


# Inclusion of minor alleles improves catalogue-based prediction of fluoroquinolone resistance in *Mycobacterium tuberculosis*

Alice E. Brankin<sup>1</sup> and Philip W. Fowler<sup>1,2\*</sup> 

<sup>1</sup>Nuffield Department of Medicine, University of Oxford, Oxford, UK; <sup>2</sup>National Institute of Health Research Oxford Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK

\*Corresponding author. E-mail: [philip.fowler@ndm.ox.ac.uk](mailto:philip.fowler@ndm.ox.ac.uk)  
 @philipwflower, @AliceBrankin

Received 16 January 2023; accepted 11 March 2023

**Objectives:** Fluoroquinolone resistance poses a threat to the successful treatment of tuberculosis. WGS, and the subsequent detection of catalogued resistance-associated mutations, offers an attractive solution to fluoroquinolone susceptibility testing but sensitivities are often less than 90%. We hypothesize that this is partly because the bioinformatic pipelines used usually mask the recognition of minor alleles that have been implicated in fluoroquinolone resistance.

**Methods:** We analysed the Comprehensive Resistance Prediction for Tuberculosis: an International Consortium (CRyPTIC) dataset of globally diverse WGS *Mycobacterium tuberculosis* isolates, with matched MICs for two fluoroquinolone drugs and allowed putative minor alleles to contribute to resistance prediction.

**Results:** Detecting minor alleles increased the sensitivity of WGS for moxifloxacin resistance prediction from 85.4% to 94.0%, without significantly reducing specificity. We also found no correlation between the proportion of an *M. tuberculosis* population containing a resistance-conferring allele and the magnitude of resistance.

**Conclusions:** Together our results highlight the importance of detecting minor resistance-conferring alleles when using WGS, or indeed any sequencing-based approach, to diagnose fluoroquinolone resistance.

## Introduction

The fluoroquinolone antibiotics levofloxacin and moxifloxacin are recommended by the WHO for the treatment of both drug-susceptible and MDR tuberculosis (TB).<sup>1</sup> It is therefore imperative that fluoroquinolone drug susceptibility testing (DST) is carried out quickly and accurately to treat patients and prevent the spread of resistant strains.

WGS can rapidly identify resistance and susceptibility to several antitubercular drugs.<sup>2,3</sup> The WHO recommends that WGS results are interpreted using a catalogue of mutations associated with resistance compiled using over 38 000 *M. tuberculosis* isolates.<sup>4</sup> The sensitivity of this catalogue for identifying resistance was >90% for the first-line drugs rifampicin and isoniazid in the dataset used to build it.<sup>5</sup> However, the sensitivity to identifying levofloxacin and moxifloxacin resistance was lower, at 84.4% and 87.7% respectively.<sup>5</sup> Because fluoroquinolone resistance is well characterized and attributed to a small number of mutations in the *gyrA* and *gyrB* genes<sup>6</sup> it is surprising that a larger proportion of fluoroquinolone resistance was not explained by the catalogue.

Mixed populations are common in *M. tuberculosis* infections and have been particularly implicated in fluoroquinolone

resistance<sup>7,8</sup>; the prevalence of mixed populations containing minor resistance-conferring alleles is estimated at ~10% of fluoroquinolone-resistant isolates.<sup>9</sup> WGS bioinformatics pipelines often use filters and thresholds to reduce the effect of sequencing errors,<sup>10</sup> which unfortunately also preclude the detection of minor alleles. Indeed, the pipeline used in processing the samples for the WHO catalogue<sup>4,5</sup> only identified a genetic variant if over 90% of reads at a genomic position supported its existence, effectively ignoring minor alleles.

Large matched WGS and phenotypic datasets, such as that compiled by the Comprehensive Resistance Prediction for Tuberculosis: an International Consortium (CRyPTIC),<sup>11</sup> provide an opportunity to study the significance of minor alleles. In this article we will investigate the extent to which catalogue-based fluoroquinolone resistance prediction is enhanced by including minor alleles containing known resistance-conferring mutations.

## Materials and methods

We assume that a population is homogeneous if 90% or more of the reads support a different nucleotide to the reference genome (i.e. fraction of read support, FRS  $\geq 0.9$ ), whereas a population is mixed if there are

more than two reads but fewer than 90% supporting a genetic variant. Our rationale is that as the error rate of Illumina sequencing is <1%, if two or more reads support an alternative allele, it is highly unlikely that this is due to sequencing error.

*M. tuberculosis* complex isolates with an associated moxifloxacin or levofloxacin MIC were obtained from the CRyPTIC FTP site ([http://ftp.ebi.ac.uk/pub/databases/cryptic/release\\_june2022/](http://ftp.ebi.ac.uk/pub/databases/cryptic/release_june2022/)). Isolates were discarded if the phenotypic measurement was annotated as 'low quality' (the three methods used to determine the MIC were not in agreement) or if they originated from a laboratory with a known quality control issue.<sup>11</sup> The genetic variants these isolates had in either the *gyrA* or *gyrB* genes, as detected by Illumina sequencing and the CRyPTIC bioinformatic pipeline,<sup>11</sup> were also extracted from the dataset. In total, 9128 isolates with WGS and MIC data for levofloxacin and 8138 isolates with WGS and MIC data for moxifloxacin were analysed.

The variant caller (Clockwork v0.8.3, <https://github.com/iqbal-lab-org/clockwork>) used by the CRyPTIC project was set up conservatively; a genetic variant had to have an FRS  $\geq 90\%$  for it to be identified, with all other potential variants being screened out.<sup>11</sup> We therefore directly parsed the variant call format files of all samples looking for isolates that had evidence of minor alleles in *gyrA* or *gyrB*. The FRS for each putative genetic variant was extracted and any resultant amino acid changes identified. Putative variants present at a position that failed the Minos<sup>12</sup> minimum sequencing depth filter were not included, and hence assumed WT for analyses. The minimum genotype confidence percentile was not used to exclude any variants because this filter is itself partially dependent on the FRS.<sup>12</sup> Finally, we assumed that genetic variants in the WHO resistance catalogue annotated as either 'resistance associated' or 'resistance associated—interim' mutations both conferred resistance to the relevant drug.

## Results

CRyPTIC isolates were classified as resistant or susceptible depending on whether their MIC (Figure 1a, b) lay above or below a published epidemiological cut-off value (ECOFF/ECV).<sup>13</sup> We then predicted genetically which of the isolates were resistant to levofloxacin and moxifloxacin assuming the populations were homogeneous (all genetic variants supported by an FRS  $\geq 0.9$ ). This approach identified the levofloxacin- and moxifloxacin-resistant isolates with 83.1% and 85.4% sensitivity, respectively, and over 90% specificity in both cases (Figure 1c, d). In total, 16.9% of levofloxacin resistance and 14.6% of moxifloxacin resistance in this dataset is therefore not explainable by catalogue mutations seen at FRS  $\geq 0.9$ .

Allowing minor alleles containing mutations in the WHO catalogue to contribute to the predictions by reducing the FRS threshold increased the sensitivity of the catalogue significantly, by 9.7% and 9.5% when identifying levofloxacin and moxifloxacin resistance, respectively (Figure 1c, d). The specificity of the predictions decreased slightly by 0.5% and 0.8% for levofloxacin and moxifloxacin, respectively, but these differences were not significant at  $P=0.05$ . For this dataset, 7.9% of levofloxacin resistance and 6.0% of moxifloxacin resistance remains unexplained by the presence of catalogue mutations.

One might expect that if a resistance-conferring allele is seen at higher FRS (i.e. is more prevalent within the mixed population) it will have a greater level of resistance because the time taken for the resistant population to present as a visible growth in antibiotic-containing wells on the plate will be shorter. Hence, we next examined whether the proportion of the population containing

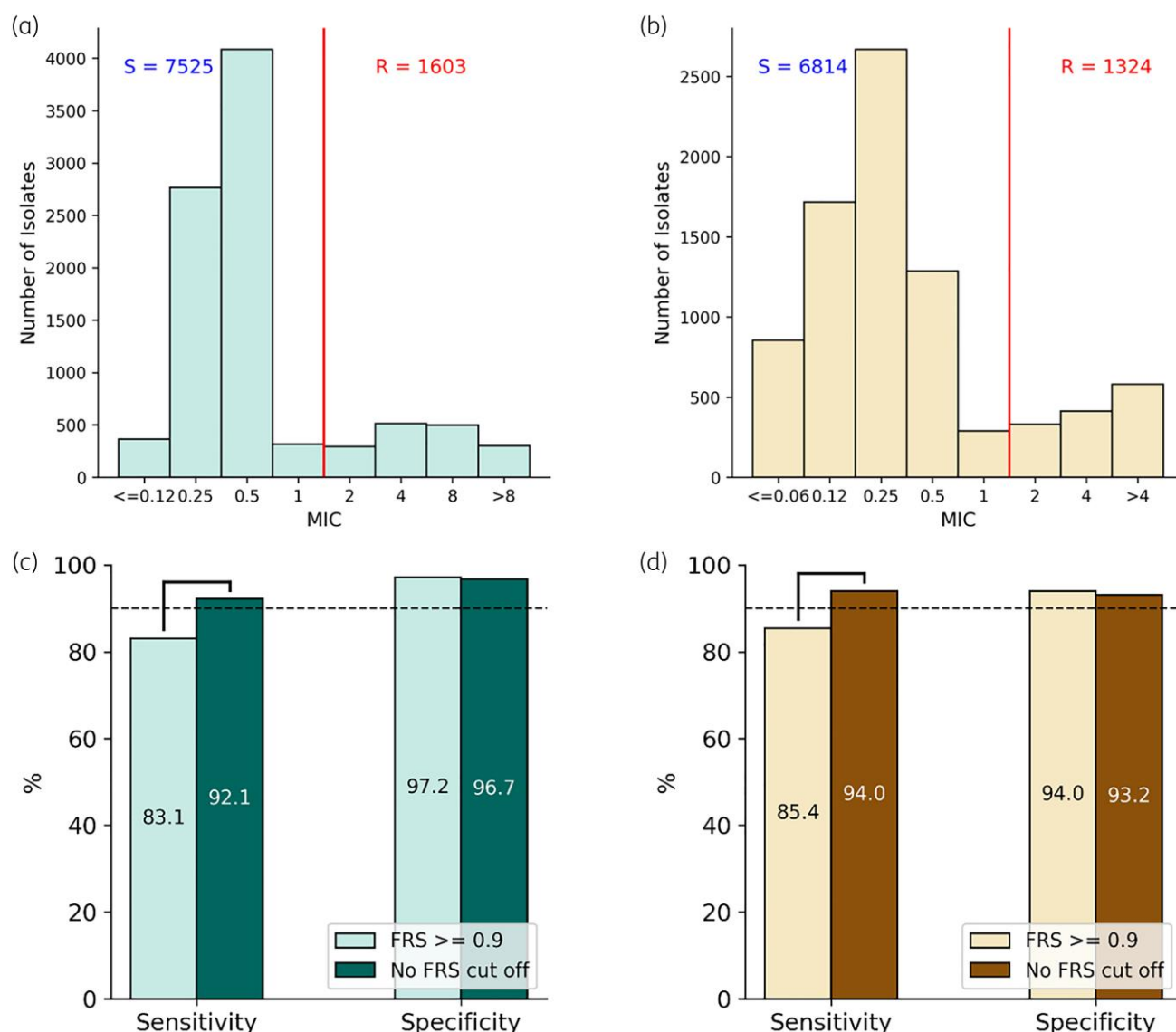
the minor resistance-conferring allele (FRS) correlated with the magnitude of the fluoroquinolone MIC after 14 days incubation on the plate. Although all catalogue resistance-associated mutations were seen as part of a mixed population in at least one sample, we only considered the two most frequent mutations seen in fluoroquinolone-resistant isolates, *gyrA* D94G and A90V. Pearson's rank correlation coefficients confirmed there is no correlation between FRS for the resistance-conferring allele and MIC to either levofloxacin or moxifloxacin after 14 days of incubation on the plate in isolates with either *gyrA* D94G or A90V. Therefore we disproved our hypothesis that a higher prevalence of a resistance-conferring allele in a population confers a greater level of resistance to selective pressure from antibiotic (Figure 2a–d).

## Discussion

We found a near 10% absolute improvement in sensitivity without a significant reduction in specificity when predicting levofloxacin and moxifloxacin resistance using the WHO mutation catalogue by including minor alleles (Figure 1c, d). Importantly, this brings the catalogue performance for detecting fluoroquinolone resistance up to the level of isoniazid and rifampicin,<sup>4,5</sup> where the determinants of resistance are also well understood. The magnitude of improvement in sensitivity correlates with the previously estimated 10% frequency of fluoroquinolone resistance conferred by alleles seen in mixed populations.<sup>10</sup> Resistance from minority populations has likewise been implicated for rifampicin, isoniazid and ethambutol,<sup>10</sup> and should be explored further.

Importantly, only when minor alleles are included does the catalogue performance exceed the minimum requirement of 90% sensitivity compared with phenotypic DST, as set out in the WHO target product profile (TPP) for next-generation sequencing technologies.<sup>14</sup> This improvement also provides indirect evidence that molecular DST tools developed to detect the catalogue mutations will exceed the WHO target of 90% sensitivity,<sup>15</sup> provided that these tests detect mutations in mixed populations. Although the small decrease in specificity observed by including mixed alleles in our study was not significant, it is important to note that the specificity compared with phenotypic DST for moxifloxacin, even without using mixed alleles, is below the minimum threshold of 95% set out in the WHO TPP (Figure 1d). This warrants further investigation but is likely due, in part, to the increased likelihood that samples containing resistance-conferring mutations in *gyrA* have an MIC below the ECOFF/ECV due to the relatively small increases in MICs observed for strains resistant to fluoroquinolones.<sup>13</sup> This effect is particularly pronounced for moxifloxacin, consistent with the lower observed specificity for this drug.

Despite the improvement in sensitivity achieved by allowing minor alleles to contribute, 7.9% of levofloxacin resistance and 6.0% of moxifloxacin resistance is not explained by the presence of catalogue mutations. Several scenarios could account for this. Firstly, there are likely additional minor resistance-conferring alleles that could not be detected in this study due to limited sequencing depth (the mean read depth for the CRyPTIC isolates<sup>11</sup> was  $74 \pm 44$ ). In addition, our use of two reads in support of an alternative allele to define a minor population is conservative. For example, at a read depth of 30x, a minor allele is



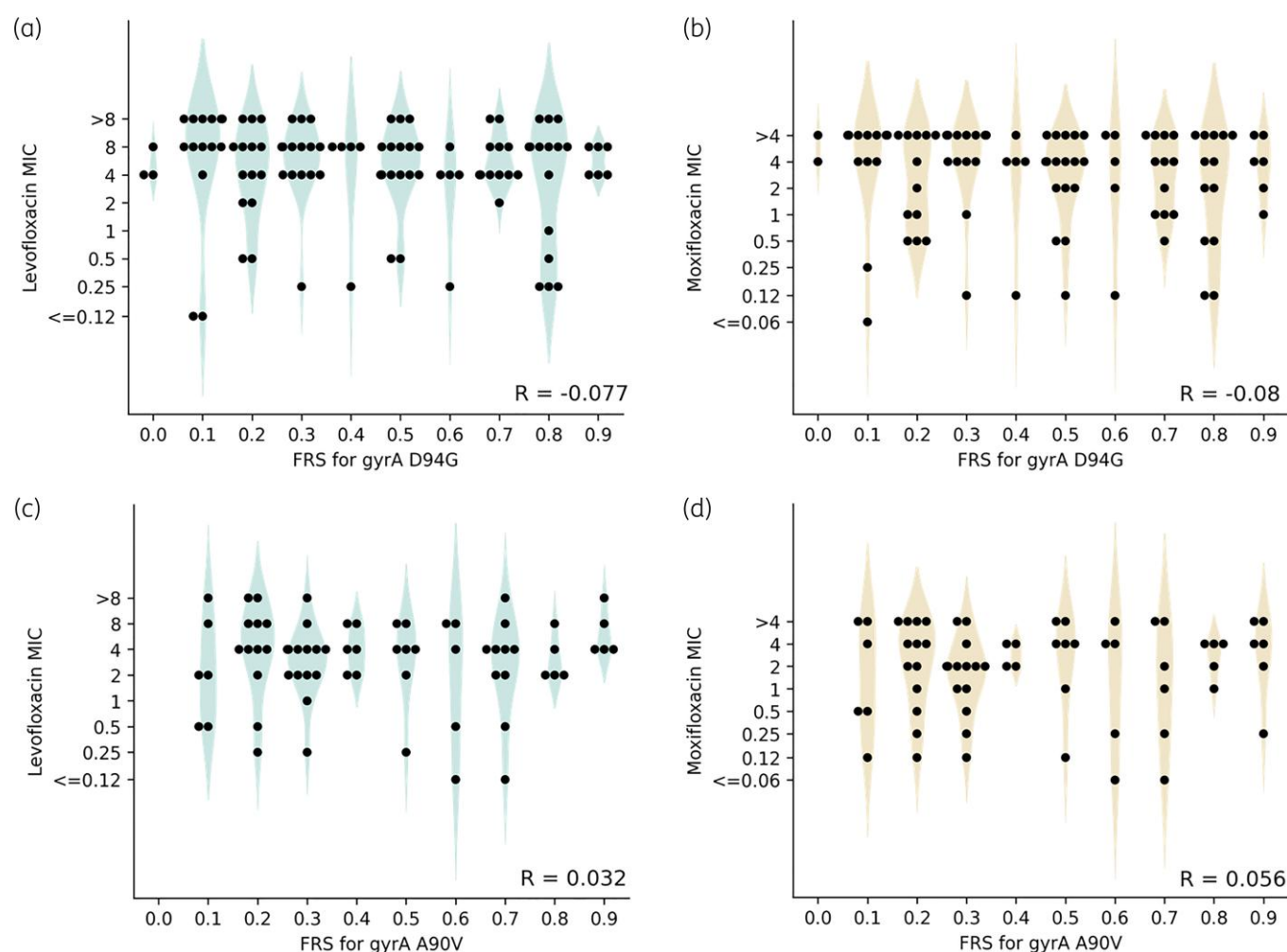
**Figure 1.** Distribution of *M. tuberculosis* isolates MICs to (a) levofloxacin and (b) moxifloxacin. The red line indicates the previously proposed epidemiological cut-off value (ECOFF/ECV)<sup>13</sup> that was used to distinguish resistant (R) and susceptible (S) isolates. Sensitivity and specificity of (c) levofloxacin and (d) moxifloxacin resistance prediction using WHO 2021 catalogue mutations with and without mixed alleles. Brackets indicate a significant difference (z-test,  $P < 0.05$ ).

only detectable by our approach if at least 6.7% of the population contains it. Secondly, the catalogue mutations are unlikely to be exhaustive because rare resistance-conferring mutations would not meet the statistical criteria required.<sup>4,5</sup> Including minor alleles when building catalogues could help improve their comprehensiveness by increasing the number of examples of rarer mutations. Finally, complex resistance mechanisms such as drug efflux may play a role.<sup>16</sup>

The lack of correlation between the magnitude of resistance and the proportion of a population containing a minor resistance-conferring allele (Figure 2a-d) is consistent with the suggestion that *gyrA* D94G and A90V do not contribute a significant fitness cost.<sup>17</sup> Isolates with *gyrA* D94G and A90V seen at very low FRS

still had high MICs to levofloxacin and moxifloxacin after 2 weeks incubation, suggesting that a small resistant population can rapidly outcompete a majority WT population under the selective pressure of fluoroquinolone treatment. We infer that it is essential that any tool used for fluoroquinolone DST detects minor resistant populations. This means that for next-generation sequencing approaches to be successful, the sequencing depths, variant callers used and any filters applied need to be carefully optimized, leading to standardized recommendations to prevent the misdiagnosis of fluoroquinolone susceptibility.

Due to the nature of Illumina sequencing performed by the CRyPTIC study, the sequencing depth is limited and highly variable between samples and genetic loci. A deep sequencing



**Figure 2.** Distribution of FRS for *gyrA* D94G and the (a) levofloxacin and (b) moxifloxacin MIC of the *M. tuberculosis* isolate; and the distribution of FRS for *gyrA* A90V and the (c) levofloxacin and (d) moxifloxacin MIC of the *M. tuberculosis* isolate. FRS was rounded to the nearest 0.1 to show the distribution of MIC values at different FRS. Both susceptible and resistant isolates containing alleles encoding the mutation at FRS < 0.9 were included, and alleles with FRS  $\geq 0.9$  were excluded to avoid overfitting to the majority homogeneous population. R shows the Pearson rank correlation coefficient between the unrounded FRS and  $\log_2$  MIC. Individual sample MICs are plotted as filled black circles and violin plots are drawn to help illustrate the distributions.

approach is necessary to confirm the importance of minority resistant populations in *M. tuberculosis* infection and to better inform the design and parameterization of bioinformatics pipelines. The Illumina WGS data analysed here provide an estimate of what could be detected by current practice and suggest that high sequencing depths at loci associated with resistance might be required to rule out fluoroquinolone resistance. In addition, it is likely that due to the culturing step the observed population may be more homogeneous than is present in the patient. Further work to reduce the number and duration of steps between sampling and sequencing is needed to both investigate and minimize the magnitude of any effect.

To conclude, identifying minority populations containing mutations known to confer resistance improved the sensitivity of fluoroquinolone resistance prediction using the 2021 WHO TB resistance catalogue. Hence, it is vital that genetics-based tools

and pipelines used for fluoroquinolone DST can resolve resistance-conferring mutations in minority populations.

## Acknowledgements

We are grateful to the CRyPTIC consortium for helpful discussions and making their data publicly available.

## Funding

A.E.B. is funded by an NDM Prize Studentship from the Oxford Medical Research Council Doctoral Training Partnership and the Nuffield Department of Clinical Medicine. This study is supported by the National Institute for Health Research (NIHR) Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance (NIHR200915), a partnership between the UK Health Security Agency

(UKHSA) and the University of Oxford. Computation used the Oxford Biomedical Research Computing (BMRC) facility, a joint development between the Wellcome Centre for Human Genetics and the Big Data Institute supported by Health Data Research UK and the NIHR Oxford Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, NIHR, UKHSA or the Department of Health and Social Care.

## Transparency declarations

The authors have nothing to declare

## References

- World Health Organization. Rapid communication: key changes to the treatment of drug-resistant tuberculosis. 2022. <https://www.who.int/publications/i/item/WHO-UCN-TB-2022-2>
- Walker TM, Kohl TA, Omar SV *et al.* Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect Dis* 2015; **15**: 1193–202. [https://doi.org/10.1016/S1473-3099\(15\)00062-6](https://doi.org/10.1016/S1473-3099(15)00062-6)
- Lam C, Martinez E, Crighton T *et al.* Value of routine whole genome sequencing for *Mycobacterium tuberculosis* drug resistance detection. *Int J Infect Dis* 2021; **113** Suppl 1: S48–54. <https://doi.org/10.1016/j.ijid.2021.03.033>
- World Health Organization. Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance. 2021. <https://www.who.int/publications/i/item/9789240028173>
- Walker TM, Miotto P, Köser C *et al.* The 2021 WHO catalogue of *Mycobacterium tuberculosis* complex mutations associated with drug resistance: a genotypic analysis. *Lancet Microbe* 2022; **3**: e265. [https://doi.org/10.1016/S2666-5247\(21\)00301-3](https://doi.org/10.1016/S2666-5247(21)00301-3)
- Nguyen L. Antibiotic resistance mechanisms in *M. tuberculosis*: an update. *Arch Toxicol* 2016; **90**: 1585–604. <https://doi.org/10.1007/s00204-016-1727-6>
- Nimmo C, Brien K, Millard J *et al.* Dynamics of within-host *Mycobacterium tuberculosis* diversity and heteroresistance during treatment. *EBioMedicine* 2020; **55**: 102747. <https://doi.org/10.1016/j.ebiom.2020.102747>
- Singhal R, Reynolds PR, Marola JL *et al.* Sequence analysis of fluoroquinolone resistance-associated genes *gyrA* and *gyrB* in clinical *Mycobacterium tuberculosis* isolates from patients suspected of having multidrug-resistant tuberculosis in New Delhi, India. *J Clin Microbiol* 2016; **54**: 2298–305. <https://doi.org/10.1128/JCM.00670-16>
- Ye M, Yuan W, Molaeipour L *et al.* Antibiotic heteroresistance in *Mycobacterium tuberculosis* isolates: a systematic review and meta-analysis. *Ann Clin Microbiol Antimicrob* 2021; **20**: 73. <https://doi.org/10.1186/s12941-021-00478-z>
- Said Mohammed K, Kibinge N, Prins P *et al.* Evaluating the performance of tools used to call minority variants from whole genome short-read data. *Wellcome Open Res* 2018; **3**: 21. <https://doi.org/10.12688/wellcomeopenres.13538.2>
- The CRYPTIC Consortium. A data compendium associating the genomes of 12,289 *Mycobacterium tuberculosis* isolates with quantitative resistance phenotypes to 13 antibiotics. *PLoS Biol* 2022; **20**: e3001721. <https://doi.org/10.1371/journal.pbio.3001721>
- Hunt M, Letcher B, Malone KM *et al.* Minos: variant adjudication and joint genotyping of cohorts of bacterial genomes. *Genome Biol* 2022; **23**: 147. <https://doi.org/10.1186/s13059-022-02714-x>
- The CRYPTIC Consortium. Epidemiological cutoff values for a 96-well broth microdilution plate for high-throughput research antibiotic susceptibility testing of *M. tuberculosis*. *Eur Resp J* 2022; **60**: 2200239. <https://doi.org/10.1183/13993003.00239-2022>
- 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. 2018. <https://apps.who.int/iris/handle/10665/274443>
- World Health Organization. Target product profile for next-generation drug-susceptibility testing at peripheral centres. 2021. <https://www.who.int/publications/i/item/9789240032361>
- Remm S, Earp JC, Dick T *et al.* Critical discussion on drug efflux in *Mycobacterium tuberculosis*. *FEMS Microbiol Rev* 2021; **46**: fuab050. <https://doi.org/10.1093/femsre/fuab050>
- Pi R, Liu Q, Takiff HE *et al.* Fitness cost and compensatory evolution in levofloxacin-resistant *Mycobacterium aurum*. *Antimicrob Agents Chemother* 2020; **64**: e00224–20. <https://doi.org/10.1128/aac.00224-20>