The Lancet Digital Health

Bringing TB genomics to the clinic: A comprehensive pipeline to predict antimicrobial susceptibility from genomic data, validated and accredited to ISO standards

--Manuscript Draft--

Manuscript Number:	tldigitalhealth-D-24-00681
Article Type:	Article (Original Research)
Keywords:	Mycobacterium tuberculosis; anti-tuberculous therapy; antimicrobial resistance; whole genome sequencing; Clinical microbiology; pathogen genomics; genomics
Corresponding Author:	Norelle L Sherry, MBBS PhD University of Melbourne Melbourne, VIC AUSTRALIA
First Author:	Kristy A Horan, PhD
Order of Authors:	Kristy A Horan, PhD
	Linda Viberg, PhD
	Susan A Ballard, PhD
	Maria Globan, BSc
	Katherine Bond, MBBS
	Jessica R Webb, PhD
	Thinley Dorji, MBBS, MPH
	Deborah A Williamson, PhD
	Michelle L Sait, PhD
	Ee Laine Tay, MPH
	Justin T Denholm, MBBS PhD
	Benjamin P Howden, MBBS PhD
	Torsten Seemann, PhD
	Norelle L Sherry, MBBS PhD
Manuscript Region of Origin:	AUSTRALIA
Abstract:	Background Whole genome sequencing (WGS) is increasingly contributing to the clinical management of tuberculosis, but accessible bioinformatic tools tailored for clinical reporting and suitable for accreditation to international standards are currently lacking. Methods We developed tbtAMR, a flexible yet comprehensive tool for analysis of Mycobacterium tuberculosis genomic data, including inference of phenotypic susceptibility and lineage calling. Validation was undertaken using local and publicly-available real-world data (phenotype and genotype) and synthetic genomic data to determine the appropriate quality control metrics and extensively validate the pipeline for clinical use. Findings tbtAMR accurately predicted lineages and phenotypic susceptibility for first- and second-line drugs, with equivalent or better computational and predictive performance compared to other bioinformatics tools currently available. tbtAMR is flexible with modifiable criteria to tailor results to users' needs.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	Interpretation The tbtAMR tool is suitable for use in clinical and public health microbiology laboratory settings, and can be tailored to specific local needs by non-programmers. We have accredited this tool to ISO standards in our laboratory, and it has been implemented for routine reporting of AMR from genomic sequence data in a clinically relevant timeframe. Reporting templates, validation methods and datasets are provided to offer a pathway for laboratories to adopt and seek their own accreditation for this critical test, to improve the management of tuberculosis globally. Funding Victorian Government Department of Health; Australian National Health and Medical Research Council and Medical Research Futures Fund.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Manuscript Click here to view linked References

Bringing TB genomics to the clinic: A comprehensive pipeline to predict antimicrobial susceptibility from genomic data, validated and accredited to ISO standards

Kristy A Horan[1], Linda Viberg[2], Susan A Ballard[1], Maria Globan[2], Katherine Bond[2,3], Jessica R Webb[4], Thinley Dorji[4], Deborah A Williamson[2,5], Michelle L Sait[1], Ee Laine Tay[6], Justin T Denholm[7,8], Benjamin P Howden[§,1,3,4,9,10] Torsten Seemann[§, 1,4,10], Norelle L Sherry[§*,1,4,9]*

§ Contributed equally as senior authors

**Corresponding authors: norelle.sherry@unimelb.edu.au; bhowden@unimelb.edu.au*

Affiliations

1. Microbiological Diagnostic Unit Public Health Laboratory (MDU-PHL), Department of Microbiology & Immunology, University of Melbourne at the Peter Doherty Institute for Infection & Immunity, Melbourne, Victoria, Australia
2. Mycobacterium Reference Laboratory, Victorian Infectious Diseases Reference Laboratory (VIDRL), at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia
3. Department of Microbiology, Royal Melbourne Hospital, Melbourne, Victoria, Australia
4. Department of Microbiology and Immunology, The University of Melbourne, at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia
5. Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia
6. Communicable Disease Epidemiology and Surveillance, Health Protection Branch, Public Health Division, Department of Health, Victoria, Australia
7. Victorian Tuberculosis Program, Melbourne Health, at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia
8. Department of Infectious Diseases, The University of Melbourne, at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia
9. Department of Infectious Diseases & Immunology, Austin Health, Heidelberg, Victoria, Australia
10. Centre for Pathogen Genomics, The University of Melbourne, Melbourne, Victoria, Australia

Abstract

Background Whole genome sequencing (WGS) is increasingly contributing to the clinical management of tuberculosis, but accessible bioinformatic tools tailored for clinical reporting and suitable for accreditation to international standards are currently lacking.

Methods We developed tbtAMR, a flexible yet comprehensive tool for analysis of *Mycobacterium tuberculosis* genomic data, including inference of phenotypic susceptibility and lineage calling. Validation was undertaken using local and publicly-available real-world data (phenotype and genotype) and synthetic genomic data to determine the appropriate quality control metrics and extensively validate the pipeline for clinical use.

Findings tbtAMR accurately predicted lineages and phenotypic susceptibility for first- and second-line drugs, with equivalent or better computational and predictive performance compared to other bioinformatics tools currently available. tbtAMR is flexible with modifiable criteria to tailor results to users' needs.

Interpretation The tbtAMR tool is suitable for use in clinical and public health microbiology laboratory settings, and can be tailored to specific local needs by non-programmers. We have accredited this tool to ISO standards in our laboratory, and it has been implemented for routine reporting of AMR from genomic sequence data in a clinically relevant timeframe. Reporting templates, validation methods and datasets are provided to offer a pathway for laboratories to adopt and seek their own accreditation for this critical test, to improve the management of tuberculosis globally.

Funding Victorian Government Department of Health; Australian National Health and Medical Research Council and Medical Research Futures Fund.

Research in context

Evidence before this study

We searched PubMed for studies using the search terms: “Mycobacterium tuberculosis”, “clinical”, “bioinformatics”, “genomics”, “drug resistance (OR antimicrobial)”, published prior to 31st July 2024 without language restrictions (n=258). We considered all studies from this search that used genomics to infer likely drug-resistance in *M. tuberculosis*. Many of these studies highlight the challenges in making meaningful interpretations from genomic data for the purpose of inferring AMR for clinical applications. Despite the development of bioinformatics tools and compilation of catalogues of resistance conferring mutations, few of these studies directly address the challenges specific to implementation of a whole genome sequencing (WGS) and bioinformatics pipelines to deliver validated and accredited results for use in clinical applications and none provide sustainable solutions.

Added value of this study

To the best of our knowledge, this study is the first to detail practical solutions to challenges to validating and accrediting the routine inference of AMR from WGS data for clinical applications for treatment of *M. tuberculosis*. We developed, validated and accredited a bioinformatics tool, tbtAMR, which is robust and flexible, using a data-driven approach to inferring resistance, allowing for simple customisation of behaviour on a large collection of in-house data supplemented with publicly available data. Furthermore, we have made this pipeline and the accompanying data available for use by the wider public health community to aid others in implementing similar programs. Consultation with infectious disease clinicians, microbiologists and epidemiologists has allowed us to develop a robust workflow, that encompasses sequencing, analysis, interpretation, reporting, discrepancy resolution and ongoing verification.

Implications of all the available evidence

This study addressed the specific needs of reporting genomic AMR for *M. tuberculosis* in a clinical and public health laboratory. Furthermore, this work makes available a pipeline and data for laboratories to utilise to implement genomic AMR for *M. tuberculosis*. Through the use and further development of publicly available and open-source bioinformatics software, we have demonstrated the feasibility of clinical reporting of genomic AMR results for *M. tuberculosis* in a low-incidence high-income setting.

82 We believe that this work lays the foundation for making reporting of genomic AMR for *M. tuberculosis*
83 more accessible to public health and reference laboratories in general.

84 # Introduction

85 *Mycobacterium tuberculosis* (Mtb) is the causative agent of tuberculosis (TB), a disease that is globally
86 distributed but predominantly affecting people in low- and middle-income communities and is the
87 leading cause of death from a single infectious agent.[1]

88 Treatment of TB infections is lengthy and expensive, requiring combination therapy with at least two
89 antimycobacterial drugs prescribed for a minimum of six months.[2,3] This is usually commenced with
90 four antimycobacterial drugs ('first-line agents'): rifampicin, isoniazid, pyrazinamide and ethambutol,
91 followed by rationalisation to two agents (rifampicin and isoniazid, if susceptible). If drug resistance is
92 detected, an alternative regimen including second- and third-line agents will be used (Supplementary
93 Figure 1 and Table 1).

94 The slow growth of Mtb *in vitro* means that clinicians treating patients with TB must choose treatment
95 regimens before phenotypic drug susceptibility test (DST) results are available, which can take up to
96 two months[4] in many settings. Whole genome sequencing (WGS) has the advantage of providing
97 comprehensive results from an Mtb isolate in a single test (rather than individual tests for species
98 identification, lineage, and drug susceptibility), potentially significantly reducing turnaround time to
99 DST for both first-line and other antimycobacterial drugs, especially when undertaken from
100 Mycobacteria Growth Indicator Tube (MGIT) culture. In addition, WGS can provide increased reliability
101 of resistance detection for some drugs, due to known challenges with phenotypic reproducibility, such
102 as ethambutol and pyrazinamide[5]. WGS data is also being used to provide support for outbreak
103 investigation and contact tracing[6] (Figure 1A). Therefore, the use of routine WGS for Mtb has the
104 potential to improve patient management in a clinical setting and public health outcomes, and new
105 Australian national guidelines recommend routine use and reporting of findings for these reasons[7].

106 Despite the urgent need for routine identification of Mtb AMR from genomic data to inform therapy
107 and the wealth of available bioinformatic resources, few clinical and public health laboratories (CPHLs)
108 globally have processes in place to routinely report AMR from WGS data for Mtb. Many factors need
109 to be considered when implementing such an AMR detection program, particularly where patient
110 management may be impacted, including performance, ease-of-use and flexibility of tools in
111 predicting drug-resistance, the impact of laboratory methodologies and existing processes, such as
112 existing bioinformatics pipelines, reporting structure, CPHL data management, and test accreditation
113 by national accreditation bodies to International Organization for Standardization (ISO) standards
114 (ISO15189:2022).[8] Laboratories seek accreditation to ISO standards to demonstrate that they are able

115 to produce robust, reproducible and high-quality laboratory services; the process of obtaining
116 accreditation, and the difficulties in accrediting genomic tools, is covered in depth elsewhere[9].

117 There are many high-quality bioinformatics tools available for identification of genomic determinants
118 of AMR in Mtb, including Mykrobe,[10] TB-Profiler[11] and ariba[12]. Additionally, a second version of the
119 WHO Mtb Mutation Catalogue[10,13] has now been released with an expanded catalogue of mutations
120 for second- and third-line drugs (Groups A-C Table 1), as well as guidance for interpreting complex
121 genomic interactions. However, use of these tools often requires additional curation and
122 interpretation of results to be suitable for reporting in a CPHL setting[14]. Furthermore, WGS in CPHLs is
123 often undertaken in large batches with existing workflows and pipelines in place, therefore
124 bioinformatics tools need to be able to be easily incorporated into these existing pipelines, be
125 straightforward to maintain and for any updates to be assessed easily by reverification.

126 Here we describe the validation of a real-time routine analysis tool for inference of phenotypic
127 resistance in Mtb, based on detection of genomic variation, in a clinical and public health microbiology
128 setting. We have subsequently gained accreditation to ISO standard 15189:2022 (Clinical Laboratories)
129 for this tool, tbtAMR, for clinical and public-health reporting of genomic AMR detection in Mtb. The
130 pipeline and validation approach are made available here for other laboratories globally to aid in the
131 transition to WGS for genomic AMR reporting.

132

Methods

tbtAMR design

Many software options exist for detection of AMR determinants in Mtb sequences, including Mykrobe[10] and TBProfiler[11], each with pros and cons. When developing a stand-alone tool, we considered that the process must be easy to incorporate into existing laboratory workflows (Figure 1A), have flexibility in key outputs, and to be easily maintained to expedite any updates, ideally by non-developers. These were noted limitations of existing tools for use in a CPHL, where software maintenance and updates can be challenging. In addition, reporting of AMR from WGS data for Mtb will directly impact patient management. Thus, the results generated should be suitable for reporting in a clinical setting with a layer of interpretation that is fit-for-purpose, not reporting raw (unfiltered) results unless required for a specific purpose. Four key parameters should be reported for each drug to assist in clinical decision making: (i) presence or absence of mutations including small insertions or deletions (indels) and SNPs, (ii) confidence that this mutation is associated with drug resistance, (iii) the level of drug resistance conferred by this mutation and (iv) the overall resistance profile based on the resistance inferred to individual drugs (e.g. based on WHO definitions).

tbtAMR (https://github.com/MDU-PHL/tbtamr), is a pure python package and can be used as a stand-alone tool, with minimal dependencies, taking an annotated variant file (vcf) to generate an inferred antibiogram and resistance profile based on user-defined criteria (Figure 1B). The recently updated WHO mutation catalogue for Mtb (version) is implemented as the default mutational catalogue within tbtAMR[13] with its accompanying guidelines documentation[15] informing criteria for classification. However, a user-defined catalogue of mutations can also be supplied, along with bespoke criteria for interpretation and classification according to a user's needs, allowing for increased flexibility and ease of update.

tbtAMR relies on four accessory files (detailed below and available at[16]): a mutational catalogue and additional files which define criteria for structure of the mutational catalogue, for interpreting likely resistance, and reporting classification of resistance profile. tbtAMR comes with default files, based on WHO catalogue of mutations version 2 and the current WHO guidelines for classification of resistance profile.

- *Mutational catalogue* – this file must contain at least the drugs, genes and variants and needs the confidence levels of each of these combinations to indicate resistance to the drug is likely.

- *db_config.json* – this file provides information about the structure of the mutational catalogue, drugs to include, the values for confidence and resistance levels, cascade reporting structure (Supplementary methods).
- *interpretation_rules.csv* – this file is the key to providing flexible interpretations to address specific needs and to be capture complex rules around interactions. Default rules are designed to be applied on a per drug basis and should address the most common desired behaviour – for example, where confidence level of a mutation is high, report that drug as resistant. At least one default rule must be supplied for all drugs the user wishes to report. Additional default rules can also be supplied, for example, to report a special comment under certain conditions. Override rules should apply to variants or groups of variants and will replace the behaviour determined by the default rules in those cases, for example: where a single specific variant is observed, report 'Low-level' resistant, or where these two variants are observed, report 'Susceptible'.
- *classification_rules.csv* – these rules determine the overall classification profile, based on the combination of resistances in the sequence. The defaults are based on the current WHO guidelines (Supplementary Figure 1).

Note that tbtAMR performance has been investigated based on the default database and interpretation criteria; performance may vary where a bespoke database and/or interpretive criteria are used. The tool has optional additional functionalities, including paired-end read alignment, variant calling and annotation, which require additional dependencies including bwa-mem[17], freebayes[18], snpEff[19], delly[20], bcftools and samtools[21] (Figure 1B). These additional functionalities are implemented in our locally-designed mutAMR tool (https://github.com/MDU-PHL/mutamr). The entire package is available as a conda package.

Validation dataset

We included three data sources in our validation dataset (Supplementary Figure 2). Firstly, we included 2005 sequences generated at Microbiological Diagnostic Unit Public Health Laboratory (MDU PHL) from 2018 – 2022 and with phenotypic (DST) data generated at the Mycobacterial Reference Laboratory (MRL) (Supplementary Methods). Secondly, 13,777 publicly available sequences were included, selected to be representative of all major global lineages and with phenotypic data available for first-line agents and a subset also having data available Group A, B and Group C drugs[22–25]. Lastly, in order to assess the accurate detection of variants by tbtAMR, we also used simulated genomic data.

195 Paired-end sequence data was generated using TreeToReads software[26–29], directly from the H37rV
196 reference strain (RefSeq accession NC_000962.3) or following introduction of 2–4 variants per 10,000
197 bases at known positions with an error-profile representative of the NextSeq500 instruments at MDU
198 PHL (n= 1830). As resistance testing aims to identify resistant subpopulations of Mtb that may emerge
199 during treatment, these simulated reads were mixed (Supplementary Figure 3) to simulate different
200 allelic frequencies (0.01 to 0.99) across a range of average genome depth (20x to 200x – incremented
201 by 20x).

202 **Measuring the accuracy of Mtb genomic sequence recovery**

203 For simulated genomes, we tested the ability of the tool to correctly identify introduced variants using
204 a minimum depth of 10X and 20X, meaning that to make a base call there was least 10 or 20 reads
205 covering each position. Reads were aligned back to the reference genome and the correctness of base
206 calling by mutAMR was assessed (Supplementary Figure 3).

207 **Analysis of phylogenetic lineage and inferred phenotypic DST from genomic data**

208 We looked at how well inferred phenotype from detected genomic resistance mechanisms (single
209 nucleotide polymorphisms [SNPs], insertions or deletions) correlated with phenotypic DST results. To
210 quantify this, we compared the phenotypes inferred from the presence of SNPs, insertions or
211 deletions in the genome sequences with actual phenotypic susceptibility profiles, and calculated
212 sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV) and accuracy
213 recorded (Supplementary Table 2). When discrepancies between inferred-resistance and the reported
214 phenotypic resistance are identified, these would ideally be resolved with additional phenotypic
215 testing. However, since this was not possible for the discrepancies observed in the publicly available
216 datasets, discrepancy resolution could not be undertaken and any such results were included as
217 discrepancies in performance assessment. tbtAMR also determines phylogenetic lineage, including
218 lineages 1–9 as well as animal adapted species/sub-species[11]. We also assessed the concordance
219 between phylogenetic lineage determined by tbtAMR and the phylogenetic lineage that was supplied
220 as part of the public dataset[25] (Supplementary Figure 3).

221 # Results

222 **Validation of variant detection**

223 *Accuracy of simulated sequence recovery by mutAMR*

224 We assessed the ability of mutAMR to recover known (introduced) SNPs in simulated sequences, and
225 the minimum average sequence depth for reliable results. Most tools used for detecting variants use

226 a minimum depth criteria of 10X[30,31], meaning that there must be at least 10 reads aligned to a
227 particular position in the reference genome. However, when we investigated SNPs identified across a
228 range of average sequence depths and allelic frequencies, several false positive (FP) SNPs (mean 6.6,
229 95% confidence interval (CI) 3.9-9.3) were observed when using the standard criteria of 10X, which
230 may result in false reporting of resistance mechanisms. Increasing the minimum depth for base calling
231 to 20X improved performance, reducing the number of FP SNPs to a mean of 1.1 (CI 0.7-1.5), with little
232 impact on the percentage of introduced SNPs recovered (Figure 3A-B). We further examined the
233 impact of varying average genome depth and allelic frequency on the ability of mutAMR to accurately
234 identify mutations in the simulated sequences. Where average genome depth was less than 40X and
235 allelic frequency was less than 10%, sensitivity of variant identification was low at only 2.1% (CI 1.48%-
236 2.66%) and False Discovery Rate (FDR) was 41.1% (CI 31.1%-51.2%). However, maintaining 20 reads
237 minimum depth for base-calling, an average genome depth of ≥40X and allele frequencies ≥10% gave
238 the best balance (Figure 3C-D), with an overall variant identification sensitivity of 96.3% (CI 95.5%-
239 97.1%) and low FDR of 0.21% (CI 0.11%-0.31%). Detection of other variant types, such as deletions, by
240 mutAMR was consistent with TB-profiler and Mykrobe, using public available sequences. Hence, the
241 minimum acceptable read depth for identification of variants was set to ≥40X and minimum allele
242 frequency of 10%.

243

244 **Accuracy of phylogenetic lineage calling by tbtAMR**

245 tbtAMR accurately identified lineages compared to lineages reported from public datasets[25], (98.5%
246 concordance 871/884). Of the 13 discordant results (Supplementary Table 1), 12 were cases where
247 tbtAMR identified two different phylogenetic lineages in the sequence. In each case, the reported
248 lineage was one of the lineages detected by tbtAMR, therefore these likely reflect mixed sequences,
249 where tbtAMR was able to identify the minor allele.

250

251 **Inference of phenotypic AMR from genomic data**

252 Many studies have explored the inference of phenotypic AMR from genomic data using various
253 tools[10,14,24,32]. We do not aim to duplicate these high-quality studies in this study, but rather to assess
254 the suitability of tbtAMR for inference and reporting of Mtb AMR to acceptable standards for clinical
255 and public health purposes.

256 We began with a dataset of 15,323 sequences from both publicly available datasets[22–25] and sequences
257 generated in-house using sequencing workflows accredited to ISO standards. Implementing the

258 quality control thresholds and limits of detection outlined above, we excluded 608 sequences based
259 on quality metrics (Supplementary Figure 4), leaving 14,715 sequences of sufficient quality to include
260 in the validation of tbtAMR for inference of phenotypic AMR.

261 The performance of tbtAMR in inferring phenotypic resistance to first-line drugs was excellent, with
262 high performance across all first-line drugs (Table 2 and Figure 3) and was consistent with the
263 performance of both TB-profiler and Mykrobe (Supplementary Figure 5A-B). Lower positive predictive
264 values for ethambutol (81.1% CI 79.4%-82.7%) were due to detection of variants in *embB* at codon
265 360 where phenotype reported was susceptible. This observation is likely due to known challenges in
266 consistency of ethambutol DST and are consistent with other studies[10,14,24,32]. Prediction of
267 pyrazinamide resistance also has lower sensitivity than other first-line drugs (80.0% CI 75.4%-84.0%)
268 and is likely due both challenges in pyrazinamide DST and less robust understanding of the molecular
269 mechanisms leading to pyrazinamide resistance[5].

270 tbtAMR predicted resistance to groups A, B and C (Table 1) with performance comparable to other
271 studies[10,14,24,32] (Supplementary Figure 5A-B). The high number of false-negative and false-positive
272 results for prediction of phenotypic resistance (Table 2) are likely to be due to incomplete
273 understanding of the mechanisms of resistance to these drugs, global paucity of data (phenotypic and
274 genotypic) and challenging or inconsistent phenotypic testing for some drugs. Although tbtAMR
275 performance was consistent with other tools and use of updated catalogue improved performance
276 compared on the previous WHO catalogue, prediction of bedaqualine, delaminid, linezolid and
277 clofazimine was not of sufficient performance to allow for routine reporting of resistance to these
278 drugs from genomic data in a CPHL setting, due to the high rates of both false positive and false
279 negative results (Supplementary Figure 6).

280 **Implementation process**

281 As AMR genotypes and inferred antibiograms for Mtb were new to most clinicians, additional
282 interpretive comments were created for inclusion in reports – for example, explaining the difference
283 between confidence level and level of resistance in the inferred phenotype (Supplementary Methods
284 and Supplementary Files 1). Clinical and laboratory colleagues were consulted about the draft report
285 design to ensure the formatting and explanations were clear and unambiguous in our local setting.
286 Once validation and report design were complete, Mtb genomic AMR inference using the tbtAMR tool
287 was implemented into existing genomic workflows, including standard operating procedure
288 documentation (SOP), staff training, incorporation into the laboratory information management
289 system, and testing of reporting and feedback loops (e.g. reporting of samples that failed QC).

290 To address any discrepancies in phenotypic and genotypic results, a multidisciplinary expert panel for
291 Mtb AMR was formed, meeting monthly to discuss individual cases and provide expert
292 recommendations to clinicians where discrepancies were identified. This provides the opportunity for
293 ongoing dialogue between the wet-lab microbiologists, clinicians, epidemiologists, medical
294 microbiologists and bioinformaticians to monitor for AMR determinants of interest, discuss individual
295 cases and discrepancies, and undertake prospective validation of novel or uncharacterised AMR
296 determinants.

297 In preparation for ISO-equivalent accreditation, revalidation and reverification processes were also
298 defined for tbtAMR (Supplementary Methods). Subsequently, the Mtb genomic AMR inference
299 workflow has now been accredited to ISO 15189:2022 standards.

300 # Discussion

301 We have reported here the development of a custom Mtb AMR analysis and reporting pipeline for use
302 in the clinical and public health microbiology setting, with easily customisable interpretative and
303 classification criteria to maximise the balance between PPV and NPV for prediction of AMR in Mtb.
304 Performance of inferred AMR using our process is consistent with or exceeds the performance of other
305 tools[10,32] (Supplementary Figure 2). This pipeline has been robustly validated, implemented and
306 accredited to ISO standards in our laboratory, with reverification strategies for ongoing improvement,
307 and reporting processes put in place.

308 The increased availability of WGS in clinical and public health labs, particularly since the COVID
309 pandemic, has resulted in a massive increase in sequencing capacity and data[33]. Mahé et al. recently
310 published a review of bioinformatics in public health laboratory setting, highlighting that there are still
311 significant skills gaps in the CPHL workforce to use, develop and/or maintain open-source free
312 software for analysis of sequence data, and that the proprietary alternatives are expensive[34]. To
313 support the development of open pathogen genomics analysis in PHLs, they recommend increasing
314 the technical workforce to ameliorate the skills gap, although, this can be difficult to achieve in
315 practice due to funding limitations, and the difficulties in upskilling new or existing staff, which can be
316 time and resource intensive.[2931]

317 While open-source software is accessible and free to users, it can be challenging to implement in a
318 PHL setting. These tools are mostly developed and maintained by independent researchers, many in
319 academia, written to suit the needs or interests of the researcher, and may not be maintained or
320 frequently updated[35]. Software may be difficult to install, difficult to run, or not meet the needs of the
321 PHL. Furthermore, maintaining and implementing such tools in an ISO-accredited laboratory structure
322 can be time-consuming and difficult, requiring significant staff time and resources[9], as laboratories
323 cannot be selective about the updates adopted if the tool is maintained elsewhere. These difficulties
324 can be exacerbated in resource-limited settings such as low- and middle-income countries[34].

325 Whilst this approach remains iterative, we have developed an approach that is independent of
326 developer interpretations, and instead relies upon user-supplied criteria which can be tailored as
327 required. This approach is beneficial in a CPHL environment, as it makes updating to new versions or
328 modification of interpretation or classification criteria a straightforward process that can be
329 undertaken by a non-developer (such as a medical microbiologist), removing the requirement for a
330 highly skilled bioinformatician to enact these changes. It also allows for flexibility to apply setting-
331 specific criteria, tailoring results to the target audience. For example, a set of criteria can be applied

332 to the same input data for clinical reporting (high certainty for patient management), whilst a different
333 set of criteria may be applied to capture mechanisms for epidemiological purposes or prospective
334 validation studies (lower certainty for non-clinical and explorative purposes), with no need to modify
335 the underlying code. In addition, due to minimal additional dependencies, tbtAMR can easily be
336 installed and run on a laptop or be incorporated into large and complex workflows easily.

337 AMR determinants are often reported as a binary 'present/absent' variable, inferring a 'resistant' or
338 'susceptible' phenotype respectively, but this does not always accurately reflect the complexity of
339 genotype-phenotype correlations. It is important to ensure that mutations used to infer resistance
340 have solid evidence supporting a role as genomic determinants of resistance. The new WHO catalogue
341 and documentation provide some information for mutations which are likely to confer 'borderline'
342 resistance such as the *rpoB* His445Asn variant[15] and confidence levels which may be used to aid
343 interpretation. The structure of tbtAMR allows us to provide meaningful reports to clinicians (based
344 on high confidence mutations), capturing the complexity of genotype-phenotype relationships, whilst
345 still being able to prospectively monitor and validate additional mutations (mutations from all
346 confidence levels). This is particularly relevant for group A, B and C drugs including bedaquiline,
347 delaminid and linezolid, where a lack of data (locally and globally) still limits our ability to adequately
348 validate inference of resistance to these drugs for reporting in a CPHL. Even for drugs which are well
349 studied, discrepancies can still occur, and further research and iterative validation will only improve
350 the performance of tbtAMR and other tools used by Mtb genomics community, ultimately leading to
351 improved patient outcomes.

352 tbtAMR represents a flexible yet robust and accurate pipeline that leverages the validation undertaken
353 here and other quality studies[14,22,24,32] to achieve the best balance between PPV and NPV for patient
354 management, and simplified outputs for clinical and public health microbiology. We have accredited
355 this pipeline to ISO-standards in our laboratory and have provided the validation methods and dataset
356 for use by other laboratories endeavouring to make a similar transition. Furthermore, the flexible
357 design of tbtAMR allows for users to supply custom catalogues and/or interpretative criteria and use
358 input generated from a variety of sequencing technologies. Our results also highlight areas where
359 further research can enhance our knowledge of AMR determinants in Mtb, with the correlation
360 between AMR determinants and phenotypes with new second-line Mtb drugs being a key priority.
361 Further enhancements of diagnostic processes, such as sequencing directly from samples, will allow
362 us to maximise the benefit of genomic technologies for patient management. Ongoing engagement
363 with stakeholders, including researchers, clinicians and public health officials, ensures that the
364 processes established can be improved and refined where needed to continue to improve patient
365 management and inform timely public health responses.

366 **Figures and tables**

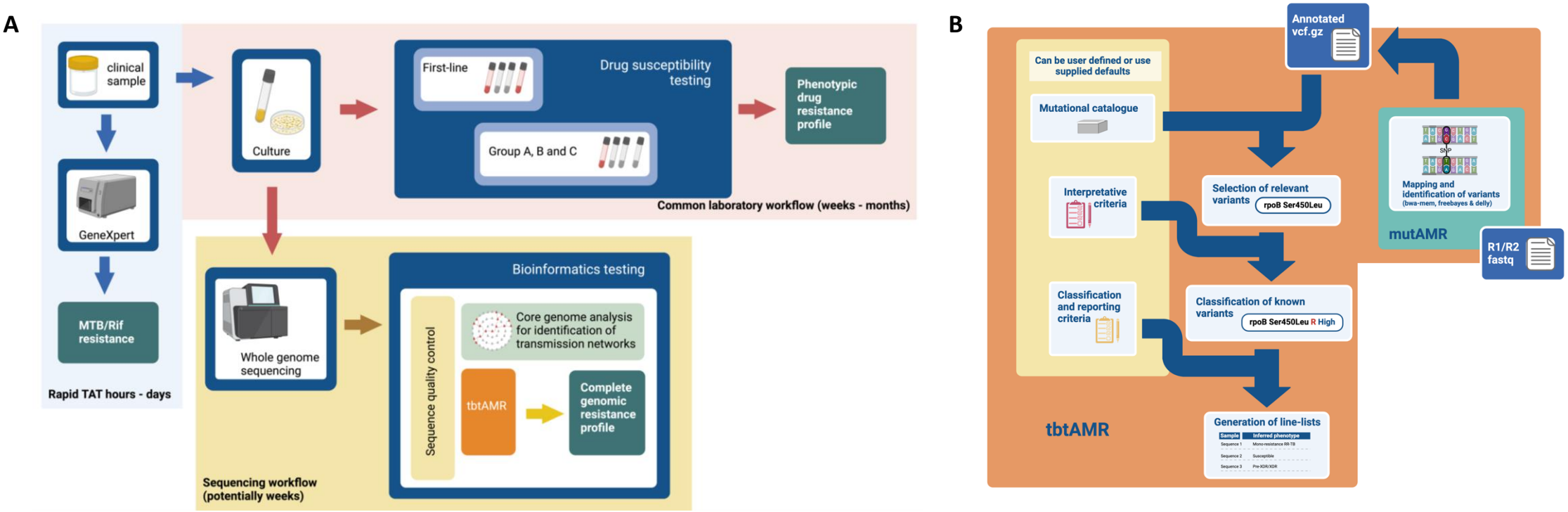
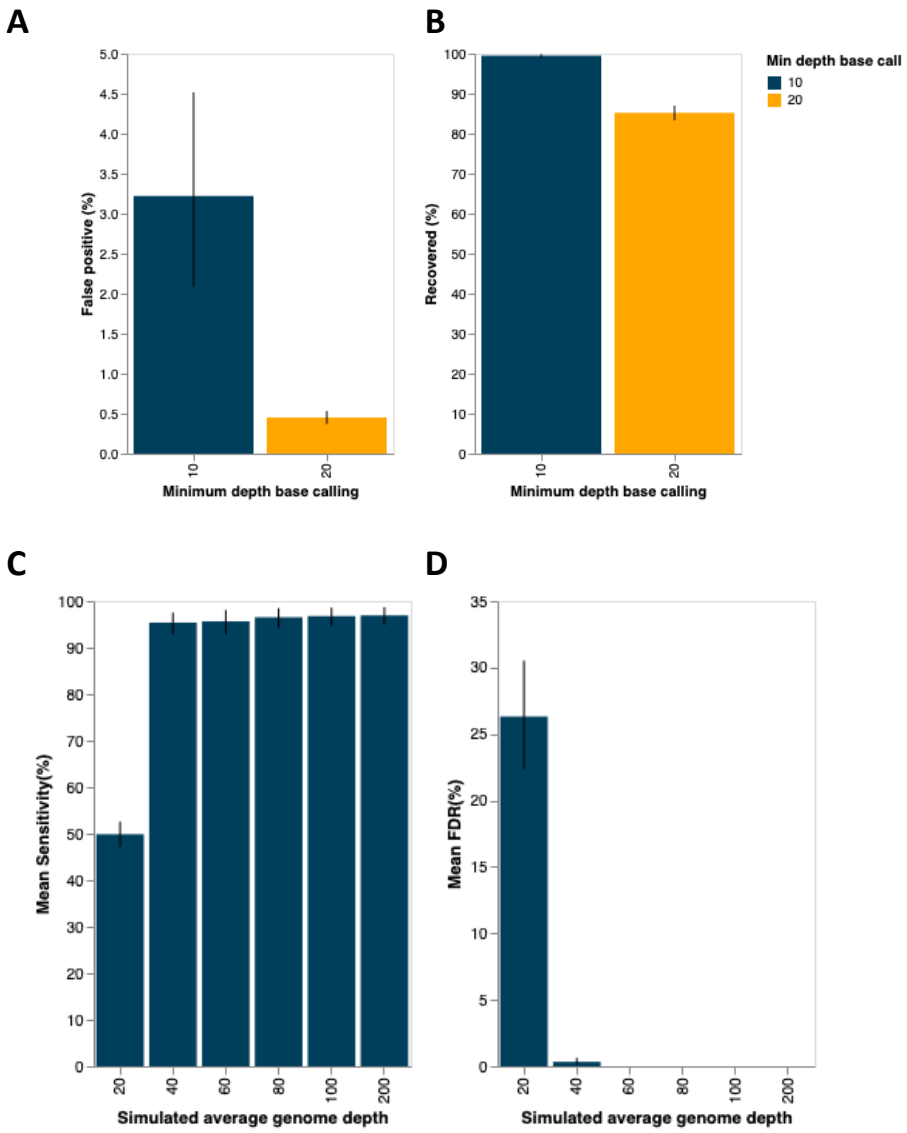


Figure 1 – Incorporation of tbtAMR into *M. tuberculosis* diagnosis and testing workflow in a reference laboratory.

A. Diagnostic testing pathways for Mtb: *M. tuberculosis* and potential rifampicin resistance can be identified rapidly using GeneXpert (closed-system PCR for most common rifampicin resistance mutations) in a diagnostic laboratory (turnaround time hours–days). Traditional culture methods require samples to be inoculated to solid media (agar slopes) and/or liquid media (most commonly MGIT). Cultures may take up to 6-8 weeks to flag positive. From positive culture, phenotypic drug susceptibility testing is then commenced, taking up to two weeks to be completed, or longer if cascade testing is used (Group A, B and C drugs only set up when resistance to first-line drugs is identified). Total turnaround times from sample to culture and phenotypic DST results is usually 6 weeks-3 months. If optimised, whole genome sequencing can potentially reduce this turnaround time and provide a comprehensive resistance profile without the need for further testing, whilst also providing important contributions to public health control efforts through contact tracing. **B. tbtAMR pipeline:** tbtAMR inputs can be either an annotated variant file (vcf format, from short-read or long-read sequences) based on H37rV v3 reference genome, or paired-end short-read data (fastq file), making tbtAMR sequencing platform agnostic. tbtAMR uses version 2 of the WHO mutational catalog as default with accompanying rules for interpretation, classification and reporting. Modified versions of these can also be supplied by the user where required.

Abbreviations: Mtb, *Mycobacterium tuberculosis*; Rif, rifampicin; TAT, turnaround time. Variant nomenclature example: “rpoB_Ser450Leu” denotes a change in the amino acid from serine to leucine at position 450 in the RpoB protein sequence.



380

381 **Figure 2 – Performance of mutAMR to recover SNPs from Mtb sequences**

382 To assess the performance of mutAMR, implemented by tbtAMR to accurately recover mutations in
383 Mtb sequences, simulated genome data across a range of average genome depths from 10X to 200X
384 at allelic frequencies 10-100% were assessed. The impact of minimum read depth for base calling at
385 an individual position in the genome (depth 10 or 20) was assessed for: the percentage of false positive
386 SNP calls (FP) identified (**Panel A**) and the percentage of introduced SNPs recovered by mutAMR
387 (**Panel B**). Using a minimum-depth criteria of 20X for individual base calling, **C.** SNP calling sensitivity
388 and **D.** false discovery rate (FDR) were assessed average genome depth using average genome depth
389 of coverage of ≥40X and allelic frequency ≥10% (error bars indicate 95% confidence interval).

390

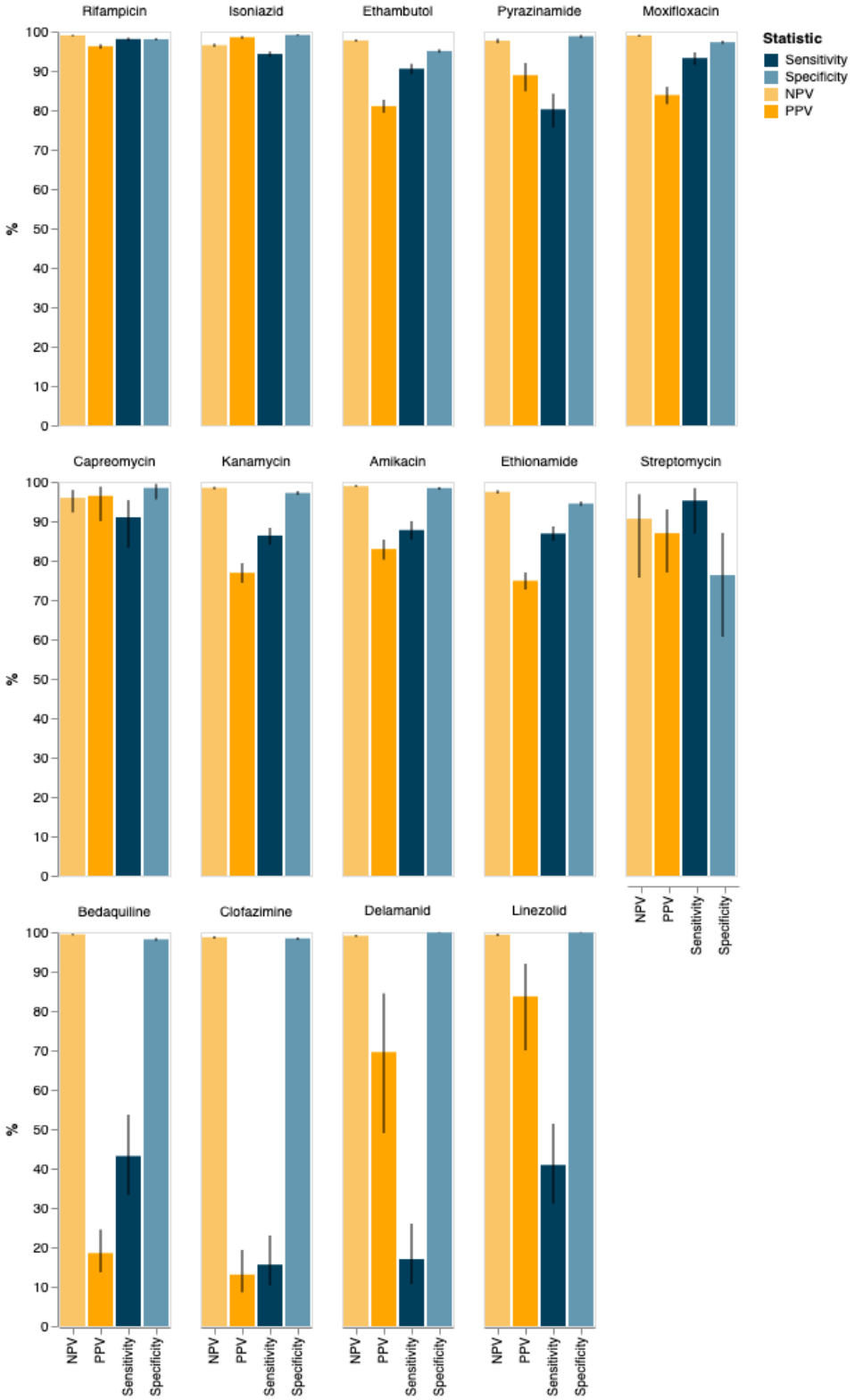


Figure 3 – Performance of tbtAMR inference of phenotypic susceptibility

Inferred phenotypes predicted by the tbtAMR pipeline were compared to phenotypic DST data, true positive (TP) was recorded when a mutation was reported and the phenotype was resistant; a true

396 negative (TN) was recorded when no mutation was reported and the phenotype was susceptible; a
397 false positive (FP) was recorded when a mutation was reported, but the phenotype was susceptible,
398 and a false negative (FN) was recorded when no mutation was reported, and the phenotype was
399 resistant. Sensitivity, specificity, positive and negative predictive values calculated and expressed as a
400 percentage (error bars indicate 95% confidence intervals). DST, direct susceptibility testing; PPV,
401 positive predictive value; NPV, negative predictive value.

402

403 **Table 1 – Antimycobacterial drugs used to treat TB**[2]

Type	Antimycobacterial drugs
First-line	Rifampicin, isoniazid, ethambutol, pyrazinamide
Second-line	
Group A for MDR TB	Fluoroquinolones (moxifloxacin or levofloxacin), bedaquiline, linezolid
Group B for MDR TB	Clofazimine, cycloserine or teridazone
Group C for MDR TB	Delaminid, imipenem-cilastatin or meropenem, ethambutol*, pyrazinamide*, amikacin or streptomycin, ethionamide, para-aminosalicylic acid
Others **(Not recommended)**	Kanamycin, capreomycin, gatifloxacin, clavulanic acid

404 * where DST indicates susceptibility

405

406

407 **Table 2 – Performance of tbtAMR pipeline to predict phenotypic susceptibility for first- and second-line antimycobacterial drugs**

408

409

Drug (n)	Sensitivity (%,95%CI)	Specificity (%,95%CI)	PPV (%,95%CI)	NPV (%,95%CI)
Rifampicin (n = 12,505)	98.1 (97.6,98.4)	98.1 (97.7,98.3)	96.3 (95.7,96.8)	99.0 (98.8,99.2)
Isoniazid (n = 12,545)	94.2 (93.5,94.9)	99.1 (98.9,99.3)	98.5 (98.1,98.8)	96.5 (96.0,96.9)
Ethambutol (n = 10,348)	90.6 (89.2,91.8)	95.0 (94.5,95.5)	81.0 (79.3,82.6)	97.7 (97.4,98.0)
Pyrazinamide (n = 3,174)	80.8 (76.4,84.5)	98.8 (98.3,99.1)	89.2 (85.4,92.2)	97.6 (96.9,98.1)
Bedaquiline (n = 9.339)	43.2 (33.3,53.6)	98.2 (97.9,98.5)	18.5 (13.8,24.4)	99.5 (99.3,99.6)
Linezolid (n = 8,017)	40.9 (31.2,51.4)	99.9 (99.8,100.0)	83.7 (70.0,91.9)	99.3 (99.1,99.5)
Moxifloxacin (n = 7,349)	93.3 (91.5,94.7)	97.3 (96.9,97.7)	84.1 (81.8,86.1)	98.9 (98.7,99.2)
Clofazimine (n = 8,591)	15.5 (10.3,22.7)	98.4 (98.1,98.7)	13.1 (8.6,19.3)	98.7 (98.4,98.9)
Amikacin (n = 10,054)	88.0 (85.6,90.0)	98.4 (98.2,98.7)	83.5 (80.9,85.8)	98.9 (98.7,99.1)
Delamanid (n = 8,882)	17.0 (10.8,25.9)	99.9 (99.8,100.0)	69.6 (49.1,84.4)	99.1 (98.9,99.3)
Ethionamide (n = 9,112)	86.9 (85.1,88.5)	94.5 (93.9,95.0)	74.9 (72.8,76.9)	97.4 (97.1,97.8)
Streptomycin (n = 123)	95.1 (88.0,98.1)	78.6 (64.1,88.3)	89.5 (81.3,94.4)	89.2 (75.3,95.7)
Capreomycin (n = 304)	92.7 (86.3,96.3)	98.5 (95.6,99.5)	97.1 (91.9,99.0)	96.0 (92.3,98.0)
Kanamycin (n= 10,397)	86.6 (84.4,88.5)	97.2 (96.8,97.5)	77.4 (74.9,79.7)	98.5 (98.2,98.7)

410

Ethics approval

Ethical approval was received from the University of Melbourne Human Research Ethics Committee (study number 1954615).

Acknowledgements

We thank all the laboratory staff at the Mycobacterial Reference Laboratory, Victorian Infectious Disease Reference Laboratory (VIDRL) and the Microbiological Diagnostic Unit Public Health Laboratory (MDU PHL) who undertook receipt, diagnostics, culture, DNA extraction and sequencing of the material used in this publication; the diagnostic laboratories who diligently provide samples for testing; and the Victorian TB Program, and Department of Health Victoria who fund these services and engage enthusiastically with the laboratories. We also thank the developers and maintainers of the open-source software used and cited in this publication. This work was funded by the Australian National Health and Medical Research Council (NHMRC), Medical Research Futures Fund (FSPGN00049), and Investigator Grant (GNT1196103) to BPH.

Data sharing

Raw sequence data generated at MDU PHL has been uploaded to the Sequence Read Archive under BioProject PRJNA857537. Accessions for all sequences generated and deposited in public repositories can be found in Supplemental File 1.

All results used for interpretation and generation of results, including phenotypic DST results and genomic results used in this study are available in Supplemental File 1 and readily available on GitHub. Further instructions for installing and running tbtamr as well as the complete code base are also available on GitHub (https://github.com/MDU-PHL/tbtamr).

Author Contributions

Kristy Horan – developed software, study design, data analysis, manuscript writing and editing.

Linda Viberg – performed DNA extractions, reviewed data, edited manuscript

Susan Ballard – contributed to study design, laboratory supervision, edited manuscript

Maria Globan – laboratory supervision, reviewed data, edited manuscript

Katherine Bond – study design, laboratory supervision, edited manuscript

Jessica R Webb – created figures, edited manuscript

Thinley Dorji – created figures, edited manuscript

Deborah Williamson – contributed to study design, edited manuscript

Michelle Sait – performed genomic sequencing, edited manuscript

Ee Laine Tay – reviewed data, edited manuscript

Justin T Denholm – contributed to study design, edited manuscript

Torsten Seemann – study design, software development, validation protocols, edited manuscript

Benjamin P Howden – funding, laboratory supervision, contributed to study design, edited manuscript

Norelle L Sherry – study design, contributed to software development, reporting templates, data analysis, manuscript writing and editing.

References

1 Bennett J, Dolin R, Blaser MJ. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases, 9th edn. .

2 WHO Consolidated Guidelines on Tuberculosis, Module 4: Treatment - Drug-Resistant Tuberculosis Treatment. https://www.who.int/publications-detail-redirect/9789240007048 (accessed March 5, 2022).

3 WHO consolidated guidelines on tuberculosis. Module 4: treatment - drug-susceptible tuberculosis treatment. Geneva: World Health Organization; https://iris.who.int/bitstream/handle/10665/353829/9789240048126-eng.pdf?sequence=1 (accessed Oct 1, 2024).

4 TB guidelines. https://www.who.int/publications/digital/global-tuberculosis-report-2021/featured-topics/tb-guidelines (accessed March 5, 2022).

5 Carrol K, Pfaller M, Landry M. Manual of Clinical Microbiology, 12th edn. Wiley, 2019.

6 Denholm JT, de Vries G, Anthony R, *et al.* Considering best practice standards for routine whole-genome sequencing for TB care and control. *IJTLD Open* 2024; **1**: 431–6.

7 Donnan EJ, Marais BJ, Coulter C, *et al.* The use of whole genome sequencing for tuberculosis public health activities in Australia: a joint statement of the National Tuberculosis Advisory Committee and Communicable Diseases Genomics Network. *Commun Dis Intell (2018)* 2023; **47**. DOI:10.33321/cdi.2023.47.8.

8 14:00-17:00. ISO 15189:2022. ISO. https://www.iso.org/standard/76677.html (accessed July 4, 2023).

9 Ballard SA, Sherry NL, Howden BP. Public health implementation of pathogen genomics: the role for accreditation and application of ISO standards. *Microbial Genomics* 2023; **9**: 001097.

10 Hunt M, Bradley P, Lapierre SG, *et al.* Antibiotic resistance prediction for Mycobacterium tuberculosis from genome sequence data with Mykrobe. *Wellcome Open Res* 2019; **4**: 191.

11 Phelan JE, O'Sullivan DM, Machado D, *et al.* Integrating informatics tools and portable sequencing technology for rapid detection of resistance to anti-tuberculous drugs. *Genome Medicine* 2019; **11**: 41.

12 ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads | Microbiology Society. https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000131 (accessed July 4, 2023).

13 Walker TM, Miotto P, Köser CU, *et al.* The 2021 WHO catalogue of Mycobacterium tuberculosis complex mutations associated with drug resistance: a genotypic analysis. *The Lancet Microbe* 2022; **3**: e265–73.

14 Shaw B, von Bredow B, Tsan A, Garner O, Yang S. Clinical Whole-Genome Sequencing Assay for Rapid Mycobacterium tuberculosis Complex First-Line Drug Susceptibility Testing and Phylogenetic Relatedness Analysis. *Microorganisms* 2023; **11**: 2538.

15 Catalogue of mutations in Mycobacterium tuberculosis complex and their association with drug resistance. https://www.who.int/publications-detail-redirect/9789240028173 (accessed Nov 23, 2022).

16 MDU-PHL/tbtamr. 2024; published online Oct 23. https://github.com/MDU-PHL/tbtamr (accessed Nov 12, 2024).

17 Vasimuddin Md, Misra S, Li H, Aluru S. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In: 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). 2019: 314–24.

18 Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012; published online July 20. http://arxiv.org/abs/1207.3907 (accessed Oct 1, 2024).

19 Cingolani P, Platts A, Wang LL, *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w ^{1118} ; iso-2; iso-3. *Fly* 2012; **6**: 80–92.

20 DELLY: structural variant discovery by integrated paired-end and split-read analysis | Bioinformatics | Oxford Academic. https://academic.oup.com/bioinformatics/article/28/18/i333/245403?login=false (accessed Oct 1, 2024).

21 Danecek P, Bonfield JK, Liddle J, *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* 2021; **10**: giab008.

22 Quantitative measurement of antibiotic resistance in Mycobacterium tuberculosis reveals genetic determinants of resistance and susceptibility in a target gene approach. *Nat Commun* 2024; **15**: 488.

23 Ezewudo M, Borens A, Chiner-Oms Á, *et al.* Integrating standardized whole genome sequence analysis with a global Mycobacterium tuberculosis antibiotic resistance knowledgebase. *Sci Rep* 2018; **8**: 15382.

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

25 Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *New England Journal of Medicine* 2018; **379**: 1403–15.

26 Huang W, Li L, Myers JR, Marth GT. ART: a next-generation sequencing read simulator. *Bioinformatics* 2012; **28**: 593–4.

27 McTavish EJ. TreeToReads. 2022; published online Sept 14. https://github.com/snacktavish/TreeToReads (accessed Nov 16, 2022).

28 Rambaut A, Grassly NC. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput Appl Biosci* 1997; **13**: 235–8.

29 Sukumaran J, Holder MT. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* 2010; **26**: 1569–71.

30 Seemann T. Snippy. 2022; published online March 19. https://github.com/tseemann/snippy (accessed March 26, 2022).

31 Phelan J. jodyphelan/TBProfiler. 2022; published online Feb 15. https://github.com/jodyphelan/TBProfiler (accessed Feb 28, 2022).

32 Mahé P, El Azami M, Barlas P, Tournoud M. A large scale evaluation of TBProfiler and Mykrobe for antibiotic resistance prediction in Mycobacterium tuberculosis. *PeerJ* 2019; **7**: e6857.

33 Black A, MacCannell DR, Sibley TR, Bedford T. Ten recommendations for supporting open pathogen genomic analysis in public health. *Nat Med* 2020; **26**: 832–41.

34 Helmy M, Awad M, Mosa KA. Limited resources of genome sequencing in developing countries: Challenges and solutions. *Applied & Translational Genomics* 2016; **9**: 15–9.

35 Siepel A. Challenges in funding and developing genomic software: roots and remedies. *Genome Biology* 2019; **20**: 147.

Click here to access/download
Supplementary Materials
Horan_et_al_2024_supp.docx

Click here to access/download
Supplementary Materials
Horan_et_al_reports_mtb.pdf

Click here to access/download
Supplementary Materials
Horan_et_al_supplemental_data_file_2.xlsx