

Use of whole genome sequencing for detection of antimicrobial resistance:

***Mycobacterium tuberculosis*, a model organism.**

Vincent Escuyer, Ph.D., Mycobacteriology Laboratory,

Wadsworth Center- New York State Department of Health, Albany, NY.

ABBREVIATIONS

NGS - Next generation sequencing, WGS - Whole genome sequencing, AMR - Anti-microbial resistance, AST- Anti-microbial susceptibility testing, TB- Mycobacterium tuberculosis, DST- Drug susceptibility testing

INDEX TERMS

Whole genome sequencing, Next-generation sequencing, Anti-microbial resistance, *Mycobacterium tuberculosis*

LEARNING OBJECTIVES

1. Explain the current challenges with antimicrobial resistance and limitations of current testing methodologies
2. Compare the differences between metagenomic next generation sequencing and isolate whole genome sequencing for anti-microbial resistance gene detection
3. Describe why Mycobacterium tuberculosis is an ideal candidate organism for whole genome sequencing
4. Discuss outcomes of Mycobacterium tuberculosis whole genome sequencing studies and their limitations

Abstract

Increasing rates of antimicrobial resistance is a public health crisis. The emergence of resistant pathogens is multi-factorial but is at least partially due to inappropriate antibiotic utilization and lack of stewardship interventions. Effective stewardship programs require timely antimicrobial resistance

testing. This can be challenging for pathogens that grow slowly or not at all in culture. Next generation sequencing (NGS) approaches, such as whole genome sequencing (WGS) of isolate, offer a more rapid alternative for such pathogens. *Mycobacterium tuberculosis* (TB) is a model organism for WGS to predict susceptibility due to its highly conserved and stable genome, extremely slow growth in culture, and increasing resistance rates to a limited armamentarium of anti-TB drugs. Studies have shown excellent concordance between conventional phenotypic susceptibility testing and use of WGS to predict susceptibility to at least 2 first-line anti-TB agents, rifampin and isoniazid. More data is needed for other agents, including a more comprehensive curated database of mutations paired with phenotypic data, before WGS can completely replace phenotypic testing.

Introduction

Antimicrobial resistance (AMR) is one of the major global public health challenges of this century. The emergence and spread of resistant microorganisms threaten most therapeutic and preventive options to manage bacterial infections, as commonly used antibiotics are no longer effective (1). This has a profound impact on patients, families, and, more broadly, on health systems, both clinically and financially (2). Causes for the rise of antibiotic resistance include overuse in human medicine, prescribing inappropriate drugs, and extensive use in agriculture, particularly with livestock. Furthermore, the development of new antibiotics has not been prioritized by the pharmaceutical industry, mostly due to lack of profit. Therefore, very few new candidate drugs are currently in clinical trials (3). This grim situation prompted public health organizations to step in to raise awareness about the need for action to contain this global crisis. In 2013, Centers for Disease Control and Prevention (CDC) published an assessment of antibacterial threats that classifies each bacterial species in three distinct categories: “urgent”,

“serious” or “concerning” (4). In 2014, the World Health Organization (WHO) released a global survey on AMR surveillance from multiple national and international networks that could lay the foundation for a comprehensive plan of action (5)

Managing bacterial infections is a complex process and one of the core elements is the availability of accurate methods to determine susceptibility of microorganisms to antimicrobials. However, most antimicrobial susceptibility testing (AST) is still based on culturing bacteria in the presence or absence of the antimicrobial and turnaround time for results can be significant, particularly for slow or poor growing organisms (6). For this reason, alternative molecular methods have been, or are being, developed for accelerated identification and detection of resistance, either on bacterial cultures or directly on clinical specimens (7). A number of amplification-based resistance gene detection assays are currently commercially available for many of the significant bacterial pathogens from positive blood cultures (8, 9). However, most of these assays focus on the detection of very specific genetic elements or chromosomal targets and rely on correlative prediction rather than determining a true phenotype. Consequently, these rapid methods may not always produce accurate results.

The recent advent of novel sequencing technologies has circumvented this shortcoming. The emergence of several scalable and fast next generation sequencing (NGS) platforms makes clinical diagnostic applications realistic as they can be utilized to interrogate a much higher number of targets at once, as compared to current commercially available methods (10). NGS can be performed on either DNA or RNA, allowing one to look for the presence of specific genes involved in drug resistance (DNA) or expression of these genes (RNA). Most of the current efforts focus on utilizing NGS directly on clinical specimens either with 1) multiplex PCR based amplicon sequencing with primers specifically designed to amplify a determined set of resistance genes (11, 12) or 2) PCR independent shotgun metagenomics, where all genomic material

present in the specimen is sequenced, followed with bioinformatics analysis to identify the resistance genes of interest (13, 14). The main restriction of the former approach is the limited number of targets that are interrogated, potentially resulting in relevant genes not being included in the assay's design. The latter approach also has several limitations. Because the respective amount of bacterial species in different specimens can greatly vary, less represented species and their resistance determinants are more likely to be undetected due to a very low amount of starting genetic material (15). Furthermore, current bioinformatics tools for metagenomics still have problems assigning resistance associated mobile genetic elements such as plasmids, bacteriophages, or transposons to their bacterial host genomes present in the specimen (15, 16, 17). Lastly, AMR prediction in metagenomes often generates false negative results mainly due to very stringent filters, including databases mostly composed only of known, well characterized and clinically important resistance genes with high thresholds for sequence homology hits (15, 18).

The use of NGS for whole genome sequencing (WGS) offers an attractive alternative as it allows for interrogation of the entire genome of an organism for the presence of any resistance determinants, including cases where more than one gene is involved, therefore increasing sensitivity and specificity. Prediction of AMR with WGS has already been extensively investigated mainly on clinically relevant bacteria such as *Escherichia coli*, *Staphylococcus aureus*, *Salmonella* species, *Streptococcus pneumoniae*, *Enterococcus* species, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, and *Neisseria gonorrhoeae* (19). A number of large, multi-species databases of resistance determinant sequences, with accompanying bioinformatics tools, are publicly available and include the Comprehensive Antibiotic Resistance Database (CARD) (20), ResFinder (21) or ARG-ANNOT (22). There are multiple applications

of WGS for AMR, from determining the best course of treatment for infected patients to global surveillance in healthcare, community, or animal health settings to detect foodborne dissemination of AMR (23).

Most of the work published on WGS AMR thus far relates to bacterial pathogens.

However, recent studies demonstrate that this approach can also be applied to clinical virology, particularly with human immunodeficiency virus (HIV) and hepatitis C virus (HCV) (24).

Stanford University created a HIV drug resistance database accessible to the general public (<https://hivdb.stanford.edu>), while University of Glasgow in the UK offers a similar service for HCV (<http://hcv.glue.cvr.ac.uk/#/home>). Yet, the use of WGS for clinical virology is still in development and more work needs to be accomplished before applying it routinely in the laboratory. The remainder of this review will focus on the application to bacterial WGS for AMR, with a particular focus on *Mycobacterium tuberculosis* (TB).

Next generation sequencing technologies and relevance to AMR

Until the early 2000's, Sanger, or chain termination, was the main sequencing method and provided high quality sequences, although its suitability for high throughput was not realistic for clinical application of WGS. The landscape changed dramatically with the emergence of the second and third generations of sequencing technologies, also referred to as “next generation sequencing” (25). Second-generation sequencing is dominated by Illumina, with different instruments offering various levels of throughput. Sequencing run generates an enormous amount of nucleotide sequences, with each genome being simultaneously sequenced multiple times in small fragments (short reads) with a low per-base error rate (generally < 0.1%) (25). Robust bioinformatics are required for post-sequencing analysis.

These short-read sequencing platforms are well suited for different applications, including

amplicon based targeted sequencing for detection of antimicrobial resistance determinants. However, they are not optimal for determining the sequence of closed (i.e. finished) bacterial genomes, as they generate fragmented genome assemblies, also referred to as contiguous sequences or contigs, and struggle with extrachromosomal elements such as plasmids. Sequencing of a completely closed bacterial chromosome requires the use of long read sequencing platforms, also referred to as third-generation sequencing (26). Until recently, the market was monopolized by Pacific Biosciences, with its PacBio RSII platform, that comes at a high price for both instrumentation and cost per sample. In addition, this system is low throughput and is more suited for core-type facilities. Recently, Oxford Nanopore Technologies developed a long-read platform, the MinION (27). The device is slightly larger than a thumb drive and can be plugged into a standard USB port on any computer. Furthermore, sequence analysis can be performed in real-time and results become available as soon as sequencing reads are generated during the run (28). This portability and real time capacity have made the MinION a particularly attractive option, especially for clinical diagnostic in the field and in lower income settings. One of the main issues with using long read platforms for AMR is a significantly higher per-base error rate compared to short read technology (5-10% vs 0.1%). This rate might be too high for accurate detection of specific single nucleotide polymorphisms (SNPs) associated with AMR but can partially be overcome by achieving higher read depth and improving base calling algorithms (27, 28, 29). Currently, the short-read Illumina systems are the predominant platforms found in clinical laboratories.

Drug resistance in *Mycobacterium tuberculosis*: Challenges

Tuberculosis is one of the leading causes of mortality around the world. In 2017, an estimated 10 million people were newly infected with TB (30). Among them, 1.6 million died from the disease.

Although the global incidence of tuberculosis is showing some decline, the emergence of drug resistance tuberculosis (DR-TB) represents a major public health crisis. Standard treatment for TB combines 4 first-line drugs: rifampin, isoniazid, pyrazinamide and ethambutol for a period of 6 months (31). Resistance to rifampin and isoniazid, the two most efficient first-line drugs results in classification of TB as multidrug-resistant (MDR-TB). Treatment of MDR-TB generally involves a combination of second-line agents: Fluoroquinolones and injectable medications, such as amikacin or capreomycin, in addition to other drugs chosen according to the resistance profile of targeted TB strain. Treatment can last up to 30 months. Strains that test resistant to fluoroquinolones and at least one injectable drug are defined as extensively drug-resistant TB (XDR-TB). Treatment of XDR cases is extremely complex, highly individualized, and often involves the use of the newest TB drugs such as bedaquiline or delamanid or repurposed drugs like clofazimine or linezolid (32).

In 2017, WHO estimated that ~ 460,000 new TB cases were MDR- TB, of which ~50,000 were XDR-TB (30). In the US, CDC classified DR-TB as a serious threat due to the complications and lower cure rates associated with long-term treatments, as well as the lack of new substitute drugs to combat drug resistant strains (4). The design and administration of an optimal drug regimen heavily relies on accurate drug susceptibility testing (DST) results.

Standard culture-based DST is impaired by the fastidious nature of TB. The very slow growth rate of the organism can significantly delay the availability of phenotypic DST results; results can take weeks to month(s) to obtain (33). Consequently, WHO proposed expanding rapid diagnosis and detection of DR-TB cases as one of the five high priority actions to address this crisis (34). WHO endorsed two rapid molecular tests for detection of mutations conferring resistance to rifampin and isoniazid directly from sputum specimens: Cepheid GeneXpert MTB/RIF and Hain line probe assays (35, 36). The target genes are *rpoB* (mutations conferring rifampin resistance), *katG* and

inhA (mutations conferring isoniazid resistance). Implementation of these tests in high-burden countries has partially answered the needs. However, the low number of resistance mutations targeted in these assays is a limitation and a negative result does not necessarily indicate susceptibility to the drug of interest. Therefore, broader genetic approaches to ensure accurate diagnosis and optimal treatment for patients are needed. The ability for WGS to interrogate the entire genome, rather than selected targets, and derive comprehensive resistance profiles makes it an attractive alternative. It can provide physicians with important information on almost all of the currently available anti-TB drugs and guide them to optimized treatment in a timely manner, a crucial necessity with MDT/XDR-TB infections.

***Mycobacterium tuberculosis*: An ideal candidate for WGS DST**

One of the main features of *Mycobacterium tuberculosis* is the extreme stability of its genome, with a very low mutation rate. There is no evidence of genetic material exchange between different strains nor horizontal transfer from other bacterial species. Consequently, transmissible mobile genetic elements such as plasmids or transposons play no role in TB drug resistance. Indeed, resistance to drugs is mostly due to the presence of mutations (single nucleotide variations, insertions or deletions) in specific genes that either code for the drug target itself or are involved in the activation of the drug, as is the case for pyrazinamide and isoniazid (37). Currently, most of the genes involved in drug resistance have been identified and multiple mutations have been reported in the literature. The main genes associated with resistance are described in Figure 1. As TB treatment always involves a multi-drug regimen for a long period of time, it is preferable to be able to analyze, simultaneously, multiple genes across the entire genome to generate the most comprehensive susceptibility profile, a task that no other targeted molecular approaches can achieve. Further, phenotypic susceptibility testing for TB can take several weeks. For these

reasons, utilization of WGS as a clinical tool for TB DST has been intensively explored the past few years and has shown great potential to improve TB clinical diagnostics by providing accurate drug susceptibility information, timely (38, 39, 40).

TB WGS workflow

In most published studies, cultured isolates are used as the starting material. However, recent publications have explored the possibility of performing WGS directly on clinical samples to improve turn-around time. Workflow for TB WGS is summarized in Figure 2 and always includes similar steps independent of the sequencing platform utilized. DNA extraction is a crucial step as the quality and quantity of the extracted DNA will determine the quality of the sequencing data (41). Several lysis methods can be used and include chemical, mechanical, thermal, or any combination of the above. Subsequently, a DNA library is prepared in which genomic DNA is fragmented and specialized adapters attached at both ends of the DNA fragments. Numerous kits to prepare sequencing libraries are available commercially from different vendors. Adapters are specific for the sequencing platform utilized and adapter ligated fragments are then amplified by polymerase chain reaction (PCR), purified and loaded into the sequencer (42). At the end of the sequencing run, an enormous amount of sequence data is generated that requires significant computational resources for analysis and storage (43). This requirement is often a major barrier for clinical laboratories interested in adding this technology in their workflow. As with the other reviews in this Focus series, there are currently no FDA approved approaches to WGS for AMR in TB or any other pathogen. Thus, there is a large burden on the adopting laboratory to undertake an expensive, complex, and time-consuming validation.

Bioinformatics tools

Prediction of resistance by WGS requires large databases of mutations associated with drug resistance and software able to interrogate these databases for the presence of these mutations to generate a drug resistance profile. A number of online tools publicly available have been developed during the past decade. The first extensive database was the Tuberculosis Drug Resistance Mutation Database (TBDRaMDB) (44). However, this database has not been updated since April 2010, but has provided the foundation for the construction of newer databases. Since 2017, The Rapid Drug Susceptibility Testing Consortium, part of the Critical Path to TB Drug Regimens (CPTR) Initiative, opened their Relational Sequencing TB Data Platform (ReSeqTB) to public (45). This tool has a user-friendly interface to access whole-genome sequencing data collected from multiple private and public databases combined with culture-based drug susceptibility information and clinical outcome data when available (46). Currently the main web-based tools specifically designed for TB sequence analysis include PhyResSE, an integral component of ReSeqTB platform, genTB and TBProfiler (47, 48, 49). In addition, two software solutions are available as well, Mykrobe Predictor TB and MTBSeq (50, 51). Criteria for selecting a particular bioinformatics platform can vary among clinical laboratories depending on their respective regulatory requirements for clinical validation.

Performances of TB WGS

Multiple studies using WGS to determine resistance profiles with TB have been published, with thousands of TB strains sequenced to date. Data from these studies showed excellent concordance with culture-based phenotypic drug susceptibility results (52, 53, 54, 55). A large study published by the CryPTIC Consortium on more than 10,000 isolates demonstrate that this is particularly true for the two main first-line drugs, rifampin and isoniazid, for which an overall specificity of 0.94

and a sensitivity of 0.94 and 0.93 respectively were found (56). More variations were observed with the other first line drugs ethambutol and pyrazinamide with specificities of 0.84 and 0.92 and sensitivities of 0.86 and 0.76 respectively. An incomplete understanding of the mechanisms underlying resistance to these drugs could partly explain this performance (37, 57). Performance data for drugs other than the first-line drugs, which are only used in case of demonstrated resistance to the standard drugs, are sparse, although WGS performs well with fluoroquinolones (58). This lack of information could be explained in part by the paucity of phenotypic data for these drugs, as their testing is mostly performed in specialized laboratories and is not broadly available.

Most of the studies published thus far were retrospective, so more prospective work is needed to fully evaluate the performance of WGS in a clinical setting. A few prospective studies have been reported and showed promising results (59, 60 61, 62). For example, in New York, Shea *et al*, tested prospectively 405 clinical isolates over one year and found susceptibility and sensitivity values similar to those reported in the retrospective studies, strongly suggesting the possibility to implement WGS as a replacement for conventional DST. Prospective data from these studies indicated a particularly high concordance between WGS prediction and susceptibility to TB drugs. Consequently, both Public Health England and New York State Department of Health recently decided to move forward and use WGS as the primary method to replace phenotypic susceptibility testing for all clinical isolates predicted to be fully susceptible by WGS. However, a similar change for isolates predicted to be resistant by WGS is not yet ready for implementation. It has been shown that some mutations detected in target genes do not confer phenotypic resistance, which could cause an isolate to be improperly reported as resistant to a particular drug.

Consequently, a physician might alter treatment by eliminating an important drug that

would have still been appropriate per conventional testing. This emphasizes the need for a better understanding of resistance mechanisms in TB as well as carefully curated mutation databases that are not yet available.

Remaining Challenges

A number of key issues, both technical and logistical, remain before being able to generalize the implementation of WGS in routine clinical testing. This starts with the necessity of standardizing procedures for sample preparation, including methodology for DNA extraction, quantification, and quality control. One of the current limitations of WGS is still the need for culture before sequencing. Consequently, turn-around time for susceptibility results is impacted by this requirement due to the fastidious nature of the organism that can take days to weeks to grow in culture. Recent studies have started to address this issue and explored the feasibility of using WGS directly on clinical specimen to predict drug resistance thereby bypassing culture (63, 64, 65, 66). However, these approaches require users to enrich TB DNA from the specimen to a level of concentration and purity suitable for WGS and involve often-costly systems and labor-intensive protocols that are impractical for routine clinical testing.

Although WGS has been shown to perform very well to predict resistance to first line drugs and fluoroquinolones, there is still a lack of data for the second line drugs and all new or repurposed drugs used to treat MDR/XDR TB such as linezolid, clofazimine or bedaquiline. Better understanding of the mechanisms of resistance to these drugs is still a work in progress and WGS performance will improve as new mutations are identified (67, 68). However, as bioinformatics pipelines progress by integrating new information, clinical laboratories will be required to revalidate these pipelines as they are updated, a process that could be cumbersome.

Conclusion

There is a crucial necessity to advance the methods of identification of AMR and DST for TB and other microbial pathogens in the clinical setting. WGS represents an attractive alternative to other conventional and molecular methods. It has the potential to transform TB DST in the clinical laboratory, thus having a profound impact on patient management and public health intervention by providing comprehensive results significantly faster than culture-based methods. Significant progress has already been made on this front, and many tools are now available to the clinical laboratories and physicians. The NGS industry is starting to evolve toward developing ease of use and lower cost technology, but affordability and complexity remain as obstacles for implementation in low- to middle-income countries. In this regard, the MinION sequencing platform could address some of these issues by offering an affordable and portable “easy to use” device. Similar ease-of-use advances would be needed for specimen and library preparation, and bioinformatics for data analysis before developing countries could realistically adopt such testing. Nevertheless, as knowledge and technology are making strides in reaching these goals, it is reasonable to anticipate that within the next decade, industrialized countries will replace conventional TB DST with WGS. Change will be more gradual in developing countries where WGS will mostly be used more sporadically as a support methodology for drug resistance surveillance and development of new rapid diagnostics.

References

1. O’Neill J. Tackling drug-resistant infections globally: Final report and recommendations. http://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf. Accessed May 19, 2019.

2. Jirka T, Hafner M, Yerushalmi E, Smith R *et al.* Estimating the economic costs of antimicrobial resistance: Model and Results. The Wellcome Trust, RR-911-WT, 2014.
https://www.rand.org/pubs/research_reports/RR911.html. Accessed May 28, 2019
3. Antibacterial agents in clinical development: An analysis of the antibacterial clinical development pipeline, including tuberculosis. Geneva: World Health Organization; 2017 (WHO/EMP/IAU/2017.12). License: CC BY-NC-SA 3.0 IGO.
https://www.who.int/medicines/areas/rational_use/antibacterial_agents_clinical_development/en/ Accessed February 25, 2019
4. Antibiotic resistance threats in the United States. 2013. Centers for Disease Control and Prevention. <https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf> . Accessed February 20, 2019
5. World Health Organization. Antimicrobial resistance: Global report on surveillance.
http://apps.who.int/iris/bitstream/handle/10665/112642/9789241564748_eng.pdf;jsessionid=1A1AE822893A3D0DA2EEFFC09F743B82?sequence=1. Accessed January 14, 2016.
6. Jorgensen, J. H. & Ferraro, M. J. Antimicrobial susceptibility testing: a review of general principles and contemporary practices. *Clin. Infect. Dis.* 2009 **49**: 1749–1755
7. Anjum MF, Zankhari E., Hasman E. Molecular methods for detection of antimicrobial resistance 5(6): doi:10.1128/microbiolspec.ARBA-0011-2017
8. Poole S, Kidd SP, Saeed K. A review of novel technologies and techniques associated with identification of bloodstream infection etiologies and rapid antimicrobial genotypic and quantitative phenotypic determination. *Expert Rev. Mol. Diagn.* 2018 18(6): 543-555
9. Dunne Jr. WM, Jaillard M, Rochas O, Van Belkum A. Microbial genomics and antimicrobial susceptibility testing. *Expert Rev. Mol. Diagn.* 2017 17(3): 257-269

10. Vincent AT, Derome N, Boyle B, Culley AI, Charette SJ. Next-generation sequencing (NGS) in the microbiological world: How to make the most of your money. *J. Microbiol. Methods*. 2017 138: 60-71
11. Stefan CP, Koehler JW, Minogue TD. Targeted next-generation sequencing for the detection of ciprofloxacin resistance markers using molecular inversion probes. *Sci Rep*. 2016 6:25904
12. Guérillot R, Li L, Baines S, Howden B *et al*. Comprehensive antibiotic-linked mutation assessment by resistance mutation sequencing (RM-seq). *Genome Med*. 2018 Aug 31;10(1):63
13. Munk P, Andersen VD, de Knecht L, Jensen MS, Knudsen BE *et al*. A sampling and metagenomic sequencing-based methodology for monitoring antimicrobial resistance in swine herds. *J. Antimicrob. Chemother*. 2017 72(2):385-392.
14. Bridier A. Exploring foodborne pathogen ecology and antimicrobial resistance in the light of shotgun metagenomics. *Methods Mol. Biol*. 2019 1918:229-245
15. Martínez JL, Coque TM, Lanza VF, de la Cruz F, Baquero F. Genomic and metagenomic technologies to explore the antibiotic resistance mobilome. *Ann. NY Acad. Sci*. 2017 1388(1): 26-41
16. Beaulaurier J, Zhu S, Deikus G, Mogno I *et al*. Metagenomic binning and association of plasmids with bacterial host genomes using DNA methylation. *Nat. Biotechnol*. 2018 Jan;36(1):61-69
17. Lanza VF, Baquero F, Martínez JL, Ramos-Ruíz R *et al*. In-depth resistome analysis by targeted metagenomics. *Microbiome*. 2018 Jan 15;6(1):11
18. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome*. 2018 Feb 1;6(1):23
19. Su M, Satola SW, Read TD. Genome-based prediction of bacterial antibiotic resistance.

- J. Clin. Microbiol.* 2019 Feb 27;57(3). e01405-18
20. Jia B, Raphenya AR, Alcock B, Waglechner N *et al.* CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 2017 Jan 4;45(D1):D566-D573
 21. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 2012 67:2640-2644
 22. Gupta SK, Padmanabhan BR, Diene SM, *et al.* ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother.* 2014;58(1):212–220
 23. Köser CU, Ellington MJ, Peacock SJ. Whole-genome sequencing to control antimicrobial resistance. *Trends Genet.* 2014;30(9):401–407
 24. Quiñones-Mateu ME, Avila S, Reyes-Teran G, Martinez MA. Deep sequencing: becoming a critical tool in clinical virology. *J Clin. Virol.* 2014;61(1):9–19
 25. Goodwin S, McPherson JD, McCombie WR. Coming of age: Ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* 2016 17: 333-351
 26. Teng JLL, Yeung ML, Chan E, Jia L, Lin CH *et al.* PacBio but not Illumina technology can achieve fast, accurate and complete closure of the high GC, complex *Burkholderia pseudomallei* two-chromosome genome. *Front. Microbiol.* 2017 Aug 2;8:1448
 27. Lu H, Giordano F, Ning Z. Oxford Nanopore MinION sequencing and genome assembly. *Genomics Proteomics Bioinformatics.* 2016 Oct;14(5):265-279
 28. Cao MD, Ganesamoorthy D, Elliott AG, Zhang H *et al.* Streaming algorithms for identification pathogens and antibiotic resistance potential from real-time MinION sequencing. *Gigascience.* 2016 5:32
 29. Rhoads A, Au KF. PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics.* 2015 13:278–289
 30. WHO. Global tuberculosis report 2018. Geneva: World Health Organization; 2018. License: CC BY-NC-SA 3.0 IGO.

31. WHO. The treatment of tuberculosis: Guidelines. 4th ed. Document WHO/HTM/TB/2009.420. Geneva, World Health Organization. 2010.
32. Ahmad N, Ahuja SD, Akkerman OW, Alffenaar JC et al. Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis. *Lancet*. 2018 Sep 8;392(10150):821-834
33. Kim SJ. Drug-susceptibility testing in tuberculosis: methods and reliability of results. 2005 *Eur. Respir. J.* 25:564–569.
34. WHO. Implementing tuberculosis diagnostics policy framework. Document WHO/HTM/TB/2015.11. Geneva, World Health Organization. 2015.
35. Boehme CC, Nabeta P, Hillemann D, Nicol MP *et al.* Rapid molecular detection of tuberculosis and rifampin resistance. *N. Engl. J. Med.* 2010 Sep 9;363(11):1005-15
36. Miotto P, Piana F, Penati V, Canducci F, Migliori GB, Cirillo DM. Use of genotype MTBDR assay for molecular detection of rifampin and isoniazid resistance in *Mycobacterium tuberculosis* clinical strains isolated in Italy. *J. Clin. Microbiol.* 2006 Jul;44(7):2485-91.
37. Almeida Da Silva PE, Palomino JC. Molecular basis and mechanisms of drug resistance in *Mycobacterium tuberculosis* classical and new drugs. *J. Antimicrob. Chemother.* 2011 66(7) 1417-1430
38. Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM. Routine use of microbial whole-genome sequencing in diagnostic and public health microbiology. *PLoS Pathog.* 2012 8:e1002824
39. Witney AA, Gould KA, Arnold A, Coleman D *et al.* Clinical application of whole-genome sequencing to inform treatment for multidrug-resistant tuberculosis cases. *J. Clin. Microbiol.* 2015 53:1473–1483
40. Witney AA, Cosgrove CA, Arnold A, Hinds J, Stoker NG, Butcher PD. Clinical use of whole genome sequencing for *Mycobacterium tuberculosis*. *BMC Med.* 2016 14:46

41. McNerney R, Clark TG, Campino S, *et al.* Removing the bottleneck in whole genome sequencing of *Mycobacterium tuberculosis* for rapid drug resistance analysis: a call to action. *Int. J. Infect. Dis.* 2017; 56: 130-135.
42. Head SR, Komori HK, LaMere SA *et al.* Library construction for next-generation sequencing: overviews and challenges. *BioTechniques* 2014; 56(2): 61-4
43. Pabinger S, Dander A, Fischer M, Snajder R *et al.* A survey of tools for variant analysis of next-generation genome sequencing data. *Brief Bioinform.* 2014 15: 256-278
44. Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Tuberculosis drug resistance mutation database. *PLoS Med.* 2009 6:e2.
45. Starks AM, Avilés E, Cirillo DM, Denkinger CM *et al.* Collaborative Effort for a Centralized Worldwide Tuberculosis Relational Sequencing Data Platform. *Clin. Infect. Dis.* 2015 Oct 15;61Suppl 3:S141-146.
46. Ezewudo M, Borens A, Chiner-Oms Á, *et al.* Integrating standardized whole genome sequence analysis with a global *Mycobacterium tuberculosis* antibiotic resistance knowledgebase. *Sci Rep.* 2018;8(1):15382
47. Feuerriegel S, Schleusener V, Beckert P, Kohl TA. PhyResSE: A Web Tool Delineating *Mycobacterium tuberculosis* Antibiotic Resistance and Lineage from Whole-Genome Sequencing Data. *J. Clin. Microbiol.* 2015 Jun;53(6):1908-1914
48. Farhat MR, Murray M, Choirat C, Prasad R, Honaker J, Durbin P, Idris MY, Crosas M, King G, genTB: Translational Genomics of Tuberculosis. 2015 <https://gentb.hms.harvard.edu>. Accessed May 30, 2019
49. Coll F, McNerney R, Preston MD, Guerra-Assunção JA *et al.* Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. *Genome Med.* 2015 May 27;7(1):51
50. Bradley P, Gordon NC, Walker TM, Dunn L *et al.* Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat. Commun.* 2015 Dec 21; 6:10063

51. Kohl TA, Utpatel C, Schleusener V, De Filippo MR, Beckert P, Cirillo DM, Niemann S. MTBseq: a comprehensive pipeline for whole genome sequence analysis of *Mycobacterium tuberculosis* complex isolates. *PeerJ*. 2018 Nov 13;6: e5895.
52. Köser CU, Bryant JM, Becq J, Török ME, Ellington MJ et al. Whole-genome sequencing for rapid susceptibility testing of *Mycobacterium tuberculosis*. *N. Engl. J. Med.* 2013 Jul 18;369(3):290-2
53. Miotto P, Tessema B, Tagliani E, 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. Published 2017 Dec 28.
54. Manson AL, Cohen KA, Abeel T, et al. Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multidrug resistance. *Nat. Genet.* 2017 49(3):395–402
55. Coll F, Phelan J, Hill-Cawthorne GA, Nair, MB et al. Genome-wide analysis of multi- and extensively drug-resistant *Mycobacterium tuberculosis*. *Nat. Genet.* 2018 50(2):307-316
56. CRyPTIC Consortium and the 100,000 Genomes Project, Allix-Béguet C, Arandjelovic I, et al. Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *N. Engl. J. Med.* 2018 379(15):1403–1415
57. Whitfield MG, Soeters HM, Warren RM, et al. A Global Perspective on Pyrazinamide Resistance: Systematic Review and Meta-Analysis. *PLoS One.* 2015 10(7): e0133869. Published 2015 Jul 28.
58. Papaventsis D, Casali N, Kontsevaya I, Drobniewski F, Cirillo DM, Nikolayevskyy V. Whole genome sequencing of *Mycobacterium tuberculosis* for detection of drug resistance: A systematic review. *Clin. Microbiol. Infect.* 2017 Feb;23(2):61-68
59. Shea J, Halse TA, Lapierre P, Shudt M 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):1871–1882.

60. Quan TP, Bawa Z, Foster D, *et al.* Evaluation of whole genome sequencing for mycobacterial species identification and drug susceptibility testing in a clinical setting: a large-scale prospective assessment of performance against line-probe assays and phenotyping. *J. Clin. Microbiol.* 2017 Nov 22;56: e01480-17
61. 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
62. Manson AL, Abeel T, Galagan JE *et al.* Mycobacterium tuberculosis whole genome sequences from Southern India suggest novel resistance mechanisms and the need for region-specific diagnostics. *Clin. Infect. Dis.* 2017 64(11):1494–1501
63. Doughty EL, Sergeant MJ, Adetifa I, Antonio M, Pallen MJ. Culture-independent detection and characterisation of *Mycobacterium tuberculosis* and *M. africanum* in sputum samples using shotgun metagenomics on a benchtop sequencer. *PeerJ.* 2014 2:e585. Published 2014 Sep 23
64. Brown AC, Bryant JM, Einer-Jensen K, *et al.* Rapid whole-genome sequencing of *Mycobacterium tuberculosis* isolates directly from clinical samples. *J. Clin. Microbiol.* 2015 53(7):2230–2237
65. Votintseva AA, Bradley P, Pankhurst L, *et al.* Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. *J. Clin. Microbiol.* 2017 55(5):1285–1298.
66. Doyle RM, Burgess C, Williams R, *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):e00666-18. Published 2018 Jul 26
67. Ismail NA, Omar SV, Joseph L, *et al.* Defining bedaquiline susceptibility, resistance, cross-resistance and associated genetic determinants: A retrospective cohort study. *EBioMedicine.* 2018 28:136–142
68. Zhang S, Chen J, Cui P, *et al.* *Mycobacterium tuberculosis* mutations associated with reduced susceptibility to linezolid. *Antimicrob. Agents Chemother.* 2016 60(4):2542–2544

Drug	Gene involved in resistance	Gene Function
Rifampin	<i>rpoB</i>	RNA polymerase β subunit
Isoniazid	<i>katG</i> <i>inhA</i> <i>mabA</i> <i>mabA/inhA</i> promoter region	Catalase-Peroxidase NADH-dependent enoyl ACP reductase 3-Ketoacyl-acyl carrier protein reductase
Ethambutol	<i>embA, embB, embC</i>	Arabinosyl transferases
Pyrazinamide	<i>pncA</i> <i>panD</i>	Pyrazinamidase/Nicotinamidase Aspartate 1-decarboxylase
Fluoroquinolones	<i>gyrA</i> <i>gyrB</i>	DNA gyrase subunit A DNA gyrase subunit B
Kanamycin/Amikacin	<i>rrs</i> <i>eis</i>	16S ribosomal RNA Aminoglycoside N-acetyltransferase
Capreomycin	<i>rrs</i> <i>tlyA</i>	16S ribosomal RNA rRNA methyltransferase
Ethionamide	<i>inhA</i> <i>ethA</i>	NADPH enoyl ACP reductase FAD-containing monooxygenase
Streptomycin	<i>rrs</i> <i>rpsL</i>	16S ribosomal RNA Ribosomal protein S12
Clofazimine	<i>Rv0678</i>	Transcriptional repressor of efflux pump Mmp15
Bedaquiline	<i>Rv0678</i> <i>atpE</i>	Transcriptional repressor of efflux pump Mmp15 ATP synthase subunit F0

Figure 1- Principal genes involved in resistance to anti-tuberculous drugs

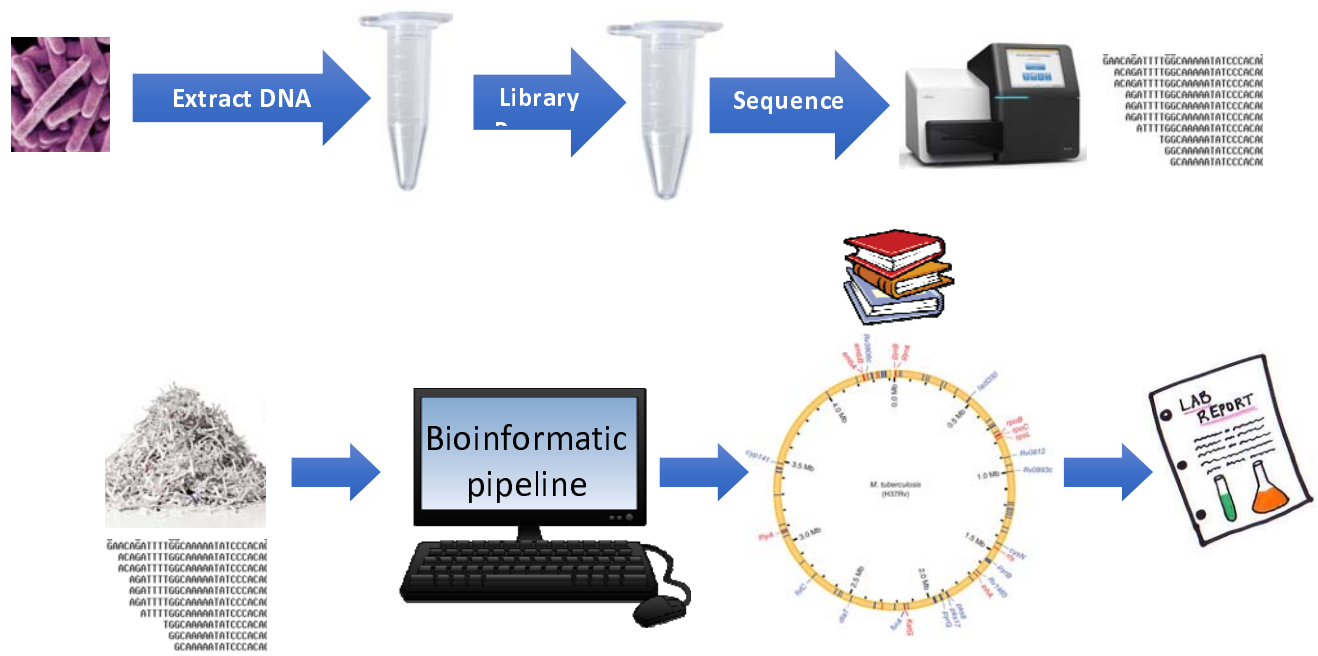


Figure 2- General workflow for TB WGS