



Published in final edited form as:

Nat Genet. 2016 May ; 48(5): 544–551. doi:10.1038/ng.3548.

## Genomic and functional analyses of *Mycobacterium tuberculosis* strains implicate *ald* in D-cycloserine resistance

Christopher A. Desjardins<sup>1,11</sup>, Keira A. Cohen<sup>1,2,3,11</sup>, Vanisha Munsamy<sup>3</sup>, Thomas Abeel<sup>1,4</sup>, Kashmeel Maharaj<sup>3</sup>, Bruce J. Walker<sup>1</sup>, Terrance P. Shea<sup>1</sup>, Deepak V. Almeida<sup>3,5</sup>, Abigail L. Manson<sup>1</sup>, Alex Salazar<sup>1,4</sup>, Nesri Padayatchi<sup>6</sup>, Max R. O'Donnell<sup>6,7,8</sup>, Koleka P. Mlisana<sup>9,10</sup>, Jennifer Wortman<sup>1</sup>, Bruce W. Birren<sup>1</sup>, Jacques Grosset<sup>3,5</sup>, Ashlee M. Earl<sup>1,12</sup>, and Alexander S. Pym<sup>3,12</sup>

<sup>1</sup>Broad Institute of MIT and Harvard, Cambridge, MA, USA <sup>2</sup>Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA <sup>3</sup>KwaZulu-Natal Research Institute for TB and HIV (K-RITH), Durban, South Africa <sup>4</sup>Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands <sup>5</sup>Center for Tuberculosis Research, Johns Hopkins School of Medicine, Baltimore, MD, USA <sup>6</sup>Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa <sup>7</sup>Division of Pulmonary, Allergy, and Critical Care Medicine, Columbia University College of Physicians and Surgeons, New York, USA <sup>8</sup>Department of Epidemiology, Columbia Mailman School of Public Health, New York, USA <sup>9</sup>School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban, South Africa <sup>10</sup>National Health Laboratory Service, Durban, South Africa

### Abstract

A more complete understanding of the genetic basis of drug resistance in *Mycobacterium tuberculosis* is critical for prompt diagnosis and optimal treatment, particularly for toxic second-line drugs like D-cycloserine. Here, we used whole-genome sequences from 498 strains of *M. tuberculosis* to identify novel resistance-conferring genotypes. By combining association and

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Correspondence should be addressed to: A.M.E. (; Email: [aearl@broadinstitute.org](mailto:aearl@broadinstitute.org)) or A.S.P. (; Email: [alex.pym@k-rith.org](mailto:alex.pym@k-rith.org))

<sup>11</sup>These authors contributed equally.

<sup>12</sup>These authors jointly supervised this work.

**URLs.** Hain Genotype MTBDR<sub>plus</sub>, <http://www.hain-lifescience.de/en/products/microbiology/mycobacteria/genotype-mtbdplus.html>; Hain Genotype MTBDR<sub>sl</sub>, <http://www.hain-lifescience.de/en/products/microbiology/mycobacteria/genotype-mtbdsl.html>; South African Policy Guidelines for the Treatment of Drug-Resistant Tuberculosis, <http://www.health-e.org.za/wp-content/uploads/2014/06/MDR-TB-Clinical-Guidelines-Updated-Jan-2013.pdf>; Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis, [http://apps.who.int/iris/bitstream/10665/130918/1/9789241548809\\_eng.pdf?ua=1&ua=1](http://apps.who.int/iris/bitstream/10665/130918/1/9789241548809_eng.pdf?ua=1&ua=1); Picard, <http://broadinstitute.github.io/picard>; BayesTraits, <http://www.evolution.rdg.ac.uk/BayesTraits.html>

### Author Contributions

A.M.E., A.S.P., B.W.B., C.A.D. and K.A.C. conceived and designed the project. N.P., M.R.O., K.P.M. and A.S.P. provided the clinical isolates. K.A.C., V.M., K.M. and D.V.A. performed the wet lab experiments. C.A.D., K.A.C., T.A., T.P.S., A.L.M. and A.S. analyzed the data. B.J.W. contributed analytic tools. C.A.D., K.A.C., A.M.E. and A.S.P. wrote the paper. J.W., B.W.B., A.M.E. and A.S.P. supervised and coordinated the project.

### Competing financial interests

The authors declare no competing financial interests.

correlated evolution tests with strategies for amplifying signal from rare variants, we found that loss-of-function mutations in *ald* (Rv2780), encoding L-alanine dehydrogenase, were associated with unexplained drug resistance. Convergent evolution of this loss-of-function was observed exclusively among multidrug-resistant strains. Drug susceptibility testing established that *ald* loss-of-function conferred resistance to D-cycloserine, and susceptibility to the drug was partially restored by complementation of *ald*. Clinical strains with mutations in *ald* and *alr* exhibited increased resistance to D-cycloserine when cultured *in vitro*. Incorporation of D-cycloserine resistance in novel molecular diagnostics could allow for targeted utilization of this toxic drug among patients with susceptible infections.

## Introduction

Molecular diagnostic testing has transformed pathogen identification and subsequent drug susceptibility testing (DST) in the field of clinical microbiology<sup>1</sup>. In particular, the recent adaptation of rapid genotypic diagnostics for *Mycobacterium tuberculosis*<sup>2,3</sup> has revolutionized the detection of *M. tuberculosis* and determination of drug resistance by assaying for the presence of specific resistance-conferring mutations. To minimize morbidity and mortality of drug-resistant *M. tuberculosis*, it is critical to treat patients with antitubercular drugs to which their infecting strains are susceptible<sup>4,5</sup>, and also to prescribe the minimally toxic regimen that is capable of curing the disease<sup>6</sup>. However, current genotypic diagnostics for *M. tuberculosis* are limited in that they identify only a subset of previously known resistance mutations, and the catalog of mutations that confer resistance is incomplete. For example, the Hain Genotype MTBDR<sub>s</sub>/test<sup>3</sup> (see URLs) can provide insight into second-line drug resistance in *M. tuberculosis* strains, but only interrogates a limited number of resistance mutations and fails to diagnose one in three to four cases of extensively drug-resistant tuberculosis (XDR TB)<sup>7</sup>. As many second-line drugs included in MDR regimens are highly toxic<sup>8–12</sup> incomplete data provided by current molecular tests may result in needless exposure to toxic agents, for example in the case of rifampicin mono-resistance detected up by GeneXpert<sup>2</sup> that is treated as MDR.

D-cycloserine is one such highly toxic second-line drug that is not evaluated by current molecular diagnostics. The World Health Organization (WHO) has classified D-cycloserine as a category 4 drug for the treatment of TB, indicating that it is an oral bacteriostatic second-line agent. D-cycloserine remains a cornerstone of MDR and XDR-TB treatment despite severe psychiatric and central nervous system toxicities<sup>9,10,12</sup>. With these associated toxicities, it would be ideal to further restrict its use to patients harboring strains that are susceptible to D-cycloserine. Phenotypic DST for D-cycloserine has only been validated on Löwenstein-Jensen media; however, routine phenotypic DST is performed infrequently and is not recommended by the WHO due to technical difficulties with the assay<sup>13–15</sup>.

Given these difficulties with phenotypic resistance characterization and toxicities, it would be ideal to detect D-cycloserine resistance *via* rapid molecular diagnostics. However, the molecular basis of D-cycloserine resistance among clinical isolates of *M. tuberculosis* is poorly understood. D-cycloserine functions by competitively inhibiting two enzymes in the alanine metabolism pathway—alanine racemase (encoded by *alr*, Rv3423c)<sup>16</sup>, and D-alanine

D-alanine ligase (encoded by *ddlA*, Rv2981c)<sup>17,18</sup>—which are essential for peptidoglycan synthesis. Overexpression of *alr* and *ddlA* have been previously shown to confer resistance to D-cycloserine in *Mycobacterium smegmatis*<sup>19,20</sup>, and acquisition of a nonsynonymous mutation in *alr* in *M. tuberculosis* was shown to be correlated with the acquisition of phenotypic resistance to D-cycloserine<sup>21,22</sup>. In addition, *Mycobacterium bovis* BCG, a live attenuated vaccine strain derived from *M. bovis*, is intrinsically resistant to cycloserine<sup>23</sup>, a phenotype that is only partially explained by a mutation in *cycA* (an alanine transporter)<sup>24</sup>. However to our knowledge no mutations in *M. tuberculosis* have been experimentally validated as conferring resistance to D-cycloserine and there is uncertainty over the critical concentration that defines resistance.

Recently, whole-genome sequencing has provided more complete information about the genetic content of drug-resistant *M. tuberculosis*<sup>21,25–28</sup>, including investigations that have used signatures of convergent evolution and positive selection to search for novel genes involved in drug resistance<sup>25,26</sup>. Here, we combined two large studies of *M. tuberculosis*<sup>26,28</sup> with comparable whole-genome sequencing and phenotypic drug resistance data to identify mutations that associate with drug resistance. These two studies, from South Africa and China, included a total of 498 strains with varying degrees of drug resistance, including large numbers of MDR and XDR strains. By utilizing strategies for amplification of signal from rare variants, and by combining an association test with a correlated evolution test, we identified the gene L-alanine dehydrogenase, *ald* (Rv2780), as a candidate novel drug resistance-conferring gene. We then demonstrated that loss-of-function of *ald* confers resistance to D-cycloserine *in vitro*, and that clinical strains with *ald* loss-of-function mutations were resistant to this drug.

## Results

### Identification of candidate resistance-conferring variants

We first compiled a dataset of 498 strains of *M. tuberculosis* from South Africa and China, including whole-genome sequencing and phenotypic drug resistance data<sup>26,28</sup>. We next sought to assess how much unexplained drug resistance was present among these strains by comparing genomic variants detected in each strain to a list of resistance-associated variants for eight drugs drawn from current diagnostics and literature (Supplementary Note, and Supplementary Tables 1–2). The amount of unexplained drug resistance varied by drug, with anywhere between 8 and 21 strains without known resistance conferring mutations per drug (Table 1). Given the limited numbers of strains that were phenotypically resistant but lacked known genotypic resistance mechanisms to individual drugs, we reasoned that novel mechanisms of resistance might be rare and difficult to identify statistically based on individual variants. We therefore collapsed the genotypic data using two approaches: first by frequency in the population, and second based on functional impact (Methods). Specifically, in the second approach, we collapsed different loss-of-function mutations (defined as a nonsense mutation, frameshift, or large in-frame insertion or deletion (indel)) within the same gene, as loss-of-function of select genes are known to confer resistance to select drugs in *M. tuberculosis*<sup>29–33</sup>. Synonymous single-nucleotide polymorphisms (SNPs) were

excluded in both approaches. Using these criteria, a combined set of 11,704 genotypic features was generated for downstream analysis.

To identify novel resistance conferring mutations in this dataset, we hypothesized that resistance-conferring genotypes should follow two patterns: 1) there should be a positive correlation between presence of a resistance-conferring mutation and presence of phenotypic resistance, and 2) a resistance-conferring mutation and phenotypic resistance should evolve convergently and in concert, *i.e.*, they should evolve repeatedly and at the same times in evolutionary history. We therefore conducted two tests: a Fisher's exact test to measure the direct association between genotype and phenotype, and a test for correlated evolution<sup>34</sup> of genotype and phenotype, and then combined the p-values from both tests into an F-score (Methods).

Analysis of this combined set of genotypic features with the Fisher's exact and correlated evolution tests for each drug revealed that, in almost all cases, mutations already known to confer drug resistance were the most significant genotypic features resulting in the highest combined F-scores (Figure 1, Supplementary Table 3). However, for most drugs, genotypic features known to confer resistance to drugs other than the one being analyzed in the association tests were also often highly significant across both tests yielding high F-scores. For example, mutations in *katG*, *pncA*, *rpoB*, *embB*, and *rpsL* were significantly associated with isoniazid resistance despite the fact that *katG* is the only gene on this list which confers resistance to isoniazid. This is due to the fact that drugs to treat *M. tuberculosis* are prescribed only in combination and, therefore, resistances frequently co-occur.

Remaining high scoring features included genes encoding PE-PGRS proteins, polyketide synthases, secretion-associated proteins, and hypothetical proteins (Supplementary Table 3). However, when further examined, none of these other high scoring features could account for a large fraction of phenotypic resistance in strains without known genotypic resistance mechanisms. Rather, these features most often appeared in strains where phenotypic resistance could already be explained by known resistance-conferring mutations.

### Loss-of-function correlated with unexplained resistance

Given that the list of highest scoring features contained a significant number of genes already associated with resistance, and that the remaining high scoring features co-occurred with known resistance-conferring mutations, we hypothesized that the strong statistical signal contributed by mutations in these genes could make the identification of novel, low frequency resistance mechanisms difficult. To amplify signal from strains with potential for novel resistance genotypes, we repeated the association analysis after removing strains with known resistance-conferring mutations (Supplementary Figure 1, Supplementary Table 4).

From this analysis, we identified 21 genes with loss-of-function mechanisms that associated with phenotypic resistance to at least one of nine drugs (excluding PE/PPE genes and transposable elements; Supplementary Tables 4, 5). Eight of these drug resistance-associated genes were represented in a library of *M. tuberculosis* single gene knockout mutants with gene interruption conferred by a kanamycin resistance cassette<sup>35</sup>. We screened all available knockout mutants representing these eight genes against a panel of five antitubercular drugs

(plus the positive control kanamycin). We observed no resistance above the critical concentration to any of the tested drugs other than kanamycin, indicating that loss-of-function of each of these genes alone is insufficient to confer resistance to this particular panel of drugs (Supplementary Table 6).

One gene, *ald* (Rv2780), was notable in that it had an especially high number of unique loss-of-function mutations that appeared to evolve in a pattern consistent with convergent evolution. We identified 11 different *ald* loss-of-function mutations in 25 South African strains across four different spoligotypes, including one nonsense mutation and ten distinct indels (Figure 2A, Supplementary Table 7). When examined within a phylogenetic context, *ald* loss-of-function mutations appeared to evolve exclusively and repeatedly in MDR and XDR genetic backgrounds, suggesting a selection for loss-of-function of this gene among patients receiving second-line drug regimens (Figure 2B). Four other genes had loss-of-function mutations occur exclusively within an MDR and XDR genetic background (Rv1250, Rv2042c, Rv3197A: *whiB7*, and Rv3823c: *mmp18*), although in all cases at a much lower frequency than *ald* (Supplementary Table 5). The role of loss-of-function mutations in these genes is unclear, as Rv1250 expression is induced by drug exposure<sup>36</sup>, *mmp18* expression has been positively correlated with virulence<sup>37</sup>, and overexpression of *whiB7* has been shown to confer cross-resistance to aminoglycosides<sup>38,39</sup> while null mutants have been shown to be hypersusceptible to multiple antibiotics including macrolides<sup>40,41</sup> (Supplementary Note).

### ***ald* loss-of-function conferred resistance to D-cycloserine**

The *ald* gene encodes L-alanine dehydrogenase, a component of the alanine metabolism pathway, which includes the targets of the second-line drug D-cycloserine (Figure 3)<sup>42</sup>. D-cycloserine resistance was not prospectively phenotypically assayed in the 498 strains analyzed here; however, D-cycloserine and terizidone, a dimeric form of D-cycloserine, are routinely prescribed as part of the South African drug-resistant *M. tuberculosis* treatment guidelines (see URLs). Therefore, MDR and XDR *M. tuberculosis* in this region could have developed D-cycloserine resistance due to selective drug pressure.

To investigate whether *ald* loss-of-function conferred resistance to D-cycloserine, we determined the minimum inhibitory concentration (MIC) on Löwenstein-Jensen (LJ) media (Methods) for four strains: 1) CDC1551, a fully drug-susceptible reference strain of *M. tuberculosis* (with wild type alleles of *ald*, *alr*, *ddlA*, and *cycA*), 2) an *ald* knockout strain (*ald*), 3) the *ald* knockout complemented with wild type *ald* (*ald*-comp) from *M. tuberculosis*, and 4) *M. bovis* BCG Danish 1331 (BCG), a strain known to be naturally resistant to D-cycloserine<sup>23,43</sup>. Testing was repeated in multiple independent experiments to establish reproducibility of the assay; results are shown in Table 2. Wild type *M. tuberculosis* was the most susceptible to D-cycloserine (MIC = 15 µg/mL) while the positive control BCG was the most resistant (MIC = 40–60 µg/mL). The *ald* knockout had increased resistance relative to the wild type strain (MIC = 25–30 µg/mL), which overlaps the WHO-determined critical concentration for resistance on LJ of 30 µg/mL (see URLs). Complementation of *ald* partially restored susceptibility to D-cycloserine (MIC = 20 µg/

mL). These results demonstrate that *in vitro*, *ald* loss-of-function confers a moderate level of resistance to D-cycloserine.

These results were also confirmed in a growth advantage assay, in which all strains were cultured in a mycobacterial growth indicator tube (MGIT) system with varying concentrations of D-cycloserine (0, 7.5 and 15 µg/mL) and the resulting time to positivity in the presence of the drug was normalized to the no-drug control (Figure 4, Methods, Supplementary Figure 2). The *ald* knockout had a significant growth advantage relative to the wild type strain at 7.5 µg/mL D-cycloserine, ( $p < 0.0001$ , ANOVA), and no significant growth difference when compared to the resistant control strain, BCG, at the same drug concentration (Figure 4A). At a higher concentration of D-cycloserine (15 µg/mL), the *ald* knockout still retained a growth advantage over the wild type strain (Figure 4B;  $p < 0.0001$ , ANOVA), but grew more slowly relative to BCG ( $p < 0.0001$ , ANOVA), supporting the MIC data, which showed that *ald* loss-of-function confers less resistance than the mechanism present in BCG. At both drug concentrations, the *ald* complement partially reverted the growth advantage conferred by the *ald* knockout (Figure 4A, B;  $p < 0.0001$ , ANOVA).

Within the *M. tuberculosis* complex (MTBC), BCG is uniquely considered resistant to D-cycloserine<sup>44</sup>. This resistance was previously shown to be, at least partly, due to a unique SNP in *cycA*, which encodes an alanine transporter<sup>24,45</sup>. However, all of the RD9 lineages of MTBC from *M. africanum* to *M. bovis*, including BCG, have a frameshift in *ald* causing loss-of-function (Supplementary Figure 3)<sup>46,47</sup>. We therefore sought to evaluate the effect of the frameshift in both the BCG and more general RD9 genetic background by determining the MIC of two additional strains: 1) *M. bovis* ATCC 19210, a reference strain of *M. bovis*, and 2) *M. bovis* BCG complemented with *M. tuberculosis ald* (BCG-*ald*-comp). *M. bovis* ATCC 19210 had a D-cycloserine MIC of 25 µg/mL (Table 2), which is similar to the level of resistance observed in *ald* and greater than the level of resistance observed in the *M. tuberculosis* reference strain, CDC1551. The *ald*-complemented BCG strain had an MIC of 40 µg/mL, which was not different from the wild type BCG strain (Table 2), although in a growth advantage assay at low concentrations of D-cycloserine, BCG had a significant growth advantage when compared to the *ald*-complemented BCG strain, ( $p < 0.02$  and  $p < 0.0001$ ; Supplementary Note, Supplementary Figures 4 and 5). Taken together, these results suggest that *ald* loss-of-function may have both a minor role in the resistance of BCG to D-cycloserine, and provide *M. bovis* with an elevated resistance to D-cycloserine relative to *M. tuberculosis*.

### ***ald* and *alr* in D-cycloserine resistant clinical isolates**

To more specifically ascertain the distribution and diversity of mutations in *ald*, along with mutations in *alr*, *ddlA*, and *cycA*, we characterized all promoter, nonsynonymous, and loss-of-function mutations in these four genes across all strains in our dataset (Table 3, Supplementary Table 8). Mutations were most frequent in *ald* (25 strains), followed by *alr* (17 strains). Only seven strains were identified to have mutations in either *ddlA* or *cycA*; none of these mutations corresponded to previously characterized resistance-conferring mutations. We observed that many more strains had *ald* loss-of-function mutations than had *ald* nonsynonymous mutations (24 versus 5, respectively). By contrast, mutations in *alr* were



exclusively nonsynonymous, with the exception of one strain with a SNP in the promoter region, supporting the hypothesis that *alr* is essential<sup>48</sup>.

We next sought to verify our prediction that clinical strains with *ald* loss-of-function mutations would exhibit resistance to D-cycloserine, and to determine how that level of resistance compared to strains with mutations in *alr*. We selected 44 clinical isolates from South Africa for testing, including 22 with wild type *ald* and *alr*, 13 with mutations in *ald* and wild type *alr*, 7 with wild type *ald* and mutations in *alr* or its promoter, and 2 with mutations in both *ald* and *alr* (Supplementary Table 9). These strains had diverse drug susceptibility phenotypes and varying genetic backgrounds (Supplementary Table 9).

Clinical isolates with *ald* mutations showed significantly increased resistance to D-cycloserine relative to strains with wild type alleles of both *ald* and *alr* (Figure 5A, Supplementary Table 9;  $p < 0.0002$ , Mann-Whitney U test). Strains with *ald* loss-of-function mutations had a median MIC of 40  $\mu\text{g/mL}$ , with 8 of those 11 strains exhibiting an MIC from 30–40  $\mu\text{g/mL}$ . The two strains with nonsynonymous SNPs in *ald* had MICs of 30  $\mu\text{g/mL}$ ; both of these nonsynonymous SNPs were predicted to have a deleterious effect on protein function by PROVEAN<sup>49</sup> (Supplementary Table 9). By contrast, strains with wild type alleles in both genes had a median MIC of 20  $\mu\text{g/mL}$ , with 91% of strains exhibiting MICs  $\leq 20$   $\mu\text{g/mL}$ . These data suggest that both deleterious nonsynonymous SNPs and loss-of-function mutations in *ald* can cause resistance to D-cycloserine in clinical isolates of *M. tuberculosis* at or above the current critical concentration (30  $\mu\text{g/mL}$ ; URLs).

Clinical isolates with *alr* mutations also showed increased resistance to D-cycloserine relative to strains with wild type alleles of both *ald* and *alr* (Figure 5A, Supplementary Table 9). Of the strains with nonsynonymous mutations in *alr*, 4 of 6 exhibited high level resistance to D-cycloserine, with MICs of  $\geq 60$   $\mu\text{g/mL}$ , while one, plus the two strains with both *alr* and *ald* mutations, exhibited MICs of 40  $\mu\text{g/mL}$ . As with *ald*, all of the nonsynonymous SNPs seen in *alr* were predicted to have a deleterious effect on protein function by PROVEAN<sup>49</sup> (Supplementary Table 9). Additionally, the one strain with an *alr* promoter mutation showed high level resistance (MIC  $> 60$   $\mu\text{g/mL}$ ), consistent with previous results in *M. smegmatis* that showed upregulation of *alr* through a promoter mutation conferred resistance to D-cycloserine<sup>19</sup>. These results suggest that both nonsynonymous SNPs in *alr* and SNPs in the *alr* promoter can confer high-level resistance to D-cycloserine in clinical isolates of *M. tuberculosis*, and that *alr* and *ald* mutations are non-additive for resistance.

## Discussion

To identify novel drug resistance-conferring mechanisms in *M. tuberculosis*, we analyzed whole-genome sequencing data from two large collections of clinical isolates from South Africa and China with diverse phenotypic drug resistance profiles. Combining a direct association test with a correlated evolution test followed by functional screening of the top scoring genes with loss-of-function mutations, we identified a novel drug resistance mechanism to D-cycloserine, *ald* loss-of-function, which we then validated experimentally. We further showed that clinical strains with mutations in *ald* and *alr*, identified through our

computational analyses, were resistant to D-cycloserine. Collectively, this work expands our understanding of D-cycloserine resistance in *M. tuberculosis* and provides new opportunities to detect and manage D-cycloserine resistant TB.

The strong linkage disequilibrium and stratification present in bacterial populations has long been highlighted as a major hurdle for bacterial genome-wide association studies<sup>50</sup>. However, the problem of interdependence is not restricted to genotypes, but extends to drug resistance phenotypes as well<sup>26</sup>. For most TB drugs, mutations known to confer resistance to a single drug were significantly associated with phenotypic resistance to numerous unrelated drugs. This is because TB is treated exclusively with multi-drug regimens, which makes it difficult to collect clinical isolates of *M. tuberculosis* with independent drug resistance phenotypes. However, combining a basic association test with a correlated evolution test appeared to mitigate this phenotypic interdependence, as it could often correctly match known resistance mechanisms with the correct drug; the method did have difficulty when the resistance-conferring genotype evolved infrequently (e.g., pyrazinamide; Supplementary Note) and had other limitations (Supplementary Note). Ultimately, to robustly overcome phenotypic interdependence in association studies of drug resistant *M. tuberculosis*, further methods development is needed to build upon this and other work<sup>25,26</sup>.

After we identified loss-of-function mutations in *ald* as a candidate novel drug resistance mechanism, we showed experimentally that *ald* loss-of-function confers resistance to D-cycloserine in *M. tuberculosis*. Complementation of *ald* back into the *ald* knockout strain partially restored susceptibility to D-cycloserine (partial restoration by complementation has been observed before in *M. tuberculosis*, and could be due to altered expression levels through usage of a non-native promoter<sup>51</sup>). We hypothesize that strains lacking functional Ald are unable to convert L-alanine to pyruvate, thereby increasing the pool of available L-alanine (Figure 3). As L-alanine is the precursor to the pathway steps competitively inhibited by D-cycloserine, abundant L-alanine in the setting of non-functional Ald may allow for continued alanine racemization and D-alanine ligase activity, and therefore continued peptidoglycan production despite competitive inhibition by D-cycloserine. However, the observation in clinical strains that *ald* loss-of-function conferred a lower level of resistance than nonsynonymous or promoter mutations in *alr* (Supplementary Table 9), the direct target of D-cycloserine, suggests that the resistance conferred by *ald* could more easily be overcome through increased dosage of D-cycloserine.

*M. bovis* and other members of the MTBC that lack RD9<sup>52</sup> share a frameshift in *ald*. We sought to determine the effect of this frameshift in *M. bovis* strains traditionally classified as D-cycloserine resistant (BCG) and susceptible (all other strains)<sup>44</sup>. We observed that the *M. bovis* reference strain, ATCC 19210, had elevated levels of D-cycloserine resistance as compared to clinical isolates of *M. tuberculosis* lacking resistance-conferring mutations. Our results suggest that *M. bovis*, and potentially the entire RD9 lineage, which is generally considered D-cycloserine susceptible, may actually have elevated resistance levels (Figure 5B, URLs). However, much of the RD9 lineage (including *M. bovis*, but excluding BCG) also shares an inactivating mutation in the gene encoding pyruvate kinase (*pykA*), which provides pyruvate to Ald<sup>53,54</sup> and could, therefore, potentially confound analysis of D-cycloserine resistance among these strains. Evolutionary analyses suggest that in animal-



infecting species of the MTBC there is more ancient functional diversity of these genes than in *M. tuberculosis sensu stricto* (Supplementary Note, Supplementary Figure 3). Future experiments to disentangle the roles of *ald*, *pykA* and other related genes in drug resistance across the other species of the MTBC will be needed.

The identification of so many unique *ald* loss-of-function mutations among drug-resistant clinical strains of *M. tuberculosis* suggests that, despite conferring only a moderate level of resistance, this mechanism provides sufficient resistance to be selected for under current drug exposure pressures in South Africa (Supplementary Note). This further suggests that *ald* loss-of-function mutations confer resistance to the D-cycloserine analog, terizidone, which is routinely prescribed for treatment of drug-resistant *M. tuberculosis* in South Africa (see URLs), although this hypothesis still needs to be tested. While *ald* loss-of-function may prove advantageous during exposure to D-cycloserine or terizidone, harboring such a mutation may prove to be deleterious in other settings. Strains with non-functional Ald may be unable to increase pyruvate availability under hypoxic conditions, and further laboratory investigation should determine whether *ald* loss-of-function yields decreased survival in hypoxia. However, compared to the *M. bovis* genetic background where *pykA* is also non-functional, clinical *M. tuberculosis* strains with loss-of-function of *ald* may have greater ability to tolerate hypoxic environments (such as in a granuloma) where reliance on pyruvate (rather than aerobic metabolism) is greater.

Given the severe psychiatric and central nervous system toxicities of D-cycloserine<sup>9,10,12</sup>, it is critical to both incorporate D-cycloserine resistance mechanisms into molecular diagnostics and optimize phenotypic resistance assays to prevent the drug from being inappropriately prescribed to patients who harbor resistant strains. In order to maximize the effectiveness of phenotypic resistance assays, the range of wild type MICs for an organism must be determined to define drug resistance and to set appropriate critical concentrations<sup>55,56</sup>. Recently, the Stop TB Partnership recommended to the WHO that the critical concentration of D-cycloserine be lowered from 40 to 30 µg/mL (URLs). Assuming that the two isolates exhibiting increased resistance to D-cycloserine, but having wild type alleles in both *ald* and *alr*, harbor an unknown resistance mechanism, our results suggest a tentative epidemiological cutoff (ECOFF)<sup>57</sup> of 20 µg/mL (Figure 5B). Although phenotyping of a larger number of strains is required, our results support a lowering of the critical concentration below 30 µg/mL. Selection of a critical concentration would also need to consider whether it should be applied to *M. tuberculosis sensu stricto* or the entire MTBC including *M. bovis*.

D-cycloserine resistance is an ideal target for molecular diagnostics, given technical difficulties with reproducibility and reliability of the phenotypic resistance assay<sup>13,15</sup>. However, current SNP-based diagnostics are not amenable to capturing loss-of-function mutations, as the mutations can occur anywhere within the gene and thus are difficult to capture with a small number of targeted SNPs. Other whole-gene-based diagnostic strategies will instead need to be employed, some of which have already been developed for detection of drug-resistant TB, such as high-resolution melt (HRM) analysis<sup>58</sup> and rapid whole-genome sequencing of clinical isolates<sup>21</sup>. Incorporation of resistance mechanisms to D-cycloserine and other highly toxic drugs into novel molecular diagnostics could prevent

those drugs from being prescribed in a treatment setting where they may be ineffective and needlessly toxic.

## Online Methods

### Bioinformatic Analyses

We used two sets of strains in this study: one set of 337 strains from South Africa<sup>28</sup> (NCBI Bioproject ID numbers: PRJNA183624, PRJNA235615) and one set of 161 strains from China<sup>26</sup> (NCBI Sequence Read Archive (SRA) accession number: SRA065095). For both sets, reads were mapped onto a reference assembly of H37Rv (GenBank accession number: CP003248.2) using BWA version 0.5.9<sup>60</sup>. In cases where read coverage of the reference was greater than 200x, reads were down-sampled using Picard (see URLs) prior to mapping. Variants were identified using Pilon version 1.5 as described<sup>61</sup>. All sites with unambiguous SNPs were used to generate a phylogenetic tree using FastTree version 2.1.3<sup>62</sup>. To assess the amount of unexplained phenotypic resistance in our dataset, we identified genomic polymorphisms known to confer drug resistance in each strain using variants from current diagnostics supplemented with a literature review (Supplementary Note).

In order to identify novel variants conferring drug resistance, we hypothesized that the presence of causal genotypes in bacteria should be both positively correlated with phenotypes and not evolutionarily independent of them. We therefore conducted two tests, the first of which was a Fisher's exact test to measure the association between genotype and phenotype. The second test was a test for correlated evolution<sup>34</sup> implemented in the software package BayesTraits (see URLs). Briefly, the correlated evolution test calculates the likelihood of two traits evolving dependently and independently and then compares the two models with a likelihood ratio test. If the dependent model is significantly more likely than the independent model, the hypothesis of evolutionary independence is rejected. All test results were corrected for multiple comparisons using the Benjamini-Hochberg method<sup>59</sup> for plotting purposes. Once both the Fisher's exact test and the correlated evolution test were calculated, their uncorrected p-values were combined into an F-score by taking the harmonic mean of the negative log p-values. Uncorrected values were used in the F-score calculation to provide maximal discrimination between variants with poor p-values.

In order to compare phenotypic drug resistances to genotypic features, two matrices of these features were constructed from the variant calls of all strains. In the first matrix, rare variants (those at  $\leq 2\%$  frequency) were collapsed by gene or intergenic region, while common variants (those at  $> 2\%$  frequency) remained uncollapsed and were considered independently. In the second matrix, we assessed whether each gene had a loss-of-function mutation (defined as a nonsense mutation, frameshift, or large ( $>30$  bp) in-frame insertion or deletion (indel)), regardless of the frequency of the mutation. We then combined the two matrices into a single matrix for association analyses. All analyses were conducted twice: once with all samples, and once after removing strains with known resistance-conferring mutations, in order to amplify signal from strains with potential for novel resistance genotypes (Results).

### DST screen of transposon mutants

Single gene knockout transposon mutants were sourced from the Johns Hopkins TARGET mutant collection<sup>35</sup>. First and second-line DST was performed by the agar proportion method on 7H11 media, using the following WHO-recommended critical concentrations (in µg/mL): isoniazid 0.2, isoniazid 1.0, rifampicin 1.0, ethambutol 7.5, streptomycin 2.0, ofloxacin 2.0 and kanamycin 6.0.

### *ald* complementation into CDC1551 and BCG Danish 1331

An *ald* knockout mutant (*ald*) of *M. tuberculosis* (JHU2780-209) was obtained from the Johns Hopkins TARGET mutant collection<sup>35</sup>. This single gene knockout mutant was created in a background of wild type *M. tuberculosis* strain CDC1551 by Himar1 transposon mutagenesis with *ald* interrupted at nucleotide 209, which contains a kanamycin-resistance cassette. To complement *ald*, a 1.116kb fragment spanning the *ald* gene was PCR amplified from *M. tuberculosis* H37Rv genomic DNA using primers listed in Supplementary Table 10. The fragment was directionally cloned into pMCZ-MOP, an L5 integrating zeocin-marked mycobacterial vector to create pMCZ-MOP-*ald*, which allowed for the expression of *ald* off the MOP (mycobacterial optimized promoter)<sup>63</sup>. The pMCZ-MOP-*ald* was then transformed into *ald* to create the *ald* complemented strain, resulting in constitutive expression of *ald*. Transformants were selected using the appropriate antibiotic markers and confirmed by amplicon sequencing of a 1.319kb fragment of the integrated plasmid spanning the inserted *ald* gene. pMCZ-MOP-*ald* was also transformed into BCG Danish 1331 (Statens Serum Institut) to create the BCG-*ald*-complement.

### *M. bovis* amplicon sequencing

Amplicon sequencing of *ald* and *pncA*<sup>64</sup> in *M. bovis* ATCC 19210 confirmed the presence of the *ald* mutation (a266del), and a mutation in *pncA* (g169c) that is common to all *M. bovis* and BCG strains.

### D-Cycloserine MIC determination

Select mycobacterial strains were subjected to D-cycloserine MIC determination on Löwenstein–Jensen (LJ) medium. LJ slants were prepared freshly as had been previously described per WHO protocols, with supplementation of D-cycloserine prior to inspissation to achieve final drug concentrations of 0µg/mL, 10µg/mL, 15µg/mL, 20µg/mL, 25µg/mL, 30µg/mL, 40µg/mL, 60µg/mL, and 120µg/mL. 10<sup>3</sup> CFU of a mid to late-log liquid culture of *M. tuberculosis* was inoculated onto D-cycloserine slants at each drug concentration in duplicate. Slants were incubated at 37°C for four weeks, after which CFU were counted. The MIC to D-cycloserine for each strain was determined to be the lowest drug concentration that resulted in 90% reduction in CFU when compared to the no drug control. Differences in MIC distributions were calculated using the Mann-Whitney U test; in cases where MICs were replicates, the average interval across replicates was used for the calculation.

### D-cycloserine growth inhibition assay

All mycobacterial strains were grown under standard conditions, in Middlebrook 7H9 broth (Difco) supplemented with 0.5% glycerol, 10% oleic acid-albumin-dextrose-catalase

(OADC), and 0.05% Tween 80 with constant shaking at 37°C. On the day of the experimental set up, D-cycloserine powder (C6880, Sigma-Aldrich) that had been stored at -20°C was freshly dissolved in a 0.1M sodium phosphate buffer (pH 8.0) and sterilized with a 0.22µm filter. BD BACTEC™ MGIT™ tubes were prepared with BD BACTEC™ MGIT™ growth supplement, and 6.5mL tween-free 7H9 media supplemented with or without D-cycloserine to achieve a final drug concentration of 0 µg/mL, 7.5 µg/mL, 15 µg/mL and 30 µg/mL of D-cycloserine. 10<sup>3</sup> CFU of a mid to late-log liquid culture of *M. tuberculosis* that had been prepared from a single colony isolate of each strain was inoculated into three MGIT tubes for each drug concentration to achieve three technical replicates. Dilutions of the initial inoculum were plated for CFU in parallel to quantify the starting inoculum. The bar code of each MGIT tube was scanned into the BD BACTEC™ MGIT™ 960 Mycobacterial Detection System, and the MGIT tubes were placed in the instrument as directed. Samples remained in the MGIT instrument until they were flagged positive for growth, and the time to positivity (ttp) was recorded. Tubes that were not flagged positive by 42 days were considered to have no growth, and were assigned a ttp value of > 42 days. Data were reported for each strain and each drug concentration as a mean ttp for the three technical replicates ± the standard error of the mean (SEM). To determine relative growth inhibition of each drug concentration for each strain, the ttp of each D-cycloserine-containing MGIT was normalized to the mean ttp of no drug control for that strain. The mean ttp of the no drug control was considered to have a growth inhibition value of 1. Growth inhibition was calculated for each MGIT tube and reported as a mean growth inhibition for the three technical replicates ± the SEM. Adjusted p-values were calculated using a two-way ANOVA with Tukey's multiple comparisons test in Prism GraphPad version 6.0 (GraphPad, San Diego, CA). We also confirmed that the transposon mutagenesis did not confer a growth advantage in the presence of D-cycloserine (Supplementary Note, Supplementary Tables 11 and 12).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors would like to express their sincere gratitude to the patients in South Africa and China who provided clinical samples, without which this study would not have been possible. The authors also gratefully acknowledge the collaborative efforts of the following individuals whose significant prior contributions allowed for this study: William R. Bishai, Nonkubela Bantubani, Lucia Alvarado, Sinéad B. Chapman, Nomonde R. Mvelase, Eamon Y. Duffy, Michael G. Fitzgerald, Pamela Govender, Sharvari Gujja, Susanna Hamilton, Clinton Howarth, Jeffrey D. Larimer, Matthew D. Pearson, Margaret E. Priest and Qiangong Zeng. *M. bovis* ATCC 19210 was generously provided by Robert Warren, Stellenbosch University. The authors would also like to thank Christina Cuomo and three anonymous reviewers for providing helpful comments on the manuscript. Additionally, the authors would like to acknowledge the Broad Sequencing Platform for assistance with data acquisition.

This project has been funded in part with Federal funds from the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health, Department of Health and Human Services, under Contract No.:HHSN272200900018C, IeDEA (NIH 5U01AI069924-07) and Grant Number U19AI110818 to the Broad Institute. K.A.C. was supported by NHLBI T32HL007633. T.A. was a postdoctoral fellow of the Research Foundation-Flanders. Additional funding sources included U19 AI51794, and the U.S. Centers for Disease Control and Prevention. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

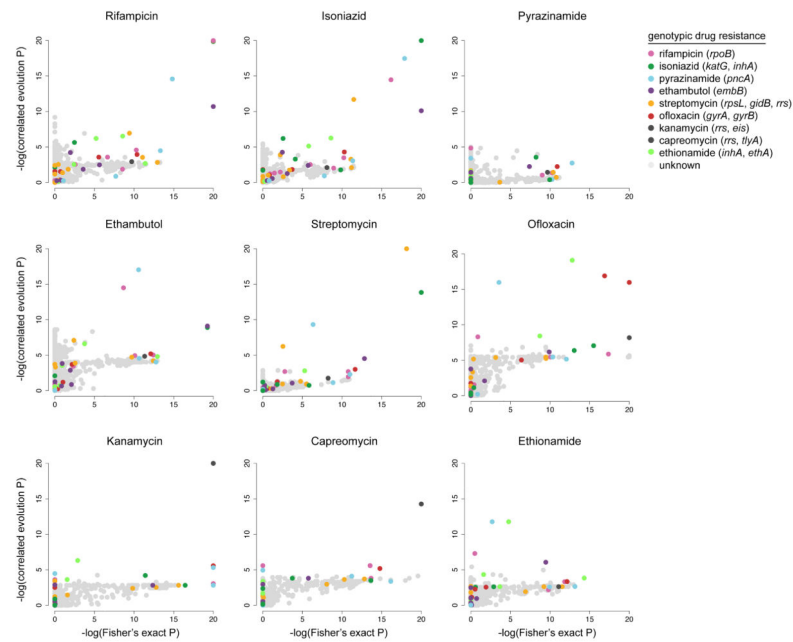
## References

1. Bauer KA, Perez KK, Forrest GN, Goff DA. Review of Rapid Diagnostic Tests Used by Antimicrobial Stewardship Programs. *Clin Infect Dis*. 2014; 59:S134–S145. [PubMed: 25261540]
2. Boehme CC, et al. Rapid Molecular Detection of Tuberculosis and Rifampin Resistance. *N Engl J Med*. 2010; 363:1005–1015. [PubMed: 20825313]
3. Kiet VS, et al. Evaluation of the MTBDRsl Test for Detection of Second-Line-Drug Resistance in *Mycobacterium tuberculosis*. *J Clin Microbiol*. 2010; 48:2934–2939. [PubMed: 20573868]
4. Bastos ML, et al. Treatment Outcomes of Patients With Multidrug-Resistant and Extensively Drug-Resistant Tuberculosis According to Drug Susceptibility Testing to First- and Second-line Drugs: An Individual Patient Data Meta-analysis. *Clin Infect Dis*. 2014; 59:1364–1374. [PubMed: 25097082]
5. Cegielski JP, et al. Extensive drug resistance acquired during treatment of multidrug-resistant tuberculosis. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 2014; 59:1049–1063.
6. Lange C, et al. Management of patients with multidrug-resistant/extensively drug-resistant tuberculosis in Europe: a TBNET consensus statement. *Eur Respir J*. 2014; 44:23–63. [PubMed: 24659544]
7. Theron G, et al. The diagnostic accuracy of the GenoType(®) MTBDRsl assay for the detection of resistance to second-line anti-tuberculosis drugs. *Cochrane Database Syst Rev*. 2014; 10:CD010705. [PubMed: 25353401]
8. Finegold SM. Kanamycin. *AMA Arch Intern Med*. 1959; 104:15–28. [PubMed: 13660522]
9. A PILOT study of cycloserine toxicity; a United States Public Health Service Cooperative Clinical Investigation. *Am Rev Tuberc*. 1956; 74:196–209. [PubMed: 13340147]
10. Bankier RG. Psychosis Associated with Cycloserine. *Can Med Assoc J*. 1965; 93:35–37. [PubMed: 14308907]
11. Drug-induced deafness. *JAMA*. 1973; 224:515–516. [PubMed: 4739610]
12. Yew WW, Wong CF, Wong PC, Lee J, Chau CH. Adverse neurological reactions in patients with multidrug-resistant pulmonary tuberculosis after coadministration of cycloserine and ofloxacin. *Clin Infect Dis Off Publ Infect Dis Soc Am*. 1993; 17:288–289.
13. Reller LB, Weinstein MP, Woods GL. Susceptibility Testing for Mycobacteria. *Clin Infect Dis*. 2000; 31:1209–1215. [PubMed: 11073754]
14. Kam KM, et al. Determination of critical concentrations of second-line anti-tuberculosis drugs with clinical and microbiological relevance. *Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis*. 2010; 14:282–288.
15. Companion Handbook to the WHO Guidelines for the Programmatic Management of Drug-Resistant Tuberculosis. (2014).
16. Lambert MP, Neuhaus FC. Mechanism of d-Cycloserine Action: Alanine Racemase from *Escherichia coli* W. *J Bacteriol*. 1972; 110:978–987. [PubMed: 4555420]
17. Strominger JL, Ito E, Threnn RH. Competitive Inhibition of Enzymatic Reactions By Oxamycin. *J Am Chem Soc*. 1960; 82:998–999.
18. Prosser GA, de Carvalho LPS. Metabolomics Reveal d-Alanine:d-Alanine Ligase As the Target of d-Cycloserine in *Mycobacterium tuberculosis*. *ACS Med Chem Lett*. 2013; 4:1233–1237. [PubMed: 24478820]
19. Cáceres NE, et al. Overexpression of the D-alanine racemase gene confers resistance to D-cycloserine in *Mycobacterium smegmatis*. *J Bacteriol*. 1997; 179:5046–5055. [PubMed: 9260945]
20. Feng Z, Barletta RG. Roles of *Mycobacterium smegmatis* d-Alanine:d-Alanine Ligase and d-Alanine Racemase in the Mechanisms of Action of and Resistance to the Peptidoglycan Inhibitor d-Cycloserine. *Antimicrob Agents Chemother*. 2003; 47:283–291. [PubMed: 12499203]
21. Köser CU, et al. Whole-Genome Sequencing for Rapid Susceptibility Testing of *M. tuberculosis*. *N Engl J Med*. 2013; 369:290–292. [PubMed: 23863072]
22. Merker M, et al. Whole Genome Sequencing Reveals Complex Evolution Patterns of Multidrug-Resistant *Mycobacterium tuberculosis* Beijing Strains in Patients. *PLoS ONE*. 2013; 8:e82551. [PubMed: 24324807]

23. Rist N, Canetti G, Boisvert H, Le Lirzin M. The BCG antibiogram. Diagnostic value of resistance to cycloserine. *Rev Tuberc Pneumol (Paris)*. 1967; 31:1060–1065. [PubMed: 4988437]
24. Chen JM, Uplekar S, Gordon SV, Cole ST. A Point Mutation in *cycA* Partially Contributes to the D-cycloserine Resistance Trait of *Mycobacterium bovis* BCG Vaccine Strains. *PLoS ONE*. 2012; 7:e43467. [PubMed: 22912881]
25. Farhat MR, et al. Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*. *Nat Genet*. 2013; 45:1183–1189. [PubMed: 23995135]
26. Zhang H, et al. Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance. *Nat Genet*. 2013; 45:1255–1260. [PubMed: 23995137]
27. Casali N, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nat Genet*. 2014; 46:279–286. [PubMed: 24464101]
28. Cohen KA, et al. Evolution of Extensively Drug-Resistant Tuberculosis over Four Decades: Whole Genome Sequencing and Dating Analysis of *Mycobacterium tuberculosis* Isolates from KwaZulu-Natal. *PLoS Med*. 2015; 12:e1001880. [PubMed: 26418737]
29. Wong SY, et al. Mutations in *gidB* Confer Low-Level Streptomycin Resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 2011; 55:2515–2522. [PubMed: 21444711]
30. Maus CE, Plikaytis BB, Shinnick TM. Mutation of *tlyA* Confers Capreomycin Resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 2005; 49:571–577. [PubMed: 15673735]
31. Morlock GP, Metchock B, Sikes D, Crawford JT, Cooksey RC. *ethA*, *inhA*, and *katG* Loci of Ethionamide-Resistant Clinical *Mycobacterium tuberculosis* Isolates. *Antimicrob Agents Chemother*. 2003; 47:3799–3805. [PubMed: 14638486]
32. Heym B, Alzari PM, Honore N, Cole ST. Missense mutations in the catalase-peroxidase gene, *katG*, are associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Mol Microbiol*. 1995; 15:235–245. [PubMed: 7746145]
33. Scorpio A, Zhang Y. Mutations in *pncA*, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nat Med*. 1996; 2:662–667. [PubMed: 8640557]
34. Pagel M. Detecting Correlated Evolution on Phylogenies: A General Method for the Comparative Analysis of Discrete Characters. *Proc R Soc Lond B Biol Sci*. 1994; 255:37–45.
35. Lamichhane G, et al. A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: Application to *Mycobacterium tuberculosis*. *Proc Natl Acad Sci*. 2003; 100:7213–7218. [PubMed: 12775759]
36. Li G, et al. Efflux Pump Gene Expression in Multidrug-Resistant *Mycobacterium tuberculosis* Clinical Isolates. *PLoS ONE*. 2015; 10:e0119013. [PubMed: 25695504]
37. Converse SE, et al. MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. *Proc Natl Acad Sci*. 2003; 100:6121–6126. [PubMed: 12724526]
38. Reeves AZ, et al. Aminoglycoside Cross-Resistance in *Mycobacterium tuberculosis* Due to Mutations in the 5' Untranslated Region of *whiB7*. *Antimicrob Agents Chemother*. 2013; 57:1857–1865. [PubMed: 23380727]
39. Köser CU, Bryant JM, Parkhill J, Peacock SJ. Consequences of *whiB7* (Rv3197A) Mutations in Beijing Genotype Isolates of the *Mycobacterium tuberculosis* Complex. *Antimicrob Agents Chemother*. 2013; 57:3461–3461. [PubMed: 23761426]
40. Morris RP, et al. Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A*. 2005; 102:12200–12205. [PubMed: 16103351]
41. Warit S, et al. Genetic characterisation of a *whiB7* mutant of a *Mycobacterium tuberculosis* clinical strain. *J Glob Antimicrob Resist*. 2015; 3:262–266.
42. Wiame JM, Piérard A. Occurrence of an L(+)-Alanine-Dehydrogenase in *Bacillus subtilis*. *Nature*. 1955; 176:1073–1075. [PubMed: 13272750]
43. Durek C, Rüsche-Gerdes S, Jocham D, Böhle A. Sensitivity of BCG to modern antibiotics. *Eur Urol*. 2000; 37(Suppl 1):21–25. [PubMed: 10575268]

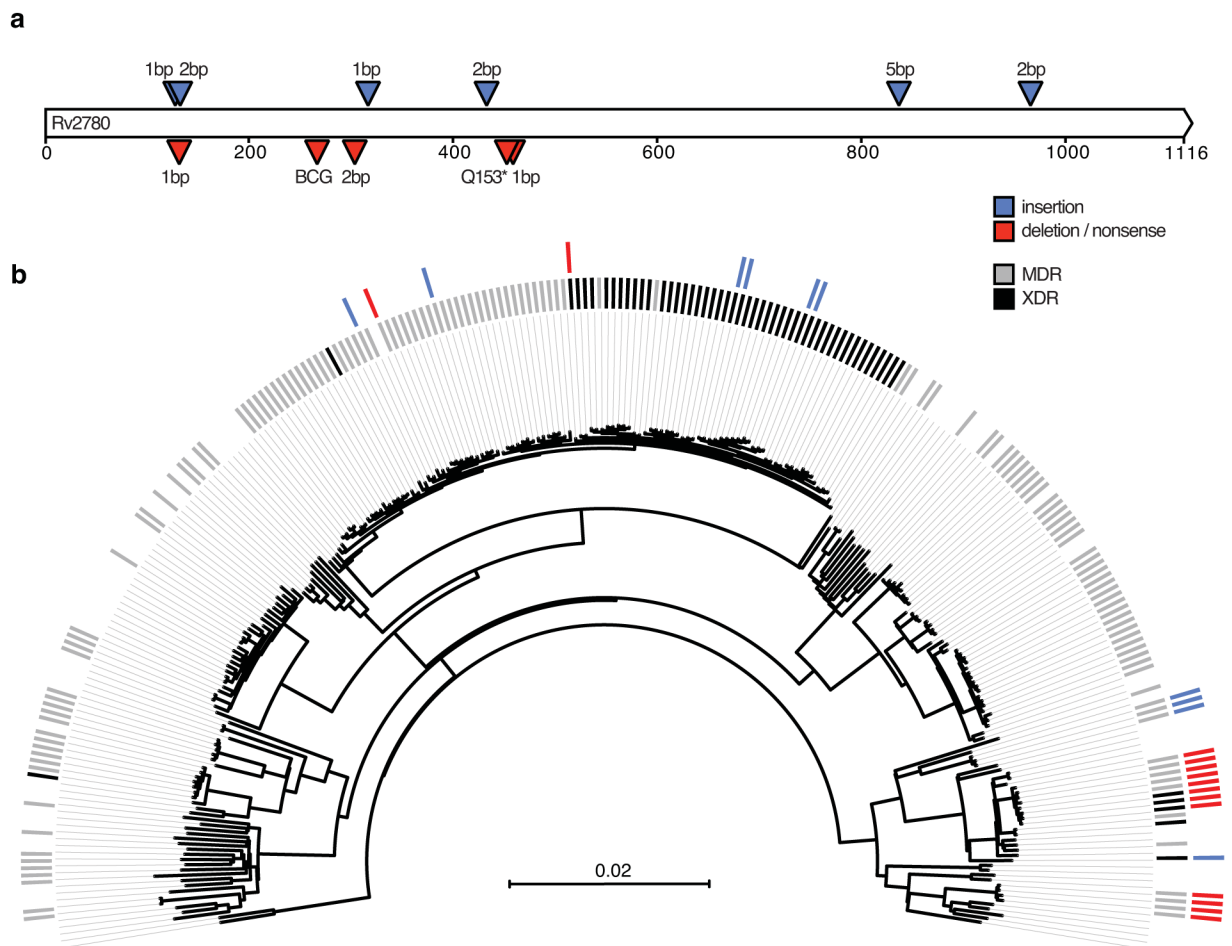


44. Goh KS, Rastogi N. Rapid preliminary differentiation of species within the *Mycobacterium tuberculosis* complex: proposition of a radiometric method. *Res Microbiol.* 1991; 142:659–665. [PubMed: 1961977]
45. Pelayo MCG, et al. A Comprehensive Survey of Single Nucleotide Polymorphisms (SNPs) across *Mycobacterium bovis* Strains and *M. bovis* BCG Vaccine Strains Refines the Genealogy and Defines a Minimal Set of SNPs That Separate Virulent *M. bovis* Strains and *M. bovis* BCG Strains. *Infect Immun.* 2009; 77:2230–2238. [PubMed: 19289514]
46. Chen JM, Alexander DC, Behr MA, Liu J. *Mycobacterium bovis* BCG Vaccines Exhibit Defects in Alanine and Serine Catabolism. *Infect Immun.* 2003; 71:708–716. [PubMed: 12540549]
47. Garnier T, et al. The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci.* 2003; 100:7877–7882. [PubMed: 12788972]
48. Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol.* 2003; 48:77–84. [PubMed: 12657046]
49. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE.* 2012; 7:e46688. [PubMed: 23056405]
50. Falush D, Bowden R. Genome-wide association mapping in bacteria? *Trends Microbiol.* 2006; 14:353–355. [PubMed: 16782339]
51. Shell SS, et al. DNA Methylation Impacts Gene Expression and Ensures Hypoxic Survival of *Mycobacterium tuberculosis*. *PLoS Pathog.* 2013; 9:e1003419. [PubMed: 23853579]
52. Brosch R, et al. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci.* 2002; 99:3684–3689. [PubMed: 11891304]
53. Chavadi S, et al. Global Effects of Inactivation of the Pyruvate Kinase Gene in the *Mycobacterium tuberculosis* Complex. *J Bacteriol.* 2009; 191:7545–7553. [PubMed: 19820096]
54. Keating LA, et al. The pyruvate requirement of some members of the *Mycobacterium tuberculosis* complex is due to an inactive pyruvate kinase: implications for in vivo growth. *Mol Microbiol.* 2005; 56:163–174. [PubMed: 15773987]
55. Ängeby K, Jurén P, Kahlmeter G, Hoffner SE, Schön T. Challenging a dogma: antimicrobial susceptibility testing breakpoints for *Mycobacterium tuberculosis*. *Bull World Health Organ.* 2012; 90:693–698. [PubMed: 22984314]
56. Torrea G, et al. Bedaquiline susceptibility testing of *Mycobacterium tuberculosis* in an automated liquid culture system. *J Antimicrob Chemother.* 2015; 70:2300–2305. [PubMed: 25977401]
57. Kahlmeter G. The 2014 Garrod Lecture: EUCAST – are we heading towards international agreement? *J Antimicrob Chemother.* 2015; 70:2427–2439. [PubMed: 26089441]
58. Pholwat S, et al. Integrated Microfluidic Card with TaqMan Probes and High-Resolution Melt Analysis To Detect Tuberculosis Drug Resistance Mutations across 10 Genes. *mBio.* 2015; 6:e02273–14. [PubMed: 25714709]
59. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol.* 1995; 57:289–300.
60. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 2009; 25:1754–1760. [PubMed: 19451168]
61. Walker BJ, et al. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS ONE.* 2014; 9:e112963. [PubMed: 25409509]
62. Price MN, Dehal PS, Arkin AP. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE.* 2010; 5:e9490. [PubMed: 20224823]
63. George KM, Yuan Y, Sherman DR, Barry CE. The Biosynthesis of Cyclopropanated Mycolic Acids in *Mycobacterium tuberculosis*. Identification and functional analysis of CMAS-2. *J Biol Chem.* 1995; 270:27292–27298. [PubMed: 7592990]
64. Streicher EM, et al. Rapid Sequencing of the *Mycobacterium tuberculosis pncA* Gene for Detection of Pyrazinamide Susceptibility. *J Clin Microbiol.* 2014; 52:4056–4057. [PubMed: 25165081]



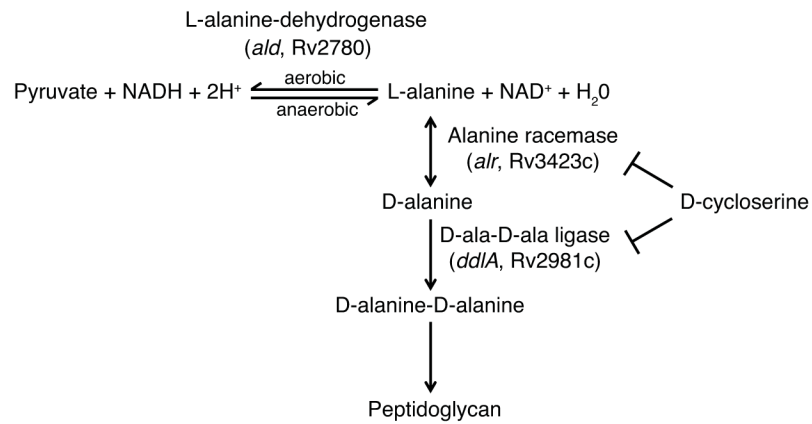
**Figure 1.**

Significance of associations between genotypic variants and drug resistance phenotypes. Each circle represents a genotypic feature that is plotted at the intersection of the negative log p-values from Fisher's exact and correlated evolution tests for each drug. P-values were corrected for multiple comparisons using the Benjamini-Hochberg method<sup>59</sup>. Genotypic variants known to confer resistance are colored according to the drug to which they confer resistance, and genotypes with no known effect on drug resistance are grey. Genotypes scoring well in both tests appear in the upper right quadrants.

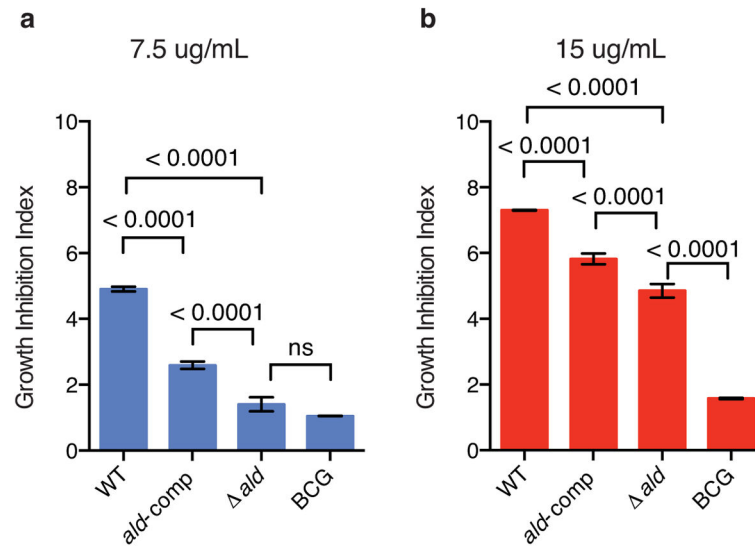


**Figure 2.**

Convergent evolution of loss-of-function mutations in *ald* in MDR and XDR *M. tuberculosis*. A) Positions of insertions, deletions, and nonsense mutations in *ald*. Insertions are marked as blue triangles while deletions and nonsense mutations are marked as red triangles. Size of the indel or details of nonsense mutation are shown above or below each triangle. The location of the frameshift in BCG is also included. B) Phylogeny of lineage 4 strains. Inner tick marks denote MDR (gray) and XDR (black) *M. tuberculosis*. Outer tick marks denote insertions (blue) and deletions or nonsense mutations (red).

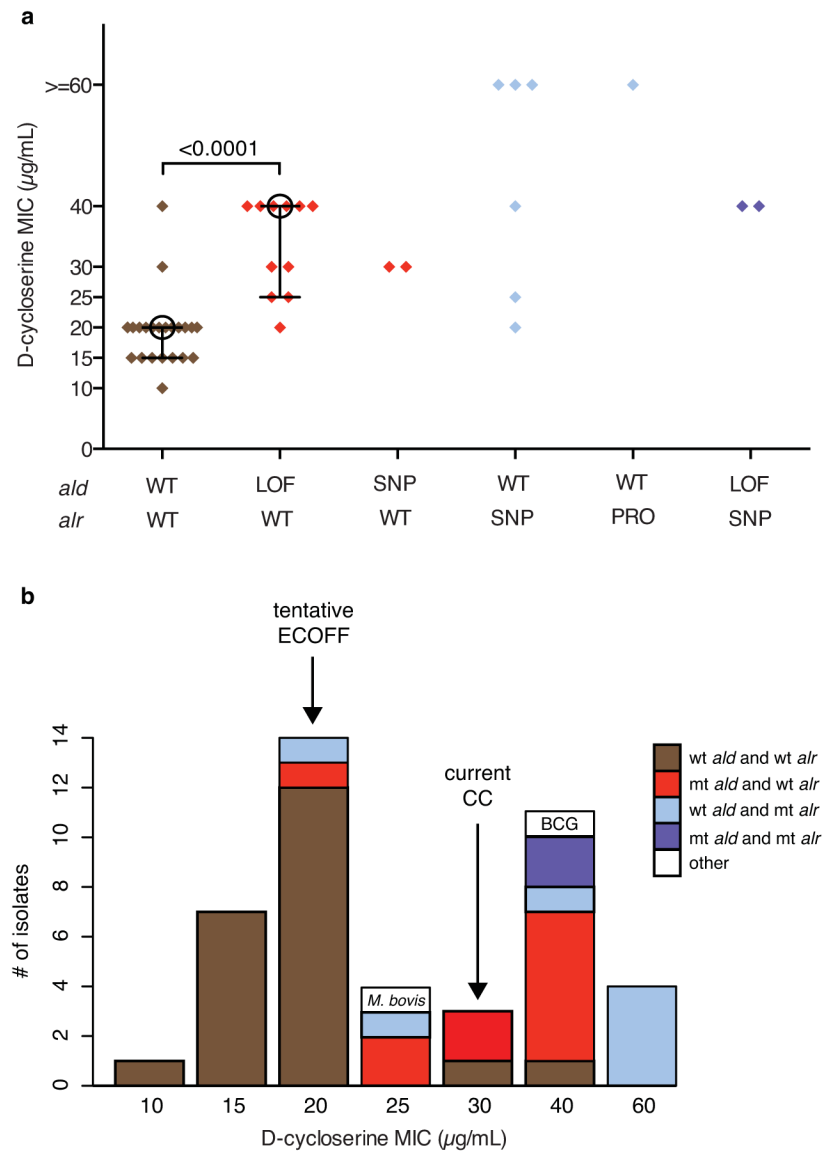
**Figure 3.**

Alanine metabolism pathway in *M. tuberculosis*. L-alanine-dehydrogenase catalyzes the NAD-dependent interconversion between L-alanine and pyruvate. Under aerobic conditions this reaction allows for the utilization of L-alanine as a nitrogen source, whereas under hypoxic conditions this reaction allows for NADH recycling to NAD<sup>+</sup>. Separately, the initial step of peptidoglycan synthesis involves conversion of L-alanine to D-alanine by alanine racemase. Two D-alanine molecules are joined by D-ala-D-ala ligase to produce the dipeptide D-alanine-D-alanine, which is subsequently incorporated into the pentapeptide chain of the peptidoglycan cell wall. Both alanine racemase and D-ala-D-ala ligase are competitively inhibited by D-cycloserine, an analog of D-alanine.



**Figure 4.**

A single gene knockout of *ald* confers a growth advantage relative to wild type *M. tuberculosis* when cultured in the presence of D-cycloserine. Four laboratory strains—wild type *M. tuberculosis* (WT), *ald* knockout (*ald*), *ald* complemented back into *ald* (*ald-comp*) and BCG—were cultured in triplicate in 0, 7.5 and 15  $\mu\text{g/mL}$  of D-cycloserine. The resulting time to positivity (ttp) in MGIT was recorded as days since inoculation, and the resulting ttp was normalized to the no drug control for each strain to calculate the growth inhibition index for each drug concentration. The mean growth inhibition index is plotted with error bars to represent the SEM. P-values were calculated using a two-way ANOVA.

**Figure 5.**

Clinical strains with mutations in L-alanine dehydrogenase (*ald*) and alanine racemase (*alr*) exhibit increased resistance to D-cycloserine. A set of 44 clinical strains of *M. tuberculosis* with differing genotypes with respect to *ald* and *alr* were selected, representing wild type alleles (WT), loss-of-function mutations (LOF), nonsynonymous mutations (NSY) and promoter mutations (PRO). The minimum inhibitory concentration (MIC) to D-cycloserine was determined for each strain on Löwenstein–Jensen media. For wild type strains and strains with *ald* loss-of-function mutations, the median is shown as an open circle and quartiles are shown as connected bars. When compared to WT isolates, strains with *ald* loss-of-function had significantly increased resistance to D-cycloserine ( $p < 0.0002$ , Mann-Whitney U test).



**Table 1**

Unexplained drug resistance in 498 strains of *M. tuberculosis* from South Africa and China. The total number of strains assayed for phenotypic resistance to each drug is shown, along with the number of phenotypically resistant (R) and phenotypically susceptible (S) strains. The number of strains with phenotypic drug resistance but without a match to known resistance-associated mutations is also given.

Drug	# of phenotyped strains			# strains w/unexplained R
	Total	R	S	
Rifampicin	498	337	161	8
Isoniazid	498	350	148	21
Pyrazinamide	142	72	70	12
Ethambutol	394	164	230	15
Streptomycin	498	256	242	21
Ofloxacin	494	162	332	15
Kanamycin	491	104	387	20
Capreomycin	300	66	234	15
Ethionamide	288	102	186	16

**Table 2**

Loss-of-function of L-alanine dehydrogenase (*ald*) confers increased resistance to D-cycloserine *in vitro*. The minimum inhibitory concentration (MIC) to D-cycloserine was determined for each strain on Löwenstein–Jensen media. Testing was done in three independent experiments (assays #1–#3) to establish reproducibility. Dashes indicate those strains were not tested during that experiment.

Strain	D-cycloserine MIC (µg/mL)		
	Assay #1	Assay #2	Assay #3
wild type <i>M. tuberculosis</i>	15	15	-
<i>ald</i>	25	30	25
<i>ald</i> -comp	20	20	20
BCG	40	40	60
BCG- <i>ald</i> -comp	40	-	40
<i>M. bovis</i>	-	25	25

**Table 3**

Counts of clinical isolates with mutations in *ald*, *alr*, *ddlA*, and *cycA*. Mutation types include wild type (no mutations; WT) loss-of-function (LOF), nonsynonymous (NSY), and promoter SNPs (PRO); synonymous SNPs are excluded. Specific strains and mutations are shown in Supplementary Table 8.

# of Clinical Isolates	<i>ald</i>	<i>alr</i>	<i>ddlA</i>	<i>cycA</i>
449	WT	WT	WT	WT
20	LOF	WT	WT	WT
5	NSY	WT	WT	WT
4	LOF	NSY	WT	WT
12	WT	NSY	WT	WT
1	WT	PRO	WT	WT
1	WT	WT	NSY	WT
1	WT	NSY	WT	NSY
5	WT	WT	WT	NSY