva. Su reloj evolutivo es lento y dan información filogenética. Los polimorfismos no sinónimos, que constituyen aproximadamente las 2 terceras partes de la variabilidad genética observada en *M. tuberculosis*, tienen potencialmente consecuencias funcionales, están tratados mediante presión selectiva y existen evidencias de que pueden influir en características importantes como la transmisibilidad, la virulencia, la resistencia a fármacos, la respuesta inmunitaria o el cuadro clínico¹.

Sreevatsan et al. en 1997² comparan las secuencias de 26 genes estructurales y detectan 2 SNP sinónimos. Uno en el gen que codifica para la catalasa peroxidasa (katG463) y otro en la subunidad A de la ADN girasa (gyrA95). Estos SNP permiten clasificar a *M. tuberculosis* en 3 grupos genéticos. En el grupo 1 se incluyen las cepas de *Mycobacterium bovis* (*M. bovis*), *Mycobacterium microti* (*M. microti*), *Mycobacterium africanum* (*M. africanum*) y las cepas de *M. tuberculosis* evolutivamente más antiguas, mientras que los grupos 2 y 3 están formados por cepas de *M. tuberculosis*. Los autores estudian unas 6.000 cepas de *M. tuberculosis* aisladas en Nueva York y Houston, y observan que la mayoría de los casos de tuberculosis agrupados (cadenas de transmisión) corresponden a cepas del grupo 1 o 2, mientras que el grupo 3 (el más reciente evolutivamente) daría lugar a casos esporádicos, lo que les hace concluir que *M. tuberculosis* está evolucionando hacia una menor transmisibilidad o virulencia. Estos 3 grupos pueden ser divididos en 6 grupos filogenéticos (con 5 subgrupos) mediante el estudio de 212 SNP³. Existe una fuerte asociación entre estos grupos y el origen geográfico de las cepas o el lugar de nacimiento de los pacientes.

Polimorfismos de secuencia larga (long sequence polymorphism)

El estudio de la delección de 20 secuencias largas distribuidas a lo largo del cromosoma, regiones de delección (RD)⁴, permite establecer el marco evolutivo del complejo *M. tuberculosis*, así como la identificación de las especies del complejo. El complejo *M. tuberculosis* es genéticamente homogéneo pero muy diverso con relación a su hábitat. Así, *M. tuberculosis*, *M. africanum* y *Mycobacterium canettii* tienen un hábitat exclusivamente humano, *M. microti* es un patógeno de los roedores, *Mycobacterium pinnipedi* de las focas, *Mycobacterium caprae* de las cabras, mientras que *M. bovis* tiene un rango de hospedadores más amplio. El estudio de las RD ha permitido establecer que la especie *M. tuberculosis* constituye el ancestro más antiguo a partir del cual han evolucionado las demás. Dentro de la especie *M. tuberculosis*, las cepas pueden clasificarse en antiguas o modernas según tengan o no la región TbD1. La delección de la región TbD1 probablemente ocurrió antes del siglo XVIII y la aparición de las cepas modernas de *M. tuberculosis* pudo contribuir a la enorme expansión de la tuberculosis (TB) a partir del siglo XVIII. Gagneux et al.⁵ analizan estas regiones RD en 875 cepas procedentes de 80 países. Las cepas se subdividen en 4 linajes (indo-oceánico, Asia del este, África del este-India, Euro-América) y 2 linajes clásicamente identificados como *M. africanum* (oeste de África I y oeste de África II).

En un estudio en San Francisco, donde coexisten comunidades de orígenes geográficos diversos, así como cepas de los distintos linajes descritos, se observó una estrecha asociación entre linajes y poblaciones de un determinado origen geográfico⁵. Es decir, existiría una adaptación o compatibilidad entre el patógeno y una determinada población del huésped.

Los estudios basados tanto en SNP como en LSP son muy congruentes, lo cual no es de extrañar dada la naturaleza clonal de la población de *M. tuberculosis*, y coinciden en señalar que existe una asociación entre linaje y origen geográfico, especulando sobre la influencia que puedan tener estos linajes sobre el rendimiento de pruebas diagnósticas, la resistencia a los antituberculosos o la respuesta a vacunas⁶.

Marcadores para estudios de epidemiología molecular

Tipificación mediante oligonucleótidos complementarios de secuencias espaciadoras (spoligotyping: spacer oligo typing)

En el cromosoma de *M. tuberculosis* existen múltiples copias de la secuencia DR localizadas en un único locus cromosómico (región DR), el cual constituye un *hot spot* de integración para la IS6110. Las secuencias repetidas DR de 36 pb se encuentran separadas por secuencias espaciadoras diferentes, esto es, no repetidas, de 34–41 pb. Cuando se comparan cepas de *M. tuberculosis*, esta región es polimórfica debido tanto a una recombinación homóloga entre DR como a cambios provocados por la localización de IS6110 en esta zona. En la técnica del *spoligotyping*⁷ se analiza este polimorfismo. Para ello, en una primera fase se procede a una amplificación que utiliza como iniciadores 2 oligonucleótidos complementarios de los extremos de la secuencia DR y que, por consiguiente, amplificarán las distintas secuencias espaciadoras existentes en la cepa. Los iniciadores están marcados con biotina. En una segunda fase, se procede a la detección de las distintas secuencias espaciadoras amplificadas mediante una hibridación del producto de PCR sobre un filtro en el que se han inmovilizado oligonucleótidos complementarios de las distintas secuencias espaciadoras descritas. La unión de los productos de PCR con sus secuencias específicas se detecta mediante peroxidasa marcada con estreptavidina que se une a la biotina presente en los productos de PCR. Existen 94 secuencias

espaciadoras, aunque en la mayoría de los protocolos analíticos se utilizan 43. Se han desarrollado técnicas alternativas a la hibridación sobre membrana, como una técnica basada en microesferas y tecnología láser⁸ o en la espectrometría de masas (MALDI-TOF)⁹, que facilitan la técnica y acortan su duración. Al ser la primera fase de la técnica una reacción de amplificación, se necesita poca cantidad de ADN para llevarla a cabo. De hecho, el *spoligotyping* puede utilizarse directamente a partir de la muestra clínica¹⁰. Los resultados se expresan como un número binario o tras una conversión simple en un número de 8 dígitos. Este hecho hace que los resultados sean fácilmente exportables y existan bases de datos globales (SpolDB4; www.pasteur-guadeloupe.fr/tb/bd_myco.html)

La principal limitación del *spoligotyping* es un poder discriminativo inferior al del IS6110-RFLP o a unidades repetidas intercaladas micobacterianas-repeticiones en tándem de número variable (MIRU-VNTR). El *spoligotyping* estaría a caballo entre aquellas técnicas que dan una información filogenética y las técnicas más discriminativas que permiten detectar cadenas de transmisión reciente. Por ello, se la considera más útil en estudios con fines filogenéticos (asignación de linajes, sublinajes o familias) que para la identificación veraz de cadenas de transmisión, con fines de epidemiología molecular.

Patrones de restricción-hibridación (IS6110-RFLP, PGRS-RFLP)

Cuando se digiere una molécula de ADN con un enzima de restricción se obtiene un número de fragmentos equivalente a las veces que se encuentra repetido el lugar de restricción a lo largo de la molécula. El tamaño de los distintos fragmentos de restricción traduce la separación entre 2 lugares de restricción contiguos. La comparación de cepas para detectar el polimorfismo del tamaño de los fragmentos de restricción (RFLP) se consigue separando los fragmentos obtenidos mediante una electroforesis en un gel de agarosa. Para simplificar la interpretación de los RFLP del ADN total, podemos centrarnos en el análisis de una secuencia repetida en el genoma en lugar de estudiar todo el cromosoma. Para ello, se realiza una transferencia de los fragmentos de restricción del ADN total obtenidos a una membrana (Southern-blot) que se analiza mediante hibridación con una sonda marcada. El revelado de la hibridación tan solo pondrá en evidencia aquellos fragmentos que contengan la totalidad o parte de la secuencia complementaria de la sonda escogida, obteniéndose perfiles con pocas bandas. El número de bandas estará en relación con el número de copias de la secuencia existentes en el genoma, mientras que su polimorfismo estará en relación con su distribución a lo largo del genoma, así como con las variaciones en los lugares de restricción dentro de esta secuencia y en las regiones más próximas. Tanto la transferencia de Southern como la hibridación complican notablemente los aspectos técnicos del análisis.

Dentro de las limitaciones de estas técnicas está la necesidad de disponer de gran cantidad de ADN, lo que impide utilizarlas directamente sobre la muestra clínica, debiéndose trabajar a partir de cultivo. Asimismo, en las cepas en las que la secuencia escogida se repite un número pequeño de veces (5 o menos en el caso de IS6110-RFLP) la capacidad discriminativa de la técnica disminuye. Por otra parte, los resultados se obtienen como un patrón de bandas sujeto a las distorsiones electroforéticas, lo que hace que la comparación de resultados entre laboratorios sea difícil. A pesar de ello, los patrones de restricción-hibridación asociados a IS6110 constituyeron durante años la técnica de referencia para el estudio de la epidemiología molecular de la TB, tanto por su gran poder discriminativo como por su estabilidad (reloj molecular estimado entre 3,2 y 8,7 años¹¹). La utilización de un protocolo estandarizado¹² con marcadores de peso incluidos en cada canal de electroforesis

permitió el establecimiento de bases de datos a nivel mundial, que mejoraron el conocimiento de la epidemiología global de la TB.

Unidades repetidas intercaladas micobacterianas-repeticiones en tándem de número variable (mycobacterial interspersed repetitive units-variable number tandem repeats [MIRU-VNTR])

A lo largo del cromosoma de *M. tuberculosis* existen unas secuencias repetidas en tándem que son polimorfas en las distintas cepas al presentar un número variable de repeticiones (VNTR). Estas estructuras son parecidas a los minisatélites observados en las células eucariotas. Las MIRU corresponden a uno de estas VNTR. Son repeticiones en tándem de 46 a 101 bp, encontrándose unos 41 MIRU dispersas por el cromosoma. La tipificación por MIRU-VNTR utilizaba inicialmente 12 *loci* (los más polimorfos), pero dada la menor discriminación observada comparando con otros marcadores, se estandarizó el análisis de 24 *loci*, que es el método utilizado en la actualidad¹³. Para ello, se realiza una PCR específica para cada *locus* en la que los iniciadores utilizados están marcados con un fluorocromo. Posteriormente, se determina el tamaño del fragmento amplificado mediante electroforesis capilar. Este tamaño dependerá del número de repeticiones de la secuencia *core* presente en este *locus*. El número de repeticiones constituye el dígito para este *locus* en el código VNTR, que tendrá consecuentemente 24 dígitos.

La tipificación basada en MIRU-VNTR es rápida, relativamente sencilla desde un punto de vista técnico, muy reproducible y discriminativa (comparable a IS6110 RFLP). Sus resultados se expresan en un código de 24 dígitos, por lo que son fácilmente exportables, existiendo bases de datos globales que permiten comparar cepas a nivel mundial (<http://www.miru-vntrplus.org>). Una de estas bases incluye los datos de MIRU-VNTR y *spoligotyping* de unas 62.000 cepas de *M. tuberculosis* aisladas de unos 153 países (<http://www.pasteur-guadeloupe.fr:8081/SITVIT.online>). Por todo ello, MIRU-VNTR se ha convertido en la técnica de referencia para la tipificación de *M. tuberculosis*. La información obtenida permite inferir tanto hipótesis filogenéticas, como discriminar a nivel de cepa, permitiendo detectar cadenas de transmisión.

A pesar de ello, la capacidad discriminativa de los *loci* utilizados puede variar en función del linaje de *M. tuberculosis* estudiado. Así, dada la homoplasia existente entre los miembros de la familia Beijing, el MIRU-VNTR de 24 *loci* estandarizado tiene sus limitaciones^{14,15}. Para el análisis de esta familia se ha propuesto incluir *loci* que no están presentes en el esquema estandarizado (VNTRs3232, 3280 y 4120). Esto es, para determinados estudios puede ser necesario adaptar el esquema MIRU-VNTR a los linajes estudiados.

Secuenciación del genoma completo (whole genome sequencing [WGS])

Hemos visto cómo, hasta ahora, los marcadores moleculares utilizados analizan las variaciones existentes en regiones «hipervariables», que representan tan solo una ínfima parte del genoma total. Cuando se utiliza la secuenciación de genoma completo (WGS) se dispone de información sobre la microevolución de todo el genoma. La utilización de WGS en la epidemiología molecular de la TB ha sido posible tanto por el desarrollo tecnológico de las plataformas de secuenciación masiva, como por la aparición de preparados comerciales que facilitan la preparación de las librerías genómicas para la secuenciación. Obviamente, la principal dificultad de esta técnica es el análisis de la información generada. En los últimos años se han desarrollado numerosas herramientas informáticas que permiten no solo calcular el número de SNP existentes entre genomas, sino también deducir, a partir del genoma completo, tanto los marcadores moleculares utilizados para la tipificación (IS6110, MIRU-VNTR, *spoligotyping*), como

los que proporcionan información filogenética (SNP, LPS), así como el antibiograma molecular (analizando las mutaciones asociadas a resistencia)¹⁶.

Como es de esperar, el análisis del genoma completo es más discriminativo que los marcadores tradicionalmente utilizados para detectar cadenas de transmisión. El parámetro utilizado es el cálculo de los SNP existentes entre las secuencias comparadas. Diversos estudios analizan los SNP existentes entre cepas aisladas secuencialmente en el mismo enfermo o cepas pertenecientes a cadenas de transmisión bien definidas, tanto por epidemiología molecular como por estudios de campo. El objetivo de estos estudios es estimar cual es la deriva genética de *M. tuberculosis* a corto plazo para establecer un *cutoff* que permita diferenciar lo relacionado de lo no relacionado epidemiológicamente. La comparación de los genomas de las cepas pertenecientes a una cadena de transmisión casi nunca excede los 3 a 5 SNP^{17,18}. Por otra parte, el cálculo del cambio promedio en la secuencia de ADN del genoma de *M. tuberculosis* se ha establecido en 0,5 SNP por genoma y año¹⁸. Los árboles filogenéticos obtenidos mediante el estudio de SNP del genoma completo correlacionan mejor con los datos epidemiológicos de campo que los árboles obtenidos con los marcadores tradicionales. Es decir, el estudio de los SNP del genoma permite separar los miembros de una cadena de transmisión de las cepas relacionadas pero no pertenecientes a dicha cadena. Esto es de especial ayuda para el análisis de brotes complejos^{18,19}.

Por otra parte, los datos de la WGS permiten determinar la distribución en el tiempo y en el espacio de los casos de TB al permitir diferenciar el caso índice de los casos secundarios en una cadena de transmisión. En los árboles filogenéticos, los casos secundarios (variantes clonales con muy pocos SNP de diferencia) darán lugar a estructuras en forma de estrella que tendrá como núcleo el caso índice. Asimismo, en una epidemia duradera en el tiempo, el acúmulo unidireccional de SNP permite asociar de forma más clara los nuevos casos a los casos precedentes^{18,19}.

Aplicaciones de las estrategias de genotipificación y de análisis genómico de *M. tuberculosis*

Desde un punto de vista estricto, las técnicas de genotipificación de *Mycobacterium tuberculosis* (MTB) encuentran su nicho de aplicabilidad natural en los estudios de epidemiología molecular, con el fin de identificar pacientes en clúster, infectados por una misma cepa, y que por tanto participan en una misma cadena de transmisión. Estos casos agrupados se diferencian de los que no forman parte de clústeres, o casos huérfanos, que por estar infectados por una cepa con genotipo único en esa población, se consideran candidatos a ser reactivaciones de exposiciones ocurridas en un tiempo pasado, más o menos remoto.

En esta revisión, con el fin de cubrir el mayor espectro de aplicabilidad de las técnicas de genotipificación de MTB hemos optado por expandirnos más allá del concepto estricto de aplicación de estas técnicas con fines de epidemiología molecular, para abarcar otros entornos de análisis donde estos mismos métodos han mostrado su utilidad. Por tanto, repasaremos una serie de preguntas a las que la caracterización de los diferentes marcadores genéticos de MTB ha podido ofrecer una respuesta.

¿Puedo identificar un falso diagnóstico de tuberculosis?

Desde el laboratorio, la sospecha de una posible contaminación cruzada (CC) surge cuando se identifican 2 o más cultivos positivos en muestras que se procesaron en el mismo día. La alarma de CC se afianza cuando además confluyen los siguientes aspectos: solo uno de los casos cuenta con otras muestras adicionales positivas, únicamente la muestra de uno de los casos implicados en la

alerta es tinción positiva y el crecimiento de la muestra del caso sospechoso de ser un falso positivo demoró más tiempo en alcanzar la positividad que el promedio. Sin embargo, para documentar fehacientemente que nos encontramos entre un evento de CC de laboratorio, debemos de demostrar que los cultivos de los casos implicados en la sospecha corresponden a la misma cepa. Para ello, la técnica más recomendable es la de MIRU-VNTR, por su mayor rapidez en ofrecer resultados. Además, puesto que en este caso no es necesario determinar el genotipo preciso de las cepas, sino que es suficiente un análisis cualitativo de identidad (demuestra la CC) o diferencia (descarta la CC) entre las mismas, se han desarrollado modalidades rápidas de análisis cualitativo que permiten la documentación rápida de identidad/diferencia entre genotipos mediante la comparación de sus patrones de movilidad en electroforesis convencional de agarosa, tanto a partir de productos de PCR triplex (técnica MLP3)²⁰, como dúplex²¹.

¿Me encuentro ante una microepidemia o brote?

Antes del desarrollo de las técnicas de genotipificación de MTB, la sospecha de brotes o microepidemias se basaba exclusivamente en la identificación de vínculos epidemiológicos entre varios casos de TB. Sin embargo, la constatación de brotes requiere de la confirmación de identidad genotípica entre las cepas que infectan a los casos teóricamente vinculados. La documentación de estos brotes se ha abordado inicialmente con RFLP-IS6110 y en los últimos años por MIRU-VNTR. Gracias a la aplicación bajo demanda de estas técnicas se ha podido desvelar que parte de las microepidemias, que resultaban obvias desde el punto de vista del estudio epidemiológico de los casos, implicaban cepas diferentes^{22,23}, lo que supuso la primera indicación de que la investigación epidemiológica debe extenderse a entornos de contacto que trasciendan el ámbito laboral, familiar o domiciliario.

De modo equivalente a como pasaba en la documentación de contaminaciones de laboratorio, el análisis epidemiológico molecular dirigido a brotes sospechados no requiere en principio la obtención del genotipo preciso de las cepas, sino únicamente establecer identidad o diferencia entre los aislados implicados, por lo que la técnica de MLP3 o derivados^{20,21} puede ser suficiente para laboratorios con recursos limitados. Sin embargo, la obtención del genotipo mediante análisis completo convencional, antes mediante RFLP, hoy basado en MIRU-VNTR, permite identificar a contactos reales con otros casos de la población que ayuden a redefinir la magnitud del brote.

¿Cómo puedo identificar infecciones mixtas o complejas?

Uno de los campos diferentes a la epidemiología molecular donde las técnicas de genotipificación han supuesto un avance notable es en la identificación de casos infectados por más de una cepa, tanto en un mismo episodio (infecciones mixtas, cuando implican a cepas diferentes, o policlonales, en el caso de variantes clonales que proceden de una misma cepa parental), como en episodios sucesivos (reinfecciones).

Las primeras documentaciones moleculares de reinfecciones²⁴ alertaban de una proporción inesperadamente alta (75%) de casos con TB recurrente, en los que la cepa aislada en un segundo episodio era diferente de la primera, si bien se trataba de entornos de elevada incidencia, donde la sobreexposición a casos infecciosos era más probable. En estos entornos incluso se demostró que la reinfección puede ocurrir poco tiempo tras la finalización del tratamiento del primer episodio²⁵. Con posterioridad, se documentó que la reinfección era la causa de una parte no desdeñable de las recurrencias, incluso en entornos de moderada o baja incidencia²⁶⁻²⁸.

En cuanto a la presencia simultánea de más de una cepa o variante clonal en un mismo paciente, es solo con la llegada de

la técnica de MIRU-VNTR cuando se ha avanzado notablemente en este terreno^{29,30}. Esto es debido a que las estrategias que la precedían, *spoligotyping* o RFLP, eran incapaces de identificar la presencia de más de una cepa. En estos casos, obteníamos un patrón genotipificado ficticio, que era el resultado del solapamiento de los patrones de cada una de las cepas, sin que ese patrón aditivo nos alertara de que correspondía a la presencia de 2 cepas o variantes clonales. Sin embargo, el hecho de que con la técnica de MIRU-VNTR se obtenga un único producto de amplificación por cada *locus* analizado hace que, en el caso de que observemos más de un producto, únicamente pueda ser debido a la presencia de más de una cepa (en caso de un número suficientemente alto de *loci* con más de un alelo) o variante clonal (en el caso de 1–2 *loci* con más de un alelo). Por tanto, la aplicación sistemática de la técnica de MIRU-VNTR nos ha permitido conocer con más precisión la proporción de infecciones mixtas o policlonales en una población^{31,32} y poder determinar que la urgencia de variantes clonales no está restringida a situaciones extremas, como prolongadas demoras diagnósticas, sino que pueden surgir en el seno de infecciones convencionales³³. La identificación eficaz de infecciones complejas tiene un indudable valor epidemiológico; sin embargo, su no consideración puede conducirnos asimismo a dificultades en el tratamiento de los pacientes, en el caso de infecciones mixtas con cepas que difieren en su patrón de sensibilidad³⁴. Por último, la presencia de una infección compleja puede conllevar complicaciones en el diagnóstico microbiológico, en aspectos tan aparentemente alejados como el análisis de contaminaciones de laboratorio³⁵.

¿Cómo abordar las dinámicas de transmisión de tuberculosis en mi población?

A diferencia de la documentación de brotes, que está al alcance de cualquier laboratorio con capacidad de implementar las técnicas de genotipificación, el análisis de transmisión en una población requiere asentar unas bases estratégicas de mayor solidez. La identificación eficaz de cadenas de transmisión exige abordar una genotipificación universal; es decir, debemos de contar con una muestra con base poblacional que nos garantice que accedemos a todos los casos de TB y que, por tanto, no van a quedar «eslabones» de la cadena de transmisión sin caracterizar. Además, dado el tiempo variable entre la exposición de un caso y el desarrollo de la enfermedad, la identificación rigurosa de cadenas de transmisión requiere de un mínimo tiempo de caracterización sistemática de todos los casos, superior a 2–3 años. Los primeros estudios de epidemiología molecular con base poblacional descansaron sobre la técnica de RFLP; sin embargo, la mayor discriminación de la técnica MIRU-VNTR, su menor demora en obtener resultados y la facilidad de intercambiar resultados entre los laboratorios, al ofrecer una genotipificación en código numérico, han hecho que desbanque completamente al RFLP.

Los estudios de epidemiología molecular universal nos permiten conocer el porcentaje de casos de TB que son resultado de cadenas de transmisión reciente, lo que constituye un excelente indicador de la eficacia de los programas de control, que deberían conducir a la reducción en estos porcentajes. Asimismo, han facilitado desvelar las dinámicas de transmisión en escenarios epidemiológicos complejos, como los derivados de poblaciones con un gran porcentaje de inmigración, en los que se ha podido identificar entornos con una importante transmisión cruzada entre casos autóctonos e inmigrantes^{36,37}. El mantenimiento de sistemas de genotipificación sistemático nos permite identificar las cepas prevalentes que son responsables de una buena parte de la carga de enfermedad de una población³⁸. Asimismo, estos programas de vigilancia molecular prolongada facilitan la disección detallada del impacto que puede tener la importación de una cepa inédita en una población,

que años después puede ser responsable de una parte relevante de los casos de TB en la misma³⁹.

¿Es posible abordar una caracterización de excelencia de la transmisión de tuberculosis?

A pesar de la valiosa información obtenida de la aplicación sistemática del RFLP, en los primeros momentos, y de MIRU-VNTR, en los últimos años, debemos de ser conscientes de las limitaciones inevitables derivadas de estar determinando relaciones de identidad o diferencia entre cepas cuando estamos atendiendo al análisis de un región mínima (< 0,1%) del cromosoma de MTB. Esto significa que podemos estar considerando como idénticas cepas que difieren en zonas del cromosoma que no estamos observando y, por tanto, podríamos estar sobrestimando una parte de los clústeres de transmisión que estamos proponiendo.

Afortunadamente, la reducción marcada en los costes y la simplificación de los procedimientos bioinformáticos han supuesto la entrada de la WGS en la caracterización de MTB. Esto significa acceder a una gran cantidad de información del genoma, lo que permite asumir una capacidad de discriminación entre cepas muy superior, idónea para garantizar máxima precisión en los estudios de epidemiología molecular, que comienzan a ser sustituidos por una epidemiología genómica.

La aplicación de WGS a los clústeres definidos por genotipificación convencional conduce a que parte de ellos, como era de esperar, se subdividan en agrupaciones más pequeñas, tras identificar diferencias entre cepas que con los niveles de resolución anteriores aparecían como asociadas en clúster⁴⁰. Esto significa que la información derivada de la epidemiología genómica resulta más acorde con la distribución geográfica de los casos y con sus vínculos epidemiológicos¹⁹. La aplicación sistemática de WGS sobre situaciones de difícil control ha revelado la enorme complejidad de algunos de estos escenarios, como la coexistencia de diferentes brotes solapantes que hasta el momento habían sido interpretados como uno solo⁴¹.

Ya en el terreno de la epidemiología con base poblacional, son aún escasos los entornos en los que se está pudiendo aplicar el análisis por WGS de modo sistemático, destacando en este sentido Inglaterra, que ha activado un sistema de análisis de todos los casos de TB, iniciado en la zona de Midlands⁴² y que recientemente se ha expandido a buena parte del país. Estos esfuerzos de aplicar sistemáticamente WGS están yendo paralelos al desarrollo de metodología que permita acelerar la obtención de resultados, aplicando directamente WGS sobre cultivos primarios, intentando eliminar la interferencia derivada del ADN humano presente en la muestra⁴³. Asimismo, comienza a haber esfuerzos encaminados a la obtención directa de información genómica a partir de la muestra clínica⁴⁴, con resultados variables, aún de infinita menor calidad que los obtenidos de cultivo, pero que pudieran ofrecer una primera línea de análisis que aporte información de utilidad.

La mayor accesibilidad del análisis por WGS ha hecho que no se restrinja estrictamente a los estudios epidemiológicos, sino que ha buscado respuestas a algunos de los interrogantes antes expuestos, que habían sido analizados por genotipificación convencional. La documentación del porcentaje de reinfecciones en una población ha sido recientemente abordada mediante la aplicación de WGS⁴⁵. Asimismo, la capacidad de la secuenciación de nueva generación para discriminar secuencias diferentes que coexisten en una misma muestra ha permitido identificar casos con infecciones complejas⁴⁶. Pudiera considerarse que la WGS aporta un nivel de resolución innecesaria para la resolución de algunos de estos problemas, máxime cuando un primer rastreo basado en técnicas de baja discriminación puede ser capaz de identificar diferencias entre cepas, sin necesidad de ir más allá. Sin embargo, es cierto que dada la reducción en sus costes le convierten en una alternativa

competitiva, que además de resolver estos interrogantes nos aporta, con el mismo esfuerzo y coste, multitud de información genotipificada adicional de interés, más allá de las necesidades epidemiológicas.

¿Puedo encontrar alternativa al camino abierto por la epidemiología genómica?

A pesar de que la capacidad de discriminación y el volumen de información de utilidad epidemiológica, terapéutica y filogenética ofrecida por el análisis por WGS son imbatibles, su aplicación sistemática aún no está al alcance de muchos contextos. Especialmente, si hablamos de entornos con escasos recursos, donde además se acumula la mayor carga de enfermedad, estas estrategias genómicas parecen de difícil implantación.

Es en esta fisura que surge entre la alta resolución ofrecida por el análisis basado en WGS y las necesidades de vigilancia de entornos con retos de transmisión que aún no pueden ser abordables por epidemiología genómica donde ha surgido una línea de avance alternativa. Esta consiste en desarrollar metodologías moleculares sencillas, de bajo coste y fácil implementación, para abordar una vigilancia dirigida de determinadas cepas que protagonicen los problemas de transmisión más relevantes en cada población.

La vigilancia específica de cepas singulares mediante técnicas moleculares sencillas no es una propuesta novedosa. Ya en la década de los 90 se desarrollaron PCR diseñadas para trazar la transmisión de la cepa W, causante de un brote de gran magnitud en Nueva York, que se dirigían a rasgos genéticos peculiares de esta cepa⁴⁷. El inconveniente de estas propuestas es que solo era posible diseñar estas PCR tras asegurar un estudio genético exhaustivo de las cepas, y siempre que este condujera a la identificación de algún rasgo singular. Ahora, con la accesibilidad del análisis por WGS, podemos identificar rasgos distintivos de cualquier cepa, mediante la secuenciación completa de un número reducido de representantes de ese aislado, que nos conduzca a identificar SNP específicos de esa cepa.

Basadas en esta premisa, se han desarrollado PCR alelo específicas que se dirigen a la vigilancia de cepas que son prevalentes en alguna población⁴⁸. Asimismo, esta filosofía de análisis se ha aplicado para permitir el rastreo masivo y rápido de colecciones retrospectivas de aislados con el fin de identificar la presencia en una población de una cepa de alto riesgo epidemiológico, por haber estado implicada en un brote de gran magnitud⁴⁹. Por último, la posibilidad de integrar esta estrategia para dar una respuesta precoz frente a una alerta real ha sido demostrada. El diagnóstico de 2 casos de TB XDR importada de Rusia a España activó un sistema que condujo a la instauración rápida de sendas PCR específicas para SNP marcadores de estas cepas, desvelados por análisis de WGS y comparación con bases de datos de SNP globales, cuya aplicación prospectiva local sobre muestras respiratorias permitió descartar precozmente casos secundarios debidos a estas cepas de alto riesgo⁵⁰.

¿Cómo se aborda la distribución de cepas de tuberculosis en un entorno global?

Hasta ahora, nos hemos focalizado exclusivamente en la utilidad de las diferentes aproximaciones de epidemiología molecular, y genómica, para abordar interrogantes en poblaciones definidas, o en pacientes determinados. Sin embargo, no debemos de olvidar que la TB es un fenómeno mundial, con una dimensión aún más global como resultado de los movimientos migratorios internacionales. Esto hace que haya preguntas de índole «macropoblacional» a las que también hay que dar respuesta. En este sentido, como se comentó en la primera parte de esta revisión, hay un nivel de catalogación de MTB que no desciende a la discriminación de los

marcadores que pretenden discriminar a nivel de cepa. Al contrario, persiguen una discriminación menor para catalogar los aislados de modo más genérico en cada uno de los linajes descritos para el complejo TB. El análisis de la distribución de linajes ofrece una información de gran relevancia no solo desde un punto de vista global, permitiendo conocer la distribución internacional de los mismos, sino asimismo desde un punto de vista evolutivo o filogenético.

Los primeros estudios de distribución global de linajes, así como filogenéticos, descansaron en la espoligotipificación⁵¹, cuya capacidad de discriminación reducida y tendencia a la homoplasia no la hacen recomendable para la aplicación epidemiológica de campo pero que aún resulta adecuado para el análisis de la distribución de linajes, familias o subfamilias de MTBC. Sin embargo, los datos obtenidos a partir del análisis de WGS también ofrecieron una alternativa para los estudios de este tipo. Se propuso un juego de SNP marcadores para los diferentes linajes que demostraron ser consistentes y que simplificaban la metodología a aplicar⁵². Diferentes propuestas se dirigieron a la identificación simplificada de estos SNP aprovechando la tecnología más habitualmente disponible en el laboratorio, abarcando aproximaciones diversas, como la PCR en tiempo real⁵², las PCR alelo-específicas⁵³, así como otros formatos de análisis más restringidos a laboratorios con mayores recursos, como Luminex⁵² o SNaPshot⁵³.

Comentarios finales

M. tuberculosis presenta una estabilidad genética que ha obligado a desarrollos metodológicos específicamente adaptados a esta singularidad. Los esfuerzos para optimizar técnicas adecuadas para este patógeno nos han ofrecido una diversidad de métodos que abarcan un amplio rango de capacidad de discriminación. Por ello, contamos con diferentes aproximaciones ajustables, con el fin de adaptarnos a las necesidades de cada estudio. Disponemos de análisis de baja discriminación para abordar estudios de índole filogenética, evolutiva o de distribución global de cepas, así como de métodos de discriminación elevada para discernir con precisión las cadenas de transmisión activa en una población determinada, identificar contaminaciones de laboratorio, caracterizar recurrencias o documentar microepidemias. La reducción de costes en el análisis de cromosomas completos, así como el desarrollo de sistemas de análisis bioinformáticos más asequibles, han supuesto la transición natural de la epidemiología molecular a una nueva epidemiología genómica de la TB. Apoyados en estas estrategias, nuestra capacidad de discriminación es hoy máxima, tanto para la identificación precisa de cadenas de transmisión con un sólido soporte epidemiológico como para determinar la cronología de dichas cadenas. Este mismo análisis de WGS nos conduce a identificar SNP marcadores de gran consistencia, tanto con fines filogenéticos como para ser utilizados como marcadores de cepas relevantes. Precisamente, estos SNP marcadores de cepa, identificados mediante WGS, pueden protagonizar una nueva vía de avance posgenómica apoyada en el desarrollo de PCR específicas, de bajo coste y transferibles, dirigidas a la vigilancia priorizada de cepas de alto riesgo en cada población. Este formato de trabajo podría dar pie a un nuevo modelo de la vigilancia de la transmisión de la TB multinodal y descentralizado.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Niemann S, Supply P. Diversity and evolution of *Mycobacterium tuberculosis*: Moving to whole-genome-based approaches. *Cold Spring Harb Perspect Med.* 2014;4:a021188.

2. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, et al. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci USA*. 1997;94:9869–74.
3. Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbon MH, Bobadilla del Valle M, et al. Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: Insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. *J Bacteriol*. 2006;188:759–72.
4. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeyer K, et al. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci USA*. 2002;99:3684–9.
5. Gagneux S, deRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, et al. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA*. 2006;103:2869–73.
6. Gagneux S, Small PM. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect Dis*. 2007;7:328–37.
7. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol*. 1997;35:907–14.
8. Cowan LS, Diem L, Brake MC, Crawford JT. Transfer of a *Mycobacterium tuberculosis* genotyping method, spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system. *J Clin Microbiol*. 2004;42:474–7.
9. Honisch C, Mosko M, Arnold C, Gharbia SE, Diel R, Niemann S. Replacing reverse line blot hybridization spoligotyping of the *Mycobacterium tuberculosis* complex. *J Clin Microbiol*. 2010;48:1520–6.
10. Cafrune PI, Possuelo LG, Ribeiro AW, Ribeiro MO, Unis G, Jarczowski CA, et al. Prospective study applying spoligotyping directly to DNA from sputum samples of patients suspected of having tuberculosis. *Can J Microbiol*. 2009;55:895–900.
11. Warren RM, van der Spuy GD, Richardson M, Beyers N, Borgdorff MW, Beh MA, et al. Calculation of the stability of the IS6110 banding pattern in patients with persistent *Mycobacterium tuberculosis* disease. *J Clin Microbiol*. 2002;40:1705–8.
12. Van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: Recommendations for a standardized methodology. *J Clin Microbiol*. 1993;31:406–9.
13. Jagielski T, Minias A, van Ingen J, Rastogi N, Brzostek A, Zaczek A, et al. Methodological and clinical aspects of the molecular epidemiology of *Mycobacterium tuberculosis* and other mycobacteria. *Clin Microbiol Rev*. 2016;29:239–90.
14. Mokrousov I. *Mycobacterium tuberculosis* Beijing genotype and mycobacterial interspersed repetitive unit typing. *J Clin Microbiol*. 2006;44, 1614 [author reply 1614–1615].
15. Hanekom M, van der Spuy GD, Streicher E, Ndadambi SL, McEvoy CR, Kidd M, et al. A recently evolved sublineage of the *Mycobacterium tuberculosis* Beijing strain family is associated with an increased ability to spread and cause disease. *J Clin Microbiol*. 2007;45:1483–90.
16. Faksri K, Tan JH, Chairprasert A, Teo YY, Ong RT. Bioinformatics tools and databases for whole genome sequence analysis of *Mycobacterium tuberculosis*. *Infection, Genetics and Evolution: J Mol Epidemiol Evol Genet Infect Dis*. 2016;45:359–68.
17. Schurch AC, Kremer K, Kiers A, Daviena O, Boeree MJ, Siezen RJ, et al. The tempo and mode of molecular evolution of *Mycobacterium tuberculosis* at patient-to-patient scale. *Infection, Genetics and Evolution: J Mol Epidemiol Evol Genet Infect Dis*. 2010;10:108–14.
18. Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: A retrospective observational study. *Lancet Infect Dis*. 2013;13:137–46.
19. Roetzler A, Diel R, Kohl TA, Ruckert C, Nubel U, Blom J, et al. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: A longitudinal molecular epidemiological study. *PLoS Medicine*. 2013;10:e1001387.
20. Sisilema-Egas F, Ruiz-Serrano MJ, Bouza E, García-de-Viedma D. Qualitative analysis to ascertain genotypic identity of or differences between *Mycobacterium tuberculosis* isolates in laboratories with limited resources. *J Clin Microbiol*. 2013;51:4230–3.
21. Yasmin M, le Moullec S, Siddiqui RT, de Beer J, Sola C, Refregier G. Quick and cheap MIRU-VNTR typing of *Mycobacterium tuberculosis* species complex using duplex PCR. *Tuberculosis*. 2016;101:160–3.
22. Martin A, Inigo J, Chaves F, Herranz M, Ruiz-Serrano MJ, Palenque E, et al. Reanalysis of epidemiologically linked tuberculosis cases not supported by IS6110-RFLP-based genotyping. *Clin Microbiol Infect*. 2009;15:763–9.
23. Verver S, Warren RM, Munch Z, Richardson M, van der Spuy GD, Borgdorff MW, et al. Proportion of tuberculosis transmission that takes place in households in a high-incidence area. *Lancet*. 2004;363:212–4.
24. Van Rie A, Warren R, Richardson M, Victor TC, Gie RP, Enarson DA, et al. Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. *N Engl J Med*. 1999;341:1174–9.
25. Uys P, Brand H, Warren R, van der Spuy G, Hoal EG, van Helden PD. The risk of tuberculosis reinfection soon after cure of a first disease episode is extremely high in a hyperendemic community. *PLoS One*. 2015;10:e0144487.
26. Caminero JA, Pena MJ, Campos-Herrero MI, Rodriguez JC, Alfonso O, Martin C, et al. Exogenous reinfection with tuberculosis on a European island with a moderate incidence of disease. *Am J Respir Crit Care Med*. 2001;163 3 Pt 1:717–20.
27. Bandera A, Gori A, Catozzi L, degli Esposti A, Marchetti G, Molteni C, et al. Molecular epidemiology study of exogenous reinfection in an area with a low incidence of tuberculosis. *J Clin Microbiol*. 2001;39:2213–8.
28. García de Viedma D, Marin M, Hernandezgomez S, Diaz M, Ruiz Serrano MJ, Alcalá L, et al. Tuberculosis recurrences: reinfection plays a role in a population whose clinical/epidemiological characteristics do not favor reinfection. *Arch Intern Med*. 2002;162:1873–9.
29. García de Viedma D, Alonso Rodríguez N, Andres S, Ruiz Serrano MJ, Bouza E. Characterization of clonal complexity in tuberculosis by mycobacterial interspersed repetitive unit-variable-number tandem repeat typing. *J Clin Microbiol*. 2005;43:5660–4.
30. Shamputa IC, Jugheli L, Sadradze N, Willery E, Portaels F, Supply P, et al. Mixed infection and clonal representativeness of a single sputum sample in tuberculosis patients from a penitentiary hospital in Georgia. *Respir Res*. 2006;7:99.
31. Navarro Y, Herranz M, Perez-Lago L, Martínez Lirola M, Indal TB, Ruiz-Serrano MJ, et al. Systematic survey of clonal complexity in tuberculosis at a population level and detailed characterization of the isolates involved. *J Clin Microbiol*. 2011;49:4131–7.
32. Perez-Lago L, Herranz M, Lirola MM, Group I-T, Bouza E, García de Viedma D. Characterization of microevolution events in *Mycobacterium tuberculosis* strains involved in recent transmission clusters. *J Clin Microbiol*. 2011;49:3771–6.
33. Perez-Lago L, Rodríguez Borlado AI, Comas I, Herranz M, Ruiz-Serrano MJ, Bouza E, et al. Subtle genotypic changes can be observed soon after diagnosis in *Mycobacterium tuberculosis* infection. *J Med Microbiol*. 2016;306:401–5.
34. Perez-Lago L, Lirola MM, Navarro Y, Herranz M, Ruiz-Serrano MJ, Bouza E, et al. Co-infection with drug-susceptible and reactivated latent multidrug-resistant *Mycobacterium tuberculosis*. *Emerg Infect Diseases*. 2015;21:2098–100.
35. Perez-Lago L, Herranz M, Navarro Y, Ruiz Serrano MJ, Miralles P, Bouza E, et al. Clonal complexity in *Mycobacterium tuberculosis* can hamper diagnostic procedures. *J Clin Microbiol*. 2017;55:1388–95.
36. Alonso Rodríguez N, Andrés S, Bouza E, Herranz M, Ruiz Serrano MJ, García de Viedma D, et al. Transmission permeability of tuberculosis involving immigrants, revealed by a multicentre analysis of clusters. *Clin Microbiol Infect*. 2009;15:435–42.
37. Borrell S, Espanol M, Orcau A, Tundo G, March F, Cayla JA, et al. Tuberculosis transmission patterns among Spanish-born and foreign-born populations in the city of Barcelona. *Clin Microbiol Infect*. 2010;16:568–74.
38. Lillebaek T, Andersen AB, Rasmussen EM, Kamper-Jorgensen Z, Pedersen MK, Bjorn-Mortensen K, et al. *Mycobacterium tuberculosis* outbreak strain of Danish origin spreading at worrying rates among greenland-born persons in Denmark and Greenland. *J Clin Microbiol*. 2013;51:4040–4.
39. Pena MJ, Caminero JA, Campos-Herrero MI, Rodríguez-Gallego JC, García-Laorden MI, Cabrera P, et al. Epidemiology of tuberculosis on Gran Canaria: A 4 year population study using traditional and molecular approaches. *Thorax*. 2003;58:618–22.
40. Gurjav U, Outhred AC, Jelfs P, McCallum N, Wang Q, Hill-Cawthorne GA, et al. Whole genome sequencing demonstrates limited transmission within identified *Mycobacterium tuberculosis* clusters in New South Wales, Australia. *PLoS One*. 2016;11:e0163612.
41. Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodtkin E, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med*. 2011;364:730–9.
42. Walker TM, Lalor MK, Broda A, Saldana Ortega L, Morgan M, Parker L, et al. Assessment of *Mycobacterium tuberculosis* transmission in Oxfordshire, UK, 2007–12, with whole pathogen genome sequences: An observational study. *Lancet Respir Med*. 2014;2:285–92.
43. Votintseva AA, Pankhurst LJ, Anson LW, Morgan MR, Gascoyne-Binzi D, Walker TM, et al. Mycobacterial DNA extraction for whole-genome sequencing from early positive liquid (MGIT) cultures. *J Clin Microbiol*. 2015;53:1137–43.
44. Votintseva AA, Bradley P, Pankhurst L, del Ojo Elias C, Loose M, Nilgiriwala K, et al. Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J Clin Microbiol*. 2017;55:1285–98.
45. Guerra-Assuncao JA, Houben RM, Crampin AC, Mzembe T, Mallard K, Coll F, et al. Recurrence due to relapse or reinfection with *Mycobacterium tuberculosis*: A whole-genome sequencing approach in a large, population-based cohort with a high HIV infection prevalence and active follow-up. *J Infect Dis*. 2015;211:1154–63.
46. Ssengooba W, de Jong BC, Joloba ML, Cobelens FG, Meehan CJ. Whole genome sequencing reveals mycobacterial microevolution among concurrent isolates from sputum and blood in HIV infected TB patients. *BMC Infect Dis*. 2016;16:371.
47. Plikaytis BB, Marden JL, Crawford JT, Woodley CL, Butler WR, Shinnick TM. Multiplex PCR assay specific for the multidrug-resistant strain W of *Mycobacterium tuberculosis*. *J Clin Microbiol*. 1994;32:1542–6.
48. Perez-Lago L, Martínez Lirola M, Herranz M, Comas I, Bouza E, García-de-Viedma D. Fast and low-cost decentralized surveillance of transmission of tuberculosis based on strain-specific PCRs tailored from whole genome sequencing data: A pilot study. *Clin Microbiol Infect*. 2015;21, 249 e241–249.
49. Perez-Lago L, Herranz M, Comas I, Ruiz-Serrano MJ, Lopez Roa P, Bouza E, et al. Ultrafast assessment of the presence of a high-risk *Mycobacterium tuberculosis* strain in a population. *J Clin Microbiol*. 2016;54:779–81.
50. Perez-Lago L, Martínez-Lirola M, García S, Herranz M, Mokrousov I, Comas I, et al. Urgent Implementation in a hospital setting of a strategy to rule out secondary cases caused by imported extensively drug-resistant *Mycobacterium tuberculosis* strains at diagnosis. *J Clin Microbiol*. 2016;54:2969–74.
51. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, et al. *Mycobacterium tuberculosis* complex genetic diversity: Mining the fourth

- international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol.* 2006;6:23.
52. Stucki D, Malla B, Hostettler S, Huna T, Feldmann J, Yeboah-Manu D, et al. Two new rapid SNP-typing methods for classifying *Mycobacterium tuberculosis* complex into the main phylogenetic lineages. *PloS One.* 2012;7:e41253.
53. Carcelén MA, Abascal E, Herranz M, Santantón S, Zenteno R, Ruiz Serrano MJ, et al. Optimizing and accelerating the assignation of lineages in *Mycobacterium tuberculosis* using novel alternative single-tube assays. *PloS One.* 2017;12, e0186956.