

1 Whole transcriptome and genomic analysis of extensively drug-resistant *Mycobacterium*  
2 *tuberculosis* clinical isolates identifies downregulation of *ethA* as a mechanism of  
3 ethionamide resistance

4

5 Lynne de Welzen<sup>1</sup>, Vegard Eldholm<sup>2</sup>, Kashmeel Maharaj<sup>1#</sup>, Abigail L. Manson<sup>3</sup>, Ashlee  
6 M. Earl<sup>3</sup> and Alexander S. Pym<sup>1,4</sup>

7

8 Africa Health Research Institute (AHRI), School of Laboratory Medicine &

9 Medical Sciences, University of KwaZulu-Natal, KwaZulu-Natal, South Africa<sup>1</sup>;

10 Infectious Disease Control and Environmental Health, Norwegian Institute of Public

11 Health, Oslo, Norway<sup>2</sup>; Broad Institute of MIT & Harvard, Cambridge, Massachusetts,

12 USA<sup>3</sup>; University College London (UCL), Bloomsbury, London, United Kingdom<sup>4</sup>

13

14 XDR-TB Transcriptomics and Ethionamide resistance

15

16 Address correspondence to Dr. Alexander S. Pym, Alex.Pym@ahri.org

17 #Present Address: Kashmeel Maharaj, Alere Healthcare (Pty) Ltd, Kempton Park, South

18 Africa

19

20 Abstract: 220 words

21 Text (Excluding Figure Legends and References): 4165 words

22

23 **ABSTRACT**

24 Genetic based drug susceptibility testing has improved the diagnosis of drug-resistant  
25 tuberculosis, but is limited by our lack of knowledge of all resistance mechanisms. Next  
26 generation sequencing has assisted in identifying the principal genetic mechanisms of  
27 resistance for many drugs, but a significant proportion of phenotypic drug resistance is  
28 unexplained genetically. Few studies have formally compared the transcriptome of  
29 susceptible and resistant *M. tuberculosis*. We carried out comparative whole genome  
30 transcriptomics on extensively drug-resistant (XDR) clinical isolates using RNA-  
31 sequencing (RNAseq) to find novel transcriptional mediated mechanisms of resistance.  
32 We identified a t-11c promoter mutation that reduces expression of a monooxygenase  
33 (EthA) that activates ethionamide. Using a flow-cytometry based reporter assay, we show  
34 that reduced transcription of *ethA* is not due to transcriptional repression by *ethR*.  
35 Clinical strains harbouring this mutation were resistant to ethionamide. Other *ethA*  
36 promoter mutations were identified in a global genomic survey of resistant *M.*  
37 *tuberculosis* strains. These results demonstrate a new mechanism of ethionamide  
38 resistance that can cause high-level resistance when combined with other ethionamide  
39 resistance conferring mutations. Our study revealed many other genes which were highly  
40 up or down regulated in XDR strains, including a toxin-antitoxin module (*mazF5 mazE5*)  
41 and tRNAs (*leuX* and *thrU*). This suggests more global transcriptional modifications have  
42 also occurred in XDR strains that could contribute to resistance or maintaining bacterial  
43 fitness.

44

45

46 **INTRODUCTION**

47 *Mycobacterium tuberculosis*, the causative agent for tuberculosis (TB), has progressively  
48 developed resistance to the most effective first and second-line anti-tuberculosis drugs.(1)  
49 Patients infected with extensively drug-resistant (XDR) strains (resistant to the  
50 fluoroquinolones and aminoglycosides in addition to rifampicin and isoniazid that define  
51 multi-drug resistance [MDR]) have extremely high mortality despite long and intensive  
52 treatment regimens.(2, 3) Ultimate control of drug resistance will require multiple  
53 interventions, one of which will be individualized therapy based on rapid comprehensive  
54 drug susceptibility testing (DST).

55

56 **Current** molecular genetic based tests, such as the Gene<sup>®</sup>Xpert MTB/RIF and  
57 GenoType<sup>®</sup> MTBDRplus, have accelerated the clinical detection of known mutations  
58 causing rifampicin (RIF) and/or isoniazid resistance.(4, 5) These and other genetic tests  
59 only detect MDR TB and a limited number of mutations associated with resistance to  
60 second-line drugs.(6) Whole genome sequencing (WGS) has the potential to rapidly  
61 detect all possible drug resistance conferring mutations.(7) However recent studies have  
62 demonstrated that genotypic DST using WGS lacks sensitivity for the detection of many  
63 second-line resistances including to fluoroquinolones.(8–11) Improving the sensitivity of  
64 genetic susceptibility testing will only be possible with a more comprehensive  
65 understanding of the genetic determinants of drug resistance.

66

67 **Our** current understanding of drug resistance in *M. tuberculosis* has developed through  
68 studying resistance mutants isolated *in vitro* and the accumulation of mutations in  
69 resistant clinical isolates.(12) These studies have identified various genetic mechanisms

70 of resistance including target modification, loss of enzymatic function required to activate  
71 prodrugs, and altered drug efflux.(13, 14)

72

73 **In** addition to intragenic mutations there is increasing evidence that alterations to gene  
74 transcription are an important mechanism of conferring drug resistance. Promoter  
75 mutations which result in upregulation of *inhA*, that encodes the target for isoniazid, were  
76 the first to be described.(15) Pyrazinamide (PZA) resistance has been associated with  
77 mutations in the regulatory region upstream of *pncA*, the enzyme responsible for  
78 activating PZA.(16–18) Aminoglycoside cross-resistance in *M. tuberculosis* can arise due  
79 to mutations in the regulatory region of *whib7* (encoding a transcriptional activator)  
80 which results in increased expression of *eis* (which acetylates and inactivates kanamycin),  
81 as well as *tap* (which encodes an efflux pump that extrudes streptomycin).(19) *eis*  
82 promoter mutations have also been described. Recently, cross-resistance between  
83 clofazamine (CFZ) and bedaquiline (BDQ) was shown to be due to mutations within  
84 *Rv0678* (20, 21), a transcriptional repressor, which results in derepression and  
85 upregulation of the multi-substrate efflux pump *mmpL5*.

86

87 **Despite** the discovery of these varied transcriptionally driven mechanisms of resistance,  
88 there have been few systematic whole genome transcriptional comparisons of suitably  
89 matched susceptible and resistant *M. tuberculosis* strains, and none to date using RNA-  
90 sequencing (RNAseq). In this study, we therefore selected phylogenetically closely  
91 related susceptible and resistant clinical strains and subjected them to comparative  
92 transcriptomics using RNAseq to identify novel mechanisms of resistance.

93

94 **RESULTS**

95 **Comparative Transcriptomics**

96 **In** order to identify novel mechanisms of resistance mediated at the level of transcription,  
97 we subjected drug resistant and drug susceptible strains of *M. tuberculosis* to comparative  
98 transcriptomics using RNA sequencing (**Table 1**). We reasoned that strains with highly  
99 complex resistance profiles were most likely to have acquired mutations resulting in  
100 transcriptional changes. Using a whole genome based phylogenetic analysis, we  
101 identified 3 XDR clinical isolates from a well-documented outbreak in KwaZulu-Natal  
102 and a closely related drug susceptible strain to act as a control.(1) All strains were from  
103 the LAM4 branch of Lineage 4. In pairwise comparisons, the 3 XDR strains varied by 7  
104 or less SNPs from each other (**Figure 1A**). The maximum SNP difference between the  
105 drug susceptible and a XDR strain was 76 SNPs, of which 6 occurred in known drug  
106 resistance conferring genes.

107

108 To determine if there were global transcriptional differences between our strains, we first  
109 carried out hierarchical clustering of their transcriptional profiles. This separated the  
110 expression profiles of the three drug-resistant strains from the susceptible control (**Figure**  
111 **1B**). To identify genes either up or down regulated in the XDR strains, we performed  
112 pairwise comparisons for each resistant strain with the drug susceptible control. In the  
113 resistant strains relative to the susceptible control, up to 40 genes were significantly over  
114 or under expressed at the 95% confidence level ( $p\text{-value} \leq 0.05$ ), and up to 10 genes at  
115 the 99% level ( $p\text{-value} \leq 0.01$ ) (**Table S1**). Importantly, in all three pairwise  
116 comparisons, *inhA* showed a greater than 8-fold up-regulation in gene expression in the  
117 resistant strains at the 99% confidence interval. All three resistant strains harboured a t-8a

118 mutation in the promoter region of *fabG1*, which is known to cause up-regulation of *inhA*.  
119 The detection of this transcriptional change therefore acted as an internal validation of  
120 our approach. Apart from *inhA* there were no other genes that were significantly  
121 upregulated in all three comparisons. There were two genes, *fabG1* (also in the *inhA*  
122 operon) and Rv1761c (a gene of unknown function), that had expression levels  
123 significantly different in two strains relative to the susceptible control.

124

125 **After** defining differential gene expression at the statistically significant levels (95% and  
126 99% confidence intervals), we extended our analysis to all genes that had a high mean  
127 fold change ( $\geq 7$  fold up/down) in transcripts relative to the susceptible control (**Figure**  
128 **1C and Table S2**). In addition to *fabG1* and *inhA*, we found that 5 other genes fell into  
129 this classification: *mazF5*, *mazE5* encoding a toxin-antitoxin module; two tRNAs (*leuX*  
130 ,*thrU*) and *ethA*. *ethA* was of particular interest as it encodes a monooxygenase required  
131 for the activation of the prodrug ethionamide (22, 23), a key component of MDR  
132 treatment. Loss of function mutations in *ethA* result in ethionamide resistance.(23, 24)  
133 *ethA* was significantly downregulated following Benjamini Hochberg correction in one of  
134 our pairwise comparisons described above.

135

### 136 **Comparative genome-transcriptome analysis**

137 **In** order to understand the genetic basis of the transcriptional changes defined by our  
138 RNAseq experiments we used comparative genomics to identify mutations located in  
139 intergenic regions associated with genes that were highly over or under expressed in our  
140 resistant strains relative to our susceptible control. This analysis identified an intergenic  
141 mutation at position -11 (t to c) relative to the start codon of *ethA*. The detected mutation

142 was located within the promoter region of *ethA* as well as within the binding domain of  
143 the divergently expressed transcriptional regulator *ethR* that is known to repress *ethA* (25)  
144 (**Figure 2**). The location of the mutation suggested it could lead to down-regulation of  
145 EthA by: i) directly reducing *ethA* transcription independent of *ethR* regulation ii)  
146 increasing *ethR* transcription leading to repression of *ethA*, or iii) affecting binding of  
147 *ethR* leading to increased repression of *ethA* transcription.

148

#### 149 **Functional characterisation of *ethA* and *ethR* promoters**

150 **To** functionally determine if the t-11c mutation influenced either *ethA* or *ethR*  
151 transcription, we used a dual-colour fluorescent protein promoter assay. The episomal  
152 construct pLDW-DC\* has a constitutively expressed TagRFP and a promoterless  
153 Emerald GFP protein in front of which promoters with or without mutations can be  
154 cloned. Promoter activity is expressed as the ratio of green to red fluorescence  
155 normalizing for any variability in plasmid number. To validate our approach, we used the  
156 *fabG1-inhA* promoter with and without the g-17t mutant promoter sequence of *inhA*. The  
157 construct harbouring the g-17t mutant promoter sequence of *inhA* resulted in a 3.4 fold  
158 increase in the ratio of the mean fluorescent intensity (MFI) relative to wild-type (**Figure**  
159 **3A**).

160

161 **We** then assayed constructs with the wild-type 250 bp upstream region of *ethA* or *ethR*,  
162 and 2 matched mutant constructs with either the t-11c mutation (relative to *ethA*) or the  
163 corresponding t- 65c in the *ethR* construct (**Figure 3B**). We observed no significant  
164 change in the MFI ratio between the two *ethR* promoter constructs. In contrast the t-11c  
165 mutant promoter resulted in an MFI ratio that was significantly lower than with the wild-

166 type control. These results suggest that the t-11c intergenic mutation does not affect  
167 transcription of *ethR*, but does diminish expression of *ethA* to levels that could result in  
168 ethionamide resistance.

169

#### 170 ***ethA* expression in clinical isolates of *M. tuberculosis***

171 **To** confirm the transcriptional changes identified by RNAseq in strains harbouring the t-  
172 11c mutation, we used quantitative qRT-PCR to measure the expression levels of *ethA* in  
173 clinical isolates (**Table S3**). In the 5 strains tested with an *ethA* t-11c promoter mutation,  
174 relative normalized expression levels of the monoxygenase were significantly lower or  
175 close to zero compared to control strains. Strains tested with the *inhA* promoter  
176 mutations all had increased relative normalised expression of *inhA* compared to those  
177 without the mutation (**Figure 4**).

178

#### 179 ***ethA* promoter mutations and ethionamide resistance in clinical isolates**

180 **To** determine if *ethA* promoter mutations were associated with clinical resistance, we  
181 tested a panel of clinical isolates, which based on genome analyses harboured putative  
182 ethionamide resistance conferring mutations, for quantitative ethionamide susceptibility  
183 (**Table 2**). The panel included three strains with the t-11c intergenic mutation that had no  
184 other mutations previously associated with clinical ethionamide resistance (*inhA*  
185 promoter mutations and intragenic mutations in *ethA*, *ethR*, *ndh* and *mshA*).<sup>(26)</sup> Recently,  
186 loss of function mutations in another *M. tuberculosis* monoxygenase *mymA* (Rv3083)  
187 have been proposed as an additional resistance mechanism (27). Interestingly during our  
188 selection of strains, we were able to identify a group of isolates with a deletion spanning  
189 *mymA* (**Figure S1**). Sequence confirmation in five of these strains showed an identical



190 deletion of 2891bp, indicating a unique event polymorphism suggestive of clonal  
191 expansion. Strains with this mutation were included in our analysis.

192

193 **The** MIC for strains that only had the *ethA* t-11c promoter mutation ranged from 5 – 20  
194 mg/L, showing a low-level resistance to ETH (**Table 2**). However, in combination with  
195 the t-8a *inhA* promoter mutation, we observed higher levels of resistance suggesting the  
196 phenotypic consequences of the two promoter mutations are additive. The two strains  
197 with only *mymA* deletions had MICs of 2.5 and 5 mg/L, only marginally elevated relative  
198 to the MICs of strains without any ethionamide drug resistance conferring mutations  
199 (1.25 – 2.5 mg/L).

200

#### 201 **Global Distribution of *ethA* promoter mutations**

202 **In** order to determine how widespread *ethA* promoter mutations are among clinical *M.*  
203 *tuberculosis* isolates, we exploited a recent genome analysis of globally isolated drug-  
204 resistant and susceptible strains (28). From a total of 5310 strains, we identified 402 with  
205 a mutation relative to H37Rv within the *ethA-ethR* intergenic region (**Table S4**). One  
206 hundred and thirty-nine of these were the t-11c mutation, all of which were identified in  
207 South African derived lineage 4 isolates. The most common mutation was the a-7g found  
208 in 212 strains, nearly all of which (205) were from Eastern Europe. Eleven other  
209 infrequently occurring mutations were identified, but one of these also mapped to the -11  
210 site (t-11g). Using parsimony to define independent mutational events across the  
211 phylogeny (29), we found that the a-7g mutation had independently evolved at least 32  
212 times, suggesting this mutation was under selective pressure and supporting a role in  
213 conferring drug resistance. In contrast the t-11c was predicted to have evolved only once,

214 which is compatible with ongoing transmission and clonal expansion of XDR strains  
215 from South Africa in which the mutation was found.

216

## 217 **DISCUSSION**

218 **The** aim of our study was to use a comparative whole genomic-transcriptomic approach  
219 to identify novel mechanisms of resistance mediated at the transcriptional level. We were  
220 able to identify a promoter mutation upstream of *ethA* that we confirmed by our dual-  
221 colour promoter assay and quantitative RT-PCR leads to reduced transcription of *ethA*,  
222 which encodes a monooxygenase that activates the prodrug ethionamide.(22, 23) Strains  
223 harbouring only this t-11c mutation and no other ethionamide resistant determining  
224 genotypes were resistant to ethionamide (MICs ranging from 5 – 20mg/L) indicating this  
225 mutation should be included in genetic based diagnostic tests. In support of this, a recent  
226 genome wide association analysis also reported an association between the t-11c and  
227 ethionamide resistance.(8)

228

229 **In** our analysis of the global distribution of *ethA-ethR* intergenic mutations, we identified  
230 the t-11c mutation solely in South African strains. This data set however, consisted of  
231 isolates from only 43 countries and notably lacked representation from several regions of  
232 the world where TB is epidemic, such as South America. Nonetheless a previous study,  
233 using direct sequencing of drug resistance loci, detected five different variants in the  
234 promoter region of *ethA*, one of which was a t-11c mutation found in an isolate from Peru  
235 (30). There was however no clinical data available to rule out whether the patient in  
236 question had any travel history to South Africa. We can therefore not confirm whether  
237 this mutation is geographically restricted. The a-7g mutation was more dispersed, but the

238 majority of strains harbouring this mutation were from Eastern Europe. A previous study  
239 reported the phenotype of 172 strains with the a-7g mutation and only 56 of these were  
240 resistant to ethionamide.(31) This could be due to inconsistencies in drug susceptibility  
241 testing, or the level of resistance conferred lying close to the break point used in  
242 susceptibility testing but suggests that not all mutations within the intergenic *ethA-ethR*  
243 result in similar levels of resistance that we observed with the t-11c mutation. (30)

244

245 **Few** studies have used quantitative drug susceptibility testing to correlate ethionamide  
246 MIC with genotype (32) so it is unclear to what extent individual mutations contribute to  
247 resistance and how they might interact. Although there may be additional mechanisms of  
248 ethionamide resistance, yet to be identified, our results suggest that the t-11c mutation  
249 causes a modest increase in ethionamide MIC but in combination with an *inhA* promoter  
250 mutation (considered to cause low-level ethionamide resistance (24)) leads to high level  
251 resistance. Among the other ethionamide resistant strains assayed, most had more than  
252 one mutation potentially contributing to their increased MIC. The pathway to clinical  
253 ethionamide resistance may therefore be the step-wise accumulation of multiple  
254 mutations rather than the selection of a single high-level resistance conferring mutations  
255 as seen with some other anti-tuberculosis drugs.

256

257 **In** the panel of clinical isolates we selected to evaluate the phenotype of the t-11c  
258 mutation, we identified polymorphisms in other genes implicated in ethionamide  
259 resistance. We found four mutations at three positions in the *inhA* promoter region all of  
260 which have been previously described.(33) One of these strains had a t-8g *inhA* promoter  
261 mutation in combination with a non-synonymous mutation (V18A) in *ndh* which encodes

262 a type II NADH dehydrogenase. Mutations in *ndh* can result in increased levels of NADH  
263 and reduce binding of the isoniazid and ethionamide NAD adducts to their target  
264 InhA.(33) However this strain had a low MIC suggesting neither of these mutations  
265 causes high-level resistance.

266

267 **We** cannot rule out the existence of other ethionamide resistance mechanisms in our  
268 strains. EthA is one of 30 other monooxygenases within the *M. tuberculosis* genome (24)  
269 and a recently characterized monooxygenase, *mymA* (Rv3083) (27) was proposed as an  
270 additional enzyme responsible for the activation of ethionamide. We identified strains  
271 with a 2891 bp deletion spanning *mymA*, *lipR* and half of *Rv3085* (**Figure S1**). Two of  
272 these strains had no other known mutations associated with ethionamide resistance but  
273 were susceptible to ethionamide employing a standard MIC cut-off (**Table 2**), suggesting  
274 *mymA* is not important for drug resistance in clinical isolates.

275

276 **Our** initial comparative transcriptional analysis only identified a limited number of genes  
277 whose expression was statistically different from the control. This may have been due to  
278 increased variability associated with propagating clinical isolates in culture media. We,  
279 therefore, looked at genes whose expression was highly divergent in the resistant strains  
280 in all pair wise comparisons. In addition to *ethA* this identified *mazF5* and *mazE5*, which  
281 encode a toxin antitoxin system, one of nine MazEF homologues in *M. tuberculosis*. A  
282 triple null mutant of *mazF3*, *F6* and *F9* was less able to survive antituberculosis drugs  
283 (34) so potentially these systems could be involved in mediating resistance, although it is  
284 unclear how downregulation of *mazE5* would influence drug susceptibility. Two tRNAs,  
285 *leuX* and *thrU*, were amongst the genes most highly upregulated in the resistant strains.

286 Beyond a fundamental role in translation, tRNAs, and their degradation products have  
287 been shown to regulate stress responses and adaptive changes in translation.(35) It is  
288 therefore conceivable that the upregulation of these two tRNAs may be a manifestation of  
289 more global regulatory changes that have occurred during the evolution of drug  
290 resistance. Future studies comprising strains from different outbreaks and lineages are  
291 however needed to determine whether these transcriptional changes are limited to the  
292 XDR outbreak from KwaZulu-Natal in 2005 (1).

293

294 **The** treatment of MDR-TB is currently undergoing a revolution with the introduction of  
295 new drugs and regimens.(36) WHO has recently approved the use of a 9-month short  
296 course of therapy, and the 4-month intensive phase of this regimen includes ethionamide  
297 (or its analogue prothionamide). Although the contribution of individual drugs to  
298 treatment efficacy is unclear, it is recommended that short course treatment should be  
299 withheld from MDR-TB patients with pre-existing resistance to any individual drug. Pre-  
300 treatment screening for ethionamide resistance is therefore critical for the implementation  
301 of short course MDR treatment. However phenotypic susceptibility testing for  
302 ethionamide is notoriously difficult.(37) Our results contribute to the development of a  
303 genetic based resistance test, but further studies are required to define the interaction of  
304 diverse mutations and drug resistance conferring loci as well as establishing a clinical  
305 relevant critical concentration for ethionamide.

306

## 307 **MATERIALS AND METHODS**

### 308 **Strains and growth conditions**

309 **Three** XDR and 1 fully drug susceptible clinical isolate from the LAM4 (KZN)  
310 spoligotype of *M. tuberculosis* (**Table 1**) were obtained from archived cultures that had  
311 been previously genome sequenced from single colonies.(1) Cultures were grown in  
312 triplicate at 37°C in BD Difco™ Middlebrook 7H9 broth supplemented with BBL™  
313 Middlebrook OADC enrichment media, 0.5% glycerol and 0.01% Tween 80, with  
314 continuous shaking at 200rpm. Additional strains were selected from the same collection  
315 based on specific *ethA*, *ethR*, *inhA*, and *mymA* genotypes (**Table 2**).(1)

316

### 317 **RNA extraction and quality control**

318 **RNA Extraction.** RNA was harvested from 25ml cultures grown to an OD<sub>600</sub> of between  
319 0.5 – 0.8, using a modified TRIzol method.(38) Briefly the cultures were centrifuged at  
320 4000 rpm for 20 minutes at 25°C and the pellet re-suspended in 1ml of TRIzol® reagent  
321 (Invitrogen, USA). Thereafter approximately 100 µl of 0.1mm Zirconia/Silica glass beads  
322 (BioSpec Products, USA) were added and the cultures subjected to four pulses of bead  
323 beating using the Roche MagNA Lyser at 7000rpm for 60s, with two minute intermittent  
324 incubations on ice. Immediately after bead beating, 200 µl of chloroform was added  
325 followed by centrifugation at 15 000 rpm for 15 minutes at 4°C and separation of the  
326 aqueous phase. The RNA was precipitated with 500µl of 100% isopropanol and  
327 incubated at -20°C for 1 hour. After centrifugation at 15 000 rpm for 10 minutes at 4°C  
328 the RNA pellet was washed with 1 ml 75% ethanol, centrifuged at 10 000 rpm for 5  
329 minutes at 4°C and air-dried. The RNA pellet was then dissolved in 30ul of RNase-free  
330 water.

331

332 ***DNase treatment and purification.*** The RNA was subjected to DNase treatment using  
333 the DNase I, RNase-free kit (Thermo Scientific, USA), as per manufacturer's  
334 instructions. The RNA was then purified using the RNeasy Mini Kit (Qiagen, Germany),  
335 during which a second round of DNase digestion took place utilizing the RNase-free  
336 DNase Set (Qiagen, Germany). The integrity of the RNA samples was confirmed using  
337 an 23S/16S ratio (above  $\geq 1.2$ ) determined by an Experion™ Std Sens Analysis Kit (Bio-  
338 Rad, USA).

339

#### 340 **RNA sequencing and bioinformatics analysis**

341 ***RNaseq library preparation.*** The Qubit™ RNA Assay kit (Invitrogen, USA) was used  
342 with the Qubit®2.0 Fluorometer to quantify the RNA. Following RNA quantification,  
343 rRNA was depleted using the Ribo-Zero Magnetic Kit (Illumina, USA). Enriched mRNA  
344 was analysed on a RNA specific E-gel EX 2% (Invitrogen, USA) to confirm rRNA  
345 removal. After purification of the mRNA using the RNeasy Mini Kit (Qiagen, Germany)  
346 RNA sequencing libraries were constructed using the NEBNext® Ultra™ Directional  
347 RNA Library Prep Kit for Illumina (New England BioLabs Inc, USA). The prepared  
348 libraries were indexed with NEBNext Multiplex Oligos for Illumina (New England  
349 BioLabs Inc. USA) and sequenced with 50 bp single end reads on an Illumina HiSeq  
350 2000 platform at the Norwegian Sequencing Centre, Oslo, Norway.

351

352 ***Bioinformatics.*** The sequence reads were aligned to the *M. tuberculosis* H37Rv genome  
353 (NCBI accession NC\_000962.2) using SeqMan NGen from the DNASTAR Lasergene 11  
354 software. Transcripts for each sample were quantified and normalized as reads per  
355 kilobase per million reads (RPKM). The three replicate RPKM values for each sample

356 were standardised based on their mean transcript values and were used to assess gene  
357 expression and fold change differences between isolates using ArrayStar (DNASTAR).  
358 Pairwise comparisons between strains were conducted with confidence intervals and  
359 statistics determined using a Students T-test, with multiple testing corrections using the  
360 Benjamini Hochberg correction to reduce False Discovery Rate (FDR). Intergenic SNPs  
361 present only in the three XDR strains and not in the drug susceptible strain were  
362 identified from whole genome sequencing data from a previous study (1). A  
363 transcriptomic-genomic analysis was then conducted to identify promoter SNPs  
364 associated with at least a 4-fold upregulation of the downstream gene.

365

#### 366 **Whole-genome phylogeny**

367 **Sequence** reads for the four KZN strains were downloaded from SRA (run accessions  
368 SRR832991, SRR833024, SRR833121 and SRR924700). Reads were aligned to the  
369 H37Rv genome (NC\_000962.3) using SeqMan NGen (DNASTAR), resulting in median  
370 alignment depths ranging from 184× to 330× for individual isolates. SNPs were called  
371 and filtered as previously described.(39) The concatenated SNPs were used to create a  
372 distance-based neighbour-joining tree.

373

#### 374 **qRT-PCR**

375 **RNA** was reverse transcribed into cDNA using the iScript cDNA Synthesis Kit (Bio-Rad,  
376 USA). Quantitative Real-Time PCR was conducted using the iTaq™ Universal SYBR®  
377 Green Supermix (Bio-Rad, USA) with forward and reverse primers for selected genes of  
378 interest. Primers were designed for *ethA*, *inhA*, *ethR* and a housekeeping gene *sigA*  
379 (**Table S5**). Expression levels were normalized to the reference gene *sigA*.



380

381 **Flow cytometry Promoter Reporter Assay**

382 To create a dual colour reporter the Multisite Gateway<sup>®</sup> Three-Fragment Vector  
383 Construction method (Invitrogen, USA) was used. The *ethA-ethR* intergenic region,  
384 mycobacterial codon optimized Emerald GFP, and mycobacterial codon optimized  
385 TagRFP constitutively expressed by the promoter pUV15, were individually cloned into  
386 entry vectors. These were combined with a destination vector based on an episomal  
387 mycobacterial vector containing a kanamycin resistance cassette (*aph*), mycobacterial  
388 origin of replication and *E. coli* origin of replication. Four separate *ethA-ethR* intergenic  
389 regions were used corresponding to the wild-type and mutant sequences (generated by  
390 PCR using genomic DNA from resistant clinical isolates) upstream of *ethA*, and the same  
391 pair in the reverse orientation corresponding to the sequences upstream of *ethR* (**Table**  
392 **S6**). Additional plasmids were constructed with the *inhA* promoter with and without an g-  
393 17t mutation and a non promoter region (intragenic *katG* sequence) cloned in front of the  
394 GFP (**Figure S2**). Promoter sequences for each construct were sequenced confirmed.  
395 Respective plasmids were transformed into H37Rv using standard protocols.(40)

396

397 **Strains** harbouring the dual-colour reporters were grown up to mid-log phase (OD<sub>600</sub> of  
398 0.5 - 0.8) in 7H9 media containing 25mg/L kanamycin. 1mL of each strain was then  
399 filtered through a 10 micron filter and acquired on the BD FACS Aria III using BD DIVA  
400 software. 100 000 events were recorded with single cell acquisition set at a threshold rate  
401 of ~ 5000-7000 events per second. Green and red fluorescence were detected using the  
402 FITC and PI filters respectively. The gating strategy employed during acquisition and  
403 software analysis, using FlowJo V10, differentiated single cells/events based on the

17

404 relationship between cell size (forward scatter – FSC) and granularity (side scatter –  
405 SSC). Secondary gating was done using FlowJo on events with red fluorescent signal to  
406 ensure only cells containing expression vectors were included in our analysis. Median  
407 fluorescent intensity (MFI) of red and green fluorescent signals were extracted. MFI of  
408 green fluorescence was normalized to MFI of red fluorescence for each replicate before  
409 calculating mean and standard deviation. A two sided t-test was used to determine  
410 statistical significance.

411

#### 412 **Drug susceptibility testing**

413 **100ul** of three dilutions of each strain, including  $1 \times 10^6$ ,  $1 \times 10^4$  and  $1 \times 10^3$  cells, were  
414 plated out onto quadrant plates containing BD Difco™ Middlebrook 7H10 agar with  
415 varying drug concentrations of ethionamide (1.25, 2.5, 5, 10, 20, 40 and 80mg/L) and  
416 counted for CFU after 3 weeks incubation at 37°C.

417

#### 418 **Global Distribution of *ethA* promoter mutations**

419 A global dataset of 5310 *M. tuberculosis* strains from five continents (28) was searched  
420 for all instances of *ethA* promoter mutations. To identify individual mutation arisal  
421 events across the phylogeny, we performed parsimony-based analysis using the PAUP  
422 software, version 4.0b10 (29) as in Manson et al (28).

423

#### 424 **FUNDING INFORMATION**

425 **This** project received funding from the Africa Health Institute (AHRI) and has also been  
426 funded in part with Federal funds from the National Institute of Allergy and Infectious  
427 Diseases, National Institutes of Health, Department of Health and Human Services under

428 Grant No. U19AI110818, and Contract No. HHSN272200900018C to the Broad Institute.  
429 The funders had no role in study design, data collection and interpretation, or the decision  
430 to submit the work for publication.

431

#### 432 **ACKNOWLEDGEMENTS**

433 **We** would like to acknowledge and thank M.R Farhat and M. Murray for providing us  
434 with additional information on previous work they had conducted surrounding the *ethA*-  
435 *ethR* locus, A.R Baulard for his helpful comments and suggestions, K.A Cohen for her  
436 insights and information regarding the clinical strains used in this study and V.  
437 Munsamy-Govender for her assistance in locating and propagating some of these strains.

438

#### 439 **REFERENCES**

- 440 1. Cohen KA, Abeel T, Manson McGuire A, Desjardins CA, Munsamy V, Shea TP,  
441 Walker BJ, Bantubani N, Almeida D V., Alvarado L, Chapman SB, Mvelase NR,  
442 Duffy EY, Fitzgerald MG, Govender P, Gujja S, Hamilton S, Howarth C, Larimer  
443 JD, Maharaj K, Pearson MD, Priest ME, Zeng Q, Padayatchi N, Grosset J, Young  
444 SK, Wortman J, Mlisana KP, O'Donnell MR, Birren BW, Bishai WR, Pym AS,  
445 Earl AM. 2015. Evolution of Extensively Drug-Resistant Tuberculosis over Four  
446 Decades: Whole Genome Sequencing and Dating Analysis of Mycobacterium  
447 tuberculosis Isolates from KwaZulu-Natal. PLOS Med 12:e1001880.
- 448 2. O'Donnell MR, Padayatchi N, Kvasnovsky C, Werner L, Master I, Horsburgh CR.  
449 2013. Treatment outcomes for extensively drug-resistant tuberculosis and HIV Co-  
450 infection. Emerg Infect Dis 19:416–424.
- 451 3. Pietersen E, Ignatius E, Streicher EM, Mastrapa B, Padanilam X, Pooran A, Badri

- 452 M, Lesosky M, van Helden P, Sirgel FA, Warren R, Dheda K. 2014. Long-term  
453 outcomes of patients with extensively drug-resistant tuberculosis in South Africa: a  
454 cohort study. *Lancet (London, England)* 383:1230–1239.
- 455 4. Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, Gler MT,  
456 Blakemore R, Worodria W, Gray C, Huang L, Caceres T, Mehdiyev R, Raymond  
457 L, Whitelaw A, Sagadevan K, Alexander H, Albert H, Cobelens F, Cox H, Alland  
458 D, Perkins MD. 2011. Feasibility, diagnostic accuracy, and effectiveness of  
459 decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and  
460 multidrug resistance: a multicentre implementation study. *Lancet (London,  
461 England)* 377:1495–1505.
- 462 5. Hillemann D, Rusch-Gerdes S, Richter E. 2007. Evaluation of the GenoType  
463 MTBDRplus assay for rifampin and isoniazid susceptibility testing of  
464 *Mycobacterium tuberculosis* strains and clinical specimens. *J Clin Microbiol*  
465 45:2635–2640.
- 466 6. Hillemann D, Rusch-Gerdes S, Richter E. 2009. Feasibility of the GenoType  
467 MTBDRsl assay for fluoroquinolone, amikacin-capreomycin, and ethambutol  
468 resistance testing of *Mycobacterium tuberculosis* strains and clinical specimens. *J*  
469 *Clin Microbiol* 47:1767–1772.
- 470 7. Brown AC, Bryant JM, Einer-Jensen K, Holdstock J, Houniet DT, Chan JZM,  
471 Depledge DP, Nikolayevskyy V, Broda A, Stone MJ, Christiansen MT, Williams  
472 R, McAndrew MB, Tutill H, Brown J, Melzer M, Rosmarin C, McHugh TD,  
473 Shorten RJ, Drobniewski F, Speight G, Breuer J. 2015. Rapid whole-genome  
474 sequencing of *mycobacterium tuberculosis* isolates directly from clinical samples.  
475 *J Clin Microbiol* 53:2230–2237.

- 476 8. Desjardins CA, Cohen KA, Munsamy V, Abeel T, Maharaj K, Walker BJ, Shea  
477 TP, Almeida D V, Manson AL, Salazar A, Padayatchi N, O'Donnell MR, Mlisana  
478 KP, Wortman J, Birren BW, Grosset J, Earl AM, Pym AS. 2016. Genomic and  
479 functional analyses of *Mycobacterium tuberculosis* strains implicate *ald* in D-  
480 cycloserine resistance. *Nat Genet* 48:544–551.
- 481 9. Walker TM, Kohl TA, Omar S V., Hedge J, Del Ojo Elias C, Bradley P, Iqbal Z,  
482 Feuerriegel S, Niehaus KE, Wilson DJ, Clifton DA, Kapatai G, Ip CLC, Bowden  
483 R, Drobniewski FA, Allix-Béguec C, Gaudin C, Parkhill J, Diel R, Supply P,  
484 Crook DW, Smith EG, Walker AS, Ismail N, Niemann S, Peto TEA, Davies J,  
485 Crichton C, Acharya M, Madrid-Marquez L, Eyre D, Wyllie D, Golubchik T,  
486 Munang M. 2015. Whole-genome sequencing for prediction of *Mycobacterium*  
487 tuberculosis drug susceptibility and resistance: A retrospective cohort study.  
488 *Lancet Infect Dis* 15:1193–1202.
- 489 10. Coll F, Mc Nerney R, Preston MD, Guerra-Assunção JA, Warry A, Hill-Cawthorne  
490 G, Mallard K, Nair M, Miranda A, Alves A, Perdigão J, Viveiros M, Portugal I,  
491 Hasan Z, Hasan R, Glynn JR, Martin N, Pain A, Clark TG. 2015. Rapid  
492 determination of anti-tuberculosis drug resistance from whole-genome sequences.  
493 *Genome Med* 7:51.
- 494 11. Zhang H, Li D, Zhao L, Fleming J, Lin N, Wang T, Liu Z, Li C, Galwey N, Deng  
495 J, Zhou Y, Zhu Y, Gao Y, Wang T, Wang S, Huang Y, Wang M, Zhong Q, Zhou  
496 L, Chen T, Zhou J, Yang R, Zhu G, Hang H, Zhang J, Li F, Wan K, Wang J,  
497 Zhang X-E, Bi L. 2013. Genome sequencing of 161 *Mycobacterium tuberculosis*  
498 isolates from China identifies genes and intergenic regions associated with drug  
499 resistance. *Nat Genet* 45:1255–60.

- 500 12. Cohen KA, Bishai WR, Pym AS. 2014. Molecular Basis of Drug Resistance in  
501 *Mycobacterium tuberculosis*. *Microbiol Spectr* 2.
- 502 13. Siddiqi N, Das R, Pathak N, Banerjee S, Ahmed N, Katoch VM, Hasnain SE.  
503 2004. *Mycobacterium tuberculosis* isolate with a distinct genomic identity  
504 overexpresses a tap-like efflux pump. *Infection* 32:109–111.
- 505 14. Louw GE, Warren RM, van Pittius NC, Leon R, Jimenez A, Hernandez-Pando R,  
506 McEvoy CRE, Grobbelaar M, Murray M, van Helden PD, Victor TC. 2011.  
507 Rifampicin reduces susceptibility to ofloxacin in rifampicin-resistant  
508 *Mycobacterium tuberculosis* through efflux. *Am J Respir Crit Care Med*  
509 184:269–276.
- 510 15. Mdluli K, Sherman DR, Hickey MJ, Kreiswirth BN, Morris S, Stover CK, Barry  
511 CE. 1996. Biochemical and genetic data suggest that *InhA* is not the primary target  
512 for activated isoniazid in *Mycobacterium tuberculosis*. *J Infect Dis* 174:1085–90.
- 513 16. Scorpio A, Lindholm-Levy P, Heifets L, Gilman R, Siddiqi S, Cynamon M, Zhang  
514 Y. 1997. Characterization of *pncA* mutations in pyrazinamide-resistant  
515 *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 41:540–543.
- 516 17. Juréen P, Werngren J, Toro JC, Hoffner S. 2008. Pyrazinamide resistance and  
517 *pncA* gene mutations in *Mycobacterium tuberculosis*. *Antimicrob Agents*  
518 *Chemother* 52:1852–1854.
- 519 18. Yüksel P, Tansel Ö. 2009. Characterization of *pncA* mutations of pyrazinamide-  
520 resistant *Mycobacterium tuberculosis* in Turkey. *New Microbiol* 32:153–158.
- 521 19. Reeves AZ, Campbell PJ, Sultana R, Malik S, Murray M, Plikaytis BB, Shinnick  
522 TM, Posey JE. 2013. Aminoglycoside cross-resistance in *Mycobacterium*  
523 *tuberculosis* due to mutations in the 5' untranslated region of *whiB7*. *Antimicrob*

- 524 Agents Chemother 57:1857–65.
- 525 20. Hartkoorn RC, Uplekar S, Cole ST. 2014. Cross-resistance between clofazimine  
526 and bedaquiline through upregulation of mmpL5 in mycobacterium tuberculosis.  
527 Antimicrob Agents Chemother 58:2979–2981.
- 528 21. Andries K, Villellas C, Coeck N, Thys K, Gevers T, Vranckx L, Lounis N, De  
529 Jong BC, Koul A. 2014. Acquired resistance of Mycobacterium tuberculosis to  
530 bedaquiline. PLoS One 9:1–11.
- 531 22. Baulard a R, Betts JC, Engohang-Ndong J, Quan S, McAdam R a, Brennan PJ,  
532 Locht C, Besra GS. 2000. Activation of the pro-drug ethionamide is regulated in  
533 mycobacteria. J Biol Chem 275:28326–31.
- 534 23. DeBarber AE, Mdluli K, Bosman M, Bekker LG, Barry CE. 2000. Ethionamide  
535 activation and sensitivity in multidrug-resistant Mycobacterium tuberculosis. Proc  
536 Natl Acad Sci U S A 97:9677–82.
- 537 24. Morlock GP, Metchock B, Sikes D, Crawford JT, Cooksey RC. 2003. ethA, inhA,  
538 and katG loci of ethionamide-resistant clinical Mycobacterium tuberculosis  
539 isolates. Antimicrob Agents Chemother 47:3799–3805.
- 540 25. Engohang-Ndong J, Baillat D, Aumercier M, Bellefontaine F, Besra GS, Locht C,  
541 Baulard AR. 2004. EthR, a repressor of the TetR/CamR family implicated in  
542 ethionamide resistance in mycobacteria, octamerizes cooperatively on its operator.  
543 Mol Microbiol 51:175–188.
- 544 26. Vilchèze C, Jacobs JR. WR. 2014. Resistance to Isoniazid and Ethionamide in  
545 Mycobacterium tuberculosis: Genes, Mutations, and Causalities. Microbiol Spectr  
546 2:1–21.
- 547 27. Grant SS, Wellington S, Kawate T, Desjardins CA, Silvis MR, Wivagg C,

- 548 Thompson M, Gordon K, Kazyanskaya E, Nietupski R, Haseley N, Iwase N, Earl  
549 AM, Fitzgerald M, Hung DT. 2016. Baeyer-Villiger Monooxygenases EthA and  
550 MymA Are Required for Activation of Replicating and Non-replicating  
551 Mycobacterium tuberculosis Inhibitors. *Cell Chem Biol* 23:666–677.
- 552 28. Manson AL, Cohen KA, Abeel T, Desjardins CA, Armstrong DT, Barry CE 3rd,  
553 Brand J, Chapman SB, Cho S-N, Gabrielian A, Gomez J, Jodals AM, Joloba M,  
554 Jureen P, Lee JS, Malinga L, Maiga M, Nordenberg D, Noroc E, Romancenco E,  
555 Salazar A, Ssengooba W, Velayati AA, Winglee K, Zalutskaya A, Via LE, Cassell  
556 GH, Dorman SE, Ellner J, Farnia P, Galagan JE, Rosenthal A, Crudu V,  
557 Homorodean D, Hsueh P-R, Narayanan S, Pym AS, Skrahina A, Swaminathan S,  
558 Van der Walt M, Alland D, Bishai WR, Cohen T, Hoffner S, Birren BW, Earl AM.  
559 2017. Genomic analysis of globally diverse Mycobacterium tuberculosis strains  
560 provides insights into the emergence and spread of multidrug resistance. *Nat*  
561 *Genet* 49:395–402.
- 562 29. Swofford D. 1991. PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.
- 563 30. Farhat MR, Sultana R, Iartchouk O, Bozeman S, Galagan J, Sisk P, Stolte C,  
564 Nebenzahl-Guimaraes H, Jacobson K, Sloutsky A, Kaur D, Posey J, Kreiswirth  
565 BN, Kurepina N, Rigouts L, Streicher EM, Victor TC, Warren RM, van Soolingen  
566 D, Murray M. 2016. Genetic Determinants of Drug Resistance in Mycobacterium  
567 tuberculosis and Their Diagnostic Value. *Am J Respir Crit Care Med* 194:621–  
568 630.
- 569 31. Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I,  
570 Corander J, Bryant J, Parkhill J, Nejentsev S, Horstmann RD, Brown T,  
571 Drobniewski F. 2014. Evolution and transmission of drug-resistant tuberculosis in



- 572 a Russian population. *Nat Genet* 46:279–286.
- 573 32. Rueda J, Realpe T, Mejia GI, Zapata E, Rozo JC, Ferro BE, Robledo J. 2015.  
574 Genotypic Analysis of Genes Associated with Independent Resistance and Cross-  
575 Resistance to Isoniazid and Ethionamide in *Mycobacterium tuberculosis* Clinical  
576 Isolates. *Antimicrob Agents Chemother* 59:7805–7810.
- 577 33. Vilcheze C, Weisbrod TR, Chen B, Kremer L, Hazbon MH, Wang F, Alland D,  
578 Sacchettini JC, Jacobs WRJ. 2005. Altered NADH/NAD<sup>+</sup> ratio mediates  
579 coresistance to isoniazid and ethionamide in mycobacteria. *Antimicrob Agents*  
580 *Chemother* 49:708–720.
- 581 34. Tiwari P, Arora G, Singh M, Kidwai S, Narayan OP, Singh R. 2015. MazF  
582 ribonucleases promote *Mycobacterium tuberculosis* drug tolerance and virulence in  
583 guinea pigs. *Nat Commun* 6:6059.
- 584 35. Kirchner S, Ignatova Z. 2015. Emerging roles of tRNA in adaptive translation,  
585 signalling dynamics and disease. *Nat Rev Genet* 16:98–112.
- 586 36. 2016. WHO treatment guidelines for drug- resistant tuberculosis 2016.
- 587 37. Kaniga K, Cirillo DM, Hoffner S, Ismail NA, Kaur D, Lounis N, Metchock B,  
588 Pfyffer GE, Venter A. 2016. A Multilaboratory, Multicountry Study To Determine  
589 MIC Quality Control Ranges for Phenotypic Drug Susceptibility Testing of  
590 Selected First-Line Antituberculosis Drugs, Second-Line Injectables,  
591 Fluoroquinolones, Clofazimine, and Linezolid. *J Clin Microbiol* 54:2963–2968.
- 592 38. Alli, Terry O a., Mangan J a., Al E. 2009. Optimization of Rna Extraction in  
593 *Mycobacterium Tuberculosis* for Studying Intracellular Gene Expression . *African*  
594 *J Clin Exp Microbiol* 10:64–79.
- 595 39. Eldholm V, Pettersson JH-O, Brynildsrud OB, Kitchen A, Rasmussen EM,

596 Lillebaek T, Rønning JO, Crudu V, Mengshoel AT, Debech N, Alfsnes K, Bohlin  
597 J, Pepperell CS, Balloux F. 2016. Armed conflict and population displacement as  
598 drivers of the evolution and dispersal of Mycobacterium tuberculosis. Proc Natl  
599 Acad Sci U S A 113:201611283.  
600 40. van Kessel JC, Hatfull GF. 2007. Recombineering in Mycobacterium tuberculosis.  
601 Nat Methods 4:147–152.

602

### 603 **FIGURE LEGENDS**

604 **FIG 1:** (A) Phylogenetic tree representing the distribution of the 4 strains (shown in red  
605 and boxed) selected for RNAseq. (B) Hierarchical gene clustering of these strains, based  
606 on their relative gene expression, shows that the drug susceptible strain clusters  
607 separately from the others. (C) Venn diagram representing genes differentially expressed  
608 7-fold or greater relative to the susceptible control. The blue, red and green circles  
609 represent pairwise comparisons with TKK-01-0033, TKK-01-0025 and TKK 01-0040  
610 respectively

611

612 **FIG 2:** Representation of the intergenic region between *ethA* and *ethR*. The location of  
613 the single nucleotide polymorphism (SNP) is found 11 base pairs upstream of *ethA* and is  
614 indicated in red. The *ethR* binding region is indicated by the black box (25)

615

616 **FIG 3:** Analysis of promoter activity between wild type and mutant constructs. The left  
617 panels represent the ratio of the mean fluorescence intensity (MFI) of green fluorescent  
618 protein (GFP) to red fluorescent protein (RFP), as well as statistical differences between  
619 the wild type and mutant constructs for (A) *inhA* promoter and (B) *ethA* and *ethR*

620 promoters. p - values are indicated on the bar charts. The panels on the right represent  
621 single cell counts from flow cytometry. RFP expression is represented on the y-axis as  
622 the PI-H channel and GFP expression is represented on the x-axis as the FITC-H channel.  
623

624 **FIG 4:** Relative gene expression of *ethA*, *ethR* or *inhA* in clinical strains of *M.*  
625 *tuberculosis* (Table S5). Gene expression levels were normalised to *sigA* for each strain.  
626 Relative normalised expression represents the fold change in normalised expression of  
627 each strain compared to the drug susceptible clinical strain 84. Light blue bars represent  
628 strains that do not contain t-11c *ethA* promoter mutations and dark blue bars represent  
629 strains that have the t-11c *ethA* promoter mutation. Light pink bars represent strains  
630 without *inhA* promoter mutations and dark pink bars represent strains with *inhA* promoter  
631 mutations. TKK strain numbers are abbreviated to their last two digits. E.g. 62 represents  
632 TKK-01-0062. Statistical significance of relative normalised expression for *ethA* and  
633 *inhA* was derived using unpaired t-tests between each strain and the clinical drug  
634 susceptible strain 84 (shown in black). In addition, statistical significance of relative  
635 normalised expression for *ethA* was derived using unpaired t-tests between each strain  
636 and strain 62, which does not harbour a t-11c *ethA* promoter mutation (shown in red).  
637 1551 corresponds to the laboratory strain CDC1551 and was excluded from this analysis.  
638 \*\* = p-value  $\leq 0.01$ , \* = p-value  $\leq 0.05$ , NS = not significant

639

640 **TABLE 1:** Strain details including resistance mutations and RNA sequencing coverage

641

642 **TABLE 2:** Minimal inhibitory concentrations (MICs) of clinical strains to ethionamide  
643 (ETH)



FIG 1

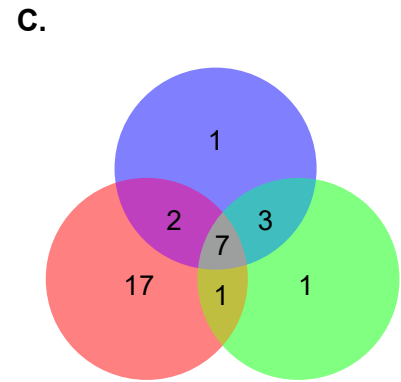
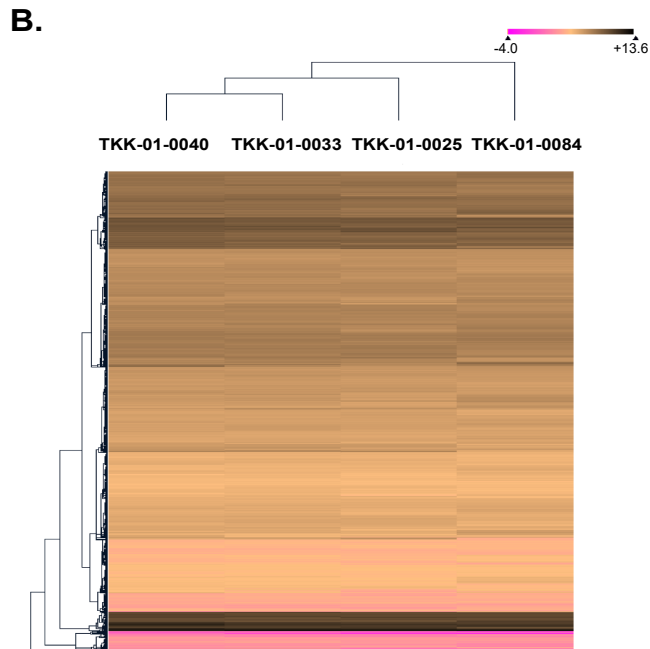
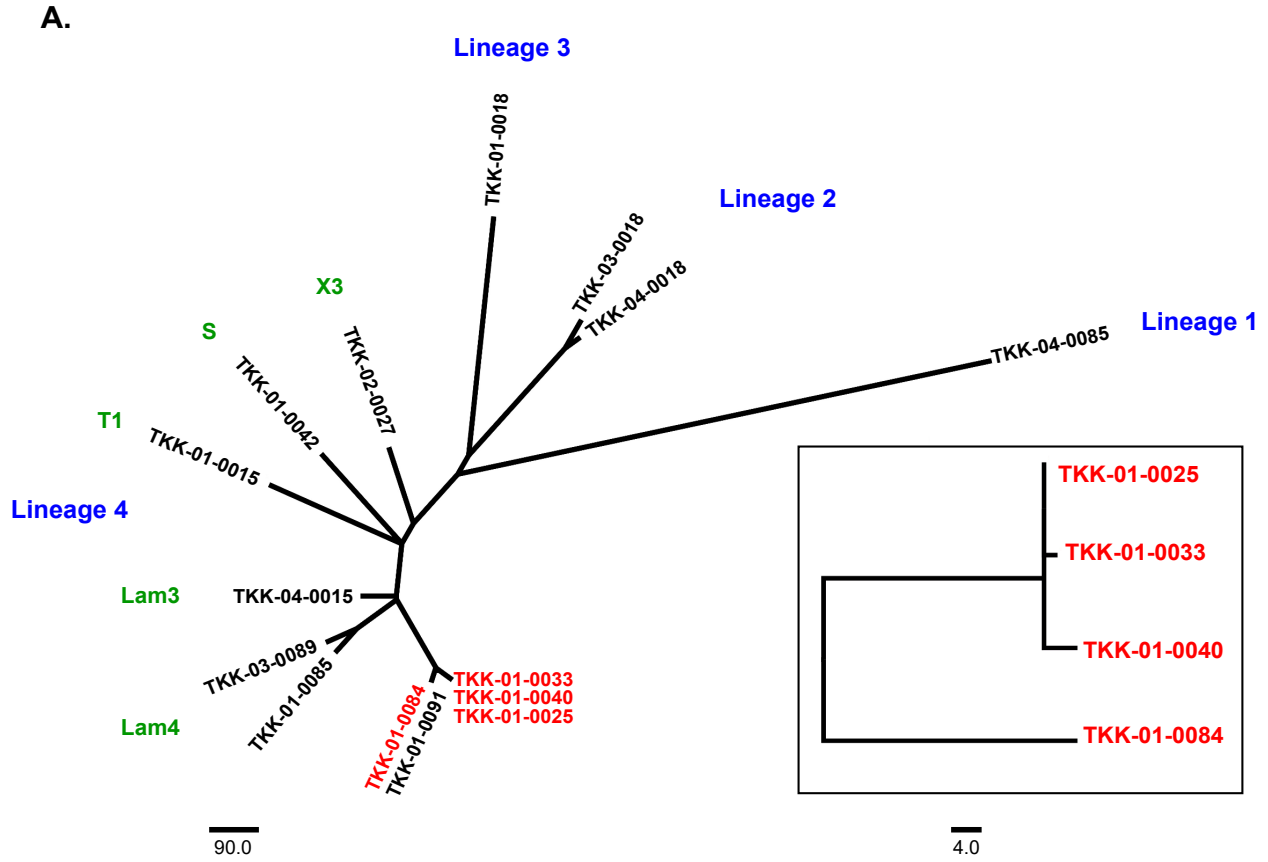


FIG 2

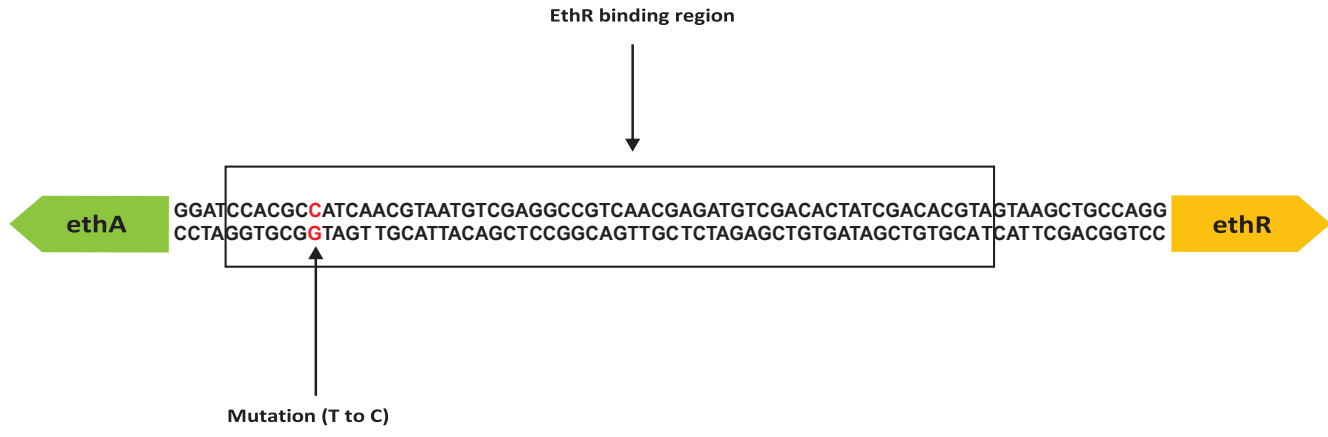
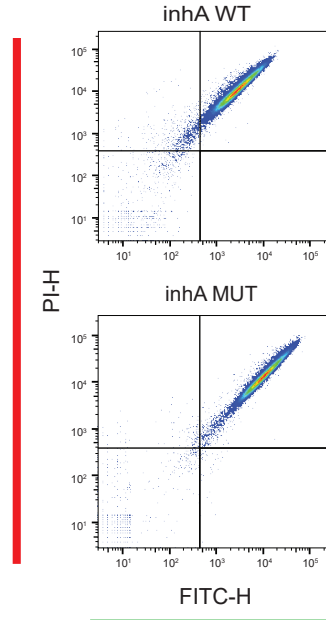
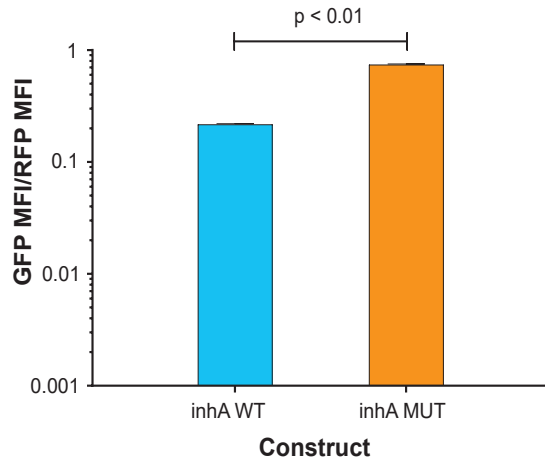


FIG 3

A.



B.

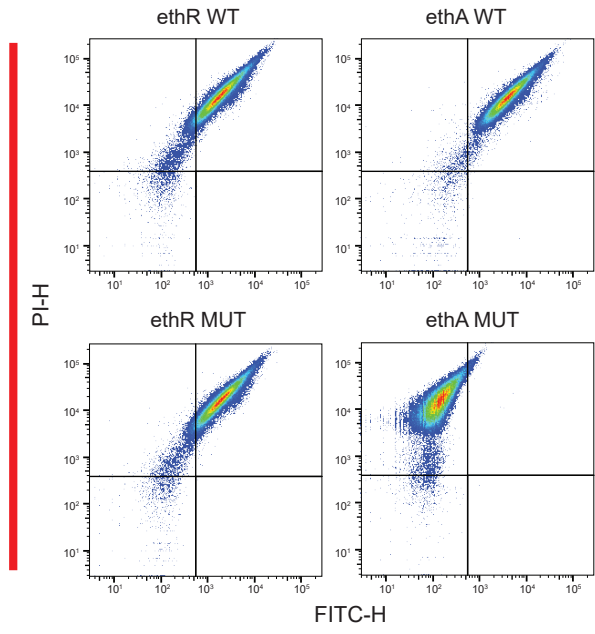
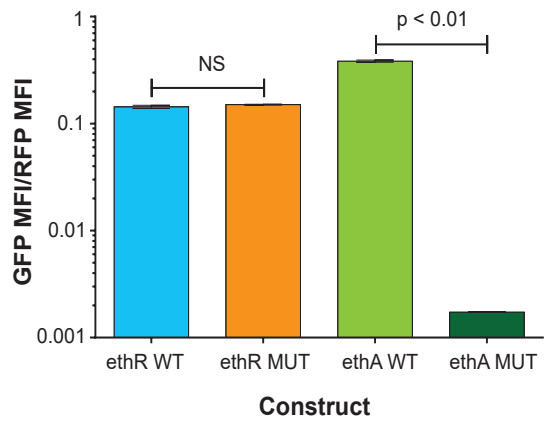
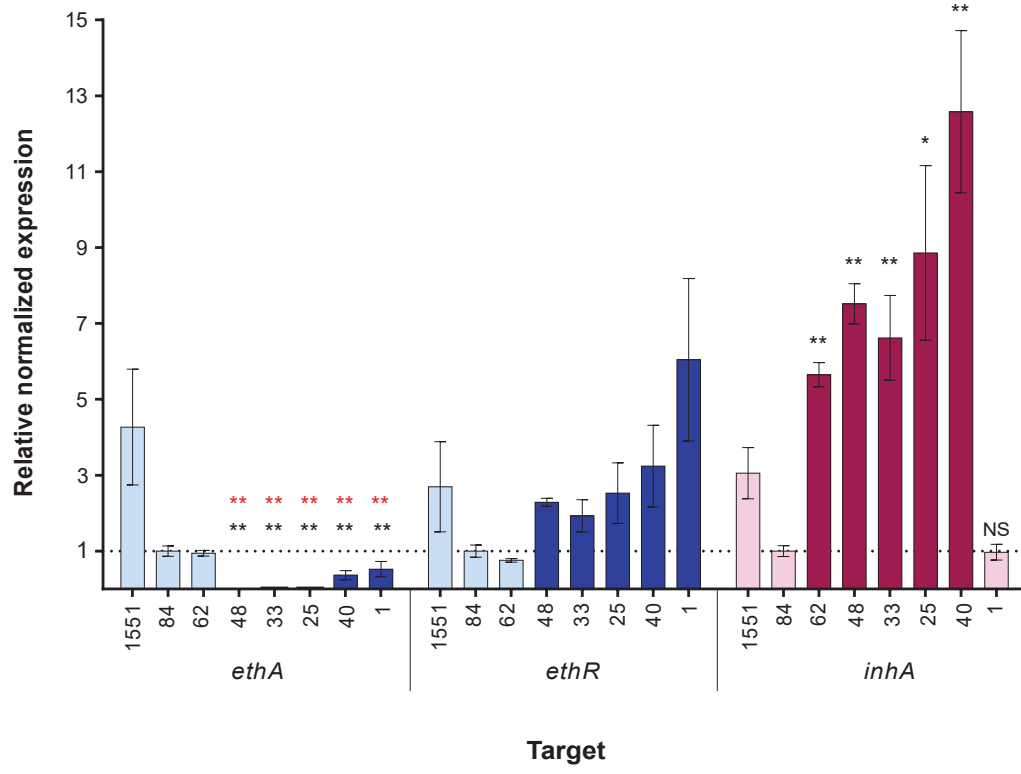


FIG 4





**TABLE 1** Strain details including resistance mutations and RNA sequencing coverage

Strain	Spoligo- type	Resistance mutations for each corresponding drug							RNAseq Coverage (x)
		INH	RIF	STR	EMB	KAN	ETH	OFL	
TKK-01-0084	LAM4								288.63
TKK-01-0025	LAM4	<i>inhA</i> t-8a <i>katG</i> :S315 T	<i>rpoB</i> :L452 P D435G	<i>gidB</i> :L16 R <i>gidB</i> :del	<i>embB</i> :M306 V	<i>rrs</i> :A1401 G	<i>inhA</i> t-8a	<i>gyrA</i> :A90 V	214.34
TKK-01-0033	LAM4	<i>inhA</i> t-8a <i>katG</i> :S315 T	<i>rpoB</i> :L452 P D435G	<i>gidB</i> :L16 R <i>gidB</i> :del	<i>embB</i> :M306 V	<i>rrs</i> :A1401 G	<i>inhA</i> t-8a	<i>gyrA</i> :A90 V	239.13
TKK-01-0040	LAM4	<i>inhA</i> t-8a <i>katG</i> :S315 T	<i>rpoB</i> :L452 P D435G	<i>gidB</i> :L16 R <i>gidB</i> :del	<i>embB</i> :M306 V	<i>rrs</i> :A1401 G	<i>inhA</i> t-8a	<i>gyrA</i> :A90 V	269.24

INH = Isoniazid, RIF = Rifampicin, STR = Streptomycin, EMB = Ethambutol, KAN = Kanamycin, ETH = Ethionamide, OFL = Ofloxacin

**TABLE 2** Minimal inhibitory concentrations (MICs) of clinical strains to ethionamide (ETH)

Strain	DST	Spoligo	Putative ETH resistance conferring mutations								MIC (mg/L)	
			<i>ethA</i> promoter	<i>inhA</i> promoter	<i>inhA</i> intragenic	<i>ethA</i> intragenic	<i>ethR</i> intragenic	<i>mymA</i>	<i>mshA</i>	<i>ndh</i>		
TKK-01-0001	MDR	KZN	t-11c									20
TKK-01-0035	MDR	KZN	t-11c									10
TKK-01-0075	MDR	KZN	t-11c									5
TKK-01-0025	XDR	KZN	t-11c	t-8a								80
TKK-01-0040	XDR	KZN	t-11c	t-8a								> 80
TKK-01-0048	MDR	KZN	t-11c	t-8a								20
TKK-01-0033	XDR	KZN	t-11c	t-8a								80
TKK-02-0001	XDR	BEIJING		c-15t	I194V			P94P		A189T		40
TKK-02-0046	POLY (P/N + RIF) POLY	BEIJING		c-15t	I194V			P94P		A189T		20
TKK-01-0005	(STR+ETH)	BEIJING		c-15t						A189T		80
TKK-01-0062	PXDR	BEIJING		g-17t		A381P						20
TKK-01-0032	MDR	S		t-8g							V18A	5
TKK-01-0013	XDR	BEIJING				Y276H				A189T		5
TKK-02-0018	MDR	T3				V202L			del			15
TKK-02-0019	PXDR					V202L			del			20
TKK-01-0026	MDR	T3							del			2,5
TKK-02-0069	PXDR								del			5
TKK-01-0081	DS	KZN										2,5
TKK-01-0084	DS	KZN										2,5
TKK-01-0047	DS	BEIJING										2,5
TKK-01-0027	DS	BEIJING										2,5
H37Rv	DS	n/a										2,5
CDC 1551	DS	n/a										1,25