Deciphering drug resistance in Mycobacterium tuberculosis using whole-genome sequencing
Progress, promise, and challenges

Cohen, Keira A.; Manson, Abigail L.; Desjardins, Christopher A.; Abeel, Thomas; Earl, Ashlee M.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

REVIEW   **Open Access**

Deciphering drug resistance in *Mycobacterium tuberculosis* using whole-genome sequencing: progress, promise, and challenges

Keira A. Cohen[1*], Abigail L. Manson[2], Christopher A. Desjardins[2], Thomas Abeel[2,3] and Ashlee M. Earl[2*]

Abstract

Tuberculosis (TB) is a global infectious threat that is intensified by an increasing incidence of highly drug-resistant disease. Whole-genome sequencing (WGS) studies of *Mycobacterium tuberculosis*, the causative agent of TB, have greatly increased our understanding of this pathogen. Since the first *M. tuberculosis* genome was published in 1998, WGS has provided a more complete account of the genomic features that cause resistance in populations of *M. tuberculosis*, has helped to fill gaps in our knowledge of how both classical and new antitubercular drugs work, and has identified specific mutations that allow *M. tuberculosis* to escape the effects of these drugs. WGS studies have also revealed how resistance evolves both within an individual patient and within patient populations, including the important roles of de novo acquisition of resistance and clonal spread. These findings have informed decisions about which drug-resistance mutations should be included on extended diagnostic panels. From its origins as a basic science technique, WGS of *M. tuberculosis* is becoming part of the modern clinical microbiology laboratory, promising rapid and improved detection of drug resistance, and detailed and real-time epidemiology of TB outbreaks. We review the successes and highlight the challenges that remain in applying WGS to improve the control of drug-resistant TB through monitoring its evolution and spread, and to inform more rapid and effective diagnostic and therapeutic strategies.

Background

Mycobacterium tuberculosis is the causative agent of tuberculosis (TB), which is most often spread person-to-person via cough aerosols. Although many individuals who are exposed to *M. tuberculosis* never develop active disease, the World Health Organization (WHO) estimated 10 million new cases of active TB and 1.3 million deaths in 2017 alone [1].

Since its initial documentation in the 1940s [2], drug-resistant TB has threatened public health control efforts. In 2016, there were an estimated 490,000 new cases of multidrug-resistant (MDR) TB, which is defined by phenotypic resistance to both isoniazid and rifampicin

[3]. Approximately 10% of MDR-TB cases globally can be classified as extensively drug-resistant (XDR), indicating that there is concomitant resistance to quinolones (such as the fluoroquinolones, levofloxacin, and moxifloxacin) and to a second-line injectable agent (amikacin, kanamycin, or capreomycin) [3]. As expected, drug-resistance patterns predict treatment outcome; in 2015, TB treatment success overall was 83%, whereas the success rate was 54% for MDR-TB or rifampicin-resistant-TB (RR-TB) and only 30% for XDR-TB [4].

Culture-based techniques remain the current reference standard for both diagnosis and drug-susceptibility testing of TB, but these processes are time-consuming and require specialized laboratory capacity. More recently, the use of rapid molecular tests for the diagnosis of TB has increased globally, particularly the use of Xpert MTB/RIF (Cepheid, Sunnyvale, CA), a PCR-based assay

* Correspondence: kcohen8@jhmi.edu; aearl@broadinstitute.org
[1]Division of Pulmonary and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, MA 21205, USA
[2]Broad Institute of Harvard and Massachusetts Institute of Technology, 415 Main Street, Cambridge, MA 02142, USA
Full list of author information is available at the end of the article

Cohen *et al. Genome Medicine*   (2019) 11:45

Page 2 of 18

that simultaneously detects the presence of *M. tuberculosis* and resistance to rifampicin.

Current recommendations for the treatment of drug-susceptible TB include a 6-month course of a multi-drug regimen of rifampicin, isoniazid, pyrazinamide, and ethambutol. Historically, treatment of MDR- or XDR-TB involved the long-term use of second-line drugs, including injectable agents [5]. More recently, the MDR-TB treatment landscape has changed with the introduction of multiple novel second-line drugs that can be administered orally (Table 1). In 2012, bedaquiline, a diarylquinolone, became the first TB drug from a novel drug class to receive US Food and Drug Administration (FDA) approval in over 40 years [48, 49] (Table 1). Another oral agent, delamanid, a nitro-dihydro-imidazooxazole derivative, has also shown promise for TB treatment [50, 51].

In 2018, the WHO published updated treatment guidelines for MDR/RR-TB [47], recommending fully oral MDR regimens for many patient groups. Recommended treatment strategies include both shorter, standardized MDR regimens (for 9 to 12 months) and longer, individualized treatment regimens (for 18 to 20 months). The updated guidelines group antitubercular drugs on the basis of how they should be combined to create individualized, longer MDR-TB regimens [47] (Table 1).

Despite advances in both diagnostics and therapeutics for TB, challenges remain. Obstacles for rapid *M. tuberculosis* diagnosis include: (i) the imperfect sensitivity of molecular tests for the detection of this pathogen, particularly in the case of paucibacillary TB (where there is a lower bacterial burden); (ii) lack of comprehensive molecular assays due to incomplete knowledge of all resistance mutations in TB; and (iii) technical limitations to the numbers of mutations that can be included on diagnostic molecular platforms. Furthermore, the rollout of rapid diagnostic platforms to low-resource settings has been a challenge. Remaining treatment challenges include: prolonged treatment courses, leading to greater drug exposure, toxicity, and non-compliance; unacceptable side-effect profiles; logistics of drug access; and reinfection [52].

The dawning of the new age of genome sequencing began to revolutionize our approach to human diseases, including TB. In 1998, Cole et al. [53] reported the complete genome sequence of the *M. tuberculosis* reference strain H37Rv, which was approximately 4.41 million base pairs in length and encoded approximately 4000 genes. The first sequencing of a clinical reference strain, CDC1551, quickly followed [54]. An accompanying editorial optimistically stated: "After several decades in the slow lane of classical microbiology, *M. tuberculosis* is once again at the cutting edge of science" [55].

However, even at the time of these breakthroughs, there was recognition that translating these genomic data into clinical benefit would prove challenging [55]. Despite these challenges, it is clear, more than 20 years later, that *M. tuberculosis* genomic data have been remarkably useful in improving our understanding of how drug-resistant TB evolves and spreads and in helping to inform diagnostics and therapies.

In this review, we discuss the molecular epidemiologic and diagnostic advances made by sequencing *M. tuberculosis*, with a focus on drug-resistant TB. We do not review the practice of whole-genome sequencing (WGS) of *M. tuberculosis* as this has been reviewed recently [56]. Key findings that are discussed include the use of WGS to identify drug-resistance determinants in *M. tuberculosis* and to elucidate the evolution and spread of drug-resistant organisms, and the clinical applications of this technology (Table 2).

Identifying *M. tuberculosis* drug-resistance determinants

Drug resistance in *M. tuberculosis* is the result of chromosomal mutations in existing genes that are passed along through vertical descent, that is, passed from mother to daughter cells. Unlike many other bacterial pathogens, *M. tuberculosis* rarely recombines via lateral exchange of DNA [83] and also lacks plasmids. Many of the resistance determinants were discovered before the sequencing of the *M. tuberculosis* genome was completed. By 1998, resistance mechanisms had already been discovered for classical first- and second-line TB drugs including isoniazid (alterations in genes *katG* and *inhA*); rifampicin (in *rpoB*); streptomycin (in *rrs* and *rpsL*); pyrazinamide (in *pncA*); ethambutol (in *embB*); quinolones (in *gyrA*); and kanamycin (in *rrs*) (reviewed in Ramaswamy and Musser [84]) (Table 1). However, the targeted amplification and sequencing of known or suspected resistance genes revealed that these mechanisms were insufficient to explain all phenotypic resistance [85, 86], and resistance mechanisms for several newer drugs—including pretomanid, bedaquiline, and delamanid—were discovered over the next eight years during a period when WGS was becoming routine. Together, in the past 20 years, WGS-based approaches, focused on both laboratory-derived and naturally circulating populations of drug-resistant *M. tuberculosis*, have provided a more complete account of the genomic features that cause treatment resistance, enabling the identification of novel resistance mechanisms for existing drugs, and the determination of the mechanisms of action of newly discovered drugs.

Identifying resistance determinants in laboratory-derived mutants

Drug-resistant mutants can be derived in vitro by growing drug-susceptible *M. tuberculosis* strains in drug-

Table 1 Antitubercular drug-resistance mechanisms[a]

WHO category	Drug or drug class	Resistance genes	Rv number	Gene function	Mechanism of drug resistance	Reference(s)
First-line agents	Rifamycins (for example, rifampicin)	rpoB	Rv0667	RNA polymerase	Target modification	[6]
		ponA1	Rv0050	Probable bifunctional penicillin-binding protein	Unknown	[7]
	Isoniazid	katG	Rv1908c	Catalase-peroxidase enzyme	Decreased drug activation	[8]
		inhA	Rv1484	NADH-dependent enoyl-acyl carrier protein	Target amplification or modification	[9, 10]
	Pyrazinamide[b]	pncA	Rv2043c	Pyrazinamidase	Decreased drug activation	[11, 12]
		panD	Rv3601c	Aspartate decarboxylase	Unknown	[13]
		rpsA	RRv1630	Ribosomal protein S1	Target modification	[14]
	Ethambutol[b]	embCAB operon	Rv3793-5	Arabinosyltransferase	Target modification	[15, 16]
		ubiA	Rv3806c	Arabinogalactan synthesis	Gain-of-function	[15]
Group A	Levofloxacin Moxifloxacin	gyrA	Rv0006	DNA gyrase A	Target modification	[17, 18]
		gyrB	Rv0005	DNA gyrase B	Target modification	[18]
	Bedaquiline	atpE	Rv1305	ATP synthase	Target modification	[19]
		pepQ	Rv2535c	Putative Xaa-Pro aminopeptidase	Unknown	[20]
		Rv0678	Rv0678	Transcriptional regulator of mmpL5	Drug efflux	[21, 22]
	Linezolid	Rrl	NA	23S rRNA	Target modification	[23]
		rplC	Rv0701	50S ribosomal protein L3	Target modification	[24]
Group B	Clofazimine	pepQ	Rv2535c	Putative Xaa-Pro aminopeptidase	Drug efflux	[20]
		Rv0678	Rv0678	Transcriptional regulator of mmpL5	Drug efflux	[21]
	Cycloserine Terizidone	Ald	Rv2780	L-alanine dehydrogenase	Substrate shunting	[25]
		alr	Rv3423c	Alanine racemase	Target modification	[26, 27]
		ddl	Rv2981c	D-alanine-D-alanine ligase	Target modification	[27]
		cycA	Rv1704c	Bacterial D-serine/L-and D-alanine/glycine/D-cycloserine proton symporter	Mechanism not confirmed	[28]
Group C	Delamanid Pretomanid	ddn	Rv3547	Oxidative stress	Decreased drug activation	[29]
		fgd1	Rv0407	Glucose-6-phosphate oxidation	Decreased drug activation	[29]
	Imipenem/cilastatin	crfA	Rv2421c-Rv2422 intergenic	Unknown	Drug inactivation	[30]
	Amikacin, Capreomycin, Kanamycin[c]	Rrs	NA	16S rRNA	Target modification	[31]
	Streptomycin	rpsL	Rv0682	12S ribosomal protein	Target modification	[32–35]
		rrs	NA	16S rRNA	Target modification	[36]
		gidB	Rv3919c	7-Methylguanosine methyltransferase	Target modification	[37]
	Ethionamide Prothionamide	ethA	Rv3854c	Mono-oxygenase	Decreased drug activation	[38, 39]
		ethR	Rv3855	Transcriptional regulatory repressor protein (TetR)	Decreased drug activation	[39]
		inhA	Rv1484	NADH-dependent enoyl-acyl carrier protein	Target amplification or modification	[10]
	Para-aminosalicylic acid (PAS)	folC	Rv2447c	Folate pathway	Decreased drug activation	[40]
		dfrA	Rv2763c	Dihydrofolate reductase	Target amplification	[40]

Cohen *et al. Genome Medicine*   (2019) 11:45

Page 4 of 18

Table 1 Antitubercular drug-resistance mechanisms[a] *(Continued)*

WHO category	Drug or drug class	Resistance genes	Rv number	Gene function	Mechanism of drug resistance	Reference(s)
		thyA	Rv2764c	Thymidylate synthase	Target modification	[41, 42]
		thyX	Rv2754c	Catalyzes dTMP and tetrahydrofolate	Mitigating target inhibition	[43]
		ribD	Rv2671	Enzyme in riboflavin biosynthesis	Mitigating target inhibition	[40, 44]
Other medicines[c]	Kanamycin	*Eis*	Rv2416c	Aminoglycoside acetyltransferase	Inactivating mutation	[45]
	Capreomycin	*tlyA*	Rv1694	rRNA methyltransferase	Target modification	[46]

Abbreviations: MDR-TB multidrug-resistant tuberculosis, *NA* not applicable, *RR-TB* rifampicin-resistant tuberculosis, *WHO* World Health Organization
[a]Antitubercular drugs are listed by the 2018 WHO grouping of medicines recommended for use in longer, individualized MDR-TB regimens [47]. For each drug or drug class, the specific genes in which drug-resistance mutations are commonly identified are listed with their gene name, gene number (Rv number), gene function, and the confirmed or putative mechanisms of resistance. [b]Pyrazinamide and ethambutol are first-line TB drugs that also are categorized as Group C medicines for the treatment of longer MDR-TB regimens. [c]Kanamycin and capreomycin are no longer recommended to be included in longer, individualized MDR/RR-TB regimens

containing media, and selecting for mutants that are able to grow in the presence of the drug. Sequencing laboratory-derived resistant mutants has played a critical role in identifying both the mechanism of action of new TB drug classes, including diarylquinolines (for example, bedaquiline) [19] and nitroimidazopyrans (for example, PA-824) [19, 29], and rare resistance mechanisms for established antitubercular drugs, including ethambutol [15], pyrazinamide [13], carbapenems [30], cycloserine [87], clofazimine, and bedaquiline [20]. For example, WGS of laboratory mutants identified drug efflux as a mechanism of resistance to clofazimine and bedaquiline [20–22], and this approach continues to be a mainstay for identifying the mechanism of action of compounds that are in development for TB [88].

Although laboratory-derived mutants are helpful in elucidating novel resistance mechanisms, mutations that have evolved in laboratory settings do not always match those in clinical isolates of drug-resistant *M. tuberculosis* [89, 90], for reasons that are largely unknown. Studies by Ford et al. [91, 92] suggested that these mismatches could not be explained by differences in the mutation rate in these settings, because the in vitro mutation rate of *M. tuberculosis* correlates well with the in vivo mutation rate in both humans and in a macaque model. Differences in the relative fitness of specific mutants grown in in vitro compared to in vivo conditions are probably responsible for these mismatches, but more work is needed. Regardless of the reason, if the goal is to identify a full complement of resistance mutations on which to base molecular diagnostics, isolates from clinical collections must be studied as these bacteria have evolved their resistance within the host.

Quantifying and identifying resistance determinants in clinical strains

Among the larger studies exploring resistance in natural populations, Walker et al. [58] analyzed the genomes of 3651 drug-resistant and -susceptible *M. tuberculosis* isolates for associations between phenotypic resistance to eight first- and second-line drugs, and then predicted genotypic resistance on the basis of a compiled catalog of 232 resistance mutations in 23 candidate resistance genes. Resistance to most drugs could be predicted accurately, with a mean sensitivity of 92% and specificity of 98%, suggesting that the majority of resistance—particularly for first-line drugs—is explained by known mechanisms and mutations (Table 1). Numerous other studies have found similar results using smaller datasets [7, 25, 57, 69, 93, 94]. This result was echoed in a more recent study by the Comprehensive Resistance Prediction for Tuberculosis (CRYPTIC) Consortium and the 100,000 Genomes Project that focused solely on first-line drugs, which included analysis of 10,209 globally diverse *M. tuberculosis* isolate genomes against a database of mutations identified in a literature search [60]. Notably, predictions for mutations that are associated with resistance to pyrazinamide were greatly improved over earlier predictions; this study achieved 91.3% sensitivity in predicting resistance to this drug compared to 57% sensitivity in Walker et al. [58]. Although the news has been encouraging with respect to completing the catalog of mutations that cause resistance to first-line drugs, few studies have attempted to predict resistance to second-line drugs [95]. Some of these drugs, such as D-cycloserine, pyrazinamide, and para-aminosalicylic acid (PAS), are more difficult to assay because they have been reported to have variable drug phenotypes in clinical microbiology laboratories [96] (discussed later).

To fill gaps in the catalog of drug-resistance mechanisms, genome-wide association study (GWAS) approaches, originally designed for use on human genomic data, have been adapted for non-recombining microbes such as *M. tuberculosis* and used to predict novel resistance mechanisms [97, 98] (Table 3). The majority of GWAS predictions remain experimentally unverified,

Cohen *et al. Genome Medicine*   (2019) 11:45

Page 5 of 18

Table 2 Spotlight on whole-genome sequencing studies of drug-resistant *M. tuberculosis*

Reference	Description	Advances
Identifying M. tuberculosis drug-resistance determinants		
Farhat et al. 2013 [7]	Large-scale WGS project: sequencing of 116 genomes from around the globe	Developed a phylogenetic convergence test, PhyC, to identify resistance associations; validated *ponA1* mutations that increase MIC for rifampicin
Zhang et al. 2013 [57]	Large-scale WGS project: sequencing of 161 genomes from China	Identified genes that are under positive selection and have increased mutation frequencies in drug-resistant isolates
Walker et al. 2015 [58]	Analysis of 23 candidate resistance genes from 3651 clinical isolates	Demonstrated that drug-resistance in *M. tuberculosis* can be predicted with high sensitivity and specificity
Desjardins et al. 2016 [25]	Use of a combination of the correlated evolution test and a GWAS framework to identify drug-resistance-associated mutations in 498 genomes from China and South Africa	Identified *ald* loss-of-function as a novel mechanism of D-cycloserine resistance
Coll et al. 2018 [59]	GWAS study of 6465 *M. tuberculosis* clinical isolates from more than 30 countries	Identified new resistance-associated mutations in *ethA* and the *thyX* promoter
The Cryptic Consortium and the 100,000 Genomes Project [60]	Prediction of first-line-drug susceptibility in a dataset of 10, 209 clinical isolates from 16 countries	Predicted drug-susceptibility phenotypes with high sensitivity and specificity using WGS in a large global dataset
Within-patient evolution of resistance		
Eldholm et al. 2014 [61]	WGS of nine serial isolates cultured from a single patient over a 42-month period	First documented case of the evolution of susceptible TB into XDR-TB in a single patient in response to selective drug pressure
Trauner et al. 2017 [62]	Very deep WGS of serial sputum specimens from patients receiving treatment for TB	Demonstrated that the combination of multiple active drugs prevented fixing and dominance of transient mutants. The fewer drugs used, the more likely it was that resistance would develop and become fixed
Transmission versus de novo evolution of resistance		
Nikolayevskyy et al. 2016 [63]	Literature review including meta-analysis of 12 studies published between 2005 and 2014	Showed that WGS studies have higher discriminatory power than fingerprinting techniques and can more sensitively detect transmission events
Ioerger et al. 2010 [64]	WGS of 14 phenotypically diverse strains within the Beijing lineage in South Africa	Showed that resistance mutations arose independently multiple times, and that XDR-TB isolates may be less fit and less able to transmit
Shah et al. 2017 [65]	Sequencing of more than 400 strains from South Africa	The majority of cases of XDR-TB in KwaZulu-Natal were due to transmission rather than de novo evolution
Manson et al. 2017 [66]	WGS of a set of 5310 isolates, with diverse geographical origin, genetic background, and drug-resistance profiles	Demonstrated that both de novo evolution and transmission contribute to drug-resistance worldwide
Geographic spread of multidrug-resistance		
Cohen et al. 2019 [67]	Further analysis of geographic trends in MDR strains within the set of 5310 strains from Manson et al. [66]	Revealed extensive worldwide spread of MDR-TB clades between countries of varying TB burden
Nelson et al. 2018 [68]	Sequencing of 344 patients with XDR-TB, combined with global positioning system coordinates	Identified many cases of probable person-to-person transmission (≤ 5 SNPs) between people living a median of 108 km apart, suggesting that drivers of XDR-TB transmission include migration between urban and rural areas
Order of acquisition of resistance mutations		
Cohen et al. 2015 [69]	WGS and drug-susceptibility testing on 337 clinical isolates collected in Kwazulu-Natal, South Africa	Showed that stepwise accumulation of mutations leading to XDR-TB in Kwazulu-Natal occurred over decades. Established the order of acquisition of drug-resistance mutations leading to XDR-TB, showing that isoniazid resistance almost always evolved prior to rifampicin resistance
Eldholm et al. 2015 [70]	WGS of all 252 available clinical isolates from an outbreak in Argentina	Showed stepwise accumulation of mutations leading to the development of MDR-TB in Argentina
Manson et al. 2017 [66]	WGS of 5310 isolates with diverse geographical origin, genetic background, and drug-resistance profiles	Established that a clear order of acquisition of resistance mutations holds globally: isoniazid resistance overwhelmingly evolves prior to rifampicin resistance across all geographies, lineages, and all time periods (including decades after rifampicin introduction)
Evolution of compensatory and stepping-stone mutations		
Fonseca et al. 2015 [71]	Review paper	Discussed the evolution of compensatory mutations that can ease fitness effects caused by resistance
Comas et al. 2012 [72]	Comparison of the genome sequences of ten clinical rifampicin-resistant isolates to those of the corresponding rifampicin-susceptible isolates from the same individual at an earlier timepoint	Identified compensatory mutations in *rpoB* that conferred high competitive fitness in vitro and were also found frequently in clinical populations
Casali et al. 2014 [73]	Large-scale analysis of 1000 strains from Russia	Examined strains with primary rifampicin-resistance mutations in *rpoB*, and identified accompanying compensatory mutations in *rpoA* and *rpoC*
Cohen et al. 2015 [69]	WGS and drug-susceptibility testing of 337 clinical isolates collected in Kwazulu-Natal, South Africa	Identified putative rifampicin compensatory mutations in *rpoA*, *rpoB*, and *rpoC*
Merker et al. 2018 [74]	Sequencing of highly resistant TB strains from Central Asia	Showed that the presence of rifampicin compensatory mutations are associated with transmission success and

Cohen *et al. Genome Medicine*   (2019) 11:45

Page 6 of 18

Table 2 Spotlight on whole-genome sequencing studies of drug-resistant *M. tuberculosis* (Continued)

Reference	Description	Advances
		higher drug-resistance rates
Coll et al. 2018 [59]	GWAS study of 6465 *M. tuberculosis* clinical isolates from more than 30 countries	Identified putative compensatory mutations for pyrazinamide and PAS resistance
Safi et al. 2018 [15]	Genetically and biochemically characterized strains selected in vitro for ethambutol resistance	Showed that multi-step selection is required to achieve the highest levels of ethambutol resistance
Understanding mixed infections and spatial heterogeneity within a patient		
Köser et al. 2013 [75]	WGS for rapid drug-susceptibility testing of a patient with XDR-TB	Determined that the patient carried two different XDR-TB Beijing strains with differing resistance mutations
Liu et al. 2015 [76]	Deep WGS of serial sputum isolates within a patient	Identified three dominant subclones differing by 10–14 SNPs within a single patient, with different resistance patterns and probably different anatomical distributions
Lieberman et al. 2016 [77]	Sequencing of samples from post-mortem biopsies from different body sites	Observed sublineages evolving within a patient, as well as distinct strains from mixed infections that were differentially distributed across body sites
Dheda et al. 2018 [78]	Sequencing of samples biopsied from seven different body sites, as well as pre-treatment and serial sputum samples	Showed that drug concentrations at different sites were inversely correlated with bacterial MICs. Sequencing and comparison to sputum samples suggested ongoing acquired resistance
Sobkowiak et al. 2018 [79]	Assessed methods for detecting mixed infections using WGS data from in vitro and in silico artificially mixed *M. tuberculosis* samples	Frequency of mixed infections in the Karonga Study in Mali is approximately 10% and only associated with year of diagnosis, not with age, sex, HIV or prior TB infection. Computational methods can identify mixed infections using WGS data
Bench to bedside with WGS		
Pankhurst et al. 2016 [80]	Prospective study evaluating the use of WGS for diagnosis	Compared WGS of positive liquid cultures to routine laboratory workflows. Illumina MiSeq-based bioinformatics classification of species and drug resistance was faster (by a median of 21 days) and cheaper (by 7%), yet offered similar accuracy to routine techniques
Doughty et al. 2014 [81]	Sequencing-based detection without culturing	Proof-of-concept culture-free metagenomics detection of *M. tuberculosis* from sputum samples using Illumina MiSeq
Votintseva et al. 2017 [82]	Evaluation of Oxford Nanopore sequencing for diagnostic or surveillance purposes	Proof-of-concept detection of *M. tuberculosis* DNA in sputum samples using a portable sequencer

Abbreviations: *GWAS* genome-wide association study, *MDR* multidrug-resistant, *MIC* minimum inhibitory concentration, *PAS* para-aminosalicylic acid, *SNP* single nucleotide polymorphism, *TB* tuberculosis, *XDR* extensively drug-resistant

but several new resistance-associated genotypes have been validated. Farhat et al. [7] sequenced 116 *M. tuberculosis* genomes and developed a phylogenetic convergence test, 'PhyC', to identify resistance associations. They identified a mutation in *ponA1* (c.1095G>T) and showed that it conferred a minimum inhibitory concentration (MIC) to rifampicin that was twofold higher than that of wild-type bacteria. Zhang et al. [57] sequenced 161 genomes from China and searched for genes that appeared to be under positive selection and more frequently mutated in drug-resistant isolates. Resistance-associated polymorphisms in two intergenic regions upstream of the known resistance genes *thyA*-Rv2765 and *thyX*-*hsdS*.1 were shown to cause increased gene expression of a *lacZ* construct in *Mycobacterium smegmatis*, suggesting that these mutations may mediate PAS resistance through the overexpression of downstream genes.

Desjardins et al. [25] used a combination of the correlated evolution test [104] (to test for correlated evolution of genotype and phenotype) and a simple GWAS framework to search for novel drug-resistance mechanisms in 498 genomes from South Africa and China. Of note, they combined all variants within each gene that were predicted to inactivate gene function, and used these combinations as the input into the association test to

increase statistical power in the detection of genomic features that are associated with resistance. They found that loss-of-function mutations in *ald* (Rv2780), which is predicted to encode an alanine dehydrogenase, correlated with unexplained resistance [25]. They also confirmed experimentally that these mutations conferred increased resistance of laboratory and clinical isolates to D-cycloserine [25], a key drug in MDR- and XDR-TB regimens that has severe psychiatric and central nervous system toxicities.

Hicks et al. [105] used the algorithm phyOverlap to perform a GWAS on 549 clinical isolates from China, in which they identified mutations that disproportionately occurred in isoniazid-resistant isolates. In addition to known resistance and compensatory mutations for first- and second-line drugs, they identified an association with *prpR* (Rv1129c). They then went on to characterize *prpR* as a transcriptional regulator of propionate metabolism which, instead of drug resistance, confers tolerance to multiple antibiotics in a macrophage model of infection.

In one of the largest GWAS published to date, Coll et al. [59] combined PhyC with a GWAS approach within a mixed-regression framework to detect determinants of resistance to 14 drugs in a large dataset of 6465

Cohen *et al. Genome Medicine* (2019) 11:45

Page 7 of 18

Table 3 Publicly available software packages implementing microbial GWAS methods for identifying drug-resistance-associated genetic variants in bacteria

Method	Details of approach	Key recent studies and advances achieved in identifying drug-resistance-associated genetic variants	Availability	Reference(s)
bugwas	Uses linear mixed models with a correction for population stratification. Uses SNPs identified through mapping to a reference	Applied to identify resistance to 17 drugs across 3144 isolates from four diverse species of bacteria, including *M. tuberculosis* [99]. Confirmed that some major known resistance determinants could be recovered. The method was recently extended in a kmer-based method based on bugwas [100]	https://github.com/sgearle/bugwas	[99, 100]
SEER	Uses logistic and linear regression with a correction for population stratification. Uses SNPs identified through mapping to a reference	Initially applied to *Streptococcus*. To date, has not been applied to *M. tuberculosis*	https://github.com/johnlees/seer/wiki	[101]
treeWAS	Uses a phylogenetic test to identify convergent evolution using kmers, which can detect both individual variants and gene presence or absence agnostic of a reference	Initially applied to *Neisseria meningitidis*. Has not yet been applied to *M. tuberculosis*	https://github.com/caitiecollins/treeWAS	[102, 103]
phyC	Uses phylogenetic tests to identify convergent evolution, using SNPs identified through mapping to a reference	Identified 39 genomic regions that are potentially involved in resistance, and confirmed a rifampicin-conferring mutation in *ponA1* [7]. Used within a mixed-regression framework to detect resistance determinants to 14 drugs in a dataset of 6465 global clinical isolates. Identified new ethionamide-resistance codons in *ethA* and PAS-resistance mutations in the thyX promoter [59]	https://bitbucket.org/rpetit3/visa-gwas	[7, 59, 102]

Abbreviation: GWAS genome-wide association study, SNP single nucleotide polymorphism

global *M. tuberculosis* clinical isolates. Although no functional experiments were performed to validate the predictions, new resistance-associated mutations were identified, including new codons in *ethA* (a gene known to activate ethionamide, which is a prodrug) that are associated with ethionamide resistance, and mutations in the *thyX* promoter associated with PAS resistance. Mutations in the promoter of *thyX* have been previously shown to upregulate *thyX* [43, 57, 106].

Predicting susceptibility and drug resistance in M. tuberculosis

As the list of suspected resistance determinants grows, there has been a need to establish well-curated databases of drug-resistance single nucleotide polymorphisms (SNPs) [107]. Initially, SNP databases, including TBDB [108] and PATRIC [109], were created to bring together genome annotation data and other functional data. Unfortunately, some of the pioneering databases of drug-resistance-associated mutations in *M. tuberculosis*, including TBDReamDB [110], have not been maintained to include emerging data.

Software and web-based tools have also been developed to enable the community to infer TB drug resistance from WGS data. These tools include CASTB [111], KVarQ [112], MyKrobe Predictor TB [113], PhyResSE [114], TBProfiler [115], and TGS-TB [116]. Studies have compared the sensitivity and specificity of these tools in predicting drug resistance [117–119], and have found that they tend to perform quite well for first-line drugs

but less optimally for second-line drugs. In addition to tools, there have been improvements to databases, including the creation of the Relational Sequencing TB Database Platform (ReSeqTB) [120, 121] and efforts from the CRyPTIC Consortium [122], which seeks to develop a curated database of clinically relevant drug-resistance mutations.

Continued refinement of these drug-resistance databases and prediction tools is necessary. Miotto et al. [123] performed a systematic review in which they assigned a confidence level to associations of individual and groups of mutations with phenotypic drug resistance. Importantly, they identified that certain mutations that are included in current commercial diagnostic tests, including *eis* c-2a, do not have a convincing association with drug resistance. Input from ongoing large sequencing projects will be needed to optimize the inference of resistance phenotypes from sequence data, especially for mutations that are present at low frequency in natural populations.

Challenges in uncovering the remaining resistance elements

Although WGS approaches have been successful in identifying resistance mechanisms, there are computational and experimental challenges that hamper efforts to complete the catalog of TB drug resistance. For example, for non-recombining organisms such as *M. tuberculosis*, interpretation of GWAS output can be complicated because non-causal variation can be tightly

linked to causal variation [124]. Furthermore, as a result of frequent multidrug resistance, resistance mutations for one drug can appear to be highly associated with phenotypic resistance to multiple drugs [25], and confirmatory wet lab studies, which are non-trivial in *M. tuberculosis*, are often necessary to identify causal resistance mutations correctly. In addition, genotype–phenotype associations are largely dependent on accurate phylogenies, and phylogenetic reconstruction can be challenging in *M. tuberculosis* because of its slow rate of evolution [92, 125–128], which gives rise to relatively few SNPs in clinical isolates.

When defining phenotypic resistance, different studies often use different drug concentration cutoffs and test in different media, complicating the meta-analysis of multiple datasets. In addition, phenotypic resistance testing of some antitubercular drugs, including pyrazinamide and D-cycloserine, is notoriously challenging and unreliable [129], introducing phenotypic inaccuracies that can confound analyses. Furthermore, the dichotomous classification of phenotypic resistance as 'resistant' or 'susceptible' will fail to identify drug-resistance mutations that result only in minimal increases in MIC, and there is emerging evidence that such mutations may be clinically relevant. TB relapse following treatment has been found to occur more commonly in individuals who harbored *M. tuberculosis* isolates that were susceptible to, yet had minimally increased MIC values for, either isoniazid or rifampicin [130]. Future study designs that address phenotypic resistance as a spectrum, rather than a binary value, will be needed to identify such mutations.

Heteroresistance, defined as the coexistence of pathogen populations that have differing nucleotides at a specific drug-resistance locus [131], can also confound genotype–phenotype comparisons [132–134]. A bacterial culture in which only a small fraction of the population is resistant can appear to be resistant when tested on media containing a drug, yet when grown on drug-free media for genome sequencing, the sensitive fraction can dominate, resulting in a genotypic prediction of sensitivity [132]. The problem of heteroresistance seems to be particularly common with fluoroquinolone resistance [135].

Last, innate characteristics of the *M. tuberculosis* genome—namely, highly repetitive DNA sequences and the high guanine-cytosine (GC) content of the genome (65.6%) [53]—present technical difficulties for both WGS and bioinformatic analyses. GC-rich regions can be troublesome for library PCR amplification and sequencing, and reads that represent highly repetitive regions of the genome can confound alignments by mapping to multiple regions of the genome and hampering accurate de novo assemblies. In addition, approximately 10% of the coding regions in *M. tuberculosis* are

dedicated to two repetitive protein families that are unique to mycobacteria (the PE and PPE families), which have conserved Pro-Glu (PE) and Pro-Pro-Glu (PPE) motifs [53]. Even with WGS investigation [136], the function of the PE and PPE genes has remained elusive, although recent studies have suggested that they may play a role in virulence [137]. Their association with drug resistance remains largely unexplored because bioinformatic studies of *M. tuberculosis* often exclude these genes [138, 139]. In the future, long-read sequencing technology may allow these regions to be sequenced successfully in order to assess if they have a role in drug resistance.

Understanding the evolution and spread of drug resistance in *M. tuberculosis*

Prior to WGS, the diversity and epidemiology of resistant *M. tuberculosis* were examined using DNA fingerprinting techniques, like IS6110 restriction fragment length polymorphism (RFLP) typing [140], spoligotyping (spacer oligonucleotide typing, a method of typing strains according to the distinct hybridization patterns of their spacer oligonucleotides) [141], and mycobacterial interspersed repetitive units-variable number of tandem repeats (MIRU-VNTR) typing [142–145]. These techniques enabled assessments of the diversity of resistant strains in specific geographic regions [146–149] and, when combined with the genetic profiling of resistance mutations, allowed strain-level monitoring of patients on TB therapy [150].

The dramatic increase in resolution afforded by WGS has extended the sensitivity and resolution with which the diversity and evolution of drug-resistant *M. tuberculosis* can be assessed. This has resulted in the more confident identification of cases of recent transmission [151] and re-infection [152], and has provided insights into the evolution of resistance within individual patients and across populations. WGS has also enabled more sensitive differentiation of de novo acquisition of resistance (where resistance mutations emerge within a host) from person-to-person transmission of resistance, a critical capability given that these two scenarios require different health-system responses in order to stem resistance.

Within patient evolution of drug resistance

Despite the slow evolutionary rate of *M. tuberculosis*, estimated at 0.3–0.6 SNPs/genome/year [69, 92, 125–128], experimental data suggest that drug resistance can evolve within an individual patient during TB treatment. Eldholm et al. [61] described the first documented case of XDR evolution of *M. tuberculosis* from a fully susceptible ancestor within a single patient, by sequencing nine serial isolates collected over a 42-month period. During

this time, seven known resistance mutations emerged in a stepwise fashion after the clinical use of each corresponding drug, revealing how TB drug pressures can rapidly shape *M. tuberculosis* populations in vivo.

However, the evolution of drug resistance within a host is not always linear, and instead can involve a complex interplay of heterogeneous *M. tuberculosis* populations [153, 154]. In particular, transient genetic diversity can exist before a dominant clone emerges. In addition, as the size of the transmission bottleneck (the number of bacteria transmitted during an infection event) in *M. tuberculosis* is not well understood [155], it is difficult to estimate the relative contribution of diversity that is transmitted to the patient versus diversity that evolves within the patient. Numerous WGS studies, performed either on isolates or directly on DNA extracted from serially collected sputum samples, have revealed substantial transient genetic diversity in pathogen populations within patients, particularly within resistance genes [61, 62, 106, 156–159]. This diversity was observed to endure months before a single variant became fixed in the population (the situation when only a single variant remains). In the study by Eldholm et al. [61] mentioned above, the seven resistance-conferring mutations that eventually dominated were from amongst 35 mutations observed in total throughout the sampling period [61, 160]. They joined eight other mutations that were not resistance-associated but that also became fixed in the population, probably as the result of a phenomenon called 'hitchhiking' in which non-adaptive mutations are selected because of their linkage and physical proximity to consequential mutations.

The relative fitness cost of drug-resistance mutations often determines which mutations become fixed within a host. While multiple mutations that confer resistance to a specific drug can evolve repeatedly, mutations conferring no or little fitness cost are typically selected, resulting in fixed dominant mutations [61, 156]. Compensatory mutations (discussed in more detail later), which serve to counterbalance the deleterious effects of acquired resistance, have also been shown to emerge during treatment [156].

WGS has also revealed how combination chemotherapy effectively prevents the emergence of drug resistance during treatment for TB. In a study of very deep WGS of serial sputum specimens from patients receiving treatment for TB, Trauner et al. [62] demonstrated that the combined action of multiple active drugs prevented transient mutants from fixing within a population and becoming dominant. The fewer the drugs that were applied, the more likely it was that resistance would develop and become fixed.

Population views of drug-resistance evolution

A number of careful WGS studies have empirically established SNP-based criteria to discriminate cases of recent transmission from unrelated infections—usually using the criterion that recently transmitted strains differ by fewer than 6–12 total SNPs across the *M. tuberculosis* genome [63, 125, 126, 161]. In a 2016 review, Nikolayevskyy and colleagues [63] systematically compared WGS to fingerprinting techniques for detecting transmission, including a meta-analysis of 12 studies published between 2005 and 2014. They concluded that results from WGS studies not only have higher discriminatory power, but they also enable more sensitive detection of transmission events that may have been missed by epidemiologic methods.

Although traditional spoligotyping analyses suggested that drug-resistant strains were diverse, WGS of clinical isolates began to reveal the full breadth of diversity in resistant *M. tuberculosis*. The TB epidemic in South Africa over the past two decades has been well-studied in this regard. In an early WGS investigation, Ioerger et al. [64] examined 14 phenotypically diverse strains within the Beijing lineage and showed that resistance mutations arose independently multiple times, and that XDR isolates may be less fit and less able to transmit. WGS studies across larger sets of strains from the same region in South Africa suggested that, although de novo resistance is indeed common, highly resistant strains (including MDR and XDR strains) have the ability to spread broadly by person-to-person transmission. This includes the ongoing transmission of a circulating XDR clone in South Africa that is linked to the infamous Tugela Ferry XDR outbreak [162] that brought XDR-TB to the world stage in 2005. A more recent large-scale study confirmed that XDR strains have been broadly transmitted person-to-person in KwaZulu-Natal [65].

The patterns observed in South Africa hold for many other parts of the world. Recent studies have shown that patterns of both de novo evolution and person-to-person spread of drug resistance in *M. tuberculosis* also occur in Belarus, Russia, England, and Malawi [73, 139, 159, 163, 164]. In a composite analysis of over 5000 *M. tuberculosis* isolates from patients from around the globe, Manson et al. [66] confirmed that both de novo evolution and person-to-person transmission are important factors for the rise and spread of drug-resistant TB worldwide. The emergence of MDR and XDR *M. tuberculosis* was found to be a frequent occurrence that is distributed fairly evenly across the globe [66]. This analysis also predicted that 37% of MDR isolates in this study had spread person-to-person, which is probably a vast underestimate of how frequently MDR is transmitted once evolved [66].

Geographic movement of people is also an important consideration with regard to person-to-person transmission. Further examination of the MDR clades from Manson et al. [66] revealed that they included widespread

Cohen *et al. Genome Medicine* (2019) 11:45

Page 10 of 18

international, and even intercontinental, dissemination of strains that were separated by as few as four SNPs, probably due to spread via international travel [67]. Even within a single province in South Africa, Nelson et al. [68] showed, using genomic sequence data and global positioning system coordinates, that many cases of person-to-person transmission (with ≤ 5 SNPs) of XDR-TB occur between people living a median of 108 km apart, pointing to migration between urban and rural areas as a driver of TB spread. Collectively, these studies reinforce the idea that the geographic movement of people must be taken into consideration in any strategy for controlling the spread of TB resistance.

Ordering of the acquisition of resistance and compensatory mutations

Recent WGS studies have helped to illuminate the steps or 'fitness landscape' through which *M. tuberculosis* develops and compensates for drug resistance. Several studies [66, 69, 70] have shown that the order of acquisition of drug-resistance mutations in complex resistance cases is partly constrained in clinical *M. tuberculosis*. For example, in MDR-TB, isoniazid resistance (most often involving a *katG* S315T mutation) overwhelmingly evolves prior to resistance to rifampicin and second-line drugs. This was first shown using regional datasets from South Africa [69] and Argentina [70], and recently confirmed by Manson et al. [66] using a global dataset of 5310 strains. In the study by Manson et al. [66], this ordering was shown to hold true over 95% of the time, even for distinct global regions and time frames, including times when both rifampicin and isoniazid were in use, suggesting that the earlier introduction of isoniazid in the 1950s was not the major contributor to this effect. It was also shown that *inhA* promoter mutations that confer isoniazid resistance (such as those observed by Perdigão et al. [165] in Portugal) were acquired earlier than rifampicin mutations, although the number of samples harboring these mutations was much smaller. Further studies are necessary to determine whether isoniazid preventive monotherapy, which is one of the treatments for latent tuberculosis, may account for some of this effect, as this could result in a background level of increased isoniazid monoresistance.

Compensatory mutations that potentially ease fitness effects caused by resistance often occur after the evolution of primary resistance. This phenomenon was reviewed by Fonseca et al. [71], and examples include mutations in the *ahpC* promoter region and the *rpoC*/*rpoA* genes, which act as compensatory mutations for isoniazid and rifampicin resistance, respectively. Newer WGS work has pointed to several novel compensatory mutations in *M. tuberculosis*, particularly for rifampicin resistance. Comas et al. [72] identified a set of compensatory mutations in the *rpoB* gene that conferred high competitive fitness in vitro and were also found frequently in clinical populations. In a large-scale analysis of 1000 strains from Russia, Casali et al. [73] examined strains with primary resistance mutations in *rpoB* and identified accompanying compensatory mutations in *rpoA* and *rpoC*. Cohen et al. [69] identified putative rifampicin compensatory mutations that are present in South African strains by searching for *rpoA*, *rpoB*, and *rpoC* mutations that evolved only after or concurrent with rifampicin resistance-conferring mutations. A recent study of highly resistant *M. tuberculosis* strains from Central Asia confirmed that the presence of compensatory mutations, particularly those compensating for the fitness cost of mutations that confer rifampicin resistance, is associated with transmission success and higher drug-resistance rates [74]. Beyond rifampicin resistance compensation, Coll et al. [59] identified mutations in *pncB2* that may compensate for pyrazinamide resistance conferred by *pncA*, and similarly, mutations in *thyX-hsdS.1* (the *thyX* promoter) that may compensate for PAS resistance conferred by *thyA*; however, experimental validation of these potential compensatory relationships is needed. Even fewer studies have identified stepping-stone mutations in *M. tuberculosis*, which emerge prior to higher-level resistance mutations. Cohen et al. [69] found that *ubiA* mutations emerge in a stepping-stone fashion prior to more classic *embB* mutations that confer ethambutol resistance. Safi et al. [15] also showed in vitro that multi-step selection involving *ubiA*, *aftA*, *embB*, and *embC* is required to achieve the highest levels of ethambutol resistance.

The challenge of mixed infections

Although WGS approaches have great sensitivity in detecting cases of recent transmission, reconstructing the details of transmission networks with WGS [166–168] can be difficult. Transmission network mapping is highly dependent on sampling density and studies rarely, if ever, comprehensively sample an outbreak or the extent of within-host diversity. It is also becoming clear, from the prevalence of very close relationships between isolates from patients who have no other direct epidemiological connections, that transmission may largely result from casual contact in community settings [169]. In addition, the phylogenetic reconstruction of transmission networks can be especially challenging, particularly because of the very close relationships between strains and the slow rate of evolution of *M. tuberculosis* [92, 125–128].

Mixed infections represent a major challenge for understanding drug-resistance evolution within patients [153, 158, 159]. It can be straightforward to disambiguate co-infections of strains from different lineages, but

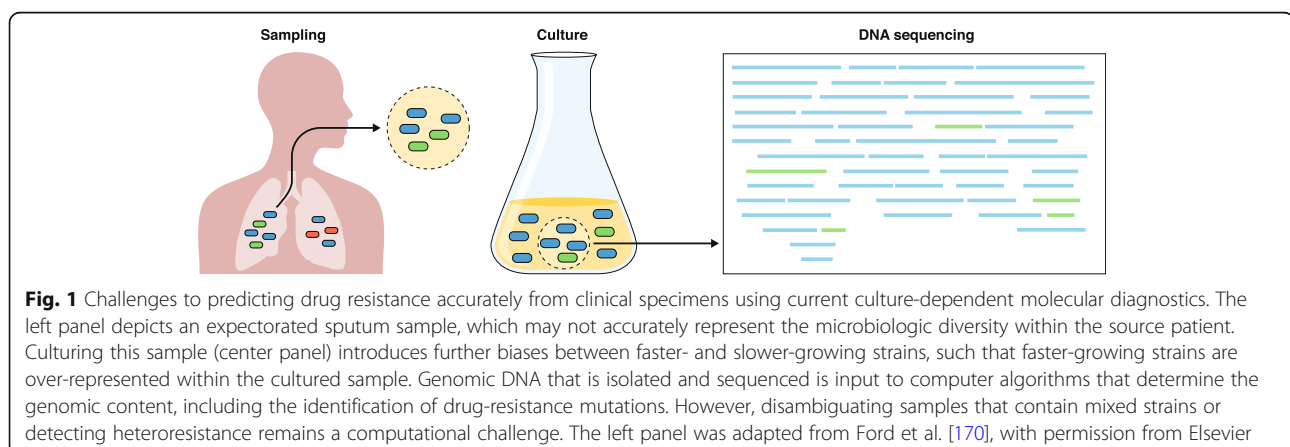Cohen *et al. Genome Medicine* (2019) 11:45

Page 11 of 18

mixed infections involving strains that have few genetic differences can also occur, making these strains difficult to distinguish. Köser et al. [75] used WGS for rapid drug-susceptibility testing of a patient with XDR-TB, and determined that the patient carried two different XDR-TB Beijing strains with differing resistance mutations. In a study by Liu et al. [76], three dominant subclones differing by 10–14 SNPs were detected within a single patient, each with different resistance patterns and probably different anatomical distributions. Also, co-infection by strains with differing resistance patterns may yield misleading composite views of resistance; for example, co-infection with two MDR-TB strains—one with quinolone resistance and the other with aminoglycoside resistance—may be mistaken for infection with a single XDR-TB strain.

Furthermore, newer data suggest that there can be genetic heterogeneity among *M. tuberculosis* isolates from different parts of the body, potentially leading to incomplete views of drug resistance within a patient (Fig. 1). In a study by Lieberman et al. [77], the authors observed evidence for both within-host evolution and mixed infection by piecing together the genetic variation observed among *M. tuberculosis* isolates from multiple post-mortem biopsies from the same patient. Another recent study by Dheda et al. [78] showed that drug concentrations at seven body sites were inversely correlated with the MIC of the bacteria isolated from these sites. Sequencing and comparison to pre-treatment and serial sputum samples suggested ongoing acquired resistance and differential evolution across sites [78]. These findings underscore the limitations of diagnosing or studying the evolution of drug-resistant *M. tuberculosis* using a single patient specimen. However, they also show the promise of WGS for informing interventions related to drug delivery, dosing, and diagnostics, thereby helping to prevent the development of acquired resistance within a patient. More research in this area is needed to determine the breadth and scope of mixed infections among patients with active TB, their contribution to changing drug-resistance patterns over time, and the role of spatial heterogeneity in the evolution of drug resistance.

From bench to bedside: promise and challenges

Given that the failure to identify and treat patients who have drug-resistant TB leads to increased mortality, spread of resistant strains, and gain of additional drug resistance [171], there is a critical need to diagnose resistant *M. tuberculosis* in patients rapidly. Several important molecular diagnostic platforms have been established for the identification of *M. tuberculosis* and drug resistance within this organism, but they are limited to the identification of a defined subset of resistance mutations [172], do not always include the earliest-arising mutations that precede MDR [66], and do not provide knowledge that is useful in determining whether a patient has been re-infected, whether the patient has a recurrent or mixed infection, or whether a particular infection represents a transmission event. WGS holds significant potential to modernize the TB laboratory and improve upon TB management [173], and this topic has been reviewed previously [173, 174]. To date, WGS has been primarily applied as a clinical tool to achieve two goals: first, to detect *M. tuberculosis* within a clinical sample, and second, to detect resistance mutations and predict resistance patterns so that appropriate treatment can be provided. In order to provide clinically useful information, a diagnostic platform must be rapid. Historically, WGS has relied upon an input of pure mycobacterial cultures, which is time-consuming (requiring multiple weeks) and therefore of less clinical utility. Several investigations have attempted to address this issue by using earlier culture inputs or by attempting culture-independent, direct sequencing from clinical specimens [80, 82, 175]. In a rapid, yet still culture-dependent method, Pankhurst et al. [80] prospectively compared real-time WGS of "early positive liquid cultures" to routine *M. tuberculosis* diagnostics, and found

Fig. 1 Challenges to predicting drug resistance accurately from clinical specimens using current culture-dependent molecular diagnostics. The left panel depicts an expectorated sputum sample, which may not accurately represent the microbiologic diversity within the source patient. Culturing this sample (center panel) introduces further biases between faster- and slower-growing strains, such that faster-growing strains are over-represented within the cultured sample. Genomic DNA that is isolated and sequenced is input to computer algorithms that determine the genomic content, including the identification of drug-resistance mutations. However, disambiguating samples that contain mixed strains or detecting heteroresistance remains a computational challenge. The left panel was adapted from Ford et al. [170], with permission from Elsevier

Cohen *et al. Genome Medicine*   (2019) 11:45

Page 12 of 18

that WGS achieved a faster time to diagnosis at a lower cost.

Although the advances achieved using WGS are promising, several hurdles must be overcome before it can be put into practical use in the clinic (Fig. 1). Requirements for costly equipment, technical expertise, and substantial computational resources present challenges to implementation [173]. Direct sequencing of patient samples has revealed that the vast majority of DNA present is from the patient or from non-mycobacterial prokaryotes, with variable quantities of mycobacterial DNA present. Doughty et al. [81] performed a pilot study demonstrating the feasibility of direct sequencing using a benchtop sequencer (Illumina MiSeq, San Diego, CA) and sputum samples from eight patients. Although they were able to identify the presence of *M. tuberculosis*, the low depth of sequencing coverage of the genome (0.002 to 0.7x) prevented drug-susceptibility prediction. Separately, Brown et al. [176] performed an enrichment step with biotinylated RNA baits prior to direct sequencing of sputum, resulting in higher quality data (> 20x depth and > 90% coverage) that allowed the identification of resistance mutations.

Using a targeted DNA enrichment strategy to study 43 individuals with active pulmonary TB, Doyle et al. [177] compared WGS directly from sputum with mycobacterial growth indicator tube (MGIT) WGS. Although direct sputum sequencing was able to identify drug resistance much more rapidly than MGIT WGS, only 74% of sputum samples yielded interpretable WGS data (vs 100% from MGIT); thus, additional optimization of these methods is needed to increase the sensitivity of this approach. Similarly, in a recent study, the use of pyrosequencing of a concentrated sputum sediment (rather than from sputum directly), dramatically shortened the time to initiation of an MDR-treatment regimen [178].

One promising technology that could change clinical WGS is long-read sequencing using the Oxford Nanopore Technologies (ONT; Oxford, UK) platform. An advantage of ONT is the ability to allow sequencing to continue until sufficient coverage of the genome has been obtained, potentially solving the problem of low or variable amounts of *M. tuberculosis* in clinical samples [82]. Early ONT studies have shown promise in identifying antimicrobial-resistance genes in different bacterial species [179]. Unfortunately, at present, both the high error rate of ONT MinION and potential difficulties with GC-rich regions limit the utility of this technology; thus, improvements in accuracy are necessary to enable the identification of resistance associated with point mutations [179]. ONT metagenomic sequencing has been successfully applied to improve pathogen detection and antimicrobial-resistance testing in other clinical settings [180]; however, to date, applications of this technology

to *M. tuberculosis* have been limited to pre-clinical research [82].

Despite these challenges, WGS offers several advantages over the technologies that are currently employed for diagnosis and epidemiological monitoring of TB. Using WGS directly on patient sputum could reduce the turnaround time for diagnosis and determination of antibiotic-resistance status from weeks to hours [61, 159], and would prevent the introduction of culture-induced biases. The depth of information provided by WGS could be used to identify whether an individual harbors multiple co-infecting strains [106, 160] and to distinguish recurrent infection as either relapse or re-infection [174, 181]. In addition, WGS could provide real-time epidemiological information that could be useful for understanding patterns of drug resistance and for establishing chains of transmission [174]. Encouragingly, the high levels of concordance observed between the genotypes and phenotypes of clinical samples indicate that WGS can provide high accuracy for both diagnosing TB and informing treatment options [113]. Finally, WGS of patient samples would provide a high level of convenience, by combining diagnosis, resistance profiling, and epidemiological analysis into a single test [85]. Given these advantages, the WHO has recently published a technical guide for the implementation of next-generation sequencing (NGS) technologies for the detection of drug resistance in *M. tuberculosis* [182].

Routine whole-genome sequencing of mycobacterial isolates

In 2017, England became the first nation to launch routine WGS of all prospectively identified *M. tuberculosis* clinical isolates [183]. Sponsored by Public Health England (PHE), prospective WGS is being performed on all positive mycobacterial cultures. Within 5–7 days of receipt of the culture from the reference lab, data will be provided on the mycobacterial species, the predicted drug susceptibility, and the molecular epidemiology of the strains. If, from the sequence analysis, a strain is predicted to be fully susceptible to first-line antitubercular drugs, phenotypic drug-susceptibility testing (DST) will no longer be performed routinely. However, if drug resistance to any first-line drug is identified, then phenotypic DST will follow. Beyond drug-susceptibility prediction, these efforts will have profound implications for TB control because WGS data can be used for real-time molecular epidemiology in this context.

Given the high sensitivity of WGS in detecting drug resistance to first-line TB drugs [60], similar algorithms utilizing WGS to predict susceptibility (rather than to identify drug resistance) for first-line drugs, in lieu of phenotypic DST, have been endorsed in the Netherlands and in New York [60]. It seems highly likely that these

kinds of efforts would be helpful in higher-burden TB settings than those mentioned here, but the feasibility of this approach has not yet been established, from either a practical or economic standpoint, in settings where the numbers of drug-resistant TB cases are high.

Conclusions and future directions

Since the first applications of WGS to *M. tuberculosis* in 1998, WGS techniques have greatly accelerated our understanding of drug-resistance mechanisms in this pathogen. Importantly, WGS studies now indicate that, for many drugs, the vast majority of resistance is explained by known mutations. The increasing availability of whole-genome sequences from phenotypically diverse *M. tuberculosis*, combined with improved GWAS algorithms, is enabling the discovery of the remaining determinants of unexplained resistance. In addition, WGS has provided valuable insight into how resistance mutations evolve and spread. It is clear that both de novo acquisition of resistance mutations and clonal transmission are critical factors in the spread of drug-resistant TB.

Furthermore, WGS investigations have revealed that there is a specific order in which drug-resistance mutations are acquired: isoniazid resistance is almost always acquired before rifampicin resistance, which has significant implications for the design of diagnostic tests. Within individual patients, WGS studies have highlighted that mixed infections are common, and often represent important intermediates in the evolution of drug resistance.

WGS also holds great promise for revolutionizing the rapid clinical diagnosis of TB in the future. Although there are still substantial technical hurdles, WGS can be used to diagnose the presence of *M. tuberculosis* rapidly, as well as to pinpoint appropriate antibiotic treatment regimens by identifying the complement of *M. tuberculosis* drug-resistance mutations that are present within a clinical sample. Indeed, improvements in the prediction of drug susceptibility with WGS may obviate the need for phenotypic culture methods, especially for first-line drugs.

Although WGS offers many benefits, targeted NGS, in which sequence data are obtained from only a focused panel of genes or genetic regions rather than from the entire genome, is gaining momentum [184]. One of the advantages of targeted NGS over WGS is that it can be performed directly on clinical specimens and is, therefore, faster than culture-based WGS. Other advantages include reduction in both labor and computational efforts and reduced costs. The potential offered by the application of targeted NGS to the prediction of drug resistance from genomic data is self-evident. Nevertheless, it seems that WGS would have greater discriminatory power than targeted NGS for molecular epidemiology purposes.

Ultimately, the use of WGS is expected to continue to advance our understanding of *M. tuberculosis* drug resistance. Furthermore, its practical use in clinical settings holds great potential to improve public health through real-time molecular epidemiology tracking, to identify global hotspots of drug-resistance emergence, and to facilitate the development of improved approaches for the diagnosis and treatment of drug-resistant TB.

Author details
[1]Division of Pulmonary and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, MA 21205, USA. [2]Broad Institute of Harvard and Massachusetts Institute of Technology, 415 Main Street, Cambridge, MA 02142, USA. [3]Delft Bioinformatics Lab, Delft University of Technology, 2628 XE Delft, The Netherlands.

References
1. World Health Organization. Global Tuberculosis Report 2018. Geneva: WHO; 2018. http://apps.who.int/iris/bitstream/handle/10665/274453/9789241565646-eng.pdf?ua=1. Accessed 9 July 2019
2. Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment for pulmonary tuberculosis. Br Med J. 1948;2:769–82 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2091872/pdf/brmedj03701-0007.pdf. Accessed 9 July 2019.
3. World Health Organization. Global Tuberculosis Report 2016. Geneva: WHO; 2016. https://apps.who.int/iris/bitstream/handle/10665/250441/9789241565394-eng.pdf?sequence=1. Accessed 9 July 2019

4. World Health Organization. Global Tuberculosis Report 2017. Geneva: WHO; 2017. https://www.who.int/tb/publications/global_report/gtbr2017_main_text.pdf. Accessed 9 July 2019

5. World Health Organization. WHO treatment guidelines for drug-resistant tuberculosis. 2016 update. Geneva: WHO; 2016. https://apps.who.int/iris/bitstream/handle/10665/250125/9789241549639-eng.pdf?sequence=1. Accessed 9 July 2019

6. Telenti A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, et al. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. Lancet. 1993;341:647–50.

7. Farhat MR, Shapiro BJ, Kieser KJ, Sultana R, Jacobson KR, Victor TC, et al. Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*. Nat Genet. 2013;45:1183–9.

8. Heym B, Alzari PM, Honoré N, Cole ST. Missense mutations in the catalase-peroxidase gene, katG, are associated with isoniazid resistance in *Mycobacterium tuberculosis*. Mol Microbiol. 1995;15:235–45.

9. Rozwarski DA, Grant GA, Barton DH, Jacobs WR, Sacchettini JC. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. Science. 1998;279:98–102.

10. Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS, Wilson T, et al. inhA, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. Science. 1994;263:227–30.

11. Sreevatsan S, Pan X, Zhang Y, Kreiswirth BN, Musser JM. Mutations associated with pyrazinamide resistance in pncA of *Mycobacterium tuberculosis* complex organisms. Antimicrob Agents Chemother. 1997;41:636–40.

12. Scorpio A, Lindholm-Levy P, Heifets L, Gilman R, Siddiqi S, Cynamon M, et al. Characterization of pncA mutations in pyrazinamide-resistant *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 1997;41:540–3.

13. Zhang S, Chen J, Shi W, Liu W, Zhang W, Zhang Y. Mutations in panD encoding aspartate decarboxylase are associated with pyrazinamide resistance in *Mycobacterium tuberculosis*. Emerg Microbes Infect. 2013;2:e34.

14. Shi W, Zhang X, Jiang X, Yuan H, Lee JS, Barry CE, et al. Pyrazinamide inhibits trans-translation in *Mycobacterium tuberculosis*. Science. 2011;333: 1630–2.

15. Safi H, Lingaraju S, Amin A, Kim S, Jones M, Holmes M, et al. Evolution of high-level ethambutol-resistant tuberculosis through interacting mutations in decaprenylphosphoryl-β-D-arabinose biosynthetic and utilization pathway genes. Nat Genet. 2013;45:1190–7.

16. Alcaide F, Pfyffer GE, Telenti A. Role of embB in natural and acquired resistance to ethambutol in mycobacteria. Antimicrob Agents Chemother. 1997;41:2270–3.

17. Maruri F, Sterling TR, Kaiga AW, Blackman A, van der Heijden YF, Mayer C, et al. A systematic review of gyrase mutations associated with fluoroquinolone-resistant *Mycobacterium tuberculosis* and a proposed gyrase numbering system. J Antimicrob Chemother. 2012;67:819–31.

18. Takiff HE, Salazar L, Guerrero C, Philipp W, Huang WM, Kreiswirth B, et al. Cloning and nucleotide sequence of *Mycobacterium tuberculosis* gyrA and gyrB genes and detection of quinolone resistance mutations. Antimicrob Agents Chemother. 1994;38:773–80.

19. Andries K, Verhasselt P, Guillemont J, Göhlmann HWH, Neefs J-M, Winkler H, et al. A diarylquinoline drug active on the ATP synthase of *Mycobacterium tuberculosis*. Science. 2005;307:223–7.

20. Almeida D, Ioerger T, Tyagi S, Li SY, Mdluli K, Andries K, et al. Mutations in pepQ confer low-level resistance to bedaquiline and clofazimine in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2016;60:4590–9.

21. Hartkoorn RC, Uplekar S, Cole ST. Cross-resistance between clofazimine and bedaquiline through upregulation of MmpL5 in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2014;58:2979–81.

22. Andries K, Villellas C, Coeck N, Thys K, Gevers T, Vranckx L, et al. Acquired resistance of *Mycobacterium tuberculosis* to bedaquiline. PLoS One. 2014. https://doi.org/10.1371/journal.pone.0102135.

23. Hillemann D, Rüsch-Gerdes S, Richter E. In vitro-selected linezolid-resistant *Mycobacterium tuberculosis* mutants. Antimicrob Agents Chemother. 2008; 52:800–1.

24. Beckert P, Hillemann D, Kohl TA, Kalinowski J, Richter E, Niemann S, et al. rplC T460C identified as a dominant mutation in linezolid-resistant *Mycobacterium tuberculosis* strains. Antimicrob Agents Chemother. 2012;56: 2743–5.

25. Desjardins CA, Cohen KA, Munsamy V, Abeel T, Maharaj K, Walker BJ, et al. Genomic and functional analyses of *Mycobacterium tuberculosis* strains implicate ald in D-cycloserine resistance. Nat Genet. 2016;48:544–51.

26. Cáceres NE, Harris NB, Wellehan JF, Feng Z, Kapur V, Barletta RG. Overexpression of the D-alanine racemase gene confers resistance to D-cycloserine in *Mycobacterium smegmatis*. J Bacteriol. 1997;179:5046–55.

27. Feng Z, Barletta RG. Roles of *Mycobacterium smegmatis* D-alanine:D-alanine ligase and D-alanine racemase in the mechanisms of action of and resistance to the peptidoglycan inhibitor D-cycloserine. Antimicrob Agents Chemother. 2003;47:283–91.

28. Chen JM, Uplekar S, Gordon SV, Cole ST. A point mutation in cycA partially contributes to the D-cycloserine resistance trait of *Mycobacterium bovis* BCG vaccine strains. PLoS One. 2012. https://doi.org/10.1371/journal.pone.0043467.

29. Haver HL, Chua A, Ghode P, Lakshminarayana SB, Singhal A, Mathema B, et al. Mutations in genes for the F420 biosynthetic pathway and a nitroreductase enzyme are the primary resistance determinants in spontaneous in vitro-selected PA-824-resistant mutants of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2015;59:5316–23.

30. Kumar P, Kaushik A, Bell D, Chauhan V, Xia F, Stevens RL, et al. Mutation in an unannotated protein confers carbapenem resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2017. https://doi.org/10.1128/AAC.02234-16.

31. Suzuki Y, Katsukawa C, Tamaru A, Abe C, Makino M, Mizuguchi Y, et al. Detection of kanamycin-resistant *Mycobacterium tuberculosis* by identifying mutations in the 16S rRNA gene. J Clin Microbiol. 1998;36:1220–5.

32. Finken M, Kirschner P, Meier A, Wrede A, Böttger EC. Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. Mol Microbiol. 1993;9:1239–46.

33. Nair J, Rouse DA, Bai GH, Morris SL. The rpsL gene and streptomycin resistance in single and multiple drug-resistant strains of *Mycobacterium tuberculosis*. Mol Microbiol. 1993;10:521–7.

34. Honoré N, Cole ST. Streptomycin resistance in mycobacteria. Antimicrob Agents Chemother. 1994;38:238–42.

35. Meier A, Kirschner P, Bange FC, Vogel U, Böttger EC. Genetic alterations in streptomycin-resistant *Mycobacterium tuberculosis*: mapping of mutations conferring resistance. Antimicrob Agents Chemother. 1994;38:228–33.

36. Douglass J, Steyn LM. A ribosomal gene mutation in streptomycin-resistant *Mycobacterium tuberculosis* isolates. J Infect Dis. 1993;167:1505–6.

37. Wong SY, Lee JS, Kwak HK, Via LE, Boshoff HIM, Barry CE. Mutations in gidB confer low-level streptomycin resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2011;55:2515–22.

38. Morlock GP, Metchock B, Sikes D, Crawford JT, Cooksey RC. ethA, inhA, and katG loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates. Antimicrob Agents Chemother. 2003;47:3799–805.

39. Baulard AR, Betts JC, Engohang-Ndong J, Quan S, McAdam RA, Brennan PJ, et al. Activation of the pro-drug ethionamide is regulated in mycobacteria. J Biol Chem. 2000;275:28326–31.

40. Zheng J, Rubin EJ, Bifani P, Mathys V, Lim V, Au M, et al. Para-aminosalicylic acid is a prodrug targeting dihydrofolate reductase in *Mycobacterium tuberculosis*. J Biol Chem. 2013;288:23447–56.

41. Mathys V, Wintjens R, Lefevre P, Bertout J, Singhal A, Kiass M, et al. Molecular genetics of para-aminosalicylic acid resistance in clinical isolates and spontaneous mutants of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2009;53:2100–9.

42. Rengarajan J, Sassetti CM, Naroditskaya V, Sloutsky A, Bloom BR, Rubin EJ. The folate pathway is a target for resistance to the drug para-aminosalicylic acid (PAS) in mycobacteria. Mol Microbiol. 2004;53:275–82.

43. Fivian-Hughes AS, Houghton J, Davis EO, Hill M, Nw L. *Mycobacterium tuberculosis* thymidylate synthase gene thyX is essential and potentially bifunctional, while thyA deletion confers resistance to p-aminosalicylic acid. Microbiology. 2012;158:308–18.

44. Zhang X, Liu L, Zhang Y, Dai G, Huang H, Jin Q. Genetic determinants involved in p-aminosalicylic acid resistance in clinical isolates from tuberculosis patients in northern China from 2006 to 2012. Antimicrob Agents Chemother. 2015;59:1320–4.

45. Zaunbrecher MA, Sikes RD, Metchock B, Shinnick TM, Posey JE. Overexpression of the chromosomally encoded aminoglycoside acetyltransferase eis confers kanamycin resistance in *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A. 2009;106:20004–9.

46. Maus CE, Plikaytis BB, Shinnick TM. Mutation of tlyA confers capreomycin resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2005;49:571–7.

Cohen *et al. Genome Medicine*   (2019) 11:45

Page 15 of 18

47. World Health Organization. Rapid Communication: key changes to treatment of multidrug- and rifampicin-resistant tuberculosis (MDR/RR-TB). Geneva: WHO; 2018. https://www.who.int/tb/publications/2018/WHO_RapidCommunicationMDRTB.pdf?ua=1. Accessed 9 July 2019.

48. Diacon AH, Donald PR, Pym A, Grobusch M, Patientia RF, Mahanyele R, et al. Randomized pilot trial of eight weeks of bedaquiline (TMC207) treatment for multidrug-resistant tuberculosis: long-term outcome, tolerability, and effect on emergence of drug resistance. Antimicrob Agents Chemother. 2012;56:3271–6.

49. Diacon AH, Pym A, Grobusch MP, de Los Rios JM, Gotuzzo E, Vasilyeva I, et al. Multidrug-resistant tuberculosis and culture conversion with bedaquiline. N Engl J Med. 2014;371:723–32.

50. Matsumoto M, Hashizume H, Tomishige T, Kawasaki M, Tsubouchi H, Sasaki H, et al. OPC-67683, a nitro-dihydro-imidazooxazole derivative with promising action against tuberculosis in vitro and in mice. PLoS Med. 2006;3:2131–44.

51. Gler MT, Skripconoka V, Sanchez-Garavito E, Xiao H, Cabrera-Rivero JL, Vargas-Vasquez DE, et al. Delamanid for multidrug-resistant pulmonary tuberculosis. N Engl J Med. 2012;366:2151–60.

52. Bloom BR, Atun R, Cohen T, Dye C, Fraser H, Gomez GB, et al. Tuberculosis. In: Holmes KK, Bertozzi S, Bloom BR, et al., editors. Major infectious diseases. Washington DC: The International Bank for Reconstruction and Development/The World Bank; 2017. https://www.ncbi.nlm.nih.gov/books/NBK525174/. Accessed 9 July 2019.

53. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature. 1998;393:537–44.

54. Fleischmann RD, Alland D, Eisen JA, Carpenter L, White O, Peterson J, et al. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. J Bacteriol. 2002;184:5479–90.

55. Young DB. Blueprint for the white plague. Nature. 1998;393:515–6.

56. Cabibbe AM, Walker TM, Niemann S, Cirillo DM. Whole genome sequencing of *Mycobacterium tuberculosis*. Eur Respir J. 2018;52:537–45.

57. Zhang H, Li D, Zhao L, Fleming J, Lin N, Wang T, et al. Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance. Nat Genet. 2013;45:1255–60.

58. Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo EC, Bradley P, et al. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. Lancet Infect Dis. 2015;3099:1–10.

59. Coll F, Phelan J, Hill-Cawthorne GA, Nair MB, Mallard K, Ali S, et al. Genome-wide analysis of multi- and extensively drug-resistant *Mycobacterium tuberculosis*. Nat Genet. 2018;50:307–16.

60. The CRyPTIC Consortium and the 100,000 Genomes Project. Prediction of susceptibility to first-line tuberculosis drugs by DNA sequencing. N Engl J Med. 2018;379:1403–15.

61. Eldholm V, Norheim G, von der Lippe B, Kinander W, Dahle UR, Caugant DA, et al. Evolution of extensively drug-resistant *Mycobacterium tuberculosis* from a susceptible ancestor in a single patient. Genome Biol. 2014. https://doi.org/10.1186/s13059-014-0490-3.

62. Trauner A, Liu Q, Via LE, Liu X, Ruan X, Liang L, et al. The within-host population dynamics of *Mycobacterium tuberculosis* vary with treatment efficacy. Genome Biol. 2017. https://doi.org/10.1186/s13059-017-1196-0.

63. Nikolayevskyy V, Kranzer K, Niemann S, Drobniewski F. Whole genome sequencing of *Mycobacterium tuberculosis* for detection of recent transmission and tracing outbreaks: a systematic review. Tuberculosis. (Edinb.). 2016;98:77–85.

64. Ioerger TR, Feng Y, Chen X, Dobos KM, Victor TC, Streicher EM, et al. The non-clonality of drug resistance in Beijing-genotype isolates of *Mycobacterium tuberculosis* from the Western Cape of South Africa. BMC Genomics. 2010. https://doi.org/10.1186/1471-2164-11-670.

65. Shah NS, Auld SC, Brust JCM, Mathema B, Ismail N, Moodley P, et al. Transmission of extensively drug-resistant tuberculosis in South Africa. N Engl J Med. 2017;376:243–53.

66. Manson AL, Cohen KA, Abeel T, Desjardins CA, Armstrong DT, Barry CE, et al. Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multidrug resistance. Nat Genet. 2017;49:395–402.

67. Cohen KA, Manson AL, Abeel T, Desjardins CA, Chapman SB, Hoffner S, et al. Extensive global movement of multidrug-resistant *M. tuberculosis* strains revealed by whole-genome analysis. Thorax. 2019. https://doi.org/10.1136/thoraxjnl-2018-211616.

68. Nelson KN, Shah NS, Mathema B, Ismail N, Brust JCM, Brown TS, et al. Spatial patterns of extensively drug-resistant tuberculosis transmission in KwaZulu-Natal. South Africa. J Infect Dis. 2018;218:1964–73.

69. Cohen KA, Abeel T, Manson McGuire A, Desjardins CA, Munsamy V, Shea TP, et al. Evolution of extensively drug-resistant tuberculosis over four decades: whole genome sequencing and dating analysis of *Mycobacterium tuberculosis* isolates from KwaZulu-Natal. PLoS Med. 2015. https://doi.org/10.1371/journal.pmed.1001880.

70. Eldholm V, Monteserin J, Rieux A, Lopez B, Sobkowiak B, Ritacco V, et al. Four decades of transmission of a multidrug-resistant *Mycobacterium tuberculosis* outbreak strain. Nat Commun. 2015. https://doi.org/10.1038/ncomms8119.

71. Fonseca JD, Knight GM, McHugh TD. The complex evolution of antibiotic resistance in *Mycobacterium tuberculosis*. Int J Infect Dis. 2015;32:94–100.

72. Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. Nat Genet. 2012;44:106–10.

73. Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. Nat Genet. 2014;46:279–86.

74. Merker M, Barbier M, Cox H, Rasigade J-P, Feuerriegel S, Kohl TA, et al. Compensatory evolution drives multidrug-resistant tuberculosis in Central Asia. Elife. 2018. https://doi.org/10.7554/eLife.38200.

75. Köser CU, Bryant JM, Becq J, Török ME, Ellington MJ, Marti-Renom MA, et al. Whole-genome sequencing for rapid susceptibility testing of *M. tuberculosis*. N Engl J Med. 2013;369:290–2.

76. Liu Q, Via LE, Luo T, Liang L, Liu X, Wu S, et al. Within patient microevolution of *Mycobacterium tuberculosis* correlates with heterogeneous responses to treatment. Sci Rep. 2015. https://doi.org/10.1038/srep17507.

77. Lieberman TD, Wilson D, Misra R, Xiong LL, Moodley P, Cohen T, et al. Genomic diversity in autopsy samples reveals within-host dissemination of HIV-associated *Mycobacterium tuberculosis*. Nat Med. 2016;22:1470–4.

78. Dheda K, Lenders L, Magombedze G, Srivastava S, Raj P, Arning E, et al. Drug-penetration gradients associated with acquired drug resistance in patients with tuberculosis. Am J Respir Crit Care Med. 2018;198:1208–19.

79. Sobkowiak B, Glynn JR, Houben RMGJ, Mallard K, Phelan JE, Guerra-Assunção JA, et al. Identifying mixed *Mycobacterium tuberculosis* infections from whole genome sequence data. BMC Genomics. 2018. https://doi.org/10.1186/s12864-018-4988-z.

80. Pankhurst LJ, Del Ojo EC, Votintseva AA, Walker TM, Cole K, Davies J, et al. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. Lancet Respir Med. 2016;4:49–58.

81. 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. doi:https://doi.org/10.7717/peerj.585.

82. Votintseva AA, Bradley P, Pankhurst L, Del Ojo EC, Loose M, Nilgiriwala K, et al. Same-day diagnostic and surveillance data for tuberculosis via whole-genome sequencing of direct respiratory samples. J Clin Microbiol. 2017;55:1285–98.

83. Liu X, Gutacker MM, Musser JM, Fu Y-X. Evidence for recombination in *Mycobacterium tuberculosis*. J Bacteriol. 2006;188:8169–77.

84. Ramaswamy S, Musser JM. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuber Lung Dis. 1998; 79:3–29.

85. Simons SO, van Ingen J, van der Laan T, Mulder A, Dekhuijzen PNR, Boeree MJ, et al. Validation of pncA gene sequencing in combination with the mycobacterial growth indicator tube method to test susceptibility of *Mycobacterium tuberculosis* to pyrazinamide. J Clin Microbiol. 2012;50:428–34.

86. Helb D, Jones M, Story E, Boehme C, Wallace E, Ho K, et al. Rapid detection of *Mycobacterium tuberculosis* and rifampin resistance by use of on-demand, near-patient technology. J Clin Microbiol. 2010;48:229–37.

87. Chen J, Zhang S, Cui P, Shi W, Zhang W, Zhang Y. Identification of novel mutations associated with cycloserine resistance in *Mycobacterium tuberculosis*. J Antimicrob Chemother. 2017;72:3272–6.

88. Zampieri M, Szappanos B, Buchieri MV, Trauner A, Piazza I, Picotti P, et al. High-throughput metabolomic analysis predicts mode of action of uncharacterized antimicrobial compounds. Sci Transl Med. 2018. https://doi.org/10.1126/scitranslmed.aal3973.

89. Gagneux S. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. Science. 2006;312:1944–6.

90. Bergval IL, Schuitema ARJ, Klatser PR, Anthony RM. Resistant mutants of *Mycobacterium tuberculosis* selected in vitro do not reflect the in vivo mechanism of isoniazid resistance. J Antimicrob Chemother. 2009;64:515–23.

91. Ford CB, Shah RR, Maeda MK, Gagneux S, Murray MB, Cohen T, et al. *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. Nat Genet. 2013;45:784–90.

92. Ford CB, Lin PL, Chase MR, Shah RR, Iartchouk O, Galagan J, et al. Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. Nat Genet. 2011;43:482–6.

93. 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;23:61–8.

94. Gygli SM, Keller PM, Ballif M, Blöchliger N, Hömke R, Reinhard M, et al. Whole genome sequencing for drug resistance profile prediction in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2019. https://doi.org/10.1128/AAC.02175-18.

95. Cohen KA, El-Hay T, Wyres KL, Weissbrod O, Munsamy V, Yanover C, et al. Paradoxical hypersusceptibility of drug-resistant *Mycobacterium tuberculosis* to β-lactam antibiotics. EBioMedicine. 2016;9:170–9.

96. World Health Organization. Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis. Geneva: WHO; 2014. https://apps.who.int/iris/bitstream/handle/10665/130918/9789241548809_eng.pdf?sequence=1. Accessed 9 July 2019

97. Power RA, Parkhill J, de Oliveira T. Microbial genome-wide association studies: lessons from human GWAS. Nat Rev Genet. 2017;18:41–50.

98. Farhat MR, Freschi L, Calderon R, Ioerger T, Snyder M, Meehan CJ, et al. Genome wide association with quantitative resistance phenotypes in *Mycobacterium tuberculosis* reveals novel resistance genes and regulatory regions. Nat Commun. 2019. https://doi.org/10.1038/s41467-019-10110-6.

99. Earle SG, Wu C-H, Charlesworth J, Stoesser N, Gordon NC, Walker TM, et al. Identifying lineage effects when controlling for population structure improves power in bacterial association studies. Nat Microbiol. 2016. https://doi.org/10.1038/nmicrobiol.2016.41.

100. Jaillard M, Lima L, Tournoud M, Mahé P, van Belkum A, Lacroix V, et al. A fast and agnostic method for bacterial genome-wide association studies: bridging the gap between k-mers and genetic events. PLoS Genet. 2018. https://doi.org/10.1371/journal.pgen.1007758.

101. Lees JA, Vehkala M, Välimäki N, Harris SR, Chewapreecha C, Croucher NJ, et al. Sequence element enrichment analysis to determine the genetic basis of bacterial phenotypes. Nat Commun. 2016. https://doi.org/10.1038/ncomms12797.

102. Alam MT, Petit RA, Crispell EK, Thornton TA, Conneely KN, Jiang Y, et al. Dissecting vancomycin-intermediate resistance in staphylococcus aureus using genome-wide association. Genome Biol Evol. 2014;6:1174–85.

103. Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLoS Comput Biol. 2018. https://doi.org/10.1371/journal.pcbi.1005958.

104. Barker D, Meade A, Pagel M. Constrained models of evolution lead to improved prediction of functional linkage from correlated gain and loss of genes. Bioinformatics. 2007;23:14–20.

105. Hicks ND, Yang J, Zhang X, Zhao B, Grad YH, Liu L, et al. Clinically prevalent mutations in *Mycobacterium tuberculosis* alter propionate metabolism and mediate multidrug tolerance. Nat Microbiol. 2018;3:1032–42.

106. Merker M, Kohl TA, Roetzer A, Truebe L, Richter E, Rüsch-Gerdes S, et al. Whole genome sequencing reveals complex evolution patterns of multidrug-resistant *Mycobacterium tuberculosis* Beijing strains in patients. PLoS One. 2013. https://doi.org/10.1371/journal.pone.0082551.

107. Stucki D, Gagneux S. Single nucleotide polymorphisms in *Mycobacterium tuberculosis* and the need for a curated database. Tuberculosis (Edinb). 2013;93:30–9.

108. Reddy TBK, Riley R, Wymore F, Montgomery P, DeCaprio D, Engels R, et al. TB database: an integrated platform for tuberculosis research. Nucleic Acids Res. 2009;37:D499–508.

109. Gillespie JJ, Wattam AR, Cammer SA, Gabbard JL, Shukla MP, Dalay O, et al. PATRIC: the comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. Infect Immun. 2011;79:4286–98.

110. Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Tuberculosis drug resistance mutation database. PLoS Med. 2009. https://doi.org/10.1371/journal.pmed.1000002.

111. Iwai H, Kato-Miyazawa M, Kirikae T, Miyoshi-Akiyama T. CASTB (the comprehensive analysis server for the *Mycobacterium tuberculosis* complex): a publicly accessible web server for epidemiological analyses, drug-resistance prediction and phylogenetic comparison of clinical isolates. Tuberculosis (Edinb). 2015;95:843–4.

112. Steiner A, Stucki D, Coscolla M, Borrell S, Gagneux S. KvarQ. Targeted and direct variant calling from fastq reads of bacterial genomes. BMC Genomics. 2014. doi:https://doi.org/10.1186/1471-2164-15-881.

113. Bradley P, Gordon NC, Walker TM, Dunn L, Heys S, Huang B, et al. Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. Nat Commun. 2015. https://doi.org/10.1038/ncomms10063.

114. Feuerriegel S, Schleusener V, Beckert P, Kohl TA, Miotto P, Cirillo DM, et al. PhyResSE: a web tool delineating *Mycobacterium tuberculosis* antibiotic resistance and lineage from whole-genome sequencing data. J Clin Microbiol. 2015;53:1908–14.

115. Coll F, McNerney R, Preston MD, Guerra-Assunção JA, Warry A, Hill-Cawthorne G, et al. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. Genome Med. 2015. https://doi.org/10.1186/s13073-015-0164-0.

116. Sekizuka T, Yamashita A, Murase Y, Iwamoto T, Mitarai S, Kato S, et al. TGS-TB: total genotyping solution for *Mycobacterium tuberculosis* using short-read whole-genome sequencing. PLoS One. 2015. https://doi.org/10.1371/journal.pone.0142951.

117. Schleusener V, Köser CU, Beckert P, Niemann S, Feuerriegel S. *Mycobacterium tuberculosis* resistance prediction and lineage classification from genome sequencing: comparison of automated analysis tools. Sci Rep. 2017. https://doi.org/10.1038/srep46327.

118. van Beek J, Haanperä M, Smit PW, Mentula S, Soini H. Evaluation of whole genome sequencing and software tools for drug susceptibility testing of *Mycobacterium tuberculosis*. Clin Microbiol Infect. 2019;25:82–6.

119. Ngo T-M, Teo Y-Y. Genomic prediction of tuberculosis drug-resistance: benchmarking existing databases and prediction algorithms. BMC Bioinformatics. 2019. https://doi.org/10.1186/s12859-019-2658-z.

120. ReSeqTB. Relational sequencing TB data platform. https://platform.reseqtb.org. Accessed 9 July 2019.

121. Starks AM, Avilés E, Cirillo DM, Denkinger CM, Dolinger DL, Emerson C, et al. Collaborative effort for a centralized worldwide tuberculosis relational sequencing data platform. Clin Infect Dis. 2015;61:S141–6.

122. CRyPTIC. Comprehensive Resistance Prediction for Tuberculosis: an International Consortium. http://www.crypticproject.org. Accessed 9 July 2019.

123. Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, et al. A standardised method for interpreting the association between mutations and phenotypic drug resistance in *Mycobacterium tuberculosis*. Eur Respir J. 2017. https://doi.org/10.1183/13993003.01354-2017.

124. Falush D, Bowden R. Genome-wide association mapping in bacteria? Trends Microbiol. 2006;14:353–5.

125. Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. Lancet Infect Dis. 2013;13:137–46.

126. Guerra-Assunção J, Crampin A, Houben R, Mzembe T, Mallard K, Coll F, et al. Large scale population-based whole genome sequencing of *Mycobacterium tuberculosis* provides insights into transmission in a high prevalence area. Elife. 2015. https://doi.org/10.7554/eLife.05166.

127. Bryant JM, Schürch AC, van Deutekom H, Harris SR, de Beer JL, de Jager V, et al. Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. BMC Infect Dis. 2013. https://doi.org/10.1186/1471-2334-13-110.

128. Roetzer A, Diel R, Kohl TA, Rückert C, Nübel U, Blom J, et al. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. PLoS Med. 2013. https://doi.org/10.1371/journal.pmed.1001387.

129. World Health Organization. Policy guidance on drug-susceptibility testing (DST) of second-line antituberculosis drugs. Geneva: WHO; 2008. https://apps.who.int/iris/bitstream/handle/10665/70500/WHO_HTM_TB_2008.392_eng.pdf?sequence=1. Accessed 9 July 2019

130. Colangeli R, Jedrey H, Kim S, Connell R, Ma S, Chippada Venkata UD, et al. Bacterial factors that predict relapse after tuberculosis therapy. N Engl J Med. 2018;379:823–33.

131. Rinder H. Hetero-resistance: an under-recognised confounder in diagnosis and therapy? J Med Microbiol. 2001;50:1018–20.

132. Rinder H, Mieskes KT, Löscher T. Heteroresistance in *Mycobacterium tuberculosis*. Int J Tuberc Lung Dis. 2001;5:339–45.

133. Folkvardsen DB, Thomsen VØ, Rigouts L, Rasmussen EM, Bang D, Bernaerts G, et al. Rifampin heteroresistance in *Mycobacterium tuberculosis* cultures as detected by phenotypic and genotypic drug susceptibility test methods. J Clin Microbiol. 2013;51:4220–2.

134. Liang B, Tan Y, Li Z, Tian X, Du C, Li H, et al. Highly sensitive detection of isoniazid heteroresistance in *Mycobacterium tuberculosis* by DeepMelt assay. J Clin Microbiol. 2018. https://doi.org/10.1128/JCM.01239-17.

135. Zhang D, Gomez JE, Chien J-Y, Haseley N, Desjardins CA, Earl AM, et al. Genomic analysis of the evolution of fluoroquinolone resistance in *Mycobacterium tuberculosis* prior to tuberculosis diagnosis. Antimicrob Agents Chemother. 2016;60:6600–8.

136. Phelan JE, Coll F, Bergval I, Anthony RM, Warren R, Sampson SL, et al. Recombination in pe/ppe genes contributes to genetic variation in *Mycobacterium tuberculosis* lineages. BMC Genomics. 2016. https://doi.org/10.1186/s12864-016-2467-y.

137. Akhter Y, Ehebauer MT, Mukhopadhyay S, Hasnain SE. The PE/PPE multigene family codes for virulence factors and is a possible source of mycobacterial antigenic variation: Perhaps more? Biochimie. 2012;94:110–6.

138. Farhat MR, Shapiro B, Sheppard SK, Colijn C, Murray M. A phylogeny-based sampling strategy and power calculator informs genome-wide associations study design for microbial pathogens. Genome Med. 2014. https://doi.org/10.1186/s13073-014-0101-7.

139. 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. https://doi.org/10.1371/journal.pmed.1002137.

140. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol. 1993;31:406–9.

141. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol. 1997;35:907–14.

142. Supply P, Magdalena J, Himpens S, Locht C. Identification of novel intergenic repetitive units in a mycobacterial two-component system operon. Mol Microbiol. 1997;26:991–1003.

143. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J Clin Microbiol. 2006;44:4498–510.

144. Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D. MIRU-VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. Nucleic Acids Res. 2010;38:W326–31.

145. Oelemann MC, Diel R, Vatin V, Haas W, Rüsch-Gerdes S, Locht C, et al. Assessment of an optimized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing system combined with spoligotyping for population-based molecular epidemiology studies of tuberculosis. J Clin Microbiol. 2007;45:691–7.

146. Chihota VN, Müller B, Mlambo CK, Pillay M, Tait M, Streicher EM, et al. Population structure of multi- and extensively drug-resistant *Mycobacterium tuberculosis* strains in South Africa. J Clin Microbiol. 2012;50:995–1002.

147. Gandhi NR, Brust JCM, Moodley P, Weissman D, Heo M, Ning Y, et al. Minimal diversity of drug-resistant *Mycobacterium tuberculosis* strains. South Africa. Emerg Infect Dis. 2014;20:394–401.

148. Bifani PJ, Plikaytis BB, Kapur V, Stockbauer K, Pan X, Lutfey ML, et al. Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family. JAMA. 1996;275:452–7.

149. Lutfey M, Della-Latta P, Kapur V, Palumbo LA, Gurner D, Stotzky G, et al. Independent origin of mono-rifampin-resistant *Mycobacterium tuberculosis* in patients with AIDS. Am J Respir Crit Care Med. 1996;153:837–40.

150. Post FA, Willcox PA, Mathema B, Steyn LM, Shean K, Ramaswamy SV, et al. Genetic polymorphism in *Mycobacterium tuberculosis* isolates from patients with chronic multidrug-resistant tuberculosis. J Infect Dis. 2004;190:99–106.

151. Roof I, Jajou R, Kamst M, Mulder A, de Neeling A, van Hunen R, et al. Prevalence and characterization of heterogeneous variable-number tandem-repeat clusters comprising drug-susceptible and/or variable resistant *Mycobacterium tuberculosis* complex isolates in the Netherlands from 2004 to 2016. J Clin Microbiol. 2018. https://doi.org/10.1128/JCM.00887-18.

152. Witney AA, Bateson AL, Jindani A, Phillips PP, Coleman D, Stoker NG, et al. Use of whole-genome sequencing to distinguish relapse from reinfection in a completed tuberculosis clinical trial. BMC Med. 2017. https://doi.org/10.1186/s12916-017-0834-4.

153. Van Rie A, Victor TC, Richardson M, Johnson R, Van Der Spuy GD, Murray EJ, et al. Reinfection and mixed infection cause changing *Mycobacterium tuberculosis* drug-resistance patterns. Am J Respir Crit Care Med. 2005;172:636–42.

154. Warren RM, Victor TC, Streicher EM, Richardson M, Beyers N, Gey van Pittius NC, et al. Patients with active tuberculosis often have different strains in the same sputum specimen. Am J Respir Crit Care Med. 2004;169:610–4.

155. Hatherell H-A, Colijn C, Stagg HR, Jackson C, Winter JR, Abubakar I. Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review. BMC Med. 2016. https://doi.org/10.1186/s12916-016-0566-x.

156. Sun G, Luo T, Yang C, Dong X, Li J, Zhu Y, et al. Dynamic population changes in *Mycobacterium tuberculosis* during acquisition and fixation of drug resistance in patients. J Infect Dis. 2012;206:1724–33.

157. Mariam SH, Werngren J, Aronsson J, Hoffner S, Andersson DI. Dynamics of antibiotic resistant mycobacterium tuberculosis during long-term infection and antibiotic treatment. PLoS One. 2011. https://doi.org/10.1371/journal.pone.0021147.

158. Manson AL, Abeel T, Galagan JE, Sundaramurthi JC, Salazar A, Gehrmann T, 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;64:1494–501.

159. Wollenberg KR, Desjardins CA, Zalutskaya A, Slodovnikova V, Oler AJ, Quiñones M, et al. Whole-genome sequencing of *Mycobacterium tuberculosis* provides insight into the evolution and genetic composition of drug-resistant tuberculosis in Belarus. J Clin Microbiol. 2017;55:457–69.

160. Koch A, Wilkinson RJ. The road to drug resistance in *Mycobacterium tuberculosis*. Genome Biol. 2014. https://doi.org/10.1186/s13059-014-0520-1.

161. Walker TM, Lalor MK, Broda A, Saldana Ortega L, Morgan M, Parker L, et al. Assessment of *Mycobacterium tuberculosis* transmission in Oxfordshire, UK, 2007–12, with whole pathogen genome sequences: an observational study. Lancet Respir Med. 2014;2:285–92.

162. Gandhi NR, Moll A, Sturm AW, Pawinski R, Govender T, Lalloo U, et al. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. Lancet. 2006;368:1575–80.

163. Casali N, Nikolayevskyy V, Balabanova Y, Ignatyeva O, Kontsevaya I, Harris SR, et al. Microevolution of extensively drug-resistant tuberculosis in Russia. Genome Res. 2012;22:735–45.

164. Guerra-Assunção JA, Houben RMGJ, Crampin AC, Mzembe T, Mallard K, Coll F, et al. Recurrence due to relapse or reinfection with *Mycobacterium tuberculosis*: a whole-genome sequencing approach in a large, population-based cohort with a high HIV infection prevalence and active follow-up. J Infect Dis. 2015;211:1154–63.

165. Perdigão J, Silva H, Machado D, Macedo R, Maltez F, Silva C, et al. Unraveling *Mycobacterium tuberculosis* genomic diversity and evolution in Lisbon, Portugal, a highly drug resistant setting. BMC Genomics. 2014. https://doi.org/10.1186/1471-2164-15-991.

166. Lalor MK, Casali N, Walker TM, Anderson LF, Davidson JA, Ratna N, et al. The use of whole-genome sequencing in cluster investigation of a multidrug-resistant tuberculosis outbreak. Eur Respir J. 2018. https://doi.org/10.1183/13993003.02313-2017.

167. Bjorn-Mortensen K, Soborg B, Koch A, Ladefoged K, Merker M, Lillebaek T, et al. Tracing *Mycobacterium tuberculosis* transmission by whole genome sequencing in a high incidence setting: a retrospective population-based study in East Greenland. Sci Rep. 2016. https://doi.org/10.1038/srep33180.

168. Packer S, Green C, Brooks-Pollock E, Chaintarli K, Harrison S, Beck CR. Social network analysis and whole genome sequencing in a cohort study to investigate TB transmission in an educational setting. BMC Infect Dis. 2019. https://doi.org/10.1186/s12879-019-3734-8.

169. Auld SC, Shah NS, Mathema B, Brown TS, Ismail N, Omar SV, et al. Extensively drug-resistant tuberculosis in South Africa: genomic evidence supporting transmission in communities. Eur Respir J. 2018. https://doi.org/10.1183/13993003.00246-2018.

170. Ford C, Yusim K, Ioerger T, Feng S, Chase M, Greene M, et al. *Mycobacterium tuberculosis*—heterogeneity revealed through whole genome sequencing. Tuberculosis (Edinb). 2012;92:194–201.

171. Farmer P, Bayona J, Becerra M, Furin J, Henry C, Hiatt H, et al. The dilemma of MDR-TB in the global era. Int J Tuberc Lung Dis. 1998;2:869–76.

172. Xie YL, Chakravorty S, Armstrong DT, Hall SL, Via LE, Song T, et al. Evaluation of a rapid molecular drug-susceptibility test for tuberculosis. N Engl J Med. 2017;377:1043–54.

173. Lee RS, Behr MA. The implications of whole-genome sequencing in the control of tuberculosis. Ther Adv Infect Dis. 2016;3:47–62.

174. Jeanes C, O'Grady J. Diagnosing tuberculosis in the 21st century—Dawn of a genomics revolution? Int J Mycobacteriol. 2016;5:384–91.

175. Lee RS, Pai M. Real-time sequencing of *Mycobacterium tuberculosis*: are we there yet? J Clin Microbiol. 2017;55:1249–54.

176. Brown AC, Bryant JM, Einer-Jensen K, Holdstock J, Houniet DT, Chan JZM, et al. Rapid whole-genome sequencing of mycobacterium tuberculosis isolates directly from clinical samples. J Clin Microbiol. 2015;53:2230–7.

177. Doyle RM, Burgess C, Williams R, Gorton R, Booth H, Brown J, et al. Direct whole-genome sequencing of sputum accurately identifies drug-resistant *Mycobacterium tuberculosis* faster than MGIT culture sequencing. J Clin Microbiol. 2018. https://doi.org/10.1128/JCM.00666-18.

178. Lowenthal P, Lin S-YG, Desmond E, Shah N, Flood J, Barry PM. Evaluation of the impact of a sequencing assay for detection of drug resistance on the clinical management of tuberculosis. Clin Infect Dis. 2018. https://doi.org/10.1093/cid/ciy937.

179. Judge K, Harris SR, Reuter S, Parkhill J, Peacock SJ. Early insights into the potential of the Oxford Nanopore MinION for the detection of antimicrobial resistance genes. J Antimicrob Chemother. 2015;70:2775–8.

180. Schmidt K, Mwaigwisya S, Crossman LC, Doumith M, Munroe D, Pires C, et al. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. J Antimicrob Chemother. 2017;72:104–14.

181. Cohen T, van Helden PD, Wilson D, Colijn C, Mclaughlin MM, Abubakar I, et al. Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control. Clin Microbiol Rev. 2012;25:708–19.

182. World Health Organization. The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex: technical guide. Geneva: WHO; 2018. https://apps.who.int/iris/bitstream/handle/10665/274443/WHO-CDS-TB-2018.19-eng.pdf?sequence=1&isAllowed=y. Accessed 9 July 2019

183. Walker TM, Cruz ALG, Peto TE, Smith EG, Esmail H, Crook DW. Tuberculosis is changing. Lancet Infect Dis. 2017;17:359–61.

184. Colman RE, Anderson J, Lemmer D, Lehmkuhl E, Georghiou SB, Heaton H, et al. Rapid drug susceptibility testing of drug-resistant *Mycobacterium tuberculosis* isolates directly from clinical samples by use of amplicon sequencing: a proof-of-concept study. J Clin Microbiol. 2016;54:2058–67.