bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Bridging the TB data gap: *in silico* extraction of rifampicin-resistant tuberculosis diagnostic test results from whole genome sequence data

Kamela C S Ng Corresp., [1, 2], **Jean Claude S Ngabonziza** [1, 3], **Pauline Lempens** [1], **Bouke C de Jong** [1], **Frank van Leth** [2, 4], **Conor J Meehan** [1, 5]

[1] Mycobacteriology Unit, Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium

[2] Amsterdam Institute for Global Health and Development, University of Amsterdam, Amsterdam, The Netherlands

[3] Department of Biomedical Services, Rwanda Biomedical Center, Kigali, Rwanda

[4] Department of Global Health, Amsterdam UMC, Location Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

[5] School of Chemistry and Biosciences, University of Bradford, Bradford, United Kingdom

Corresponding Author: Kamela C S Ng
Email address: kng@itg.be

Background. *Mycobacterium tuberculosis* rapid diagnostic tests (RDTs) are widely employed in routine laboratories and national surveys for detection of rifampicin-resistant (RR)-TB. However, as next generation sequencing technologies have become more commonplace in research and surveillance programs, RDTs are being increasingly complemented by whole genome sequencing (WGS). While comparison between RDTs is difficult, all RDT results can be derived from WGS data. This can facilitate continuous analysis of RR-TB burden regardless of the data generation technology employed. By converting WGS to RDT results, we enable comparison of data with different formats and sources particularly for low and middle income high TB burden countries that employ different diagnostic algorithms for drug resistance surveys. This allows national TB control programs (NTPs) and epidemiologists to utilize all available data in the setting for improved RR-TB surveillance. **Methods.** We developed the Python-based MTB Genome to Test (MTBGT) tool that transforms WGS-derived data into laboratory-validated results of the primary RDTs – Xpert MTB/RIF, XpertMTB/RIF Ultra, GenoType MDRTB*plus* v2.0, and GenoscholarNTM+MDRTB II. The tool was validated through RDT results of RR-TB strains with diverse resistance patterns and geographic origins and applied on routine-derived WGS data. **Results.** The MTBGT tool correctly transformed the SNP data into the RDT results and generated tabulated frequencies of the RDT probes as well as rifampicin susceptible cases. The tool supplemented the RDT probe reactions output with the RR-conferring mutation based on identified SNPs. The MTBGT tool facilitated continuous analysis of RR-TB and Xpert probe reactions from different platforms and collection periods in Rwanda. **Conclusion.** Overall, the MTBGT tool allows low and middle income countries

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

to make sense of the increasingly generated WGS in light of the readily available RDT results, and assess whether currently implemented RDTs adequately detect RR-TB in their setting. With its feature to transform WGS to RDT results and facilitate continuous RR-TB data analysis, the MTBGT tool may bridge the gap between and among data from periodic surveys, continuous surveillance, research, and routine tests, and may be integrated within the existing national connectivity platform for use by the NTP and epidemiologists to improve setting-specific RR-TB control. The MTBGT source code and accompanying documentation is available at https://github.com/KamelaNg/MTBGT .

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1 # Bridging the TB data gap: *in silico* extraction of rifampicin-
2 resistant tuberculosis diagnostic test results from whole
3 genome sequence data
4
5
6 Kamela Charmaine S. Ng[1,2*], Jean Claude S. Ngabonziza[1,3], Pauline Lempens[1], Bouke Catherine
7 de Jong[1], Frank van Leth[2,4], Conor Joseph Meehan[1,5]
8
9 [1]Mycobacteriology Unit, Department of Biomedical Sciences, Institute of Tropical Medicine,
10 Antwerp, Belgium
11 [2]Amsterdam Institute for Global Health and Development, University of Amsterdam,
12 Amsterdam, The Netherlands
13 [3]Department of Biomedical Services, Rwanda Biomedical Center, Kigali, Rwanda
14 [4]Department of Global Health, Amsterdam UMC location Academic Medical Center,
15 Amsterdam, The Netherlands
16 [5]School of Chemistry and Biosciences, University of Bradford, Bradford, UK
17
18
19 Corresponding Author:
20 Kamela C. S. Ng
21 Nationalestraat 155, 2000 Antwerp, Belgium
22 Email address: kng@itg.be
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

40 ## Abstract
41 **Background.** *Mycobacterium tuberculosis* rapid diagnostic tests (RDTs) are widely employed
42 in routine laboratories and national surveys for detection of rifampicin-resistant (RR)-TB.
43 However, as next generation sequencing technologies have become more commonplace in
44 research and surveillance programs, RDTs are being increasingly complemented by whole
45 genome sequencing (WGS). While comparison between RDTs is difficult, all RDT results can be
46 derived from WGS data. This can facilitate continuous analysis of RR-TB burden regardless of
47 the data generation technology employed. By converting WGS to RDT results, we enable
48 comparison of data with different formats and sources particularly for low and middle income
49 high TB burden countries that employ different diagnostic algorithms for drug resistance
50 surveys. This allows national TB control programs (NTPs) and epidemiologists to utilize all
51 available data in the setting for improved RR-TB surveillance.
52 **Methods.** We developed the Python-based MTB Genome to Test (MTBGT) tool that transforms
53 WGS-derived data into laboratory-validated results of the primary RDTs – Xpert MTB/RIF,
54 XpertMTB/RIF Ultra, GenoType MDRTB*plus* v2.0, and GenoscholarNTM+MDRTB II. The
55 tool was validated through RDT results of RR-TB strains with diverse resistance patterns and
56 geographic origins and applied on routine-derived WGS data.
57 **Results.** The MTBGT tool correctly transformed the SNP data into the RDT results and
58 generated tabulated frequencies of the RDT probes as well as rifampicin susceptible cases. The
59 tool supplemented the RDT probe reactions output with the RR-conferring mutation based on
60 identified SNPs. The MTBGT tool facilitated continuous analysis of RR-TB and Xpert probe
61 reactions from different platforms and collection periods in Rwanda.
62 **Conclusion.** Overall, the MTBGT tool allows low and middle income countries to make sense of
63 the increasingly generated WGS in light of the readily available RDT results, and assess whether
64 currently implemented RDTs adequately detect RR-TB in their setting. With its feature to
65 transform WGS to RDT results and facilitate continuous RR-TB data analysis, the MTBGT tool
66 may bridge the gap between and among data from periodic surveys, continuous surveillance,
67 research, and routine tests, and may be integrated within the existing national connectivity
68 platform for use by the NTP and epidemiologists to improve setting-specific RR-TB control. The
69 MTBGT source code and accompanying documentation is available at
70 https://github.com/KamelaNg/MTBGT.
71

72 ## Introduction
73 Resistance to rifampicin (RIF), the most potent anti-tuberculosis (TB) drug, hampers TB control.
74 Rifampicin-resistant (RR)-TB persists as an urgent public health crisis as only 29% of estimated
75 RR-TB patients world-wide were detected and notified in 2017. Further, 18% of previously
76 treated TB patients were found to have RR-TB (WHO 2018). The WHO endorsed the following
77 rapid molecular RR-TB diagnostic test (RDTs) to address this concern: Xpert MTB/RIF (Xpert
78 Classic) and the new version Xpert MTB/RIF Ultra (Ultra) [Cepheid, Sunnyvale, USA] which
79 employ heminested real-time polymerase chain reaction and molecular beacon technology

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

80 (Blakemore et al. 2010); and the line probe assays - GenoType MDRTB*plus* v2.0 (hereinafter
81 called LPA-Hain) [Hain Lifescience GmbH, Nehren, Germany]; and
82 GenoscholarNTM+MDRTB II (hereinafter called LPA-Nipro) [NIPRO Corporation, Osaka,
83 Japan] which rely on multiplex amplification and reverse hybridization of target to both wild-
84 type and mutant probes on a membrane strip (Dheda et al. 2017). These tests were designed to
85 detect RR-conferring mutations within the rifampicin resistance determining region (RRDR) or
86 hotspot of the *rpoB* gene (Andre et al. 2017; Blakemore et al. 2010; Ng et al. 2018a; Ng et al.
87 2018c).
88 The Xpert assay involves binding of five short overlapping fluorescent probes to wild-type
89 regions of the RRDR. Each of the five probes corresponds to several mutations which inhibit
90 probe binding, thereby disrupting the signal from the respective probe. Thus, when a mutation is
91 detected by Xpert, at least one probe does not bind. Certain mutations which completely interfere
92 with probe binding result in 'drop-out' or absent probe, whereas mutations that allow limited
93 probe hybridization are represented by 'delayed' probes (Blakemore et al. 2010). Xpert probes
94 are therefore often considered as a proxy for circulating rifampicin resistance-conferring
95 mutations which may be common or infrequent in a particular country. Understanding their
96 frequency can be crucial for rifampicin resistance surveillance. For example, mutation
97 Ser450Leu repeatedly detected globally, (Coll et al. 2018; Walker et al. 2015), is captured by
98 'absent' Xpert probe E (Ng et al. 2018a).
99 The implementation of Xpert was a breakthrough in TB diagnosis as it revolutionized detection
100 of RR-TB worldwide, allowing for prompt identification of patients who need to undergo
101 adapted treatment. Xpert Classic is the most widely deployed RDT globally, implemented as the
102 initial diagnostic tool for all presumptive pulmonary TB patients by 32 out of 48 high TB-burden
103 countries (WHO 2018). The wide utilization of the tests in both low and high burden TB
104 countries resulted in the production of large volumes of RDT data, although comparisons within
105 and between countries can be difficult due to use of differing technologies.
106 The utility of next generation sequencing technologies has been widely studied for improved
107 detection of drug-resistant TB in diverse laboratory settings worldwide (Gardy & Loman 2018).
108 Next generation sequencing-based whole genome sequencing (WGS) of *Mycobacterium*
109 *tuberculosis* (*Mtb*) has been shown to accurately detect RR-TB by calling relevant single
110 nucleotide polymorphisms (SNPs) in the *rpoB* gene (Coll et al. 2018; Miotto et al. 2017). WGS
111 is already being widely implemented in the United Kingdom, the Netherlands, and New York,
112 aimed towards completely replacing phenotypic drug susceptibility testing in the clinic
113 (CRyPTICConsortium et al. 2018; de Viedma 2019). WGS implemented in high burden TB
114 countries was shown to accurately estimate the prevalence of DR-TB (Zignol et al. 2018). The
115 conventional periodic TB drug resistance surveys have been gathering data representative of the
116 *Mtb* population in poor resource settings, while high income countries typically apply continuous
117 surveillance (WHO 2018; Zignol et al. 2018), to help improve the choice of standard TB
118 treatment before full drug sensitivity profile is known (WHO 2018).

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

119 The increasing use of WGS in research and public health initiatives can lead to a disconnect from
120 the RDT-based data being generated routinely in the clinic, and the majority of surveys,
121 widening the *Mtb* data gap. We aim to bridge this *Mtb* data gap by transforming WGS data into
122 each of the related RDT data outputs, to facilitate analysis of RR-TB prevalence and underlying
123 mutations regardless of the switch between data generation technologies. This will allow end-
124 users to compare 'apples with apples', and analyze previous historical strains with current
125 isolates representative of the entire TB patient population in the country.
126 We present the Myc TB Genome to Test (MTBGT) tool, a Python 3-based set of scripts that
127 rapidly transform WGS data type to RDT and mutation reports. The modules automatically
128 generate frequencies of rifampicin-susceptible (RS) and RR-TB samples detected and
129 supplement the RDT probes output with the detected RR-conferring mutation in the format of
130 'wild-type amino acid-codon number-mutant amino acid'.
131
132 ## Materials & Methods
133 We developed the MTBGT tool, a Python 3-executable that converts *Mtb* WGS data in the form
134 of variant call format (VCF) or MTBseq (Kohl et al. 2018) tab files into the most likely output
135 that would be observed from Xpert Classic, Xpert Ultra, LPA-Hain, and LPA-Nipro based on
136 previously validated work (Ng et al. 2018a; Ng et al. 2018c). The MTBGT tool can be accessed
137 through https://github.com/KamelaNg/MTBGT.
138
139 **The MTBGT tool**
140 The MTBGT tool can be run on any python-enabled operating system with no additional
141 prerequisites. The MTBGT workflow is shown in Figure 1. The input is a folder of files derived
142 from a SNP calling pipeline, either in standard raw VCF format or tab format as output from
143 MTBseq. The MTBGT tool assumes the standard H37Rv NC000962.3 genome was used for
144 calling these SNPs. If not, the user may remap the genome positions to the specific RR-TB-
145 related codons using a tab delimited mapping file. An example of this tab-separated file is
146 bundled with the tool. By default, the module will run all the RDTs on the input files.
147 The generated tab-delimited output file includes the Sample name, RIF resistance or
148 susceptibility, the associated mutant codon position and mutation pattern, and a series of 0's or
149 1's indicating absence or presence of the capturing probe for the RDTs, and the RR-conferring
150 mutation (An example output is given in Table 2). A summary table with counts and proportions
151 of detected RS-TB cases, RDT probes, and RR-conferring mutations is also generated
152 (Supplemental File Table S1). These tab-separated output files can be easily imported into Excel
153 as outlined in the associated manual.
154
155 **Validation and sample application of the MTBGT modules**
156 We simulated VCF files for internal validation of the MTBGT modules. We then randomly
157 chose a VCF file and edited it in Notepad ++ 7.6.3 (https://notepad-plus-plus.org/) to contain all
158 previously tested and validated RR-conferring mutations (Andre et al. 2017; Miotto et al. 2017)

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

159 which were mapped to known RDT results (Ng et al. 2018a; Ng et al. 2018c) (Table 1). We
160 generated files with single and multiple RR mutations to ensure the tool is robust for all
161 scenarios. These simulated VCFs are provided with the tool.
162 To show real world applicability, the MTBGT tool was run on WGS of RS and RR -TB WHO
163 Tropical Disease Research (TDR) strains with diverse resistance patterns and geographic origins
164 stored in the Belgian Coordinated Collections of Microorganisms in the Institute of Tropical
165 Medicine including the 47 TDR-TB strains tested in the previous validation of the RDTs against
166 the available *rpoB* Sanger sequences of the strains (Ng et al. 2018a; Ng et al. 2018c; Vincent et
167 al. 2012). The fastQ files (ENA accession PRJEB31023) for these samples were run through the
168 MTBseq pipeline with default settings (Kohl et al. 2018) to generate the tab files for input to
169 MTBGT. Additionally, WGS from 324 phenotypically RR-TB isolates of retreatment TB
170 patients in Kinshasa, the Democratic Republic of Congo (DRC) collected in 2005-2010 (Meehan
171 et al. 2018), and 233 WGS isolated 1991- 2010 from Rwanda were subjected to the MTBseq
172 pipeline and run through the MTBGT tool.
173 We were then able to extract the associated Xpert results from the 1991 to 2010 Rwandan dataset
174 and compare it with actual Xpert results from 2012 to 2017 (Ng et al. 2018b).
175

176 Results

177 We tested the MTBGT tool on a 64-bit Windows 10 Enterprise computer with a 2.50 GHz
178 processor and a 8.00 GB of RAM. The running time was 39 milliseconds for 1 MTBseq tab file
179 and 530 milliseconds for a VCF file.
180 The MTBGT tool correctly transformed the WGS-derived SNPs in the VCF and MTBseq tab
181 files into the laboratory-validated RDT probe reactions (Table 2), and accurately detected all
182 previously validated RR-conferring mutations in the simulated VCF and MTBseq clinical WGS
183 data from the TDR, DRC (Figures 2A and 2B), and Rwanda (Figures 3A and 3B) strains,
184 including double nucleotide changes – two SNPs covering two loci or genome positions – such
185 as the CAC → AGC and CAC → TCC His445Ser mutations and combinations of any two
186 unlinked mutations. The generated frequency and proportion tables (Supplemental File Table S1)
187 showed the distribution of the RS and RR-TB strains, the RDT probe reactions, and the RR-
188 conferring mutations detected.
189 The *in silico* extracted Xpert results from the Rwandan WGS collected in 1991-2010 revealed
190 that the majority of RR-TB cases detected were linked with Xpert Classic probe E (Figure 3A).
191 This observation was consistent with the actual Xpert results gathered in 2012 to 2017. Figure
192 3C shows the distribution of 12 years worth of Xpert results in Rwanda. The MTBGT tool
193 facilitated this continuous analysis of *in silico* predicted Xpert results from 2005 to 2010 WGS
194 and actual Xpert results in 2012 to 2017, and allowed us to plot the distribution of absent probe E
195 against the notified RR-TB cases from surveillance programs and national surveys in Rwanda.
196 The low WGS sampling reflected in Figure 3C is explained by the limited TB cultures kept in
197 the freezer in 2005 to 2010. Notably, we also observed the divergence of documented probe
198 reactions from the predominant absent probe E in 2013 to 2017.

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

199 Xpert Classic probe E was also predominantly seen in samples from Kinshasa, DRC (Figure 2A,
200 Supplemental File Table S1). Mutation S450L, linked with Xpert Classic probe E, was the major
201 RR-conferring mutation observed in both settings (Figures 2B and 3B).
202

203 Discussion

204 To make sense of the increasingly produced WGS in light of the existing and readily available
205 routine RDT results, we developed, validated, and applied the Myc TB Genome to Test
206 (MTBGT), a Python-implemented tool that extracts RR-TB RDT results and RR-conferring
207 mutations from WGS-derived SNP data.
208 If a country performed prior drug resistance surveys with a different diagnostic algorithm, the
209 MTBGT tool allows for comparing data of different formats and sources, facilitating analysis of
210 previous and current RR-TB case counts, probe reactions, and mutations, such as periodic DRS
211 results conducted as cross-sectional surveys, often with different technology from the previous
212 one. Although WGS data gives a much higher resolution than the RDTs, low and middle income
213 high RR-TB burden countries are not yet capable of implementing routine WGS for all
214 presumptive TB patients due to limitations in funding and logistic challenges (Meehan et al.
215 2019). However, RDTs are used routinely in such countries, producing a wealth of RR-TB data,
216 whereas WGS is used for regular country-wide drug resistance surveys by the WHO (Zignol et
217 al. 2018), and implemented in low burden high income countries. But since RDT results cannot
218 be upscaled to WGS results, for comparison of these two data sources, WGS must be
219 downscaled to the RDT results. By downscaling, the MTBGT tool adds to a substantial history
220 of RDT probe data which amplifies surveillance analyses that can inform NTPs of low and
221 middle-icome settings. This is similar to how CRISPR-based strain typing of *M. tuberculosis*
222 (termed spoligotyping) can be predicted from WGS reads through SpoTyping (Xia et al. 2016).
223 Although spoligotyping is known to have a lower resolution than WGS for typing, it is still
224 widely used in many low and middle income TB endemic countries due to its low cost (Montoya
225 et al. 2013; Suzana et al. 2017; Tulu & Ameni 2018). We foresee our tool being used in a similar
226 manner, until WGS capabilities become more commonplace in low and middle income high RR-
227 TB burden countries.
228 We show in Figure 3C how the MTBGT tool would be beneficial in Rwanda, a low income
229 country which uses Xpert for routine diagnostics and WGS for WHO-led 5-year drug resistance
230 surveys and selected research projects. Figure 3C reflects continuous distribution of RR-TB
231 cases and Xpert absent probe E in Rwanda for more than a decade. These trends of RR-TB and
232 absent probe E boost surveillance analyses that can inform the Rwandan National TB Control
233 Program and support public health efforts for improved RR-TB control. The MTBGT tool allows
234 for such comparisons and continuous distributions to be made, enabling low and middle income
235 researchers to utilize all of their data for surveillance of RR-TB. This analysis of Rwandan
236 datasets is a proof-of-concept that the MTBGT tool provides an opportunity to compare and
237 analyze old and new data produced by different technologies bridging the RR-TB data gap. Care
238 must be taken that the comparison is valid given the populations sampled. The predominant

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

239 Xpert Classic probe E observed in Rwanda is supported by the associated RR mutation
240 Ser450Leu (Tables 1 and 2) being present in the primary circulating multidrug-resistant TB
241 clone in this setting (Ngabonziza et al. 2018), and also most frequently associated with global
242 RR-TB (Cohen et al. 2015; Coll et al. 2018; Georghiou et al. 2016).
243 Programatic surveillance is possible with only the rifampicin susceptibility status of patients'
244 samples, but at a different level. The RDT probe information which represents the underlying
245 RR-conferring mutation provides an added critical value for epidemiological studies and
246 surveillance (Ng et al. 2018b). This is exemplified by Xpert probe B and probe binding delay in
247 South Africa being associated with false RR-TB results (Berhanu et al. 2019). In this setting,
248 patients who were true rifampicin-susceptible (RS) but diagnosed as RR were treated with less
249 effective and more toxic MDR drugs relative to the standard TB treatment. Further, non-routine
250 WGS alone may give incorrect distribution of circulating mutations in the setting due to culture
251 and sampling biases, as shown by the limited stored cultures resulting in low WGS sampling in
252 Rwanda in 2005 to 2010.
253 Settings which employ different RDTs can refer to Table 1 where they can match laboratory-
254 validated Xpert Classic and Xpert Ultra signatures per RR-TB mutation type (Ng et al. 2018a;
255 Ng et al. 2018c). This would be helpful for settings such as South Africa which implements
256 Xpert Classic and concurrently transitions to Xpert Ultra (Berhanu et al. 2018), and many other
257 countries that will shift to Xpert Ultra in the near future, such as Rwanda by 2020.
258 The genome-based approach of the MTBGT tool also allows for reporting of disputed mutations
259 that confer occult RR and are frequently missed by the *Mycobacterium* Growth Indicator Tube
260 phenotypic DST (Ng et al. 2018a; Van Deun et al. 2015). For instance, disputed mutation
261 Leu452Pro, epidemiologically linked with an extensively drug-resistant TB outbreak in
262 KwaZulu-Natal, South Africa in 2005 (Cohen et al. 2015; Ioerger et al. 2009), was detected in
263 some strains from Kinshasa, DRC and Rwanda (Figures 2B and 3B, Supplemental File).
264 Mutation Leu452Pro is captured by delayed Xpert Classic probe E, denoting partially inhibited
265 probe E fluorescence (Lawn & Nicol 2011; Ng et al. 2018a), and specifically identified by the
266 unique combination of Xpert Ultra probe rpoB4B and corresponding melting temperature shift
267 (Table 1), provided sufficient *Mtb* DNA is detected. Leu452Pro was reported to be missed in
268 clinical samples by LPA-Hain due its end-probe location (Al-Mutairi et al. 2011; Rigouts et al.
269 2013), thus contributing to the RR-TB detection gap. The ability of the MTBGT tool to rapidly
270 and accurately detect disputed mutations is therefore important.
271 As a supplementary feature, the MTBGT tool picks up any RR-conferring mutation present in
272 the VCF or MTBseq tab file, and may help assess whether RDTs sufficiently detect RR-TB cases
273 in specific settings.
274 Looking into the future when whole genome sequencing will be implemented as the primary
275 diagnostic tool for all presumptive TB patients to confirm TB and/or RR-TB, the MTBGT tool
276 will bridge the TB data gap by allowing comparison of past Xpert and LPA results with current
277 *in silico* predicted results from routine whole genome sequences. These data would be
278 comparable based on a common statistical parameter and representative of the presumptive TB

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

279 population in the setting. With more whole genomes of *M. tuberculosis* sequenced, the *in silico*
280 predicted probe distribution will more accurately capture the circulation of RR-TB mutations in
281 the setting. The tool can do so by precisely revealing the divergence of current probes from a
282 previously documented predominant probe over time, implicative of other circulating mutations
283 such as that shown by actual Xpert results in 2013 to 2017 (Figure 3C).
284 The MTBGT tool was developed to aid surveillance efforts of low and middle income high RR-
285 TB burden countries with likely fragmented data collection systems (Mazumdar et al. 2019).
286 Through facilitated continuous analysis of previous historical and current RR-TB data, as well as
287 clinical data from large aggregated files and routine laboratory and periodic survey RDT results
288 combined with WGS-derived data from national surveillance programs and public data
289 repositories, the MTBGT tool may create a larger picture of the RS-TB and RR-TB burden in a
290 country. The RS and RR-TB counts and proportions and RDT probe reactions generated by the
291 MTBGT tool may contribute to an extensive global database with years' worth of data for
292 continuous statistical modeling analyses and surveillance investigations.
293 Potentially, the MTBGT tool may be integrated in the national TB diagnostic algorithm through
294 the existing connectivity platform or the newly implemented WHO cloud-based software (Dean
295 2019). The prospective application of the MTBGT tool may bridge and transform the *Mtb* data
296 gap into action points for RR-TB clinicians to provide appropriate care for the individual
297 TB/RR-TB patient, and the NTP, public health officials, and policy makers to intervene at the
298 population-level for improved and sustained RR-TB control (Gardy & Loman 2018). The
299 MTBGT modules could also be expanded to report non-RRDR RR-causing mutations and
300 include other TB drugs – e.g. isoniazid, pyrazinamide, and fluoroquinolones.
301

302 Conclusions

303 The MTBGT tool leverages improved access to next generation sequencing technologies in this
304 genomic epidemiology era of TB, and complement *Mtb* WGS by rapidly transforming WGS files
305 that store genomic sequence variations to validated outputs of the RR-TB RDTs. The prospective
306 application of the MTBGT tool within a nationwide genomic epidemiology program will bridge
307 the *Mtb* data gap among routine RDT data, research setting WGS, and periodic survey and
308 continuous surveillance *rpoB* sequencing and WGS data. This may result in facilitated
309 continuous analysis of the circulating RDT probes and underlying distribution of RR-conferring
310 mutations in the setting, and may help assess whether currently implemented RDT(s) serve the
311 detection of RR-TB cases in the country.
312

313 Acknowledgements

314 We thank Button Ricarte for his insightful feedback on the manuscript.
315

316 Additional Information & Declarations

317 Data Availability

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

318 All scripts and data for validation may be accessed through GitHub:
319 https://github.com/KamelaNg/MTBGT.
320

321 ## References

322 Al-Mutairi NM, Ahmad S, and Mokaddas E. 2011. Performance comparison of four methods for
323 detecting multidrug-resistant Mycobacterium tuberculosis strains. *Int J Tuberc Lung Dis* 15:110-
324 115.
325 Andre E, Goeminne L, Cabibbe A, Beckert P, Kabamba Mukadi B, Mathys V, Gagneux S, Niemann S,
326 Van Ingen J, and Cambau E. 2017. Consensus numbering system for the rifampicin resistance-
327 associated rpoB gene mutations in pathogenic mycobacteria. *Clin Microbiol Infect* 23:167-172.
328 10.1016/j.cmi.2016.09.006
329 Berhanu RH, David A, da Silva P, Shearer K, Sanne I, Stevens W, and Scott L. 2018. Performance of
330 Xpert MTB/RIF, Xpert Ultra, and Abbott RealTime MTB for Diagnosis of Pulmonary
331 Tuberculosis in a High-HIV-Burden Setting. *J Clin Microbiol* 56. 10.1128/JCM.00560-18
332 Berhanu RH, Schnippel K, Kularatne R, Firnhaber C, Jacobson KR, Horsburgh CR, and Lippincott CK.
333 2019. Discordant rifampicin susceptibility results are associated with Xpert((R)) MTB/RIF probe
334 B and probe binding delay. *Int J Tuberc Lung Dis* 23:358-362. 10.5588/ijtld.16.0837
335 Blakemore R, Story E, Helb D, Kop J, Banada P, Owens MR, Chakravorty S, Jones M, and Alland D.
336 2010. Evaluation of the analytical performance of the Xpert MTB/RIF assay. *J Clin Microbiol*
337 48:2495-2501. 10.1128/JCM.00128-10
338 Cohen KA, Abeel T, Manson McGuire A, Desjardins CA, Munsamy V, Shea TP, Walker BJ, Bantubani
339 N, Almeida DV, Alvarado L, Chapman SB, Mvelase NR, Duffy EY, Fitzgerald MG, Govender P,
340 Gujja S, Hamilton S, Howarth C, Larimer JD, Maharaj K, Pearson MD, Priest ME, Zeng Q,
341 Padayatchi N, Grosset J, Young SK, Wortman J, Mlisana KP, O'Donnell MR, Birren BW, Bishai
342 WR, Pym AS, and Earl AM. 2015. Evolution of Extensively Drug-Resistant Tuberculosis over
343 Four Decades: Whole Genome Sequencing and Dating Analysis of Mycobacterium tuberculosis
344 Isolates from KwaZulu-Natal. *PLoS Med* 12:e1001880. 10.1371/journal.pmed.1001880
345 Coll F, Phelan J, Hill-Cawthorne GA, Nair MB, Mallard K, Ali S, Abdallah AM, Alghamdi S, Alsomali
346 M, Ahmed AO, Portelli S, Oppong Y, Alves A, Bessa TB, Campino S, Caws M, Chatterjee A,
347 Crampin AC, Dheda K, Furnham N, Glynn JR, Grandjean L, Minh Ha D, Hasan R, Hasan Z,
348 Hibberd ML, Joloba M, Jones-Lopez EC, Matsumoto T, Miranda A, Moore DJ, Mocillo N,
349 Panaiotov S, Parkhill J, Penha C, Perdigao J, Portugal I, Rchiad Z, Robledo J, Sheen P, Shesha
350 NT, Sirgel FA, Sola C, Oliveira Sousa E, Streicher EM, Helden PV, Viveiros M, Warren RM,
351 McNerney R, Pain A, and Clark TG. 2018. Genome-wide analysis of multi- and extensively drug-
352 resistant Mycobacterium tuberculosis. *Nat Genet* 50:307-316. 10.1038/s41588-017-0029-0
353 CRyPTICConsortium, the GP, Allix-Beguec C, Arandjelovic I, Bi L, Beckert P, Bonnet M, Bradley P,
354 Cabibbe AM, Cancino-Munoz I, Caulfield MJ, Chaiprasert A, Cirillo DM, Clifton DA, Comas I,
355 Crook DW, De Filippo MR, de Neeling H, Diel R, Drobniewski FA, Faksri K, Farhat MR,
356 Fleming J, Fowler P, Fowler TA, Gao Q, Gardy J, Gascoyne-Binzi D, Gibertoni-Cruz AL, Gil-
357 Brusola A, Golubchik T, Gonzalo X, Grandjean L, He G, Guthrie JL, Hoosdally S, Hunt M, Iqbal
358 Z, Ismail N, Johnston J, Khanzada FM, Khor CC, Kohl TA, Kong C, Lipworth S, Liu Q,
359 Maphalala G, Martinez E, Mathys V, Merker M, Miotto P, Mistry N, Moore DAJ, Murray M,
360 Niemann S, Omar SV, Ong RT, Peto TEA, Posey JE, Prammananan T, Pym A, Rodrigues C,
361 Rodrigues M, Rodwell T, Rossolini GM, Sanchez Padilla E, Schito M, Shen X, Shendure J,

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

362 Sintchenko V, Sloutsky A, Smith EG, Snyder M, Soetaert K, Starks AM, Supply P, Suriyapol P,
363 Tahseen S, Tang P, Teo YY, Thuong TNT, Thwaites G, Tortoli E, van Soolingen D, Walker AS,
364 Walker TM, Wilcox M, Wilson DJ, Wyllie D, Yang Y, Zhang H, Zhao Y, and Zhu B. 2018.
365 Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *N Engl J Med*
366 379:1403-1415. 10.1056/NEJMoa1800474
367 de Viedma DG. 2019. Pathways and strategies followed in the genomic epidemiology of Mycobacterium
368 tuberculosis. *Infect Genet Evol*. 10.1016/j.meegid.2019.01.027
369 Dean A. 2019. Sequencing for the surveillance of drug-resistant TB. WHO Global TB Programme.
370 Dheda K, Gumbo T, Maartens G, Dooley KE, McNerney R, Murray M, Furin J, Nardell EA, London L,
371 Lessem E, Theron G, van Helden P, Niemann S, Merker M, Dowdy D, Van Rie A, Siu GK,
372 Pasipanodya JG, Rodrigues C, Clark TG, Sirgel FA, Esmail A, Lin HH, Atre SR, Schaaf HS,
373 Chang KC, Lange C, Nahid P, Udwadia ZF, Horsburgh CR, Jr., Churchyard GJ, Menzies D,
374 Hesseling AC, Nuermberger E, McIlleron H, Fennelly KP, Goemaere E, Jaramillo E, Low M,
375 Jara CM, Padayatchi N, and Warren RM. 2017. The epidemiology, pathogenesis, transmission,
376 diagnosis, and management of multidrug-resistant, extensively drug-resistant, and incurable
377 tuberculosis. *Lancet Respir Med*. 10.1016/S2213-2600(17)30079-6
378 Gardy JL, and Loman NJ. 2018. Towards a genomics-informed, real-time, global pathogen surveillance
379 system. *Nat Rev Genet* 19:9-20. 10.1038/nrg.2017.88
380 Georghiou SB, Seifert M, Catanzaro D, Garfein RS, Valafar F, Crudu V, Rodrigues C, Victor TC,
381 Catanzaro A, and Rodwell TC. 2016. Frequency and Distribution of Tuberculosis Resistance-
382 Associated Mutations between Mumbai, Moldova, and Eastern Cape. *Antimicrob Agents*
383 *Chemother* 60:3994-4004. 10.1128/AAC.00222-16
384 Ioerger TR, Koo S, No EG, Chen X, Larsen MH, Jacobs WR, Jr., Pillay M, Sturm AW, and Sacchettini
385 JC. 2009. Genome analysis of multi- and extensively-drug-resistant tuberculosis from KwaZulu-
386 Natal, South Africa. *PLoS One* 4:e7778. 10.1371/journal.pone.0007778
387 Kohl TA, Utpatel C, Schleusener V, De Filippo MR, Beckert P, Cirillo DM, and Niemann S. 2018.
388 MTBseq: a comprehensive pipeline for whole genome sequence analysis of Mycobacterium
389 tuberculosis complex isolates. *PeerJ* 6:e5895. 10.7717/peerj.5895
390 Lawn SD, and Nicol MP. 2011. Xpert(R) MTB/RIF assay: development, evaluation and implementation
391 of a new rapid molecular diagnostic for tuberculosis and rifampicin resistance. *Future Microbiol*
392 6:1067-1082. 10.2217/fmb.11.84
393 Mazumdar S, Satyanarayana S, and Pai M. 2019. Self-reported tuberculosis in India: evidence from
394 NFHS-4. *BMJ Glob Health* 4:e001371. 10.1136/bmjgh-2018-001371
395 Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, Farhat MR, Guthrie JL,
396 Laukens K, Miotto P, Ofori-Anyinam B, Dreyer V, Supply P, Suresh A, Utpatel C, van Soolingen
397 D, Zhou Y, Ashton PM, Brites D, Cabibbe AM, de Jong BC, de Vos M, Menardo F, Gagneux S,
398 Gao Q, Heupink TH, Liu Q, Loiseau C, Rigouts L, Rodwell TC, Tagliani E, Walker TM, Warren
399 RM, Zhao Y, Zignol M, Schito M, Gardy J, Cirillo DM, Niemann S, Comas I, and Van Rie A.
400 2019. Whole genome sequencing of Mycobacterium tuberculosis: current standards and open
401 issues. *Nat Rev Microbiol*. 10.1038/s41579-019-0214-5
402 Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, Hanna D, Kim PS, Liwski R,
403 Zignol M, Gilpin C, Niemann S, Denkinger CM, Fleming J, Warren RM, Crook D, Posey J,
404 Gagneux S, Hoffner S, Rodrigues C, Comas I, Engelthaler DM, Murray M, Alland D, Rigouts L,
405 Lange C, Dheda K, Hasan R, Ranganathan UDK, McNerney R, Ezewudo M, Cirillo DM, Schito

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

406 M, Koser CU, and Rodwell TC. 2017. A standardised method for interpreting the association
407 between mutations and phenotypic drug resistance in Mycobacterium tuberculosis. *Eur Respir J*
408 50. 10.1183/13993003.01354-2017
409 Montoya JC, Murase Y, Ang C, Solon J, and Ohkado A. 2013. A molecular epidemiologic analysis of
410 Mycobacterium tuberculosis among Filipino patients in a suburban community in the Philippines.
411 *Kekkaku* 88:543-552.
412 Ng KC, Meehan CJ, Torrea G, Goeminne L, Diels M, Rigouts L, de Jong BC, and Andre E. 2018a.
413 Potential Application of Digitally Linked Tuberculosis Diagnostics for Real-Time Surveillance of
414 Drug-Resistant Tuberculosis Transmission: Validation and Analysis of Test Results. *JMIR Med*
415 *Inform* 6:e12. 10.2196/medinform.9309
416 Ng KCS, Ngabonziza JCS, Meehan CJ, Migambi P, de Jong BC, Cobelens F, and van Leth F. 2018b.
417 Automated algorithm for early identification of rifampicin-resistant tuberculosis transmission
418 hotspots in Rwanda. *The International Journal of Tuberculosis and Lung Disease* 22:S1 – S652.
419 Ng KCS, van Deun A, Meehan CJ, Torrea G, Driesen M, Gabriels S, Rigouts L, Andre E, and de Jong
420 BC. 2018c. Xpert Ultra Can Unambiguously Identify Specific Rifampin Resistance-Conferring
421 Mutations. *J Clin Microbiol* 56. 10.1128/JCM.00686-18
422 Ngabonziza JCS, Rigouts L, Mazarati JB, Uwizeye C, Nyaruhirira AU, Torrea G, de Jong BC, and
423 Meehan CJ. 2018. Transmission drives the increase of multidrug-resistant tuberculosis in
424 Rwanda. *The International Journal of Tuberculosis and Lung Disease* 22:S1 – S652.
425 Rigouts L, Gumusboga M, de Rijk WB, Nduwamahoro E, Uwizeye C, de Jong B, and Van Deun A. 2013.
426 Rifampin resistance missed in automated liquid culture system for Mycobacterium tuberculosis
427 isolates with specific rpoB mutations. *J Clin Microbiol* 51:2641-2645. 10.1128/JCM.02741-12
428 Suzana S, Shanmugam S, Uma Devi KR, Swarna Latha PN, and Michael JS. 2017. Spoligotyping of
429 Mycobacterium tuberculosis isolates at a tertiary care hospital in India. *Trop Med Int Health*
430 22:703-707. 10.1111/tmi.12875
431 Tulu B, and Ameni G. 2018. Spoligotyping based genetic diversity of Mycobacterium tuberculosis in
432 Ethiopia: a systematic review. *BMC Infect Dis* 18:140. 10.1186/s12879-018-3046-4
433 Van Deun A, Aung KJ, Hossain A, de Rijk P, Gumusboga M, Rigouts L, and de Jong BC. 2015. Disputed
434 rpoB mutations can frequently cause important rifampicin resistance among new tuberculosis
435 patients. *Int J Tuberc Lung Dis* 19:185-190. 10.5588/ijtld.14.0651
436 Vincent V, Rigouts L, Nduwamahoro E, Holmes B, Cunningham J, Guillerm M, Nathanson CM, Moussy
437 F, De Jong B, Portaels F, and Ramsay A. 2012. The TDR Tuberculosis Strain Bank: a resource
438 for basic science, tool development and diagnostic services. *Int J Tuberc Lung Dis* 16:24-31.
439 10.5588/ijtld.11.0223
440 Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, Iqbal Z, Feuerriegel S, Niehaus
441 KE, Wilson DJ, Clifton DA, Kapatai G, Ip CLC, Bowden R, Drobniewski FA, Allix-Beguec C,
442 Gaudin C, Parkhill J, Diel R, Supply P, Crook DW, Smith EG, Walker AS, Ismail N, Niemann S,
443 Peto TEA, and Modernizing Medical Microbiology Informatics G. 2015. Whole-genome
444 sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a
445 retrospective cohort study. *Lancet Infect Dis* 15:1193-1202. 10.1016/S1473-3099(15)00062-6
446 WHO. 2018. Global tuberculosis report.
447 Xia E, Teo YY, and Ong RT. 2016. SpoTyping: fast and accurate in silico Mycobacterium spoligotyping
448 from sequence reads. *Genome Med* 8:19. 10.1186/s13073-016-0270-7

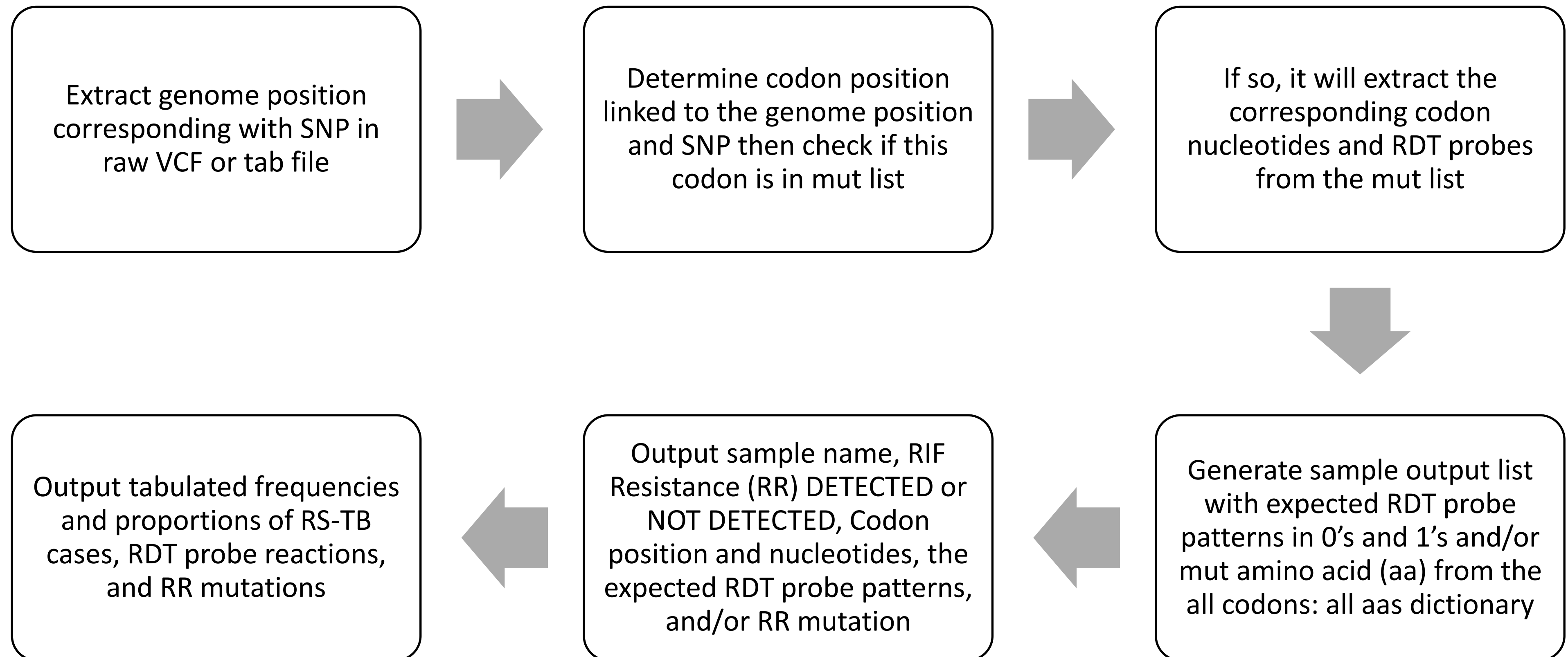
bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

449 Zignol M, Cabibbe AM, Dean AS, Glaziou P, Alikhanova N, Ama C, Andres S, Barbova A, Borbe-Reyes
450 A, Chin DP, Cirillo DM, Colvin C, Dadu A, Dreyer A, Driesen M, Gilpin C, Hasan R, Hasan Z,
451 Hoffner S, Hussain A, Ismail N, Kamal SMM, Khanzada FM, Kimerling M, Kohl TA, Mansjo
452 M, Miotto P, Mukadi YD, Mvusi L, Niemann S, Omar SV, Rigouts L, Schito M, Sela I,
453 Seyfaddinova M, Skenders G, Skrahina A, Tahseen S, Wells WA, Zhurilo A, Weyer K, Floyd K,
454 and Raviglione MC. 2018. Genetic sequencing for surveillance of drug resistance in tuberculosis
455 in highly endemic countries: a multi-country population-based surveillance study. *Lancet Infect*
456 *Dis* 18:675-683. 10.1016/S1473-3099(18)30073-2
457

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1 (on next page)

The MTBGT tool workflow.

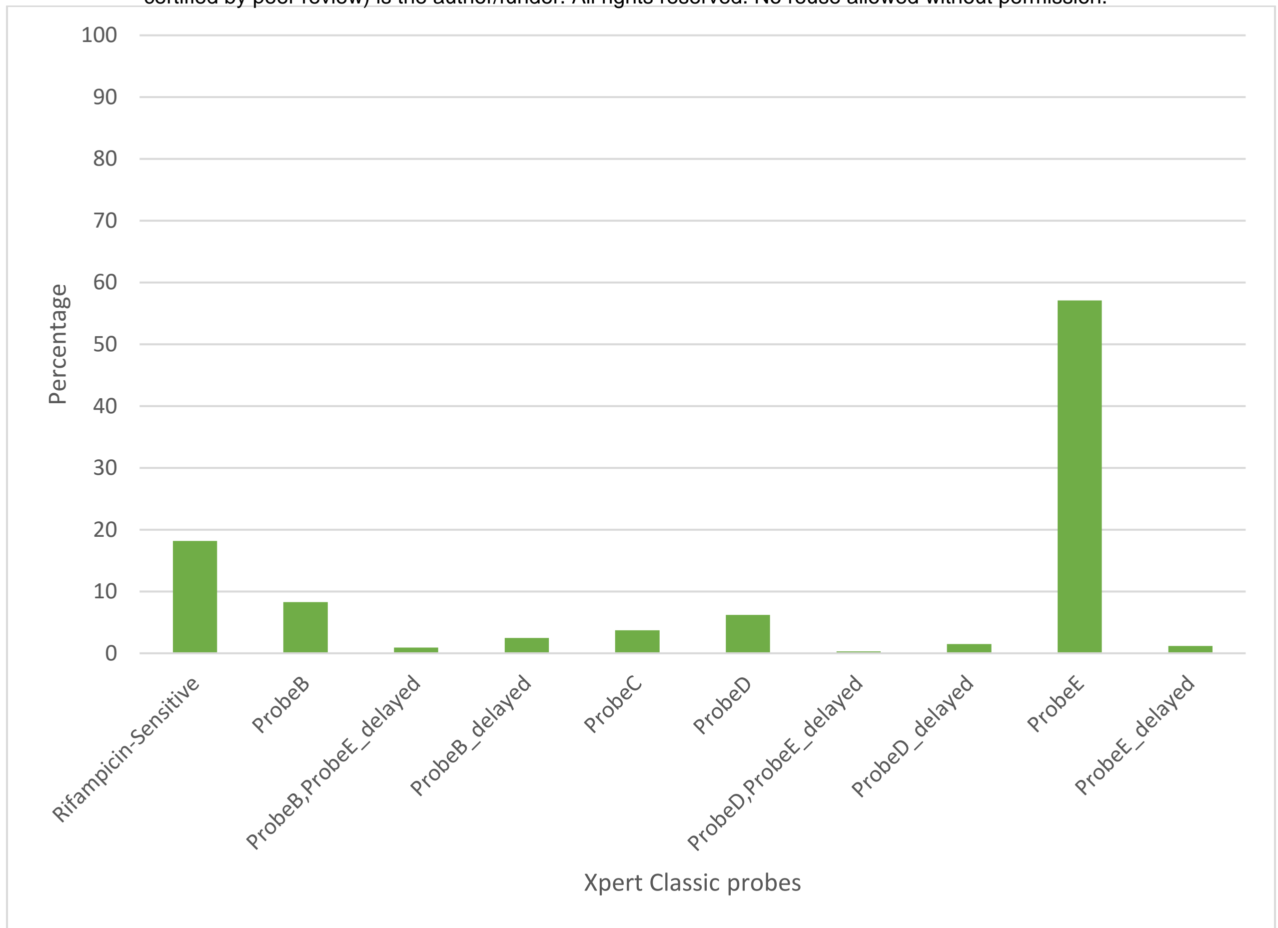

Extract genome position corresponding with SNP in raw VCF or tab file
Determine codon position linked to the genome position and SNP then check if this codon is in mut list
If so, it will extract the corresponding codon nucleotides and RDT probes from the mut list
Generate sample output list with expected RDT probe patterns in 0's and 1's and/or mut amino acid (aa) from the all codons: all aas dictionary
Output sample name, RIF Resistance (RR) DETECTED or NOT DETECTED, Codon position and nucleotides, the expected RDT probe patterns, and/or RR mutation
Output tabulated frequencies and proportions of RS-TB cases, RDT probe reactions, and RR mutations

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 2 (on next page)

Distribution of rifampicin-sensitive samples and Xpert Classic probes among rifampicin-resistant tuberculosis isolates in Kinshasa, DRC from 2005 to 2010, detected by the MTBGT tool.

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

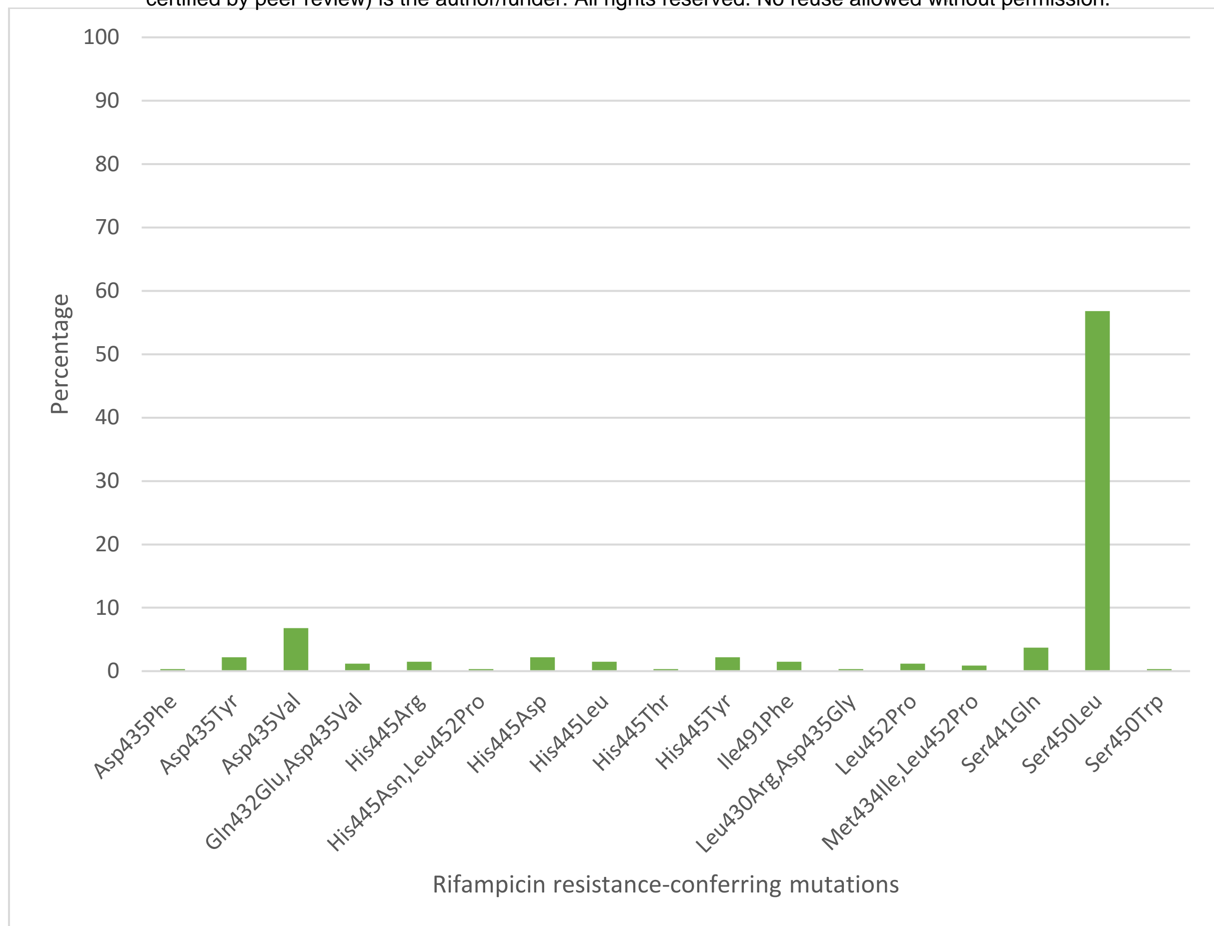


bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3 (on next page)

Distribution of rifampicin resistance-conferring mutations in Kinshasa, DRC from 2005 to 2010, detected by the MTBGT tool.

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

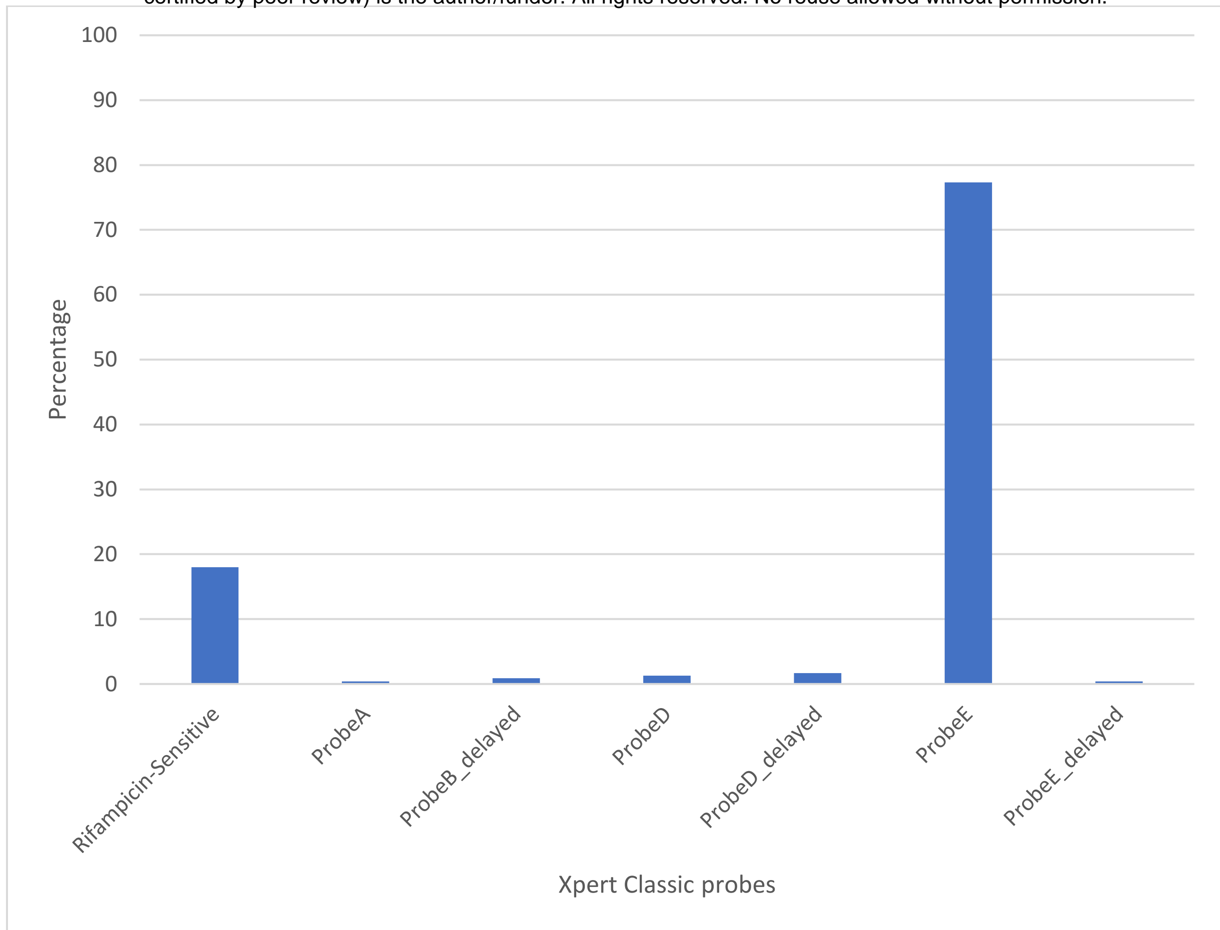


bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 4 (on next page)

Distribution of rifampicin-sensitive samples and Xpert Classic probes among rifampicin-resistant tuberculosis isolates in Rwanda from 1991 to 2010, determined by the MTBGT tool.

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

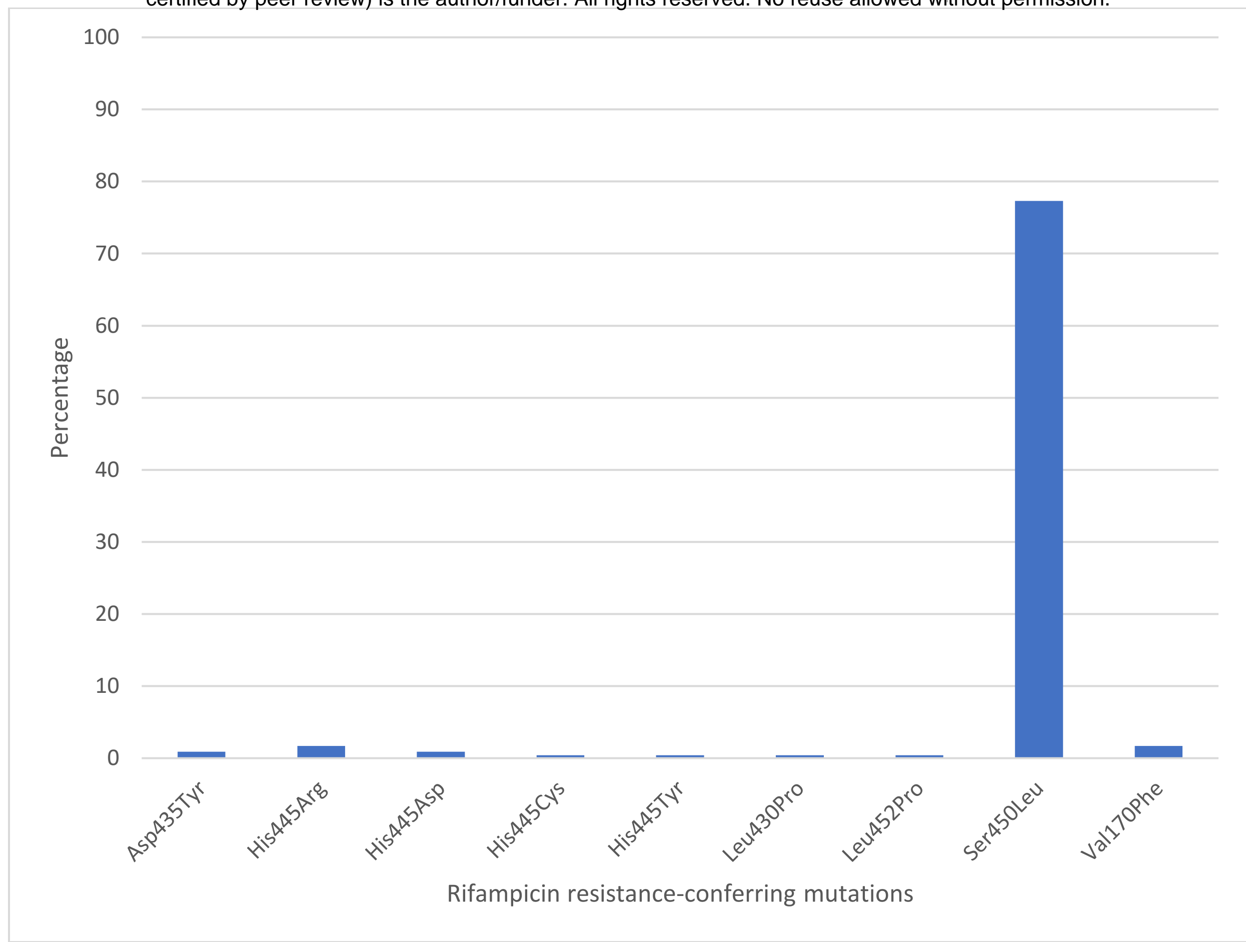


bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 5 (on next page)

Distribution of rifampicin resistance-conferring mutations in Rwanda from 1991 to 2010, detected by the MTBGT tool.

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 6

Distribution of RR-TB cases and Xpert Classic probe reactions predicted *in silico* from whole genomes collected in 2005 to 2010 in Rwanda and actual Xpert Classic results gathered in 2012 to 2017.

The distribution of Xpert Classic absent probe E was overlaid on the documented RR-TB case counts and plotted against the notified RR-TB case counts from surveillance programs and national surveys in Rwanda.

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1 (on next page)

Rifampicin resistance (RR)-conferring mutations and associated rapid diagnostic test (RDT) results, previously tested and validated, and used as basis for developing the TBGT tool.

Capturing probe, RDT probe associated with the RR-conferring mutation; Melting temperature shift, difference between mutant (MUT) and wild-type (WT) melting temperatures; ND, ‘not detected’, refers to mutations outside the RR determining region, not detected by the RR-TB RDTs.

1 Table 1. Rifampicin resistance (RR)-conferring mutations and associated rapid diagnostic test (RDT) results, previously tested and
2 validated, and used as basis for developing the TBGT tool.

RR-TB RDT							RR mutation
Xpert Classic	Xpert Ultra		LPA-Hain		LPA-Nipro		
Capturing probe	Capturing probe	Melting temperature shift	Absent (WT) probe	Developing (MUT) probe	Absent (WT) probe	Developing (MUT) probe	
ND	ND	ND	ND	ND	ND	ND	Val170Phe
Probe A	rpoB1	3.5	WT1		S1		Ser428Arg
Probe A	rpoB1	5.9-6.3	WT2		S1		Leu430Pro
Probe A, Probe B	rpoB1	2.9	WT2		S1		Ser431Gly
Probe B	rpoB1	3.4	WT2, WT3		S1		Gln432Glu
Probe B	rpoB2	3.2	WT3		S2		Met434Ile
Probe B	rpoB2	3.3	WT3		S2		Met434Thr
Probe B	rpoB1	6.3	WT3		S2		Met434Val
Probe B	rpoB2	2.8	WT3, WT4		S2		Asp435Glu
Probe B delayed	rpoB2	5.3	WT3, WT4		S2		Asp435Phe
Probe B	rpoB2	3.3	WT3, WT4		S2		Asp435Gly
Probe B	rpoB2	3.5-3.7	WT3, WT4	MUT1	S2	R2	Asp435Val
Probe B delayed	rpoB2	4.0-4.4	WT3, WT4		S2		Asp435Tyr
Probe C	rpoB2	6.4	WT4		S2		Asn437Asp
Probe C	rpoB2; rpoB3	3.0; 2.3	WT5, WT6		S3		Ser441Leu
Probe C	rpoB2; rpoB3	4.7; 2.3	WT5, WT6		S3		Ser441Gln
Probe D	rpoB3	3.7-3.9	WT7	MUT2B	S4	R4b	His445Asp
Probe D	rpoB3	4.9	WT7		S4		His445Gly
Probe D	rpoB3	3.5-3.6	WT7		S4		His445Leu
Probe D	rpoB3	3.4-3.5	WT7		S4		His445Asn
Probe D	rpoB3	3.6	WT7		S4		His445Gln (CAG)

RR-TB RDT							RR mutation
Xpert Classic	Xpert Ultra		LPA-Hain		LPA-Nipro		
Capturing probe	Capturing probe	Melting temperature shift	Absent (WT) probe	Developing (MUT) probe	Absent (WT) probe	Developing (MUT) probe	
Probe D	rpoB3	4.1	WT7		S4		His445Gln (CAA)
Probe D delayed	rpoB3	1.9	WT7		S4		His445Arg
Probe D	rpoB3	4.7	WT7		S4		His445Ser
Probe D	rpoB3	4.9	WT7		S4		His445Thr
Probe D	rpoB3	3.2-3.3	WT7	MUT2A	S4	R4a	His445Tyr
Probe D	rpoB4B	5.0	WT7		S4		Lys446Gln
Probe E	rpoB3	4.0	WT8		S5		Ser450Phe
Probe E	rpoB3; rpoB4A	2.5-2.9; 6.0-6.5	WT8	MUT3	S5	R5	Ser450Leu
Probe E	rpoB3; rpoB4A	2.3-2.7; 3.3-3.7	WT8		S5		Ser450Trp
Probe E delayed	rpoB4B	5.7-6.1	WT8		S5		Leu452Pro
ND	ND	ND	ND	ND	ND	ND	Ile491Phe

3 Capturing probe, RDT probe associated with the RR-conferring mutation; Melting temperature shift, difference between mutant
4 (MUT) and wild-type (WT) melting temperatures; ND, 'not detected', refers to mutations outside the RR determining region, not
5 detected by the RR-TB RDTs.

bioRxiv preprint doi: https://doi.org/10.1101/628099; this version posted July 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 2(on next page)

Example of combined results from the TBGT tool rapid diagnostic test (RDT) modules supplemented by the rifampicin resistance (RR)-conferring mutations detected.

1 Table 2. Example of combined results from the TBGT tool rapid diagnostic test (RDT) modules supplemented by the rifampicin
2 resistance (RR)-conferring mutations detected

Filename	RIF Resistance	Codon number	Codon	Xpert Classic		Xpert Ultra		LPA-Hain		LPA-Nipro		RR mutation
				Capturing probe	Probe pattern	Capturing probe, Melting temperature shift	Probe pattern	Capturing probe	Probe pattern	Capturing probe	Probe pattern	
DRC-052577	DETECTED	450	TTG	Probe E	1 1 1 1 0 1	rpoB3,rpoB4A; 2.5-2.9,6.0-6.5	1 1 0 0 1	WT8, MUT3	1 1 1 1 1 1 1 0 0 0 0 1	S5, R5	1 1 1 1 0 0 0 0 1	Ser450Leu
DRC-091003	DETECTED	452	CCG	Probe E delayed	1 1 1 1 1 0	rpoB4B; 5.7-6.1	1 1 1 1 0	WT8	1 1 1 1 1 1 1 0 0 0 0 0	S5	1 1 1 1 0 0 0 0 0	Leu452Pro
1993-09004	DETECTED (only by *rpoB* Sanger sequencing module)	170	TTC	NOT DETECTED		NOT DETECTED		NOT DETECTED		NOT DETECTED		Val170Phe
DRC-101308	DETECTED (only by *rpoB* Sanger sequencing module)	491	TTC	NOT DETECTED		NOT DETECTED		NOT DETECTED		NOT DETECTED		Ile491Phe

3

4