

_ Nova era na vigilância da tuberculose multirresistente em Portugal: sequenciação do genoma completo

New era on the surveillance of multidrug resistant tuberculosis in Portugal: whole genome sequencing

Rita Macedo¹, Miguel Pinto², Vítor Borges², Alexandra Nunes², Joana Isidro², Sílvia Duarte³, Luís Vieira^{3,4}, João Paulo Gomes²

rita.macedo@insa.min-saude.pt

(1) Laboratório Nacional de Referência para as Micobactérias. Departamento de Doenças Infecciosas, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal.

(2) Núcleo de Bioinformática. Departamento de Doenças Infecciosas, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal.

(3) Unidade de Tecnologia e Inovação, Departamento de Genética Humana, Instituto Nacional de Doutor Ricardo Jorge, Lisboa, Portugal.

(4) Centro de Toxicogenómica e Saúde Humana (ToxOmics), Genética, Oncologia e Toxicologia Humana da NOVA Medical School/Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisboa, Portugal.

_Resumo

A tuberculose multirresistente continua a ser um dos principais desafios no controlo da Tuberculose em Portugal. Os métodos de diagnóstico e vigilância clássicos são trabalhosos, muito demorados e nem sempre fáceis de interpretar. A introdução de metodologias moleculares permitiu ultrapassar alguns destes obstáculos, mas, sendo limitada no número de alvos genéticos que analisa, não permite uma análise completa das características das estirpes de *M. tuberculosis* isoladas. Para ultrapassar esta questão, tivemos como objectivo a implementação de uma metodologia baseada na sequenciação total do genoma que, para além de permitir um *screening* de todas as mutações conhecidas associadas a resistência, permite fazer uma vigilância molecular das estirpes com uma sensibilidade muito elevada, possibilitando a intervenção das Autoridades de Saúde de forma atempada e otimizada. O Laboratório Nacional de Referência de Micobactérias/Tuberculose do Instituto Nacional de Doutor Ricardo Jorge está, neste momento, capacitado para responder eficazmente a estas necessidades garantindo um diagnóstico e vigilância em "tempo real" de todas as estirpes de *M. tuberculosis* multirresistentes.

_Abstract

*Multidrug-resistant tuberculosis continues to be one of the main challenges for Tuberculosis' control in Portugal. Classical diagnostic and surveillance methods are laborious, time-consuming and not always easy to interpret and perform. The introduction of molecular methodologies overcame some of these obstacles, but, as they are limited in the number of genetic targets analyzed, they do not allow a complete analysis of the characteristics of the isolated *M. tuberculosis* strains. As such, we aimed to implement a methodology based on whole genome sequencing that, in addition to enable the screening of all known mutations associated with resistance, it makes it possible to carry out molecular surveillance of the strains with a very high sensitivity, allowing a timely and optimized intervention of the Health Authorities. The National Reference Laboratory of Mycobacteria / Tuberculosis is now capable of responding effectively to these needs, ensuring a "real time" diagnosis and surveillance of all multidrug resistant *M. tuberculosis* strains.*

_Introdução

A Tuberculose (TB) continua a ser uma das principais causas de morte por doença infecciosa em todo o Mundo (1,2). Em 2017, 10.4 milhões de pessoas adoeceram, 1.7 milhões morreram da doença e, embora a incidência de TB tenha vindo a decrescer cerca de 2% por ano, o problema da resistência aos antibióticos e, em particular, da multirresistência, desafia o cumprimento das metas de erradicação previstas pela Organização Mundial da Saúde (OMS) para 2030 (1). Com 460 mil casos (~5% do total de casos diagnosticados) em 2017 causados por estirpes de TB multirresistente (TB-MR, ou seja, resistentes a rifampicina e isoniazida), das quais, 8,5% são também extensivamente resistentes (TB-XDR; isto é, TB-MR com resistência adicional a qualquer fluoroquinolona e amicacina/canamicina ou capreomicina) e considerando que a transmissão pessoa-a-pessoa é a principal via de contágio, a monitorização e controlo destes casos é fundamental para um programa de controlo de TB ter sucesso (2). Em Portugal, a incidência de TB tem também vindo a diminuir nos últimos anos, com uma média de decréscimo de cerca de 5% ao ano (3). De 2013 a 2017, cerca de 10.000 novos casos de TB foram notificados à Direção-Geral da Saúde (DGS) e a proporção de doentes com TB-MR permaneceu estável, correspondendo a 1% do total de casos diagnosticados (3,4).

Tendo em consideração que a maioria das mortes por TB pode ser minimizada através do diagnóstico precoce e tratamento adequado, a OMS propôs, em 2015, a expansão do diagnóstico laboratorial rápido como uma das cinco prioridades para combater a crise global de casos de TB resistente aos

antibacilares (1). Os testes fenotípicos, usando métodos convencionais, podem levar até oito semanas para o isolamento e identificação das estirpes do complexo *Mycobacterium tuberculosis* (MTC) e duas a quatro semanas adicionais até que o perfil de resistência aos antibacilares seja conclusivo (5). Obviamente, a introdução de ferramentas de diagnóstico molecular para a detecção de resistência permite acelerar o diagnóstico da TB-M/XDR, evitando a disseminação da doença e permitindo opções terapêuticas dirigidas. Actualmente, existem já inúmeros testes moleculares que permitem a detecção rápida de mutações, mas o número de alvos genéticos pesquisados é muito limitado. Por outro lado, as intervenções direcionadas para interromper cadeias de transmissão exigem um conhecimento epidemiológico exaustivo e fundamentado que só pode ser conseguido aliando a informação genética à informação resultante da investigação epidemiológica. Por estes motivos, têm sido inúmeros os desenvolvimentos laboratoriais na área da vigilância molecular da TB e, actualmente, a sequenciação do genoma completo (*whole genome sequencing* - WGS) tem o potencial para se tornar a ferramenta de eleição para a genotipagem das estirpes do complexo MTC (6-9), fornecendo informações genéticas abrangentes, incluindo todos os possíveis alvos genéticos que podem dar informações rápidas sobre a resistência aos antibacilares e virulência das estirpes MTC (6,10-12). Para além disso, estão a ser desenvolvidos esforços no sentido de permitir aplicar métodos baseados em WGS diretamente a partir da amostra clínica, passando-se assim das 8 semanas necessárias ao crescimento das estirpes de MTC, para um conhecimento genómico completo da mesma em menos de uma semana após colheita do produto biológico (13).

Desde 2014 que, em Portugal, existem centros específicos para o diagnóstico, consultoria, acompanhamento e tratamento dos casos de TB-M/XDR. Estes centros permitem agilizar a ligação entre a investigação epidemiológica realizada pelas Autoridades de Saúde Pública e os dados laboratoriais que são obtidos de forma sistemática pelo Laboratório Nacional de Referência de Tuberculose (LNR-TB) do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA). Dado que o LNR-TB recebe as estirpes isoladas de todos os doentes com TB-MR

diagnosticada em Portugal (obrigatório desde 2007) (14), é de extrema importância definir um sistema de genotipagem molecular centralizado e robusto que mais rapidamente estabeleça a associação entre os dados genéticos e epidemiológicos, tendo em vista a deteção e monitorização em “tempo real” dos perfis de resistência e possíveis cadeias de transmissão.

_Objetivo

O presente estudo descreve a implementação de uma abordagem baseada em WGS, para a vigilância da TB e deteção de mutações associadas à resistência aos antibacilares, através da avaliação retrospectiva de todas as estirpes de TB-M/XDR isoladas nos últimos 5 anos em Portugal.

_Materiais e métodos

Do total das cerca de 10.000 estirpes de MTC isoladas em Portugal entre 2013 e 2017, 96 eram casos de TB-M/XDR. Destas, 83 (86,5%) estavam disponíveis para análise sendo que os restantes casos foram diagnosticados apenas por biologia molecular não tendo havido isolamento da estirpe. Os testes de sensibilidade aos antibacilares (1ª e 2ª linha: Isoniazida, Rifampicina, Estreptomina, Etambutol, Pirazinamida, Etionamida, Fluoroquinolonas, Canamicina, Amicacina, Linezolida, Cicloserina e PAS) e a genotipagem pelo método tradicional, *Mycobacterium Interspersed Repetitive Units* (MIRU-VNTR) foram efetuados de acordo com as indicações do fabricante. A metodologia de WGS foi realizada como descrito num estudo nosso recentemente publicado (12). Para a extração *in silico* dos perfis genéticos de resistência foi utilizado o *software* bioinformático TB profiler v0.3.0 (12,15). A análise bioinformática usada para a vigilância com base em WGS assenta numa abordagem gene-a-gene (*multilocus sequence typing*) que tira partido de um conjunto de 3656 genes de MTC (7,8) e utiliza os *softwares* de utilização livre chewBBACA (16) e PHYLOViZ disponíveis online (17).

Dado que o estabelecimento de valores de *cutoff* para proximidade genética é ainda um assunto em investigação, optámos por uma abordagem de *cutoff* relativa, a qual tem em conta o número de diferenças genéticas para um deter-

minado denominador (i.e., número de genes em análise). Por outro lado, este é um processo dinâmico pois o próprio conhecimento de links epidemiológicos contribui para a definição de se determinada distância genética é próxima ou distante.

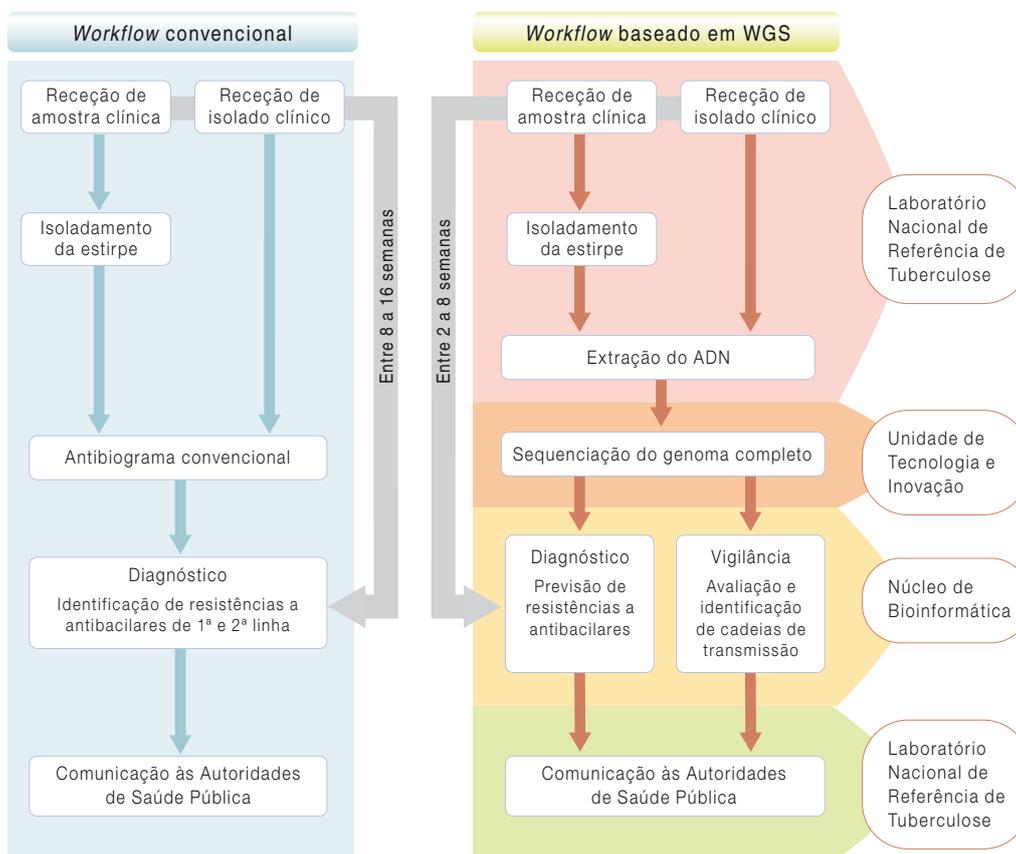
Resultados e discussão

O presente estudo descreve a implementação de uma metodologia de vigilância laboratorial baseada em WGS, tendo em conta a exigência supranacional, a curto prazo, da substituição dos métodos tradicionais de genotipagem por MIRU-VNTR por uma abordagem assente na sequenciação do genoma completo. Para além de permitir analisar a totalidade do genoma para determinar potenciais cadeias de transmissão, esta metodologia permite, simultaneamente, a deteção dos vários marcadores genéticos de resistência (figura 1).

Para este efeito, 83 estirpes de TB-M/XDR isoladas entre 2013 e 2017, foram caracterizadas por MIRU-VNTR e por WGS em simultâneo. Este é o primeiro estudo a ser realizado em Portugal com estes objetivos específicos e envolve todas as estirpes M/XDR isoladas nos últimos 5 anos.

Todas as estirpes foram analisadas relativamente à presença/ausência de mutações associadas à resistência aos antibióticos de 1ª e 2ª linha através da utilização de plataformas gratuitas e *online* que permitem usar diretamente os dados de sequenciação de nova geração. Para a validação destes dados foi efetuado um estudo piloto (12) que definiu a estratégia adotada pelo LNR-TB e a plataforma com maior congruência (TB Profiler) com os resultados fenotípicos. Observou-se uma correlação superior a 90% entre os resultados genotípicos e os resultados obtidos com os testes de suscetibilidade convencionais. Desta forma, através da sequenciação total do

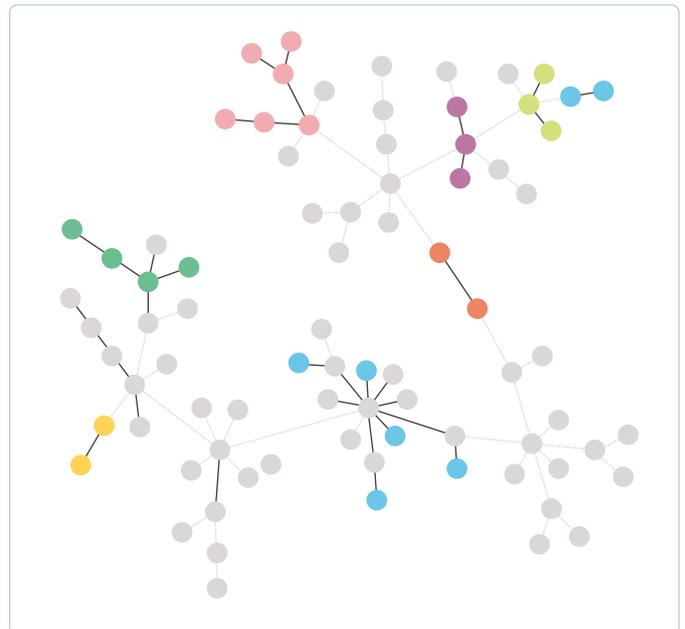
Figura 1: Fluxograma do diagnóstico laboratorial baseado em metodologias convencionais e em sequenciação do genoma completo (*whole genome sequencing - WGS*).



genoma de uma estirpe MTC é possível prever fenótipos de resistência em cerca de uma semana, acelerando o diagnóstico de um possível caso de TB-M/XDR em cerca de 10 semanas.

Neste estudo, e relativamente à definição de possíveis cadeias de transmissão verificou-se, numa primeira abordagem e tal como seria previsto, que estirpes com o mesmo perfil MIRU-VNTR revelaram maior proximidade genética, muito embora também se tenha observado a existência de ligações genéticas muito próximas entre estirpes com perfis distintos de MIRU-VNTR. Para validar a definição dos *clusters* encontrados e usando esta nova abordagem baseada em WGS, correlacionou-se os dados moleculares obtidos com a informação epidemiológica disponível acerca dos doentes. Como esperado, observou-se existir uma boa correlação entre a proximidade genética das estirpes e a força do vínculo epidemiológico dos doentes. De facto, todas as estirpes isoladas de doentes com ligações fortes (familiares, amigos e colegas de trabalho) apresentaram uma elevada proximidade genética, isto é, com um cutoff abaixo de 0,4% de diferenças alélicas entre as estirpes (figura 2). Quanto aos doentes sem ligação epidemiológica conhecida as diferenças genéticas foram superiores a o cutoff estabelecido e aos limites previamente descritos noutros estudos envolvendo fenómenos de transmissão recentes (7,8,18-20). No entanto, é importante referir que encontramos estirpes muito próximas para as quais não havia dados epidemiológicos disponíveis, levantando a hipótese de que algumas ligações entre doentes não terão sido detetadas. Esta observação ilustra a necessidade de melhorar a deteção e investigação epidemiológica precoces (por exemplo, melhorando as investigações “de campo” e fortalecendo a comunicação entre as autoridades de saúde e o LNR-TB) para aliar os benefícios de uma vigilância laboratorial focada na deteção/confirmação de relações entre doentes, mais rápida e baseada em WGS, facilitando assim o rastreio de contatos. Finalmente, considerando que esta metodologia se baseia totalmente em ferramentas bioinformáticas gratuitas e disponíveis *online*, acredita-se que a abordagem do presente estudo pode ser muito útil para todos os laboratórios com acesso a estas tecnologias, especialmente para aqueles com menos recursos.

Figura 2: ↓ Árvore filogenética simplificada de todas as estirpes analisadas, ilustrando a proximidade genética das estirpes de MTC baseada na aplicação de WGS.



Cada círculo representa um perfil genético; os círculos a cores representam grupos moleculares de estirpes com ligação epidemiológica confirmada; a cinzento – são grupos moleculares sem informação epidemiológica disponível. As estirpes com proximidade genética (i.e., abaixo do *cut-off* definido no estudo) estão ligadas com linhas contínuas.

_Conclusão

O LNR-TB do INSA está, neste momento, em condições de garantir uma vigilância laboratorial baseada em WGS que, por ser mais robusta e atempada, permite a deteção rápida de novos casos de doença, a monitorização de cadeias de transmissão existentes e a determinação célere de novas cadeias. A articulação entre as diferentes unidades do INSA (figura 1) têm garantido o sucesso da execução rápida deste *workflow* de diagnóstico, diminuindo os tempos de resposta e aumentando o poder de discriminação e confiança dos resultados obtidos. A centralização deste diagnóstico e a genotipagem molecular por WGS, associadas à melhoria da articulação com as entidades clínicas e de saúde pública, constitui a força motriz para uma vigilância mais rápida e efetiva dos casos de TB-M/XDR, favorecendo assim a previsão da resistência aos antibióticos e a deteção precoce de cadeias de transmissão.

Agradecimentos

Este trabalho foi financiado pela Fundação para a Ciência e a Tecnologia (FCT) no âmbito do projeto Centro de Toxicogenómica e Saúde Humana (UID/BIM/0009/2016) e desenvolvido no âmbito do projeto GenomaPT (POCI-01-0145-FEDER-022184), cofinanciado pelo Programa Operacional Competitividade e Internacionalização (Compete2020), Programa Operacional Regional de Lisboa (Lisboa 2020) e Programa Operacional Regional do Algarve (CRESC Algarve2020), através do Portugal 2020 e do Fundo Europeu de Desenvolvimento Regional (FEDER), e pela FCT.

Agradece-se à Doutora Cristina Furtado pela revisão do artigo.

Referências bibliográficas:

- (1) World Health Organization. Global tuberculosis report 2015. 20th ed. Geneva: WHO, 2015. <http://www.who.int/iris/handle/10665/191102>.
- (2) World Health Organization. Global tuberculosis report 2018. Geneva: WHO, 2018. https://www.who.int/tb/publications/global_report/en/
- (3) Programa Nacional para a Tuberculose. Tuberculose em Portugal: Desafios e estratégias. 2018. Lisboa: Direção-Geral da Saúde, 2018. <https://www.dgs.pt/portal-da-estatistica-da-saude/diretorio-de-informacao/diretorio-de-informacao/por-serie-963780-pdf.aspx?v=11736b14-73e6-4b34-a8e8-d22502108547>
- (4) Programa Nacional para a Infeção VIH, Sida e Tuberculose. Programa Nacional para a Infeção VIH, SIDA e Tuberculose 2017. Lisboa: Direção-Geral da Saúde, 2017. <http://www.dgs.pt/portal-da-estatistica-da-saude/diretorio-de-informacao/diretorio-de-informacao/por-serie-875387-pdf.aspx?v=11736b14-73e6-4b34-a8e8-d22502108547>
- (5) Pfyffer GE, Wittwer F. Incubation time of mycobacterial cultures: how long is long enough to issue a final negative report to the clinician? *J Clin Microbiol.* 2012;50(12):4188-9. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3502948/>
- (6) 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(2):e1001387. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3570532/>
- (7) 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(7):2479-86. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4097744/>
- (8) Kohl TA, Harmsen D, Rothgänger J, et al. Harmonized Genome Wide Typing of Tubercle Bacilli Using a Web-Based Gene-By-Gene Nomenclature System. *EBioMedicine.* 2018;34:131-8. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6116475/>
- (9) Bravo LT, Tuohy MJ, Ang C, et al. Pyrosequencing for rapid detection of Mycobacterium tuberculosis resistance to rifampin, isoniazid, and fluoroquinolones. *J Clin Microbiol.* 2009;47(12):3985-90. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2786679/>
- (10) Boehme CC, Nabeta P, Hillemann D, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med.* 2010;363(11):1005-15. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2947799/>
- (11) Satta G, Atzeni A, McHugh TD. Mycobacterium tuberculosis and whole genome sequencing: a practical guide and online tools available for the clinical microbiologist. *Clin Microbiol Infect.* 2017;23(2):69-72. Epub 2016 Sep 15. <https://doi.org/10.1016/j.cmi.2016.09.005>
- (12) Macedo R, Nunes A, Portugal I, et al. Dissecting whole-genome sequencing-based online tools for predicting resistance in Mycobacterium tuberculosis: can we use them for clinical decision guidance? *Tuberculosis (Edinb).* 2018;110:44-51.
- (13) Satta G, Lipman M, Smith GP, et al. Mycobacterium tuberculosis and whole-genome sequencing: how close are we to unleashing its full potential? *Clin Microbiol Infect.* 2018;24(6):604-609. Epub 2017 Nov 3. <https://dx.doi.org/10.1016/j.cmi.2017.10.030>
- (14) Direção-Geral da Saúde. Circular Normativa nº 1/DT de 11/01/2007. Testes de Sensibilidade aos Antituberculosos de 2ª Linha. <https://www.dgs.pt/directrizes-da-dgs/normas-e-circulares-normativas/circular-normativa-n-1dt-de-11012007.aspx>
- (15) Coll F, McNERNEY R, Preston MD, et al. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Med.* 2015;7(1):51. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4446134/>
- (16) Silva M, Machado MP, Silva DN, Rossi M, Moran-Gilad J, Santos S, Ramirez M, Carriço JA. chewBBACA: A complete suite for gene-by-gene schema creation and strain identification. *Microbial Genomics* 2018; 4(3): e000166.
- (17) Ribeiro-Gonçalves B, Francisco AP, Vaz C, et al. PHYLOViZ Online: web-based tool for visualization, phylogenetic inference, analysis and sharing of minimum spanning trees. *Nucleic Acids Res.* 2016;44(W1):W246-51. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4987911/>
- (18) Cabibbe AM, Walker TM, Niemann S, et al. Whole genome sequencing of Mycobacterium tuberculosis. *Eur Respir J.* 2018;52(5). pii: 1801163.
- (19) 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(2):137-46. Epub 2012 Nov 15. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3556524/>
- (20) Cabibbe AM, Trovato A, De Filippo MR, et al. Countrywide implementation of whole genome sequencing: an opportunity to improve tuberculosis management, surveillance and contact tracing in low incidence countries. *Eur Respir J.* 2018;51(6). pii: 1800387.