



# Clinical and public health utility of *Mycobacterium tuberculosis* whole genome sequencing

Alice Kizny Gordon<sup>a,b,\*</sup>, Ben Marais<sup>c,d</sup>, Timothy M. Walker<sup>e</sup>, Vitali Sintchenko<sup>a,b,c,d</sup>

<sup>a</sup> Centre for Infectious Diseases and Microbiology – Public Health, Westmead Hospital, Western Sydney Local Health District, Sydney, New South Wales, Australia

<sup>b</sup> Sydney Medical School, Faculty of Medicine and Health, University of Sydney, Sydney, New South Wales, Australia

<sup>c</sup> WHO Collaborating Centre for Tuberculosis, The University of Sydney, Sydney, New South Wales, Australia

<sup>d</sup> Marie Bashir Institute for Infectious Diseases and Biosecurity, The University of Sydney, Sydney, New South Wales, Australia

<sup>e</sup> Oxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam

## ARTICLE INFO

### Article history:

Received 29 January 2021

Received in revised form 18 February 2021

Accepted 25 February 2021

### Keywords:

*Mycobacterium tuberculosis*

Whole genome sequencing

Multidrug-resistance

## ABSTRACT

The World Health Organization (WHO) estimates that around 10 million people develop tuberculosis (TB) every year, with 1.5 million deaths attributed to TB in 2019 (World Health Organization, 2020). The majority of the disease burden occurs in low-income countries, where access to diagnostics and tailored treatment remains problematic. The current COVID-19 pandemic further threatens to impact global TB control by diverting resources, reducing notifications and hence significantly increasing deaths attributable to TB (World Health Organization, 2020).

Whole genome sequencing (WGS) is becoming increasingly accessible, and has particular value in the diagnosis and management of TB disease (Cabibbe et al., 2018; Meehan et al., 2019). Not only does it have the potential to give more rapid and complete information on drug-resistance, but the high discriminatory power it offers allows detection of clusters and transmission pathways, as well as likely contamination events, mixed infections and to differentiate between re-infection and relapse with much greater confidence than previous typing methods.

© 2021 The Authors. Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

The basic biology of *Mycobacterium tuberculosis*, with its lack of horizontal gene transfer, simplifies bioinformatics analyses and makes it an ideal organism to establish 'proof of principle' for the clinical and public health value of WGS. It has been adopted as an integral part of the public health strategy in several low-burden, high-income countries with the benefits previously described (Gurjav et al., 2016; Shea et al., 2017; Olaru et al., 2018). Routine WGS results will also facilitate timely cross-jurisdictional communication and data-sharing in the management of outbreaks that cross national or international borders. Initially the costs required for sequencing infrastructure, maintenance and expertise were high, but this has been considerably reduced over the last two decades. Big challenges remain however in the availability and

funding of adequately trained bioinformaticians, while the cost and sophistication of the technology, as well as the large number of tuberculosis patients, limits its use in most high TB incidence countries with limited resources. The potential benefits and challenges of using WGS in these different settings are summarised in Table 1.

Drug-resistant TB poses a major global public health threat, since it is associated with greater morbidity and mortality, and potentially drug resistance amplification and increased transmission if management is sub-optimal. It requires prolonged treatment with expensive, difficult to access and potentially more toxic second-line drugs. In 2019 there was an estimated half a million new cases of rifampicin-resistant TB; 78% of which had multidrug-resistant TB (MDR-TB) (World Health Organization, 2020). Globally, 3.3% of new cases and 17.7% of previously treated cases had drug-resistant TB, indicating active transmission of these strains within communities (World Health Organization, 2020). Large outbreaks of drug resistant strains that have gone undetected for decades when smear microscopy was the only available diagnostic tool have been well characterised with the advent of

\* Corresponding author at: Centre for Infectious Diseases & Microbiology – Public Health (CIDM-PH), Level 3, Institute of Clinical Pathology & Medical Research (ICPMR), Westmead Hospital, Cnr Hawkesbury & Darcy Roads, Westmead, New South Wales 2145, Australia.

E-mail address: [alice\\_kg@hotmail.com](mailto:alice_kg@hotmail.com) (A. Kizny Gordon).

**Table 1**  
Benefits and challenges of WGS in TB diagnostics.

WGS considerations	All settings	Specific priorities/challenges in high TB incidence settings
Benefits	<p><u>Public health</u></p> <ul style="list-style-type: none"> <li>• Drug resistance surveillance</li> <li>• Outbreak/cluster recognition</li> <li>• Identifying unknown cases in the transmission chain</li> <li>• Data-sharing/cross-jurisdictional communication</li> </ul> <p><u>Individual patient</u></p> <ul style="list-style-type: none"> <li>• Drug resistance prediction</li> <li>• Differentiating relapse from re-infection</li> <li>• Identifying mixed infections</li> <li>• Identifying laboratory contamination events</li> </ul>	<p><u>High TB incidence setting</u></p> <p>Routine WGS in every single patient not feasible:</p> <ul style="list-style-type: none"> <li>• Focus on drug resistance surveillance to inform optimal programmatic treatment options</li> <li>• Detailed transmission tracking in particular ‘hot spot’ areas</li> <li>• Monitor for resistance against newer drugs, such as bedaquiline, in MDR-TB patients</li> <li>• Monitor laboratory contamination events to identify potential ‘systemic issues’</li> </ul>
Challenges	<ul style="list-style-type: none"> <li>• Costs – infrastructure, maintenance</li> <li>• Expertise</li> <li>• Quality control/assurance</li> <li>• Current requirement for culture</li> <li>• Comprehensive genotype-phenotype databases</li> <li>• Data handling &amp; security</li> <li>• Data handling &amp; security</li> <li>• Data interoperability</li> </ul>	

TB – tuberculosis; WGS – whole genome sequencing; MDR-TB – multidrug resistant TB.

WGS (Bainomugisa et al., 2018). In 2016 it was estimated that 6.2% of MDR-TB cases had additional drug-resistance causing extensively drug-resistant TB (World Health Organization., 2017), which is even more difficult to treat.

The most commonly used microbiological methods for detecting antimicrobial resistance involve phenotypic drug susceptibility testing (DST) using liquid broth culture, mostly using the Mycobacteria Growth Indicator Tube (MGIT) system. While this has been considered the gold-standard in TB antimicrobial resistance diagnostics, it still takes a long time to obtain results as *M. tuberculosis* is a slow growing organism. Moreover, the expertise and infrastructure required is often in excess of what is available. With access to routine WGS a complete resistome may be available within days of the initial positive culture. While several studies have evaluated the correlation between certain resistance polymorphisms and phenotypic resistance (Rodwell et al., 2014; Walker et al., 2015; Miotto et al., 2017; CRyPTIC Consortium, 2018), to date this has been most complete and useful for first-line drugs. Continued research in this area will aid in improving resistance prediction and creating comprehensive genome-based drug susceptibility profiling against first- and second-line drugs.

In addition, several studies have shown accurate genomic DST prediction when sequencing directly from patient specimens, using enhanced methods for human nucleic acid depletion and target amplification, which has the potential to dramatically reduce turn-around times and the requirement for expensive biosafety facilities to first grow these drug-resistant TB strains (Brown et al., 2015; Votintseva et al., 2017; Doyle et al., 2018). WGS may also allow the detection of novel drug resistance mechanisms when performed in conjunction with phenotypic DST, and help to determine whether resistance is due to *de novo* acquisition or clonal spread, as well as inform decisions about which resistance mutations should be included on more accessible extended molecular panels (Cohen et al., 2019). Ultimately, accurate and prompt diagnosis of MDR-TB allows more effective TB control through more timely initiation of appropriate treatment of individual patients and institution of the necessary public health measures to limit its spread.

The superior resolution of WGS compared to conventional typing methods, such as mycobacterial interspersed repetitive units-variable number tandem repeats (MIRU-VNTR), provides

several key advantages. Firstly, it has given us deeper insight into the transmission dynamics of TB within communities, facilitating the detection of clusters and probable transmission pathways with much greater confidence; especially in strains that are poorly differentiated by MIRU-VNTR such as the Beijing lineage (Walker et al., 2013; Stucki et al., 2016). The implementation of routine WGS will allow detection of unsuspected transmission events, as well as highly contagious patients in order to guide improved public health containment measures. Secondly, the greater discriminatory power of WGS may allow for the distinction between cases of endogenous relapse from exogenous reinfection (Bryant et al., 2013; Witney et al., 2017), which can provide invaluable information when evaluating treatment adherence and the efficacy of current drug regimens. Thirdly, WGS has improved the ability to detect mixed infections with several different strains of *M. tuberculosis* (Tarashi et al., 2017; Wyllie et al., 2018). Mixed infections may involve strains with different susceptibility patterns, which has important implications for treatment selection and therapeutic outcomes. And lastly, the resolution offered by WGS has enabled the detection of probable laboratory contamination events with much greater precision than previous typing methods (Min et al., 2019; Wu et al., 2019; Kizny Gordon et al., 2021). This has major consequences for individual patients by preventing unnecessary exposure to prolonged courses of toxic drugs, and also has value in highlighting possible failures in institutional infection control or laboratory processes.

Despite the challenges of 2020, global travel and migration are likely to return to pre-COVID rates in the future. While cluster detection and contact tracing have been less of a priority than diagnosis and treatment in high burden countries due to constrained resources, with access to WGS these public health measures may be better targeted and more feasible. Although it remains problematic, collecting genomic data in a standardised format will allow easier data exchange and comparison across jurisdictions, facilitating early and accurate outbreak detection or even recognition of laboratory contamination events (Kizny Gordon et al., 2021; Meehan et al., 2019; Walker et al., 2018). This has important implications for an improved understanding of *M. tuberculosis* transmission dynamics providing better guidance for local and ultimate global TB control. Although developed countries are ‘leading the way’ in integrating routine WGS into

clinical and public health practice, it is important to consider strategies that would extend the benefits of routine WGS to high TB incidence settings as well.

The development of effective techniques for culture-independent sequencing would dramatically change the landscape of TB diagnostics, and targeted next generation sequencing focusing on panels of genes or genetic regions may be an alternative solution when sequencing directly from specimens (Cabibbe et al., 2018; Cohen et al., 2019). The low abundance of target *M. tuberculosis* genomes in respiratory microbiomes from patients without advance pulmonary disease presents a major challenge for the evaluation and translation of 'deep' sequencing into diagnostic practice. Perhaps WGS greatest utility in high burden settings will be in antimicrobial resistance surveillance to inform appropriate empiric treatment regimens when resistance prevalence is low and to guide tailored therapy when resistance is more prevalent, as well as guiding broader public health measures. We are experiencing a 'genomics revolution' in microbiology with major potential benefits for TB patients and TB control programs, but we need to continue to invest in and advance these technologies and ensure that high TB incidence countries also share in the benefits as the technology becomes more mature and cost efficient.

### Conflict of interest

No conflicts of interest to declare

### Funding source

TMW is a Wellcome Trust (United Kingdom) Clinical Career Development Fellow (214560/Z/18/Z).

### Ethical approval

Not required.

### Transparency declaration

This article is part of a supplement entitled Commemorating World Tuberculosis Day March 24th, 2021: "The Clock is Ticking" published with support from an unrestricted educational grant from QIAGEN Sciences Inc.

### References

- Bainomugisa A, Lavu E, Hiashiri S, Majumdar S, Honjepari A, Moke R, et al. Multiclonal evolution of multi-drug-resistant/extensively drug resistant *Mycobacterium tuberculosis* in a high-prevalence setting of Papua New Guinea for over three decades. *Microb Genom* 2018;4:15.
- Brown AC, Bryant JM, Einer-Jensen K, Holdstock J, Houniet DT, Chan JZM, et al. Rapid whole-genome sequencing of *Mycobacterium tuberculosis* isolates directly from clinical samples. *J Clin Microbiol* 2015;53(7):2230–7.
- Bryant JM, Harris SR, Parkhill J, Dawson R, Diacon AH, van Helden P, et al. Whole-genome sequencing to establish relapse or re-infection with *Mycobacterium tuberculosis*: a retrospective observational study. *Lancet Respir Med* 2013;1(10):786–92.
- Cabibbe AM, Walker TM, Niemann S, Cirillo DM. Whole genome sequencing of *Mycobacterium tuberculosis*. *Eur Respir J* 2018;52(5):1801163.
- Cohen KA, Manson AL, Desjardins CA, Abeel T, Earl AM. Deciphering drug resistance in *Mycobacterium tuberculosis* using whole-genome sequencing: progress, promise, and challenges. *Genome Med* 2019;11(1):1–18.
- CRyPTIC Consortium and the 100,000 Genomes Project, Allix-Béguec C, Arandjelovic I, Bi L, Beckert P, Bonnet M, et al. Prediction of susceptibility to first-line tuberculosis drugs by DNA sequencing. *NEJM* 2018;379(15):1403–15.
- Doyle RM, Burgess C, Williams R, Gorton R, Booth H, Brown J, et al. Direct whole-genome sequencing of sputum accurately identifies drug-resistant *Mycobacterium tuberculosis* faster than MGIT culture sequencing. *J Clin Microbiol* 2018;56(8):931.
- 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(10):e0163612.
- Kizny Gordon A, Tong SYC, Martinez E, Crichton T, Denholm JT, Sintchenko V. TB genomic surveillance and data sharing in recognizing contamination events. *Int J Tuberc Lung Dis* 2021; in press.
- Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, et al. Whole genome sequencing of *Mycobacterium tuberculosis*: current standards and open issues. *Nat Rev Microbiol* 2019;17(9):533–45.
- Min J, Kim K, Choi H, Kang ES, Shin YM, An JY, et al. Investigation of false positive *Mycobacterium tuberculosis* culture tests using whole genome sequencing. *Ann Thorac Med* 2019;14(1):90–3.
- Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, et al. A standardized method for interpreting the association between mutations and phenotypic drug resistance in *Mycobacterium tuberculosis*. *Eur Respir J* 2017;50(6):1701354.
- Olaru ID, Patel H, Kranzer K, Perera N. Turnaround time of whole genome sequencing for mycobacterial identification and drug susceptibility testing in routine practice. *Clin Microbiol Infect* 2018;24(6):659.e5–7.
- World Health Organization. Global tuberculosis report 2020. World Health Organization; 2020 Geneva. <https://www.who.int/publications/i/item/9789240013131>. [Accessed 21st December 2020].
- Rodwell TC, Valafar F, Douglas J, Qian L, Garfein RS, Chawla A, et al. Predicting extensively drug-resistant *Mycobacterium tuberculosis* phenotypes with genetic mutations. *J Clin Microbiol* 2014;52(3):781–9.
- Shea J, Halse TA, Lapiere P, Shudt M, Kohlerschmidt D, Van Roey P, et al. Comprehensive whole-genome sequencing and reporting of drug resistance profiles on clinical cases of *Mycobacterium tuberculosis* in New York State. *J Clin Microbiol* 2017;55(6):1871–82.
- Stucki D, Ballif M, Egger M, Furrer H, Altpeter E, Battegay M, et al. Standard genotyping overestimates transmission of *Mycobacterium tuberculosis* among immigrants in a low incidence country. *J Clin Microbiol* 2016;54(7):1862–70.
- Tarashi S, Fateh A, Mirsaedi M, Siadat SD, Vaziri F. Mixed infections in tuberculosis: the missing part in a puzzle. *Tuberculosis (Edinb)* 2017;107:168–74.
- 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(5):1285–98.
- Walker TM, Ip CLC, 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(2):137–46.
- Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, et al. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect Dis* 2015;15(10):1193–202.
- Walker TM, Merker M, Knoblauch AM, Helbling P, Schoch OD, van der Werf MJ, et al. A cluster of multidrug-resistant *Mycobacterium tuberculosis* among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. *Lancet Infect Dis* 2018;18(4):431–40.
- Witney AA, Bateson ALE, Jindani A, Phillips PPJ, Coleman D, Stoker NG, et al. Use of whole genome sequencing to distinguish relapse from reinfection in a completed tuberculosis clinical trial. *BMC Med* 2017;15(1):71.
- World Health Organization. Multidrug-resistant tuberculosis (MDR-TB) 2017 update. Geneva: World Health Organization; 2017 <https://www.who.int/tb/areas-of-work/drug-resistant-tb/global-situation/en/>. [Accessed 12th January 2021].
- Wu J, Yang C, Lu L, Dai W. Detection of tuberculosis laboratory cross contamination using whole-genome sequencing. *Tuberculosis (Edinb)* 2019;115:121–5.
- Wyllie DH, Robinson E, Peto T, Crook DW, Ajileye A, Rathod P, et al. Identifying mixed *Mycobacterium tuberculosis* infection and laboratory cross contamination during mycobacterial sequencing programs. *J Clin Microbiol* 2018;56(11):1176.