

Molecular typing of *Mycobacterium tuberculosis* strains: a fundamental tool for tuberculosis control and elimination

Angela Cannas,¹ Antonio Mazzarelli,¹ Antonino Di Caro,¹ Giovanni Delogu,^{1,2} Enrico Girardi¹

¹National Institute for Infectious Diseases L. Spallanzani, Rome; ²Institute of Microbiology, Sacro Cuore Catholic University, Rome, Italy

Abstract

Tuberculosis (TB) is still an important cause of morbidity and mortality worldwide. An improvement of the strategies for disease control is necessary in both low- and high-incidence TB countries. Clinicians, epidemiologists, laboratory specialists, and public health players should work together in order to achieve a significant reduction in TB transmission and spread of drug-resistant strains. Effective TB surveillance relies on early diagnosis of new cases, appropriate therapy, and accurate detection of outbreaks in the community, in order to implement proper TB control strategies. To achieve this goal, information from classical and molecular epidemiology, together with patient clinical data need to be combined. In this review, we summarize the methodologies currently used in molecular epidemiology, namely molecular typing. We will discuss their efficiency to phylogenetically characterize *Mycobacterium tuberculosis* isolates, and their ability to provide information that can be useful for disease control. We will also introduce next generation sequencing as the methodology that potentially could provide in a short time both, detection of new outbreaks and identification of resistance patterns. This could envision a potential of next generation sequencing as an important tool for accurate patient management and disease control.

Introduction

Effective TB surveillance programs rely on the capacity to promptly diagnose new cases, to provide an appropriate therapy, and to accurately detect outbreaks in the community, in order to implement proper TB control and elimination strategies.^{1,2} To achieve this goal, information from classical and molecular epi-

demology, together with patient clinical data need to be combined. In this review, we summarize the methodologies currently used in molecular epidemiology, namely molecular typing. We will discuss their efficiency to phylogenetically characterize TB isolates, and their ability to provide information that can be useful for disease control. We will also introduce next generation sequencing as the methodology that potentially could provide in a short time both, detection of new outbreaks and identification of resistance patterns. This could envision a potential of next generation sequencing as an important tool for accurate patient management and disease control.

Principles of classical genotyping

Mycobacterium tuberculosis complex (MTBC) is a monomorphic bacterium that emerged as a human pathogen from a genetic bottleneck and spread by clonal expansion globally.^{3,4} Given the high homology at genetic level, it is appropriate to refer to MTBC as a species and to *M. tuberculosis*, *M. bovis*, *M. africanum*, *M. microti* as subspecies or ecotypes.⁴ MTBC emerged about 70,000 years ago, accompanying migrations of anatomically modern humans out of Africa, and expanded as a consequence of increases in human population density during the Neolithic.⁵⁻⁸ This long co-evolutionary history is consistent with MTBC displaying characteristics indicative of both low and high host densities, with the identification of ancient and modern *M. tuberculosis* lineages, respectively.

The diversity of strains provides a means to identify instances of recent transmission of TB as well as the chains of transmission that occur among persons with TB. This diversity also helps to elucidate the patterns and dynamics of TB transmission.⁹ In addition to the possible impact of strain variation on the outcome of TB infection and disease,¹⁰ the diversity of strains is relevant for our understanding of drug resistance.¹¹ For example, the so-called *Beijing* family of strains has been associated with drug resistance in several reports.¹² Mathematical models predict that one of the most important factors influencing the future of MDR/XDR TB is the relative fitness of drug-resistant versus drug-susceptible strains. As in other bacteria, resistance-conferring mutations are often associated with a fitness cost in MTBC, which means reduction of growth rate, virulence, and transmissibility.¹³ However, in some areas of the world, such as the countries of the former Soviet Union where the prevalence of MDR is particularly high, MDR strains of MTBC are highly successful.¹⁴ The interaction between different genetic mutations, known as epistasis, probably plays a role in the

Correspondence: Angela Cannas, National Institute for Infectious Diseases "L. Spallanzani", Via Portuense 292, 00149 Rome, Italy, Tel.: +39.06.5517.0905. E-mail: angela.cannas@inmi.it

Key words: Tuberculosis; molecular typing; outbreak investigation; whole genome sequencing; multidrug resistance.

Contributions: the authors contributed equally.

Funding: this work was supported by grants from the Italian Ministry of Health "Ricerca Finalizzata 97/RF-2011-02348713" and "Ricerca Corrente INMI Spallanzani".

Conflict of interest: the authors declare no potential conflict of interest.

Received for publication: 29 April 2016. Accepted for publication: 29 April 2016.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

©Copyright A. Cannas et al., 2016
Licensee PAGEPress, Italy
Infectious Disease Reports 2016; 8:6567
doi:10.4081/idr.2016.6567

compensation of the fitness cost associated with the development of drug resistance, possibly increasing the level of resistance, or accommodating broader changes in the physiology of resistant bacteria.¹¹

These observations clearly indicate that continued inappropriate treatment, in part caused by misdiagnosis of resistance, drives the evolution of more transmissible, increasingly drug-resistant strains.¹⁵

Genotyping methods

Over the past two and half decades, genotyping has been used to support epidemiological investigation by matching isolates from patients with culture-confirmed TB, to identify TB transmission in certain communities and to define and characterize outbreaks.¹⁶ A number of methods have been used to type *M. tuberculosis* strains. Their evolution reflects the need of more sensitive, reliable, and less cumbersome techniques, in order to make typing of *M. tuberculosis* strains more feasible. Importantly, the availability of standardized techniques could help to compare results obtained in different laboratories.¹⁷

IS6110 fingerprinting

Among the first techniques that started to

be widely used as a molecular epidemiology tool, IS6110-based restriction fragment length polymorphism (RFLP) emerged as the most reliable and easier to standardize.¹⁸

IS6110-based RFLP genotyping detects variations generated by the insertion element IS6110, a MTBC-specific mobile genetic element with a highly conserved sequence, that can be found inserted in multiple copies anywhere in the genome in a process known as transposition. Strains can differ in both the number of copies of IS6110 and their positions in the bacterial genome.^{18,19} Genomic DNA isolated from a strain is cut by the restriction enzyme *PvuII*, generating hundreds of different fragments. These are then separated on an agarose gel and transferred to a membrane, that is probed with IS6110. The number of bands reflects the number of IS6110 copies in the genome and their position identifies a specific pattern that varies depending on the IS6110 position in the genome, so that patterns obtained from distantly related *M. tuberculosis* strains are different.¹⁶ IS6110-fingerprinting patterns are scanned and digitalized for computer analysis. A specific software measures genetic distances among strains, based on the number of bands with equal molecular weight. Analysis results are represented as a dendrogram that defines the genetic, and therefore epidemiological, relatedness among the *M. tuberculosis* strains analyzed. Although IS6110-DNA fingerprinting emerged as a useful adjunct to contact investigations, technical limitations (cost, laboriousness and difficulty in standardization) prevented its widespread use. Moreover, the fact that the number of IS6110 copies in *M. tuberculosis* strains can range from few copies to >30, introduced a significant bias when comparing strains with low and high IS6110 copies.²⁰

Spoligotyping

Spacer oligonucleotide typing (Spoligotyping) is a hybridization assay that detects variability in the direct repeat (DR) region in the DNA of *M. tuberculosis*. The DR region consists of multiple copies of a conserved 36-base-pair sequence separated by multiple unique spacer sequences. Different *M. tuberculosis* strains have various complements of the spacers, and these different complements form the basis of the assay.²¹ The DR region is amplified by PCR from the *M. tuberculosis* strains to be genotyped and amplicons are used as probes in a hybridization assay that contains 43 different unique spacer sequences that can be found in MTBC.

Results, based on the presence or absence of a spacer on a strain, are represented as a binary code or an octal designation, and are analyzed with the web tool SITVIT Database (www.pasteur-guadeloupe.fr:8081/

SITVITDemo). Strains are then assigned a specific clade based on the genetic diversity and grouped in a phylogenetic tree. Spoligotyping requires less genomic DNA, is less cumbersome and it is a standardized technique that has been widely used in many laboratories.²² Its main limitation is that strains may have the same pattern while not been very closely related, and therefore the technique may not be accurate enough to describe outbreaks in communities. Nevertheless, Spoligotyping is very useful to identify strains belonging to different clades or lineages, which can be used to get a snapshot of the *M. tuberculosis* strains circulating in a given geographic region or area.

MIRU-VNTR typing

Fifteen-locus and now 24-locus mycobacterial interspersed repetitive-unit-variable-number tandem-repeats (MIRU-VNTR) typing represents the current standard for *M. tuberculosis* genotyping. MIRU-VNTR methodology is based on analysis of DNA segments containing *tandem repeated* sequences in which the number of copies of the repeated sequence varies among strains. The method relies on PCR amplification and calculation of the number of repeats on the basis of the size of the amplified product. MIRU results are reported as 15 or 24-character designations, each character corresponding to the number of repeats at one of the loci, listed in a standard order. MIRU results are then analyzed by comparison with a strain database on the web-based tool MIRU-VNTR plus (www.miru-vntrplus.org).²⁰ MIRU-VNTR is easy to standardize and MIRU-VNTR typing shows a higher discriminatory power than spoligotyping, such that the combined use of the two methods provides a powerful tool to characterize *M. tuberculosis* outbreaks in the context of the TB pandemic at global level. Compared to IS6110 DNA fingerprinting, the combined use of MIRU-VNTR and spoligotyping is less cumbersome and more rapid. Moreover, the possibility to produce digital profiles that can be readily compared across laboratories offers an opportunity for standardization and comparison in dedicated databases and software.²³

Other tools in *M. tuberculosis* typing

In an effort to develop more rapid and economic tools to type *M. tuberculosis* strains, a number of DNA fingerprinting or PCR-based methodologies have been proposed. The use of probes that recognize repetitive GC-rich sequences in the *M. tuberculosis* genome (polymorphic GC-rich sequences) has been used in the RFLP DNA fingerprinting technique.²⁴ The use of PCR-based techniques, able to give results at higher speed and reduced cost, was implemented using a number of random primers or primers directed against repetitive but specific DNA sequences

in the *M. tuberculosis* genome.^{25,26} Most of these typing techniques did not warrant enough reproducibility and standardization, and none of them, included some commercial PCR-based methods, such as DiversiLab® (Biomérieux, Marcy-l'Étoile, France), were successful among TB laboratory users.²⁷

Next generation sequencing

Although the typing techniques described above target especially polymorphic genetic sequences, they interrogate less than 1% of the genome and have therefore an intrinsically restricted discriminatory power. This limitation could be overcome by the application of next generation whole genome sequencing (WGS) for genome-based epidemiology. Various platforms of WGS have been developed in the last decade: Roche 454 FLX (Roche Applied Science, Indianapolis, IN, USA), Illumina (Illumina, San Diego, CA, USA), SOLiDTM System (Applied Biosystems, Foster City, CA, USA), Heliscope™ (Helicos BioSciences Corporation, Cambridge, MA, USA), and nanopore sequencing.

The concept behind next generation sequencing (NGS) technology is simple: DNA polymerase catalyzes incorporation of fluorescently labeled deoxyribonucleotide triphosphates (dNTPs) into a DNA template strand during sequential cycles of DNA synthesis. During each cycle, at the point of incorporation, the nucleotides are identified by fluorescence excitation. The critical difference from previous techniques is that, instead of sequencing a single DNA fragment, NGS extends this process across millions of fragments in a massively parallel fashion.

NGS by Illumina (Illumina, San Diego, CA, USA) is currently the method of choice for WGS of *M. tuberculosis*. The workflow includes four basic steps: i) library preparation by random fragmentation of the DNA sample, followed by adapter ligation; ii) cluster generation by loading the library into a flow-cell, followed by amplification of each fragment of the library into distinct clonal clusters; iii) sequencing by a reversible terminator-based method that detects single bases as they are incorporated into DNA template strands; iv) data analysis and alignment of the newly identified sequence reads to a reference genome. Following alignment, a variety of analysis can be performed, such as single nucleotide polymorphism (SNP) or insertion-deletion identification, phylogenetic or metagenomic analysis, and more. This provides that NGS, in addition to be a useful tool for epidemiological studies, can add valuable information on drug resistance, virulence determinants, and genome evolution.

Whole genome sequencing for *M. tuberculosis* typing and contact tracing

In the monomorphic bacterium *M. tuberculosis*, genetic differences result from genomic rearrangements (insertion, deletion and duplication) and single nucleotide polymorphisms (SNPs). However, given the complexity to use genomic rearrangements to measure genetic and therefore epidemiological distances, variations in terms of SNPs are used to delineate genetic relatedness. Moreover, it is known that large portions of *M. tuberculosis* genome are very rich in GC, such as the genes encoding PE_PGRS and PPE_MPTR genes that are responsible for most of the genetic variability in MTBC.^{28,29} However, these regions cannot easily be resolved by NGS sequencing because of technological hurdles. Thus, these sequences are excluded, and only a portion of the genome is used for SNPs comparison.³⁰ In the last few years, a number of studies have highlighted the potential of NGS in molecular epidemiology of TB. In a retrospective cross-sectional study, WGS analysis of *M. tuberculosis* isolates from pulmonary and extra-pulmonary samples taken from the same patient was coupled with a longitudinal study where *M. tuberculosis* strains were isolated from the same patient over time. The results obtained from this study were used to estimate the rate of mutation of *M. tuberculosis* strains.⁹ In patients with relapsing TB, comparison of *M. tuberculosis* strains isolated over a period of 7-10 years resulted in up to a maximum 10 SNPs,³¹ while the cross-sectional analysis of *M. tuberculosis* strains isolated from samples obtained from the same patient and collected in a very close timeframe, resulted consistently in less than 4 SNPs. Based on these studies, it has been possible to estimate that *M. tuberculosis* evolves, in terms of single mutations, at ≈ 0.5 SNPs per year (for the part of the genome that can be reliably resolved by NGS), and that two strains are considered clustered when the difference in SNPs is < 12 .³¹ Hence, the 12 SNPs can be used as a cutoff to distinguish epidemiologically linked strains from those not related. Using a similar approach, a study conducted on strains isolated during more than two decades in two settings in Germany, demonstrated that WGS-based typing can provide epidemiologically relevant resolution of large *M. tuberculosis* outbreaks much more efficiently than classical genotyping. Genome-based analysis correlated better with contact tracing information and spatial-temporal patterns of the pathogen's spread. WGS provides a measure of *M. tuberculosis* genome evolution over time in its natural host context.³⁰

WGS was also applied to 26 *M. tuberculosis*

strains with an identical IS6110-DNA fingerprinting pattern, isolated during a TB outbreak in Hamburg. Interestingly, the results from WGS indicated a likely transmission for only 14 strains, while a number of previously linked cases was excluded.³² The results obtained with the newly available WGS analysis provided a new picture of the outbreak in Hamburg that was more consistent with the epidemiological data available.³²

Similar conclusions were drawn from a WGS analysis conducted on 247 TB strains isolated in the UK between 2007 and 2012. Thirty-nine cases were genomically linked within 13 clusters, implying 26 local transmission events. Only 11 of 26 possible transmissions had been previously identified through contact tracing.³³ Based on these evidences, WGS typing appears to offer an optimal resolution of *M. tuberculosis* isolates in molecular epidemiological studies, when compared to classical genotyping (Figure 1).

The possibility to perform WGS in many laboratories, its relative low cost compared to MIRU-VNTR and other methods, and the possibility to complete analysis in few days following isolation in culture of the *M. tuberculosis* strain, are opening the possibility for a widespread use of WGS in TB molecular epidemiology.^{16,34} Advances in software performances will soon enable sequence data analysis and interpretation to take place without the need for bioinformatics skills, although a consensus on the part of the genome that should be used for SNPs analysis is needed. The use of a defined number of genes such as in the SeqSphere+ software, based on a core genome multi locus sequence typing (cgMLST)

approach, may provide a valuable way for standardization of data analysis and for comparison among different laboratories.³²

Whole genome sequencing for tuberculosis diagnosis

The potential use of WGS in TB diagnosis and drug susceptibility testing (DST), in addition to its application to epidemiological analyses, makes this technology very attractive. Culture-based DST is currently the *gold standard*, but is technically cumbersome and time consuming. DST methods are not standardized and results may vary depending on the culture techniques employed.³⁵ In recent years, important aspects regarding the molecular basis of anti-TB drug resistance have been elucidated and molecular methods based on genomic DNA sequencing have been successfully used to detect the main mutations involved in drug resistance.³⁶ The possibility to analyze the complete genome sequence of a clinical strain obtained with WGS offers a unique opportunity to identify all the mutations potentially associated with drug resistance. Studies in which WGS data of a large number of strains collected worldwide are coupled with phenotypic DST results, drug treatment and clinical outcomes are providing key information that will be valuable to improve our understanding of drug resistance, including questions surrounding cross-resistance, heteroresistance, additive effects of compensatory mutations.³⁷ Interestingly, it is expected that detection of SNPs on certain genes associated with drug

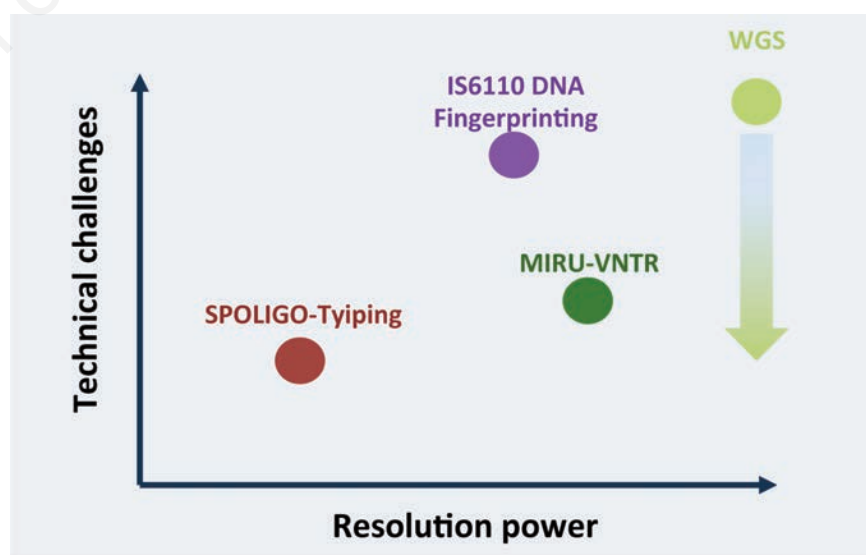


Figure 1. The different methods used in molecular typing of *Mycobacterium tuberculosis* show different features in terms of technical challenges and resolution power of the genetic distances between the strains analyzed.

resistance may provide more reliable data on the clinical susceptibility pattern of a given strain compared to traditional phenotypic DST results.

Given the enormous opportunities offered by the implementation of WGS in TB diagnosis, a comprehensive, standardized, and accurate resource for clinically relevant genetic mutations and the associated metadata (*e.g.*, geographic location, methodology, DST results) is currently under development.³⁷ A variety of advanced software tools exist for data analysis, however they require bioinformatic analysis skills and hardware infrastructure which are not always warranted in routine microbiology laboratories. To address these challenges, easy-to-use, web-based tools are needed and recently the online free tool PhyResSE (<https://bioinf.fz-borstel.de/mchips/phyresse>) has been developed to be available for larger communities.^{33,38,39} The WGS data obtained from an NGS platform, such as the widely used MySeq by Illumina, can be directly applied to the PhyResSE tool, to obtain epidemiological information (such as lineages and clades), together with highlighted SNPs associated with drug resistance, providing the subset of mutations predictive of treatment failure to any drug.³⁶ This web tool and other software that can be potentially developed in the future should be linked to databases that need to be continuously updated, so that any new mutation associated with drug resistance can be readily identified by the softwares.³⁹ Implementation of a rapid and user friendly

WGS analysis is of particular relevance in settings with high incidence of MDR-TB, where the impact of an early and effective identification of drug resistances can have clinical and public health implications.

As ongoing technological developments are rapidly decreasing costs, WGS has the potential to become the ultimate tool for typing and diagnosis purposes, and to dramatically amplify the impact of molecular diagnostics on clinical microbiology.⁴⁰

Whole genome sequencing in public health

Part of the challenge for controlling the TB pandemic requires a global political effort to address poverty and other underlying causes of the disease. However, TB control is also a local issue that requires public health policies and public health teams to interrupt the spread of this disease wherever it can be identified. The data produced by WGS could enable public health teams to target their contact investigations with greater confidence (Figure 2).^{9,31,33}

Taken together, these observations envision for the near future the potential to use such a powerful technological tools as NGS for a complete information, at the patient level, for diagnosis and decisions on therapy, and at the community level for epidemiological studies, both contributing to disease control. The

prospect of deploying drug-susceptibility testing based on WGS globally, including to settings where no phenotypic DST is currently performed, is an attractive possibility. Outbreaks could be detected from the same data at no additional cost, potentially adding to local TB control. Genomic analysis will thus allow simultaneous prediction of MTBC lineage and clade, first- and second line drug resistance, monitoring the emergence of new resistance mechanisms, and high resolution outbreak monitoring on a timescale week faster than with traditional diagnostics. Coupled with public health interventions, WGS will transform MTBC patient care and disease control.

Reference

- Lienhardt C, Glaziou P, Uplekar M, et al. Global tuberculosis control: lessons learnt and future prospects. *Nat Rev Microbiol* 2012;10:407-16.
- Sandgren A, Schepisi MS, Sotgiu G, et al. Tuberculosis transmission between foreign- and native-born populations in the EU/EEA: a systematic review. *Eur Respir J* 2014;43:1159-71.
- Gutierrez MC, Brisse S, Brosch R, et al. Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathog* 2005;1:e5.
- Galagan JE. Genomic insights into tuberculosis. *Nat Rev Genet* 2014;15:307-20.
- Hershberg R, Lipatov M, Small PM, et al. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol* 2008;6:e311.
- Gagneux S, DeRiemer K, Van T, et al. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* 2006;103:2869-73.
- Comas I, Chakravarti J, Small PM et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet* 2010;42:498-503.
- Comas I, Coscolla M, Luo T et al. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet* 2013;45:1176-82.
- Walker TM, Monk P, Smith EG et al. Contact investigations for outbreaks of *Mycobacterium tuberculosis*: advances through whole genome sequencing. *Clin Microbiol Infect* 2013;19:796-802.
- Portevin D, Gagneux S, Comas I, et al. Human macrophage responses to clinical isolates from the *Mycobacterium tuberculosis* complex discriminate between ancient and modern lineages. *PLoS Pathog*

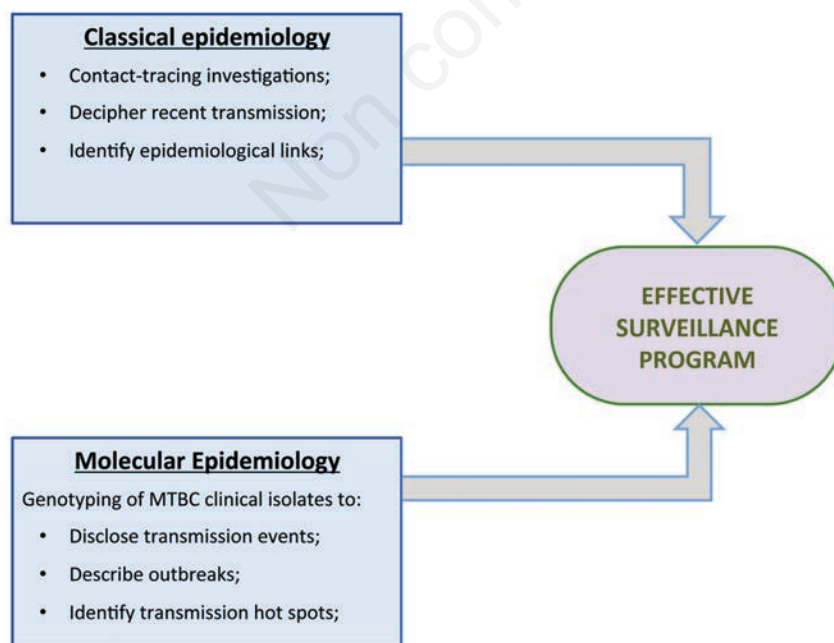


Figure 2. A concerted use of classical and molecular epidemiology tools is required to control tuberculosis at local and regional level.

- 2011;7:e1001307.
11. Borrell S, Gagneux S. Strain diversity, epistasis and the evolution of drug resistance in *Mycobacterium tuberculosis*. *Clin Microbiol Infect* 2011;17:815-20.
 12. Borrell S, Gagneux S. Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis* 2009;13:1456-66.
 13. Gagneux S, Long CD, Small PM, et al. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science* 2006;312:1944-6.
 14. Andersson DI, Hughes D. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat Rev Microbiol* 2010;8:260-71.
 15. Trauner A, Borrell S, Reither K, et al. Evolution of drug resistance in tuberculosis: recent progress and implications for diagnosis and therapy. *Drugs* 2014;74:1063-72.
 16. Niemann S, Supply P. Diversity and evolution of *Mycobacterium tuberculosis*: moving to whole-genome-based approaches. *Cold Spring Harb Perspect Med* 2014;4:a021188.
 17. Schurch AC, van SD. DNA fingerprinting of *Mycobacterium tuberculosis*: from phage typing to whole-genome sequencing. *Infect Genet Evol* 2012;12:602-9.
 18. van Embden JD, Cave MD, Crawford JT, et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* 1993;31:406-9.
 19. Sechi LA, Zanetti S, Delogu G, et al. Molecular epidemiology of *Mycobacterium tuberculosis* strains isolated from different regions of Italy and Pakistan. *J Clin Microbiol* 1996;34:1825-8.
 20. Mazars E, Lesjean S, Banuls AL, et al. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. *Proc Natl Acad Sci USA* 2001;98:1901-6.
 21. Kamerbeek J, Schouls L, Kolk A, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 1997;35:907-14.
 22. Goguet de la Salmoniere YO, Li HM, Torrea G, et al. Evaluation of spoligotyping in a study of the transmission of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1997;35:2210-4.
 23. Oelemann MC, Fontes AN, Pereira MA, et al. Typing of *Mycobacterium tuberculosis* strains isolated in community health centers of Rio de Janeiro City, Brazil. *Mem Inst Oswaldo Cruz* 2007;102:455-62.
 24. van Soolingen D, de Haas PE, Hermans PW, et al. Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1993;31:1987-95.
 25. Sola C, Horgen L, Devallois A, et al. Combined numerical analysis based on the molecular description of *Mycobacterium tuberculosis* by four repetitive sequence-based DNA typing systems. *Res Microbiol* 1998;149:349-60.
 26. Sechi LA, Zanetti S, Dupre I, et al. Enterobacterial repetitive intergenic consensus sequences as molecular targets for typing of *Mycobacterium tuberculosis* strains. *J Clin Microbiol* 1998;36:128-32.
 27. Masala S, Molicotti P, Bua A, et al. Molecular characterization of Sardinian *Mycobacterium tuberculosis* isolates by IS6110 restriction fragment length polymorphism, MIRU-VNTR and rep-PCR. *New Microbiol* 2010;33:155-62.
 28. Delogu G, Cole ST, Brosch R. The PE and PPE protein families of *Mycobacterium tuberculosis*. In: Kaufmann SH, Rubin E, eds. *Handbook of tuberculosis*. Weinheim: Wiley-VCH Verlag GmbH Co. KGaA; 2008. pp 131-150.
 29. Copin R, Coscolla M, Seiffert SN, et al. Sequence diversity in the *pe_pgrs* genes of *Mycobacterium tuberculosis* is independent of human T cell recognition. *MBio* 2014;5:e00960-13.
 30. Roetzer A, Diel R, Kohl TA, et al. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. *PLoS Med* 2013;10:e1001387.
 31. Walker TM, Ip CL, Harrell RH, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. *Lancet Infect Dis* 2013;13:137-46.
 32. Kohl TA, Diel R, Harmsen D, et al. Whole-genome-based *Mycobacterium tuberculosis* surveillance: a standardized, portable, and expandable approach. *J Clin Microbiol* 2014;52:2479-86.
 33. Walker TM, Lalor MK, Broda A, 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.
 34. Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect Dis* 2015;15:1193-202.
 35. Van DA, Wright A, Zignol M, et al. Drug susceptibility testing proficiency in the network of supranational tuberculosis reference laboratories. *Int J Tuberc Lung Dis* 2011;15:116-24.
 36. Dominguez J, Boettger EC, Cirillo D, et al. Clinical implications of molecular drug resistance testing for *Mycobacterium tuberculosis*: a TBNET/RESIST-TB consensus statement. *Int J Tuberc Lung Dis* 2016;20:24-42.
 37. Starks AM, Aviles E, Cirillo DM, et al. Collaborative effort for a centralized worldwide tuberculosis relational sequencing data platform. *Clin Infect Dis* 2015;61:S141-6.
 38. Feuerriegel S, Schleusener V, Beckert P, et al. PhyResSE: a web tool delineating *Mycobacterium tuberculosis* antibiotic resistance and lineage from whole-genome sequencing data. *J Clin Microbiol* 2015;53:1908-14.
 39. Salamon H, Yamaguchi KD, Cirillo DM, et al. Integration of published information into a resistance-associated mutation database for *Mycobacterium tuberculosis*. *J Infect Dis* 2015;211:S50-7.
 40. Pankhurst LJ, Del Ojo EC, Votintseva AA, et al. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. *Lancet Respir Med* 2016;4:49-58.