Whole genome sequencing for drug resistance determination in *Mycobacterium tuberculosis*

South Africa remains challenged with a high tuberculosis burden accompanied by an increase in drug resistant cases. We assessed the use of the Illumina MiSeq, a next-generation sequencing platform for whole genome sequencing, followed by bioinformatic analysis using a commercial software package to determine resistance to selected drugs used for *Mycobacterium tuberculosis* treatment in our setting. Whole genome sequencing shows potential as a diagnostic platform for the detection of drug resistance in *Mycobacterium tuberculosis* with the provision of information for several drugs simultaneously.

### Introduction

Drug-resistant tuberculosis poses a significant challenge to tuberculosis control programmes in high burden settings. Undiagnosed drug resistance leads to further transmission, poor patient outcomes and potential for amplification of drug resistance, impeding the World Health Organization’s (WHO) strategy to end tuberculosis by 2035. The drug-resistant tuberculosis outbreaks in Tugela Ferry and other regions of South Africa highlight the need for early and accurate diagnosis of drug resistance.

Often, comprehensive phenotypic baseline testing is not available nor is a robust surveillance programme in place to inform regimen changes appropriate to local resistance profiles. A paradigm shift is needed in the approach to diagnosis and surveillance of drug-resistant tuberculosis to ensure that new drug potential is not lost due to the evolution and spread of resistant strains. Molecular testing such as the line probe assay and Xpert MTB/RIF assay (Cepheid, Sunnyvale, California, United States) show potential superiority in overall performance over phenotypic drug susceptibility testing (DST). A targeted sequencing approach for resistance detection in *Mycobacterium tuberculosis* by application of next-generation sequencing benchtop platforms showed good performance in terms of sensitivity. With the decreasing cost of next-generation sequencing, whole genome sequencing (WGS) could be applied for this purpose as an alternative to conventional phenotypic methods. The direct benefit of WGS is its ability to provide organism identification, strain relatedness and a drug resistance profile for characterised resistance-conferring mutations. In addition, WGS may be useful for resistance determination for newer drugs lacking validated DST such as bedaquiline and delamanid, utilising information available for the genetic basis associated with resistance in vitro to these novel drugs.

We assessed the use of the Illumina MiSeq® sequencing, followed by bioinformatic analysis using a commercial software (CLC Genomics Workbench, Qiagen, Venlo, The Netherlands) for drug resistance determination at the National Tuberculosis Reference Laboratory in South Africa.

### Methods

#### Ethical considerations

Ethical approval was not required for this laboratory-based study as only anonymised isolates were used.

#### Sample selection

Twenty geographically diverse clinically isolated *M. tuberculosis* strains, with varying resistance profiles and spoligotype patterns, isolated between June 2012 and January 2013 were selected for this pilot evaluation (Table 1). Laboratory processing for culture, smear microscopy and DST were performed according to WHO guidelines. Six of the 20 isolates had discordant phenotypic results between initial and repeat testing to either the fluoroquinolones or pyrazinamide.
Routine laboratory phenotypic testing

Phenotypic DST was performed on the BACTEC Mycobacterial Growth Indicator Tube (MGIT) 960 system (Becton Dickinson Diagnostic Systems, Sparks, Maryland, United States) following the manufacturer’s recommendation. First and second-line anti-mycobacterial drugs (rifampicin, isoniazid, ofloxacin, moxifloxacin, pyrazinamide, amikacin, and kanamycin) were tested following the WHO 2012 Policy Guidelines.14 Replicate testing was performed on any isolate resistant to pyrazinamide or second-line drugs on initial testing.

Next-generation sequencing

WGS was performed using the MiSeq version 2 kit (Illumina, San Diego, California, United States). In brief, DNA was extracted using the NucliSENS easyMAG system (BioMérieux, Marcy-l’Étoile, France) from a 200 µl aliquot of heat-inactivated, MGIT-cultured isolate and concentrations quantified using the Qubit dsDNA HS (high sensitivity) assay (Life Technologies, Carlsbad, California, United States). Libraries were prepared using Nextera XT kit (Illumina, San Diego, California, United States) following the manufacturers’ protocol with one modification (Figure 1). The modification deviated at the normalisation step, where the indexed DNA libraries concentrations were quantified as described above and normalised to 4 nM by addition of Tris-Cl (10 nM, pH8.5 with 0.1% Tween20). Thereafter, the indexed libraries of all 20 isolates were pooled to a final concentration of 12 pM and loaded onto the MiSeq for sequencing.

Bioinformatic analysis of WGS data

CLC Genomics Workbench version 6.0.1 (Qiagen, Venlo, the Netherlands) was used for bioinformatic analysis. Variant tables for genetic targets associated with resistance to rifampicin, isoniazid, fluoroquinolones (ofloxacin and moxifloxacin), pyrazinamide, aminoglycosides (amikacin and kanamycin), bedaquiline and delamanid (Table 2) were generated using the Map Reads to Reference tool and Quality-based Variant Detection algorithm on CLC Genomics Workbench using the H37Rv Sanger reference genome (GenBank NC000962.3). The following cut-offs were applied to call a single nucleotide polymorphism or insertion/deletion: a minimum paired coverage depth of five times (5×), frequency of > 70% and a Phil’s Read Editor, or PHRED, quality score of ≥ Q20 (≥ 99% accuracy) at the variant position and neighbouring nucleotides within a radius of five base pairs. To ensure that an isolate was truly wild-type for a specific gene target, we further ran the Create Statistics for Target Regions on CLC Genomics Workbench to ensure that the entire length of the gene investigated was completely sequenced. Since no thresholds have been formally established for bioinformatic analysis, we utilised less stringent parameters than those previously described.15

Association of mutations as resistance predictors were identified using the TB Drug Resistance Mutation Database (TBDBReaMDB) database16 primarily. If a mutation was not listed, literature, including newer published databases such as TBProfiler and PhyResSE, was surveyed to identify the association.17,18 Putative mutations associated with the novel drugs bedaquiline and delamanid were exclusively identified using published literature.19,20 The rpoB-associated mutations were converted to the widely used Escherichia coli nomenclature (addition of 81 codon positions).20

Resolving discordant phenotypic and WGS results

Discordant results were resolved using the minimum inhibitory concentration broth microdilution method (TREK Sensititre, Thermofisher, Waltham, Massachusetts, United States) and interpreted using the critical concentrations established by Hall et al. (2012).21 In the case of pyrazinamide, the modified Wayne’s test22 was used to resolve discordance. Additionally, the GenoType MTBDRplus assay version 2 (MTBDRplus) (Hain LifeSciences, Nehren, Germany) line probe assay was performed according to the manufacturer’s instruction for the first-line drugs rifampicin and isoniazid on all isolates. Figure 1 provides an overview of the operational workflow for this study.

Results

Concordance between WGS and the phenotypic DST method for resistance determination was noted for all isolates except one phenotypically susceptible isolate for all targets explored.
## TABLE 1: Summary of performance for drug resistance determination using the MGIT960, MTBDRplus assay and whole genome sequencing.

| Isolate | Phenotype | Concordant Susceptibility | Multidrug resistant | Extensively drug resistant | Per extensively drug resistant | Discordant pyrazinamide isolates |
|---------|-----------|---------------------------|---------------------|---------------------------|-----------------------------|--------------------------------|
| 1       | S         | S                          | S                   | S                         | S                           |                                |
| 2       | S         | S                          | S                   | S                         | S                           |                                |
| 3       | S         | S                          | S                   | S                         | S                           |                                |

**Mutations**

- **mut1**: c-15t (RPOB promoter)
- **mut2a**: H526Y (INHA)
- **mut2**: S315T (GyrA)
- **mut3a**: t-8c (INHA)
- **mut3**: S531L (MutA)

**Other genotypic findings**

- **Rifampicin**
  - **Mut1**: C-15T (RPOB promoter)
  - **Mut2a**: H526Y (INHA)
  - **Mut2**: S315T (GyrA)
  - **Mut3a**: t-8c (INHA)
  - **Mut3**: S531L (MutA)
- **Ofloxacin/Moxifloxacin**
  - **Mut1**: c-15t (RPOB promoter)
  - **Mut2a**: H526Y (INHA)
  - **Mut2**: S315T (GyrA)
  - **Mut3a**: t-8c (INHA)
  - **Mut3**: S531L (MutA)
  - **Mut4**: S531L (MutA)
- **Isoniazid**
  - **Mut1**: c-15t (RPOB promoter)
  - **Mut2a**: H526Y (INHA)
  - **Mut2**: S315T (GyrA)
  - **Mut3a**: t-8c (INHA)
  - **Mut3**: S531L (MutA)
  - **Mut4**: S531L (MutA)
- **Rifampicin resistance**
  - **Mut1**: C-15T (RPOB promoter)
  - **Mut2a**: H526Y (INHA)
  - **Mut2**: S315T (GyrA)
  - **Mut3a**: t-8c (INHA)
  - **Mut3**: S531L (MutA)
  - **Mut4**: S531L (MutA)
- **Ofloxacin resistance**
  - **Mut1**: c-15t (RPOB promoter)
  - **Mut2a**: H526Y (INHA)
  - **Mut2**: S315T (GyrA)
  - **Mut3a**: t-8c (INHA)
  - **Mut3**: S531L (MutA)
  - **Mut4**: S531L (MutA)

**Phenotypic discordant isolates**

- **Isolate 1**: Phenotypically susceptible to rifampicin, ofloxacin, and pyrazinamide, resistant to isoniazid and amikacin.
- **Isolate 2**: Phenotypically susceptible to rifampicin, ofloxacin, and pyrazinamide, resistant to isoniazid and amikacin.
- **Isolate 3**: Phenotypically susceptible to rifampicin, ofloxacin, and pyrazinamide, resistant to isoniazid and amikacin.

**Whole genome sequencing**

- **MTBDRplus assay**
  - Concordant high-confidence resistance
  - Discordant in only 1 test
  - Incorrectly classified as susceptible by MTBDRplus assay
- **Whole genome sequencing**
  - Concordant high-confidence resistance
  - Discordant in only 1 test
  - Incorrectly classified as susceptible by MTBDRplus assay

**Discussion**

The application of whole genome next-generation sequencing technology for drug resistance determination in *M. tuberculosis* has been shown to be a valuable tool in this study. Despite the small sample size, the performance of WGS for predicting resistance was consistent with published studies containing subsets of South African isolates.

The use of the MiSeq® offers reduced hands-on preparation time (~6 h per batch of isolates) compared to other next-generation sequencing technologies. Bioinformatic analysis...
using the commercial software was relatively straightforward, particularly the use of workflows for automation. Once a workflow is saved, imported data are automatically analysed, producing a final output table displaying mutations. However, sequence analysis requires an understanding of the associated genetic targets and drug resistance mutations.

Concordance of WGS with initial MGIT DST was lacking for isoniazid, pyrazinamide and fluoroquinolones; however, resolution testing improved agreement between WGS and phenotypic drug susceptibility profiles for the discordant isolates, even at a coverage of 5× (paired) with acceptable quality scores.

The use of this technology for resistance determination in M. tuberculosis is currently limited due to the lack of a comprehensive tuberculosis mutation catalogue predicting susceptibility, as seen in the case of the pncAThr-15Thr mutation in this study, which was not associated with resistance. Furthermore, sequence data can only be generated from cultured isolates, creating a lag between specimen receipt and a positive culture. Despite these limitations, WGS could benefit the majority of patients by enabling them to be placed on optimum regimens sooner in comparison to phenotypic methods.

**Limitations**

The small sample size was inadequate for assessing diagnostic performance statistically. In addition, the data are based on a first attempt without any optimisation for sequence output.

**Conclusion**

WGS correctly predicted resistance or susceptibility using commercial bioinformatics software based on already identified resistance-determining mutations. Our findings suggest the system shows promise as a tool for predicting drug resistance in a short time frame for multiple drugs and multiple samples simultaneously, provided the genetic basis for resistance is well described.

**Acknowledgements**

Detailed WGS results and statistics are available from the corresponding author.

We would like to thank Professor Jasper Rees and Jonathan Feathers (Agricultural Research Council, Biotechnology Platform, Pretoria, South Africa) for their valuable input and assistance. We further would like to thank the technical staff at the Centre for Tuberculosis.

**Competing interests**

The authors declare that they have no financial or personal relationships which may have inappropriately influenced them in writing this article.

---

**Sources of support**

Funding was provided by the National Institute for Communicable Diseases, a division of the National Health Laboratory Service, South Africa.

**Authors’ contributions**

N.A.I. and S.V.O. were the project leads. S.V.O., N.A.I., L.J., H.M.S. and F.I. were responsible for the experimental and project designs. S.V.O., L.J., H.M.S., N.I. and T.L.G. performed most of the experiments. Calculations were performed by S.V.O. and N.A.I.

---

**References**

1. WHO. WHO, editor. Global tuberculosis report 2017 [homepage on the Internet]. Geneva, Switzerland: WHO Press; 2017. Available from: apps.who.int/iris/bitstream/10665/259366/1/9789241565516-eng.pdf?ua=1
2. Gandhi NR, Nunn P, Dheda K, et al. Multidrug-resistant and extensively drug-resistant tuberculosis: A threat to global control of tuberculosis. Lancet. 2010;375(9728):1830–1843. https://doi.org/10.1016/S0140-6736(10)60410-2
3. Klopper M, Warren RM, Hayes C, et al. Emergence and spread of extensively and totally drug-resistant tuberculosis, South Africa. Emerg Infect Dis. 2013;19(3):449–455. https://doi.org/10.3201/eid1903.120046
4. Muller B, Chihota VN, Pillay M, et al. Programatically selected multidrug-resistant strains drive the emergence of extensively drug-resistant tuberculosis in South Africa. PLoS One. 2013;8(8):e70919. https://doi.org/10.1371/journal.pone.0070919
5. Gurbanofo E, Mehdiyev R, Blondal K, et al. Mitigation of discordant Rifampicin-susceptibility results obtained by Xpert Mycobacterium tuberculosis/Rifampicin and Mycobacterium growth indicator tube. Microb Resist. 2017;23(8):1045–1052.
6. Van Deun A, Aung KJ, Hossain A, et al. Disputed rpoB mutations can frequently cause important rifampicin resistance among new tuberculosis patients. Int J Tuberc Lung Dis. 2015;19(2):185–190. https://doi.org/10.5588/ijtld.14.0449
7. Daum LT, Rodriguez JD, Worthy SA, et al. Next-generation ion torrent sequencing of drug resistance mutations in Mycobacterium tuberculosis strains. J Clin Microbiol. 2012;50(12):3831–3837. https://doi.org/10.1128/JCM.01893-12
8. Köser CU, Bryant JM, Becq J, et al. Whole-genome sequencing for rapid susceptibility testing of M. tuberculosis. N Engl J Med. 2013;369(3):290–292. https://doi.org/10.1056/NEJMMc1215305
9. Walker TM, Merker M, Kohl TA, Crook DW, Niemann S, Peto TE. Whole genome sequencing for MTDR tuberculosis surveillance and for resistance testing. Clin Microbiol Infect. 2017;23(3):161–166. https://doi.org/10.1016/j.clinmicinfe.2016.10.014
10. Drobniewski F, Nikolaevsky V, Maxeiner H, et al. Rapid diagnostics of tuberculosis and drug resistance in the industrialized world: Clinical and public health benefits and barriers to implementation. BMC Med. 2013;11:190. https://doi.org/10.1186/1741-7015-11-190
11. Hartkorn RC, Uplekar S, Cole ST. Cross-resistance between clofazimine and bedaquiline through upregulation of MmpL5 in Mycobacterium tuberculosis. Antimicrob Agents Chemother. 2014;58(5):2979–2981. https://doi.org/10.1128/AAC.0037-14
12. Lee RS, Belf MA. The implications of whole-genome sequencing in the control of tuberculosis. Ther Adv Infect Dis. 2016;3(2):147–62. https://doi.org/10.1177/2049936115624630
13. Initiative GL. Global laboratory initiative, editor. Mycobacteriology laboratory manual [homepage on the Internet]. 1st ed. Geneva, Switzerland: WHO Press; 2014 [cited 2015 May 25]. Available from: http://www.stoptb.org/en/gil/assets/documents/gil_mycobacteriology_lab_manual_web.pdf
14. WHO. Updated interim critical concentrations for first-line and second-line DST [as of May 2012] [homepage on the Internet]. 2012 [cited 2012 May]. Available from: http://apps.who.int/iris/bitstream/10665/259366/1/9789241565516-eng.pdf?ua=1
15. Farhat MR, Shapiro BJ, Kieser KJ, et al. Genomic analysis identifies targets of clinically relevant drug resistance in Mycobacterium tuberculosis. Nat Genet. 2015;47(10):1183–1189. https://doi.org/10.1038/ng.1177
16. Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Whole genome sequencing data. J Clin Microbiol. 2015;53(6):1908–1914. https://doi.org/10.1128/JCM.00374-14
17. Feurriegel S, Chlebusquer V, Beckert P, et al. PhyTrees: A web tool delineating Mycobacterium tuberculosis antibiotic resistance and lineage from whole-genome sequencing data. J Clin Microbiol. 2015;53(6):1908–1914. https://doi.org/10.1128/JCM.00374-14
18. 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://doi.org/10.1186/s13015-015-0164-0
19. Feurriegel S, Koser CU, Bau D, et al. Impact of Fgd1 and ddn diversity in Mycobacterium tuberculosis complex on in vitro susceptibility to PA-824. Antimicrob Agents Chemother. 2011;55(12):5718–5722. https://doi.org/10.1128/AAC.05500-11
20. Andre E, Goeminne L, Cabibbe A, et al. Consensus numbering system for the rifampicin resistance-associated rpoB gene mutations in pathogenic mycobacteria. Clin Microbiol Infect. 2017;23(3):167–172. https://doi.org/10.1016/j.cmi.2016.09.006

21. Hall L, Jude KP, Clark SL, et al. Evaluation of the sensititre MycoTB plate for susceptibility testing of the Mycobacterium tuberculosis complex against first- and second-line agents. J Clin Microbiol. 2012;50(11):3732–3734. https://doi.org/10.1128/JCM.02048-12

22. Singh P, Wesley C, Jadaun GP, et al. Comparative evaluation of Lowenstein-Jensen proportion method, BacT/ALERT 3D system, and enzymatic pyrazinamidase assay for pyrazinamide susceptibility testing of Mycobacterium tuberculosis. J Clin Microbiol. 2007;45(1):76–80. https://doi.org/10.1128/JCM.00951-06

23. 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(10):1193–1202. https://doi.org/10.1016/S1473-3099(15)00062-6

24. Louw GE, Warren RM, Donald PR, et al. Frequency and implications of pyrazinamide resistance in managing previously treated tuberculosis patients. Int J Tuberc Lung Dis. 2006;10(7):692–697.

25. Huitric E, Verhasselt P, Koul A, Andries K, Hoffner S, Andersson DI. Rates and mechanisms of resistance development in Mycobacterium tuberculosis to a novel diarylquinoline ATP-synthase inhibitor. Antimicrob Agents Chemother. 2010;54(3):1022–1028. https://doi.org/10.1128/AAC.01611-09

26. Zignol M, Cabibbe AM, Dean AS, et al. Genetic sequencing for surveillance of drug resistance in tuberculosis in highly endemic countries: A multi-country population-based surveillance study. Lancet Infect Dis. 2018;18(6):675–683. https://doi.org/10.1016/S1473-3099(17)30073-2

27. Loman NJ, Constantinidou C, Chan JZ, et al. High-throughput bacterial genome sequencing: An embarrassment of choice, a world of opportunity. Nat Rev Microbiol. 2012;10(9):599–606. https://doi.org/10.1038/nrmicro2850