Narrative review (Clinical Microbiology and Infection – CMI)

Title: *Mycobacterium tuberculosis* and whole-genome sequencing: how close are we to unleash its full potential?

Authors: Satta G[1,2], Lipman M[3,4], Smith GP[5,6], Arnold C[1,7] , Kon OM[2,8], McHugh TD[1].

1. UCL-TB and UCL Centre for Clinical Microbiology, Department of Infection, University College London, UK

2. Imperial College Healthcare NHS Trust, London, UK

3. UCL-TB and UCL Respiratory, University College London, UK

4. Royal Free London NHS Foundation Trust, London, UK

5. National Mycobacterium Reference Laboratory, Public Health England, UK

6. Heart of England NHS Foundation Trust, Birmingham, UK

7. Genomic Services and Development Unit, Public Health England, UK

8. National Heart and Lung Institute, Imperial College London, UK

Corresponding author:

Dr Giovanni Satta

Research Fellow - University College London, Centre for Clinical Microbiology, Department of Infection

Consultant in Medical Microbiology and Infectious Diseases - Imperial College Healthcare NHS Trust, London

Giovanni.satta@nhs.net

Abstract

Background: Nearly two decades after the completion of the genome sequence of *Mycobacterium tuberculosis* (MTB), and with the advent of next generation sequencing technologies (NGS), whole-genome sequencing (WGS) has been applied to a wide range of clinical scenarios. Starting in 2017, England is the first country in the world to pioneer its use on a national scale for the diagnosis of tuberculosis, detection of drug resistance and typing of MTB.

Aims: This narrative review critically analyses the current applications of WGS for MTB and explains how close we are to realizing its full potential as a diagnostic, epidemiologic and research tool.

Sources: We searched for reports (both original articles and reviews) published in English up to 31st May 2017, with combinations of the following keywords: whole genome sequencing, Mycobacterium and tuberculosis. MEDLINE, Embase and Scopus were used as search engines. We included articles that covered different aspects of whole genome sequencing in relation to MTB.

Content: This review focuses on three main themes: the role of WGS for the prediction of drug susceptibility, MTB outbreak investigation and genetic diversity, and research applications of NGS.

Implications: Many of the original expectations have been accomplished, and we believe that with its unprecedented sensitivity and power WGS has the potential to address many unanswered questions in the near future. However, caution is still needed when interpreting WGS data as there are some important limitations to be aware of, from the correct interpretation of drug susceptibilities to the bioinformatic support needed.

Introduction

The complete genome sequence of *Mycobacterium tuberculosis* (MTB) was described in 1998 [1]. Since then, whole-genome sequencing (WGS) has been applied to a wide range of clinical scenarios, with the potential to revolutionize tuberculosis (TB) diagnosis, outbreak investigation, drug and vaccine development, plus assist in understanding MTB evolution and pathogenicity. In 2017, Public Health England (PHE) introduced routine WGS in the clinical setting of the National Health Service (NHS); and England will be the first country in the world to pioneer its use at a population level for the diagnosis, detection of drug resistance and typing of MTB [2] (Scotland, Wales and Northern Ireland are not included yet).

The main aim of this narrative review is to summarize and critically analyze the current literature relating to WGS and MTB as a diagnostic, epidemiologic and research tool; and determine to what extent we have been able to utilize its full potential (search criteria are specified in **Box 1**). This review will provide the reader with an overview of the current applications (and limitations) of WGS, with a specific focus on three main themes: prediction of drug susceptibility, outbreak investigation and genetic diversity, and research applications (pathogenesis, drug discovery and vaccine development) (**Table 1**). It will interpret the present situation and provide guidance on the future direction of development.

Prediction of drug susceptibility and resistance

Extensive work on WGS and susceptibility testing has been led by the Wellcome Trust/University of Oxford (and collaborators), with the demonstration that this approach can be integrated into routine diagnostic workflows, with data generated in 9 days (compared to an average of 3 weeks for conventional susceptibility methods) (**Figure 1**) and at a similar or cheaper price than the current

diagnostic workflows [3] [4]. A recent systematic review [5] on the use of WGS for the detection of drug resistance in MTB highlighted its role as a promising alternative to existing phenotypic and molecular drug susceptibility testing methods. In particular, this high sensitivity and specificity has been demonstrated for the first line drugs isoniazid (INH) and rifampicin (RIF). However, there is still significant variation with the remaining first and second line drugs. This results from the complexity of genetic mechanisms conferring resistance, plus in some cases resistance driven by non-specific mechanisms such as efflux pumps [6]. This may be of considerable importance when using WGS to guide clinical decisions. The lack of standardization for some phenotypic tests, in particular pyrazinamide, also confounds the identification of resistance in clinical isolates and mutants [7].

The implementation of WGS in the clinical setting is hindered by some significant limitations. Not least of these is the requirement for a culture sample before DNA isolation and sequencing. Direct WGS from clinical samples has been successfully demonstrated [8] [9] [10]; though all current protocols for clinical practice still require a culture step. Additionally, before switching completely to WGS testing, a long transition period with both WGS and conventional culture susceptibility methods used in parallel will be needed to clarify the current discrepancies between genotype and phenotype and to satisfy the need for large data sets to better understand the role of rare resistance mechanisms, and the level of resistance conferred by different mutations.

On the other hand, major progress has already been made in terms of data interpretation and standardization. Several online software tools, *all-in-one* and easy to use, are available for the rapid interpretation of WGS data in MTB[11] and various online databases (ReSeqTB[12], TB Portal[13], GenTB[14]) offer support to share, interpret and link genetic results to drug resistance phenotype and other epidemiological variables, providing international harmonization of such data. At the same time, these and other databases also allow the automatic surveillance of drug resistance at a global level, influencing public health interventions.

WGS has the potential to revolutionize the definition of drug susceptibility testing (DST) of MTB in both high- and low-income settings and a growing knowledge of the genetic mechanisms of resistance, combined with an improved IT infrastructure, will facilitate its adoption and enhance its clinical utility for drug testing. A key challenge is to demonstrate that its use in the routine diagnostic service will have an impact on patient outcomes. Treatment guidelines have evolved over decades to reflect the role of clinical suspicion and not to over-rely on microbiological positive results, especially in the setting of culture negative samples or in case of extra-pulmonary disease. Recently, rapid PCR methods have allowed the rapid detection of resistance within the same day. However data from developing countries do not seem to support any improvement in outcome [15] [16] and it is currently unknown if this is also relevant to WGS.

Outbreak investigation and genetic diversity

Several studies [4] [17] [18] [19] have demonstrated the utility of WGS in public health interventions and in the detection of outbreaks and transmission events, confirming its higher resolution when compared to MIRU-VNTR typing, IS6110 RFLP typing and spoligotyping methods. Some of the latest techniques (e.g.

spoligotyping) lack discrimination [20] when determining the likely route of transmission or distinguishing isolates with minor differences. That WGS can identify Single Nucleotide Polymorphisms (SNPs) between different strains is particularly useful when trying to discriminate clinical relapse from reinfection. For example, using data from the recent RIFAQUIN trial [21], 36 patients with positive cultures before and after treatment had their strains typed using 24-loci MIRU-VNTR, in silico spoligotyping and WGS. Whilst WGS and MIRU-VNTR differentiated relapse and reinfection with a similar level of discrimination, WGS provided significant extra information. One pair had an intermediate number of SNP differences (more than 5), which was likely to result from a mixed infection with a pre-treatment minor genotype that was highly related to the post-treatment genotype. WGS reclassified this as a relapse, whilst MIRU-VNTR typing could not do this. From a global perspective, WGS has allowed the detection of genetic diversity in MTB with unprecedented resolution. It has been used to understand its evolution, including the discovery of its origin from *Mycobacterium canettii* (first described in 1969 and now considered the progenitor species from which MTB emerged) [22], and also lineages, variation at global, local and individual levels [23].

By contrast, the utility of WGS has been questioned in a recent paper analyzing a large isoniazid-resistant (INH-R) MTB outbreak in London (UK) [24]. WGS did provide increased resolution over variable number tandem repeat (VNTR)-based clustering but was still insufficient to resolve transmission networks in a very large tuberculosis outbreak. There appears to be little evidence that the routine use of VNTR-genotyping helps in controlling transmission; and it has not been shown to be effective or cost-effective in its current form [25]. Thus, uncertainty remains whether this also applies to WGS though its *all-in-one* potential (including identification, susceptibility and "typing") can provide relevant data for different purposes.

From a practical perspective, microbiology laboratories, physicians and TB nurses are familiar with the linear, visual numeric MIRU-VNTR profile that is relatively easy to compare between patients. Without adequate bioinformatics support and IT infrastructure, public health teams will struggle to interpret WGS output data, which may lead to significant delays in the identification of outbreaks and transmission events. Additionally, the delivery of this information to clinicians and TB nurses in an understandable format and in a timely manner remains a key issue. Obtaining MIRU-VNTR profiles from WGS data is technically challenging and phylogenetic trees need to be run whenever a new sample has been added. The amount of data generated can be immense and not easily supported by the current IT infrastructure. This is something that still needs to be addressed. Cloud-based systems may represent a solution and allow the discovery of nationwide transmission events.

Pathogenesis, drug targets discovery and vaccine development

The complete genome of MTB includes 4 million base pairs and 4000 genes [1]. However, up to 50% of these genes are still labeled as unknown, uncharacterized or with hypothetical function [26]. More recently, a SNP-calling program combined with WGS has allowed the creation of a catalog of virulence genes and new potentially virulent sub-lineages [27]. As illustrated in a recent paper exploring the genetic variation of a sub-cluster of the London INH-R TB outbreak [28], WGS was also able to demonstrate specific deletions and SNPs affecting different genes that were peculiar to those clinical strains and

which could potentially explain the persistence of the outbreak over years. However, the simple presence of such mutations does not necessary confirm their function – which is another limitation of WGS. Further studies (transcriptomics, knock out models, recombineering) are still needed to confirm the impact of those genes in the pathogenesis of TB and if any of the virulence genes involved by SNPs can be actually used as drug targets for the development of new compounds. Caution also needs to be used with this approach. Many of these genes may not be essential, accompanied by genetic redundancy or metabolic scavenging (overcoming nutrient insufficiency by using alternative pathways or substrates), leading us to appreciate that pharmacological target-validation may be far more reliable and a better predictor of success [29] [30]. Additionally, compensatory mutations in other part of the genome may balance the initial fitness cost and further complicate the overall picture [31].

WGS has the potential to revolutionize the process of drug target identification [32]. It has been successfully used to determine the target of Bedaquiline (BDQ) [33]. Here, the authors selected and sequenced BDQ resistant *Mycobacterium smegmatis* strains and identified mutations in the proton pump of adenosine triphosphate (ATP) synthase associated with resistance. Others [34] have selected mutants directly using MTB to evaluate combination treatment against drug resistant strains. A scalable platform for the discovery of drug targets, based on combining high-throughput screening (HTS) with whole-genome sequencing (WGS) of resistant isolates, has been proposed [35]. Using this approach, the authors have identified promising new drug targets. Surprisingly, some of the resistant bacteria had mutations on the same genes (though were resistant to structurally different compounds). When sequencing the genome of resistant mutants, researchers should remember that the mutations conferring resistance may involve different stages of drug metabolism: mutations can affect enzymes that convert the compound into an active form (and then inhibit an unknown target), the target gene, regions responsible for up- or down-regulation of targets and activating enzymes, efflux pumps or detoxifying enzymes. It is also possible that the same drug may be inactivated by different mechanisms of resistance and some of them may be common to other compounds, as already described for other anti-MTB drugs such as isoniazid and ethionamide [36]. WGS could therefore contribute to elucidating global mechanisms of action/resistance.

Understanding bacterial pathogenesis is expected to provide an instrumental contribution to vaccine development, particularly to target those pathogens (such as MTB) for which the traditional approaches have not been completely successful [37]. The comparison by genome sequencing of MTB clinical strains CDC1551 with the laboratory-adapted H37Rv has demonstrated a more extensive variability than had been anticipated [38], highlighting the importance of understanding genetic diversity in pathogenic strains and of validating candidate targets with other genomic techniques. A functional genomic approach has been developed and inhibition techniques, such as signature-tagged mutagenesis (STM) and transposon site hybridization (TraSH), have been successfully applied to MTB [39] [40] with the potential for better-attenuated vaccines.

Discussion

WGS has undoubtedly offered us a greater understanding of MTB, in particular its epidemiology and evolution. WGS has the potential to revolutionize susceptibility testing in the routine microbiology

laboratory. With the cost of sequencing likely to continue to decrease (it has already dropped from thousands to around a hundred dollars per test) and the capabilities of sequencing technologies to improve, the number of clinical and research laboratories implementing it will increase, as well as the number of trained scientists able to interpret the data. However, several important challenges still exist and it is important for clinicians and microbiologists to be aware of the limitations.

From a clinical perspective, the impact on treatment outcomes and on the reduction of onward transmission still needs to be demonstrated and other questions still remain unanswered (**Table 2**). A positive culture is still needed before DNA isolation and sequencing can be performed and this generally takes at least a couple of weeks. On the positive side, clinicians will be able to receive preliminary results weeks in advance of conventional susceptibility culture methods. PHE plans to release genotypic resistance prediction for seven antituberculous drugs (Rifampicin, Isoniazid, Ethambutol, Pyrazinamide, Quinolones group, Streptomycin and Amikacin) as soon as WGS has been performed on the isolate (unpublished data from PHE). Culture results will follow but the plan is to potentially stop routine phenotypic testing for Rifampicin, Isoniazid, Ethambutol and Pyrazinamide on the majority of those isolates that have no mutations in resistance genes for these four drugs, and to provide MICs on a wider range of drugs where there is an indication of resistance mutations, drug intolerance or other clinical indicators of poor response to first line treatment. Current molecular methods (i.e. GeneXpert and Line Probe Assays) already offer rapid detection of Rifampicin and Isoniazid resistance directly from clinical samples (smear positive and negative) [41] [42]. They objectively reduce the diagnostic time in the laboratory with also a major reduction in the delay between identification of patients at risk of MDR-TB and initiation of treatment [43]. However, they are not exempt from some major limitations, including the detection of mutations outside the target regions (GeneXpert cannot detect INH mono-resistance) and low sensitivity for other drugs (Line Probe MTBDRsl) [44]. WGS is unlikely to completely replace phenotypic DST for TB in the near future. Large datasets are needed to clarify the current discrepancies between genotype and phenotype, as well as the role of rare resistance mechanisms and the level of resistance conferred by different mutations. In this respect, the CRyPTIC Project (Comprehensive Resistance Prediction for Tuberculosis: an International Consortium) aims to collect 100,000 MTB isolates with each global site performing drug-resistance testing in parallel with WGS. The results will be assembled into a single open-access database [45] that may allow a more complete understanding of the relationship between clinical resistance and genotype. Additionally, the real potential of WGS, together with deep sequencing, will be to distinguish also individual variation (that is, relapse from reinfection) and the detection of minority variants and subpopulations with low-level drug resistance, that may be selected during treatment and contribute to the emergence of drug resistance [46].

From a laboratory perspective, various practical considerations need to be taken into account. These include the challenge in extracting MTB DNA to the need for extensive and robust IT infrastructure with cloud-based systems (or remote access to servers) and reliable high-speed internet connections. Despite different online tools available for the rapid interpretation of WGS data [11], a powerful Information Technology (IT) infrastructure is still needed for the analysis, storage and transfer of multiple samples and cloud based services and reliable internet connections are necessary. The amount of data generated by WGS is immense and laboratory computers will need to be upgraded and with a robust backup

system in place. The lack of bioinformatics expertise among clinical microbiologists is another major potential barrier for clinical adoption as the introduction of WGS will be hampered by the complexity of data and its analysis. A new generation of biomedical and clinical scientists will need to be trained and, considering the applications of WGS in many other fields (genetics and oncology in particular), bio-informaticians may become part of the pathology workforce as IT managers were in the early 1990s when laboratory information systems were first introduced [47]. Finally, accreditation (ISO 15189 and others) will be essential for clinical diagnostic laboratories performing and interpreting WGS: PHE is already working toward this goal [2].

From a research perspective, WGS should support us in unveiling the real role in virulence and pathogenicity of MTB genes and clinical strains represent the perfect samples for further analysis. In this context, we would support wide-scale investment in WGS and the creation of comparable databases of genetic variations of clinical strains, plus the addition of biological information, including fitness assays [48] and mutation rate [49], for a comprehensive picture useful in clinical practice. This will potentially allow the discovery of new drug targets and the deciphering of MTB gene function.

In conclusion, this review has summarized the current applications, achievements and limitations of WGS for MTB's diagnosis, epidemiology and research. Many of the original expectations have been accomplished, with greater understanding of its evolution and global diversity, applications in drug discovery and vaccine development and, more importantly, its introduction at a national scale for susceptibility testing and typing. With its incremental sensitivity and power WGS has the potential to address many unanswered questions in the near future. However, caution is still needed when interpreting WGS data and it is important to be aware of its limitations.

Tables and figures

Box 1: Search Strategy and selection criteria

We searched for PubMed reports published in English until 31st May 2017, with combinations of the following keywords: whole genome sequencing, WGS, Mycobacterium and tuberculosis. We searched for reports (both original articles and reviews) published in English up to 31st May 2017, with combinations of the following keywords: whole genome sequencing, Mycobacterium and tuberculosis. MEDLINE, Embase and Scopus were used as search engines. We reviewed the articles resulting from these searches and the relevant references cited in them. Additional search using an internet search engine was performed to potentially identify some additional papers. As this is narrative review, we considered to be most appropriate to focus on three main themes (prediction of susceptibility, outbreak investigation and research applications) and included articles that reflected this. Inevitably, we did not include some papers that were not relevant to the main themes and that did not provide any additional information.

Table 1: Current applications of whole genome sequencing and *Mycobacterium tuberculosis,* with achievements and limitations

1. Prediction of drug susceptibility and resistance	
Achievements	**Limitations**
- Diagnostic workflow with data generated in 9 days and at a price 7% cheaper - First line drugs (Rifampicin and Isoniazid): strong performance with high sensitivity and specificity - Potential for WGS directly from clinical samples - Online tools available for rapid data interpretation	- Significant variation for the remaining first line and other drugs - Culture still needed for DNA extraction and WGS - Bioinformatics support and IT infrastructure needed to download and analyze data - Lack of accreditation (ISO 15189 and others)
2. Epidemiological analysis	
Achievements	**Limitations**
- Higher resolution compared to MIRU-VNTR typing, IS6110 RFLP typing and spoligotyping methods - Ability to distinguish relapse from reinfection - Better understanding of evolution, lineages and genomic variation	- Still insufficient to resolve transmission networks in tuberculosis outbreaks - Clinical benefits and cost-effectiveness not demonstrated - Bioinformatics support and IT infrastructure needed to download and analyze data
3. Research	
Achievements	**Limitations**
- Demonstration of specific deletions and SNPs peculiar to clinical strains	- Further studies and techniques still needed to confirm gene function

- Discovery of mechanism of action of new drugs - Demonstration of more extensive genetic variability than original expected - Potential in vaccine development	- Mutations may be nonspecific - Need for large database to compare data at international level

Table 2: Questions still to be addressed by whole genome sequencing

1. Does the application of WGS provide better treatment outcomes?
2. Does the application of WGS help in reducing transmission?
3. Does WGS provide any benefit compared with current molecular methods?
4. Can WGS clarify the current discrepancies between genotype and phenotype, the role of rare resistance mechanisms and the level of resistance conferred by different mutations?
5. How can WGS be implemented in the current laboratory workflow for MTB diagnosis?
6. Can WGS identify minority variants and subpopulations from pooled sweeps of colonies?
7. What is the real extent of MTB global, local and individual diversity? And how can we use this information for the selection of a better vaccine?

Figure 1: Workflow or the processing of mycobacterial samples and whole genome sequencing. Genotypic susceptibilities can be available within 9 days of a positive culture, compared to an average of 3 weeks with traditional phenotypic methods. Abbreviations: AFB – Acid Fast Bacilli, RIF – Rifampicin, INH – Isoniazid.
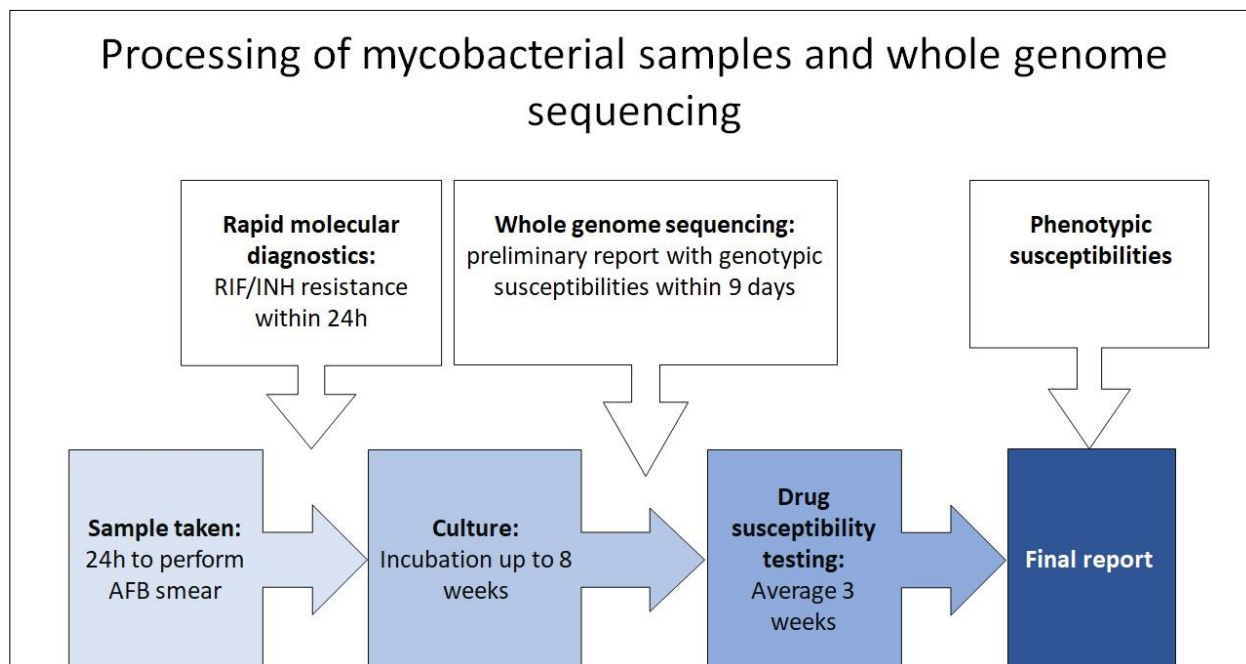
References

1.   Cole ST, Brosch R, Parkhill J, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* 1998; **393**(6685): 537-44.
2.   PHE. England world leaders in the use of whole genome sequencing to diagnose TB. 2017. https://www.gov.uk/government/news/england-world-leaders-in-the-use-of-whole-genome-sequencing-to-diagnose-tb (accessed 30th May 2017.
3.   Pankhurst LJ, del Ojo Elias C, Votintseva AA, et al. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. *The Lancet Respiratory Medicine* 2016; **4**(1): 49-58.
4.   Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *The Lancet Infectious Diseases* 2015; **15**(10): 1193-202.
5.   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; **23**(2): 61-8.
6.   Zhang Y, Yew WW. Mechanisms of drug resistance in Mycobacterium tuberculosis: update 2015. *Int J Tuberc Lung Dis* 2015; **19**(11): 1276-89.
7.   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.
8.   Brown AC, Bryant JM, Einer-Jensen K, et al. Rapid Whole-Genome Sequencing of Mycobacterium tuberculosis Isolates Directly from Clinical Samples. *Journal of Clinical Microbiology* 2015; **53**(7): 2230-7.
9.   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): 1285-98.
10.   Nimmo C, Doyle R, Burgess C, et al. Rapid identification of a Mycobacterium tuberculosis full genetic drug resistance profile through whole genome sequencing directly from sputum. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* 2017; **62**: 44-6.
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.
12.   Relational Sequencing TB Data Platform (ReSeqTB). https://platform.reseqtb.org/ (accessed 13th October 2017).
13.   TB Portal Data. https://data.tbportals.niaid.nih.gov/ (accessed 13th October 2017).
14.   Translational Genomics of Tuberculosis (genTB). https://gentb.hms.harvard.edu/ (accessed 13th October 2017).
15.   Trajman A, Durovni B, Saraceni V, et al. Impact on Patients' Treatment Outcomes of XpertMTB/RIF Implementation for the Diagnosis of Tuberculosis: Follow-Up of a Stepped-Wedge Randomized Clinical Trial. *PLOS ONE* 2015; **10**(4): e0123252.
16.   Chawla KS, Kanyama C, Mbewe A, et al. Policy to practice: impact of GeneXpert MTB/RIF implementation on the TB spectrum of care in Lilongwe, Malawi. *Transactions of The Royal Society of Tropical Medicine and Hygiene* 2016; **110**(5): 305-11.
17.   Török ME, Reuter S, Bryant J, et al. Rapid Whole-Genome Sequencing for Investigation of a Suspected Tuberculosis Outbreak. *Journal of Clinical Microbiology* 2013; **51**(2): 611-4.
18.   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 Medicine* 2013; **10**(2): e1001387.

19.	Stucki D, Ballif M, Bodmer T, et al. Tracking a tuberculosis outbreak over 21 years: strain-specific single-nucleotide polymorphism typing combined with targeted whole-genome sequencing. *The Journal of infectious diseases* 2014; **211**(8): 1306-16.

20.	Jagielski T, van Ingen J, Rastogi N, et al. Current Methods in the Molecular Typing of Mycobacterium tuberculosis and Other Mycobacteria. *BioMed Research International* 2014; **2014**: 21.

21.	Witney AA, Bateson ALE, Jindani A, et al. Use of whole-genome sequencing to distinguish relapse from reinfection in a completed tuberculosis clinical trial. *BMC Medicine* 2017; **15**(1): 71.

22.	Blouin Y, Hauck Y, Soler C, et al. Significance of the identification in the Horn of Africa of an exceptionally deep branching Mycobacterium tuberculosis clade. *PLoS One* 2012; **7**(12): e52841.

23.	Ford C, Yusim K, Ioerger T, et al. Mycobacterium tuberculosis--heterogeneity revealed through whole genome sequencing. *Tuberculosis (Edinb)* 2012; **92**(3): 194-201.

24.	Casali N, Broda A, Harris SR, Parkhill J, Brown T, Drobniewski F. Whole Genome Sequence Analysis of a Large Isoniazid-Resistant Tuberculosis Outbreak in London: A Retrospective Observational Study. *PLoS Med* 2016; **13**(10): e1002137.

25.	Mears J, Vynnycky E, Lord J, et al. The prospective evaluation of the TB strain typing service in England: a mixed methods study. *Thorax* 2015.

26.	Mazandu GK, Mulder NJ. Function Prediction and Analysis of Mycobacterium tuberculosis Hypothetical Proteins. *International Journal of Molecular Sciences* 2012; **13**(6): 7283-302.

27.	Mikheecheva NE, Zaychikova MV, Melerzanov AV, Danilenko VN. A Nonsynonymous SNP Catalog of Mycobacterium tuberculosis Virulence Genes and Its Use for Detecting New Potentially Virulent Sublineages. *Genome Biol Evol* 2017; **9**(4): 887-99.

28.	Satta G, Witney AA, Shorten RJ, Karlikowska M, Lipman M, McHugh TD. Genetic variation in Mycobacterium tuberculosis isolates from a London outbreak associated with isoniazid resistance. *BMC Med* 2016; **14**(1): 117.

29.	Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular Microbiology* 2003; **48**(1): 77-84.

30.	Lechartier B, Rybniker J, Zumla A, Cole ST. Tuberculosis drug discovery in the post-post-genomic era. *EMBO Molecular Medicine* 2014; **6**(2): 158-68.

31.	Comas I, Borrell S, Roetzer. Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet* 2011; **44**.

32.	Sala C, Hartkoorn RC. Tuberculosis drugs: new candidates and how to find more. *Future Microbiology* 2011; **6**(6): 617-33.

33.	Andries K, Verhasselt P, Guillemont J, et al. A Diarylquinoline Drug Active on the ATP Synthase of Mycobacterium tuberculosis. *Science* 2005; **307**(5707): 223-7.

34.	Gonzalo X, Casali N, Broda A, Pardieu C, Drobniewski F. Combination of amikacin and doxycycline against multidrug-resistant and extensively drug-resistant tuberculosis. *Int J Antimicrob Agents* 2015; **45**(4): 406-12.

35.	Ioerger TR, O'Malley T, Liao R, et al. Identification of New Drug Targets and Resistance Mechanisms in Mycobacterium tuberculosis. *PLOS ONE* 2013; **8**(9): e75245.

36.	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): 1417-30.

37.	Scarselli M, Giuliani MM, Adu-Bobie J, Pizza M, Rappuoli R. The impact of genomics on vaccine design. *Trends in Biotechnology*; **23**(2): 84-91.

38.	Fleischmann RD, Alland D, Eisen JA, et al. Whole-Genome Comparison of Mycobacterium tuberculosis Clinical and Laboratory Strains. *Journal of Bacteriology* 2002; **184**(19): 5479-90.

39.	Camacho LR, Unité de Génétique Mycobactérienne IP, 25 rue du Dr Roux, 75724 Paris Cedex 15, France., Ensergueix D, et al. Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis. *Molecular Microbiology* 2017; **34**(2): 257-67.

40.   Sassetti CM, Boyd DH, Rubin EJ. Comprehensive identification of conditionally essential genes in mycobacteria. 2001.

41.   Small PM, Pai M. Tuberculosis Diagnosis — Time for a Game Change. *New England Journal of Medicine* 2010; **363**(11): 1070-1.

42.   Tomasicchio M, Theron G, Pietersen E, et al. The diagnostic accuracy of the MTBDRplus and MTBDRsl assays for drug-resistant TB detection when performed on sputum and culture isolates. 2016; **6**: 17850.

43.   Singla N, Satyanarayana S, Sachdeva KS, et al. Impact of Introducing the Line Probe Assay on Time to Treatment Initiation of MDR-TB in Delhi, India. *PLOS ONE* 2014; **9**(7): e102989.

44.   Kiet VS, Lan NT, An DD, et al. Evaluation of the MTBDRsl test for detection of second-line-drug resistance in Mycobacterium tuberculosis. *J Clin Microbiol* 2010; **48**(8): 2934-9.

45.   CRyPTIC. Comprehensive Resistance Prediction for Tuberculosis: an International Consortium. 2017. http://www.crypticproject.org/.

46.   Colman RE, Schupp JM, Hicks ND, et al. Detection of Low-Level Mixed-Population Drug Resistance in Mycobacterium tuberculosis Using High Fidelity Amplicon Sequencing. *PLOS ONE* 2015; **10**(5): e0126626.

47.   Gibbon G. A brief history of LIMS. *Laboratory Automation & Information Management* 1996; **32**(1): 1-5.

48.   O'Sullivan DM, McHugh TD, Gillespie SH. Mapping the fitness of Mycobacterium tuberculosis strains: a complex picture. *Journal of Medical Microbiology* 2010; **59**(12): 1533-5.

49.   Pope CF, O'Sullivan DM, McHugh TD. A practical guide to measuring mutation rates in antibiotic resistance. *Antimicrob Agents Chemother* 2008; **52**.