eLife

Reviewed Preprint
v1 • November 11, 2024
Not revised

Reviewed Preprint
v2 • February 20, 2026
Revised by authors

A macroevolution-inspired approach to reveal novel antibiotic resistance mechanisms

Fernanda T Subtil[1], Teresa FG Machado[1,†], Holly Douglas[1,‡], Joanna M Kirkpatrick[2], Mark Skehel[2], Acely Garza-Garcia[1], Luiz Pedro S de Carvalho[1,3]

[1]Mycobacterial Metabolism and Antibiotic Research Laboratory, The Francis Crick Institute, London, United Kingdom • [2]Proteomics Scientific Technology Platform, The Francis Crick Institute, London, United Kingdom • [3]Department of Chemistry, The Herbert Wertheim UF Scripps Institute for Biomedical Innovation & Technology, Jupiter, United States

For correspondence: soriodecarval.lp@ufl.edu

Competing interests: The authors declare no competing interests.

† Present address: The Institute of Cancer Research, Centre for Cancer Drug Discovery, Sutton, United Kingdom

‡ Present address: AstraZeneca, The Discovery Centre, Cambridge Biomedical Campus, Cambridge, United Kingdom

Funding: See page 20

Reviewing editor: Bavesh D Kana, University of the Witwatersrand, South Africa

eLife Assessment

This **important** study introduces an approach to discovering antibiotic resistance determinants by leveraging diverse susceptibility profiles among related mycobacterial species, with particular relevance to high-level resistance against natural product-derived antibiotics. The research provides **convincing** evidence for the role of ADP-ribosylation enzymes in rifamycin resistance among mycobacteria, whilst also demonstrating that antibiotic susceptibility is not correlated with growth rate or intracellular compound concentration. The revision is substantively improved, but some broader claims still require additional experimental support. This work lays a significant foundation for understanding the complexity of antibiotic resistance mechanisms in mycobacteria and opens new avenues for future antimicrobial research.

https://doi.org/10.7554/eLife.101940.2.sa3

Abstract

With the continuous rise in antibiotic resistance, novel methods that can reveal currently unknown antibiotic resistance mechanisms are essential to prepare and inform health responses and novel antibiotic discovery campaigns. Here we built a library of species representative of the genus *Mycobacterium* and determined their antibiotic resistance profiles, allowing for the first time systematic multispecies comparisons. Analyzing antibiotic resistance in the context of other closely related yet diverse organisms revealed species with truly exceptional traits as well as general principles underpinning antibiotic resistance. Among these, we reveal that intrabacterial accumulation of antibiotics does not correlate with their potency, at the species level. Our data also reveals that rifamycin resistance in mycobacteria is dominantly caused by antibiotic modification, contrary to what has been observed in *M. tuberculosis*. Our data provides a solid starting point for the exploration of novel determinants of antibiotic resistance. We illustrate the utility of this species-level approach to discovery of novel traits by characterizing a previously unrecognized rifamycin-inactivating enzyme group that is present in a wide range of bacterial genera.

Introduction

The discovery and successful clinical deployment of antibiotics is one of the most important breakthroughs in medical history, as it has dramatically reduced the morbidity and mortality in infections. Conversely, the extensive use of antibiotics has promoted the selection and propagation of mechanisms that enable bacteria to thrive regardless (1,2). Not all resistance mechanisms can be attributed to human-derived antibiotic exposure, as many predate the clinical use of antibiotics and/or are observed in bacteria of no clinical importance (3). These mechanisms are derived from

the natural evolution of species and the environment they occupy, as bacteria that co-inhabit niches with antibiotic-producing organisms naturally develop ways to avoid antibiotic-caused death (4, 5). Of key importance, antibiotic resistance mechanisms do not have to evolve de novo, as they can be incorporated from other organisms via horizontal gene transfer (HGT).

Two main strategies are used to identify novel antibiotic resistance determinants. One clinical or host-biased approach ensues when a patient who is infected with a presumably antibiotic-sensitive species fails to improve upon treatment. Isolation and study of the resistant strain then leads to the identification of a novel determinant of antibiotic resistance. Several seminal discoveries of direct clinical importance have taken place in this manner (6–8). However, this approach mostly discovers incremental strain-specific mechanisms, such as single nucleotide polymorphisms (SNP) that alter antibiotic binding to its target. Another strategy involves ecological sampling and screening, using targeted approaches or metagenomics (9). The latter approach holds the promise to uncover truly novel mechanisms, although these mechanisms might never find their way into extant pathogens, and therefore will never represent a clinically relevant problem.

Mycobacterium is a genus that includes both environmentally and clinically relevant microorganisms. Mycobacteria can be found in most environments including rivers and lakes (10), soil (11), plant roots (12), moss (13), reptiles (14), amphibians (15), fish (16) and mammals (17). A typical divide of the genus is based on growth rate in a defined solid medium when sub-cultured from highly dilute inoculum (18). A mycobacterial species is “fast-growing” if it forms visible colonies within seven days; this phenotype is considered to be the ancestral state of the genus (19). A species that takes longer than seven days to form mature colonies is classified as “slow-growing”, this category includes the most devastating disease-causing species *Mycobacterium ulcerans*, *M. avium* complex (MAC), *M. leprae*, and the *M. tuberculosis* complex (MTBC). Until recently, human infections with mycobacteria other than the MTBC and *M. leprae*, i.e. non-tuberculous mycobacteria (NTM), were overshadowed by the TB burden but are now gaining increased attention due to their growing prevalence (20). Some of the key NTM species that cause disease are the fast-growers *M. abscessus* and *M. fortuitum*, and the slow-growers *M. avium*, *M. marinum*, *M. xenopi*, *M. gordonae* and *M. kansasii* (21). NTM can infect a variety of tissues including the lungs, central nervous system, lymphatic system, joints, and skin (21). As the frequency of NTM infections is increasing, so is the worry of their severity and resistance to treatment with available antibiotics. Contrasting with *M. tuberculosis* and *M. leprae*, which due to their isolation inside the host do not show HGT, NTM might acquire resistance determinants from environmental bacteria using phages, plasmids, and other mobile genetic elements (22, 23). Therefore, the potential of mycobacterial species to contain and disseminate antibiotic resistance determinants is likely as good as that of other environmental bacteria.

Here, we propose a new method to study antibiotic resistance by comparing antibiotic resistance profiles across related bacterial species (macroevolution), instead of comparing strains of the same species (microevolution) to identify previously unknown high-level antibiotic resistant species and understand general principles of antibiotic resistance in a genus. Then, taking advantage of the natural genetic diversity among species (largely encoded in their accessory genome) and their genetic similarity (encoded in their core genome), computational, molecular, and cellular approaches can more readily pinpoint resistance determinants. We demonstrate the power of this method by characterizing a previously unrecognized rifamycin-inactivating enzyme group that is widely distributed across bacterial genera.

Results

Building a diverse library of mycobacterial species

A collection of 44 tractable mycobacterial species was assembled to cover most of the mycobacterial phylogenetic tree (Fig. 1 a). To evaluate the biological diversity of our library we analyzed the ecological and genomic information available for the selected species. We determined doubling times in liquid culture in equivalent experimental conditions for 26 of the

species (Fig. 1c). The results revealed large differences within the genus. For example, *M. szulgai* and *M. flavescens* divided every 1.3 and 2.0 hours, respectively, while *M. marinum* and *M. tuberculosis*, divided every 17.1 hours. Strikingly, among slow-growers a 13-fold change difference exists between the fastest and the slowest-growing species (*M. szulgai* and *M. tuberculosis*, respectively). The genome size of the library species is also highly different, with genomes as small as 3.59 Mbp and 4.08 Mbp, for *M. triviale* and *M. koreense*, respectively, to genomes as large as 6.99 Mbp and 8.01 Mbp for *M. smegmatis* and *M. mageritense*, respectively (Fig. 1d). These genome size differences suggest that the nature of the accessory genome varies widely from species to species. Interestingly, there is only a very small difference when comparing the average genome size of all fast and slow growers, 6.1 vs 5.8 Mbp, respectively. The high guanine-cytosine content (GC%) is a defining characteristic of mycobacteria. Within our library, GC% varies from 66.9% to 68.8%, in slow growers and the *M. terrae* clade respectively (Fig. 1e). Next, we investigated the number of ribosomal RNA encoding genes (*rrn* operon), a feature that has often been linked to growth rate. Twenty-five species possess a single copy of the *rrn* operon while the other 19 possess two copies (Fig. 1f). In general, but not always, “fast growers” tend to have two copies while “slow growers” one. In the genus *Bacillus* a correlation between growth rate and *rrn* operon copy number has been experimentally disproven (24), and number of *rrn* operon copies could be related to response of resource availability (25). Furthermore, comparative pangenome analyses conducted by Bachmann and collaborators suggested that growth limitation in slow growing mycobacteria might instead be related to loss of amino acid transporters (26). Finally, we analyzed the gene ontology (GO) distribution for our library in order to evaluate functional genetic diversity at the genome level. GO distribution varies noticeably across the analyzed genomes. The number of genes associated with transcription and membrane transport (categories indicated by arrows in Fig. 1g) appeared to be particularly variable. We noted that much of the genome of several of these species is poorly annotated. Therefore, Fig. 1 supports the potential usefulness of our library as a resource to investigate mycobacterial biology, antibiotic resistance, and pathogen evolution.

Wide variation in antibiotic resistance in mycobacteria

To harness the potential of our library, we tested the antibiotic potency and the extent of its variation across the *Mycobacterium* genus to identify biologically-relevant differences to be further studied. We determined minimal inhibitory concentrations (MIC_{99}) for 15 antibiotics, spanning most of the classes employed to treat mycobacterial infections, including TB (Fig. 2a, Supplementary Table 1). We found that several species displayed at least one MIC_{99} value that is considerably different from the mean (Supplementary Fig. 1 and 2), highlighting the biological diversity of the genus with respect to antibiotic potency. As expected, the notoriously multi-drug resistant *M. abscessus* was resistant to several antibiotics (Fig. 2a) (21,27), yet *M. abscessus* was not the most resistant species studied. *M. mageritense*, *M. salmoniphilum* and *M. houstonense* were highly resistant to most of the antibiotics tested. *M. abscessus* was somewhat sensitive to amikacin (AMK) and bedaquiline (BDQ), which is consistent with other findings in the literature (28, 29). Also, the magnitude of the differences in MIC_{99} when compared to *M. tuberculosis* is remarkable, of the order of 100- to 1000-fold in some instances. Of note, when the data were re-ordered and unbiased clustered, based on the overall antibiotic sensitivity of each individual species, the species distribution were divided in three main clusters, which are dramatically different when compared to their phylogenetic positioning, as shown by the tanglegram between the two heatmaps in Fig. 2a. To illustrate the absence of a taxonomic trend in antibiotic resistance we explored in detail the clade composed by *M. holsaticum*, *M. phlei*, *M. flavescens*, *M. tusciae* and *M. moriokaense* (Supplementary Fig. 3). While most antibiotics behaved similarly across this group (i.e., MIC_{99} FC < 3-fold). *M. holsaticum* was highly sensitive to para-aminosalicylic acid (PAS) and highly resistant to BDQ. This higher sensitivity to PAS cannot be attributed to any known mechanisms. For example, the recently discovered mutations associated with lower methylene tetrahydrofolate reductase (Rv2172c) activity, leading to increased PAS sensitivity in *M. tuberculosis* (30), are not observed in the *M. holsaticum* homolog (WP_069405771.1 - Supplementary Fig. 4). High-level BDQ resistance was also observed with *M. flavescens*.

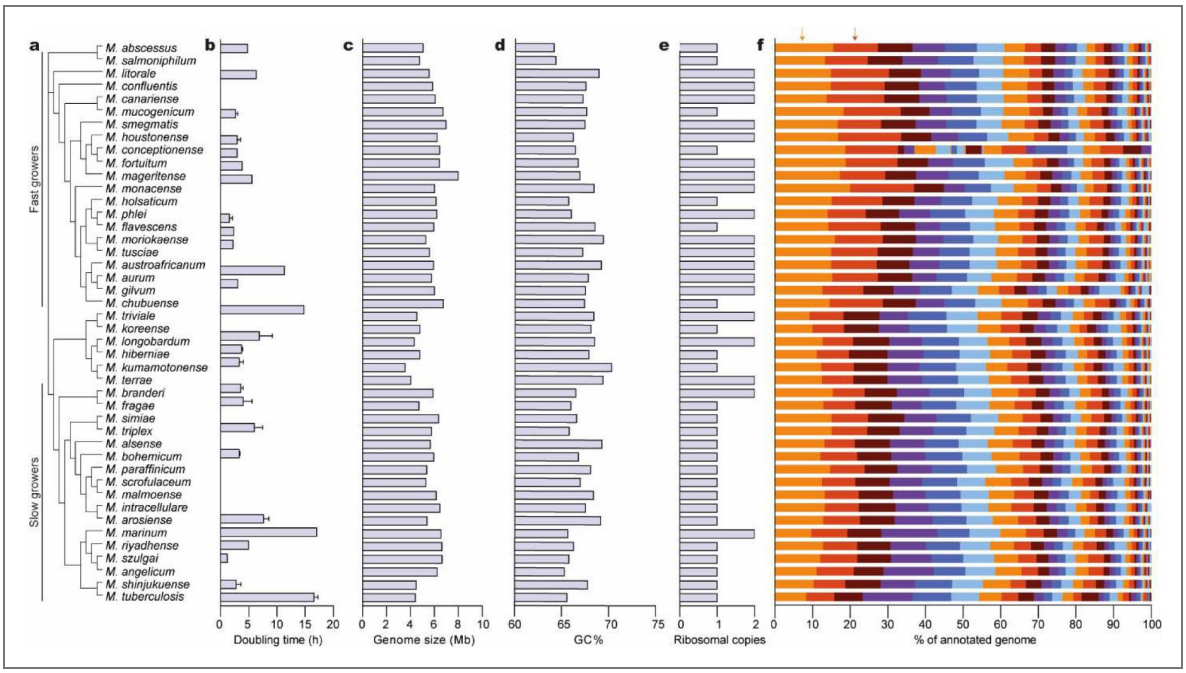


Fig. 1. A diverse species library of the genus *Mycobacterium*.

a, Phylogenetic tree of mycobacterial species in our library calculated using the bcgTree pipeline v 1.1.0 [62, 65-68]. b, Doubling time. c, Genome size. d, Guanine, and cytosine percentage (GC%). e, Ribosomal copies (rrn operon). f, Gene ontology (GO) distribution. Colors from left to right represent the following GO categories: regulation of DNA-templated transcription (highlighted with a light orange arrow), transmembrane transport (highlighted with a dark orange arrow); amino acid, lipid, carbohydrate derivative, nucleobase-containing small molecule, and carbohydrate metabolic processes; generation of precursor metabolites and energy; sulfur compound, vitamin, and tRNA metabolic processes; DNA repair, protein modification process, signaling, cell wall organization or biogenesis, cellular modified amino acid metabolic process, DNA replication, DNA recombination, ribosome biogenesis, anatomical structure development, protein catabolic process, protein-containing complex assembly, protein maturation, nitrogen cycle metabolic process, intracellular protein transport, metal ion homeostasis, cell division, protein secretion, mRNA metabolic process, DNA integration, transport, organic substance transport, defense response to other organism, organic substance biosynthetic process, organic substance metabolic process, cellular process, nitrogen compound transport, regulation of gene expression, cellular biosynthetic process, other metabolic processes.

Interestingly, M. flavescens is unusually sensitive to D-cycloserine (DCS). The mechanisms underpinning these distinct responses to antibiotics are currently unknown. We also observed that once ordered by antibiotic sensitivity, *M. tuberculosis* is positioned at the middle of the heatmap and the number of NTM more resistant to antibiotics compared to *M. tuberculosis* is nearly equal to the number of NTM that are more sensitive. Based on the data above, obtained with 44 species, NTM are not intrinsically more drug resistant to the antibiotics tested than *M. tuberculosis*. In summary, antibiotic sensitivity varies dramatically across the *Mycobacterium* genus and our data provide the first quantitative blueprint of this variation.

There are striking differences in antibiotic response across the genus that highlight the value of a genus-wide approach to inform antibiotic research efforts. For example, *M. smegmatis* is frequently used as a model organism for *M. tuberculosis* in TB antibiotic discovery (31), but it displays a completely different sensitivity profile from *M. tuberculosis*, being highly resistant to PAS, ethionamide (ETH), DCS and RIF.

Conversely, our results suggest that *M. marinum* is a better *M. tuberculosis* proxy (32) as both have a similar sensitivity profile, except to ofloxacin (OFX) (Fig. 2b). The overall distribution of antibiotic potency (MIC_{99}) against different species is shown in Fig. 2c. Cellenvelope-targeting antibiotics and PAS exhibit a weaker potency across the genus, while antibiotics that target protein synthesis, DNA gyrase and the ATP synthase on average displayed an overall lower MIC_{99}, indicating that most mycobacteria are sensitive to them. From the antibiotics that inhibit protein synthesis, linezolid (LZD) displays the lowest overall MIC_{99} and was effective against most species (Fig. 2c). To verify whether there is a correlation between doubling time and sensitivity to antibiotics we compared the doubling time of a subset of species (Fig. 1 c) with the MIC_{99} of a subset of antibiotics. No correlation is apparent between growth rate and antibiotic sensitivity in mycobacteria (Fig. 2d). Given the antibiotic response profiles observed, we selected BDQ, LZD and RIF to explore the molecular causes of these dramatic changes in antibiotic potency observed across the *Mycobacterium* genus.

Intra-bacterial antibiotic accumulation does not predict potency

We employed liquid chromatography—time-of-flight mass spectrometry (LC-MS) to determine the relative intrabacterial (IB) concentration of antibiotic ($[ABX]_{IB}$) with an antibiotic concentration in the growth medium of 6×MIC_{99}. Generally, $[ABX]_{IB}$ is a function of three parameters: antibiotic uptake, efflux, and modification. Fig. 3a shows extracted ion chromatograms (EICs) in five mycobacterial species for BDQ, LZD and RIF (Fig. 3b). Quantification of $[BDQ]_{IB}$, $[LZD]_{IB}$ and $[RIF]_{IB}$ illustrates the variability and the magnitude of the changes observed in $[ABX]_{IB}$, spanning from 2- to 200-fold (Fig. 3c). As the experiment was performed at a concentration of antibiotic proportional to its MIC_{99}, and therefore antibiotics were equally potent, we replotted these data as a function of each antibiotic MIC_{99} (Fig. 3d). Only for BDQ we could observe some correlation between antibiotic potency and $[BDQ]_{IB}$ which could be indicative of efflux playing a role in antibiotic efficacy. In the case of RlF, where no correlation between antibiotic potency and its accumulation in mycobacteria is observed (Fig. 3d), factors other than uptake and efflux must be at play.

A minor role for pre-existing target modification in RIF resistance

Considering the importance of rifamycins for the treatment of TB, leprosy, Buruli ulcer, MAC and *M. kansasii* infections, we focused on RIF resistance mechanisms operating in mycobacteria. Fig. 4a highlights the diversity in RIF potency across our library, ranging from an MIC_{99} of more than 100 µg/mL to less than 0.2 µg/mL. Arranging species by decreasing MIC_{99} value highlights that there are species better suited for the identification of target-mediated resistance mechanisms (dark orange), and species that are better suited for the identification of non-target-based resistance mechanisms (dark purple). At this stage, we focused our work on four species, all of which are resistant (MIC_{99} = 12.5 µg/mL) or super-resistant (MIC_{99} = 100.0 µg/mL) to RIF, compared to *M. tuberculosis* (MIC_{99} = 0.9 µg/mL): *M. smegmatis* and *M. flavescens* (MIC_{99} = 12.5 µg/mL), *M. houstonense* (M1C99 = 25.0 µg/mL), and *M. conceptionense* (MIC_{99} = 100.0 µg/mL) (Fig. 4b).

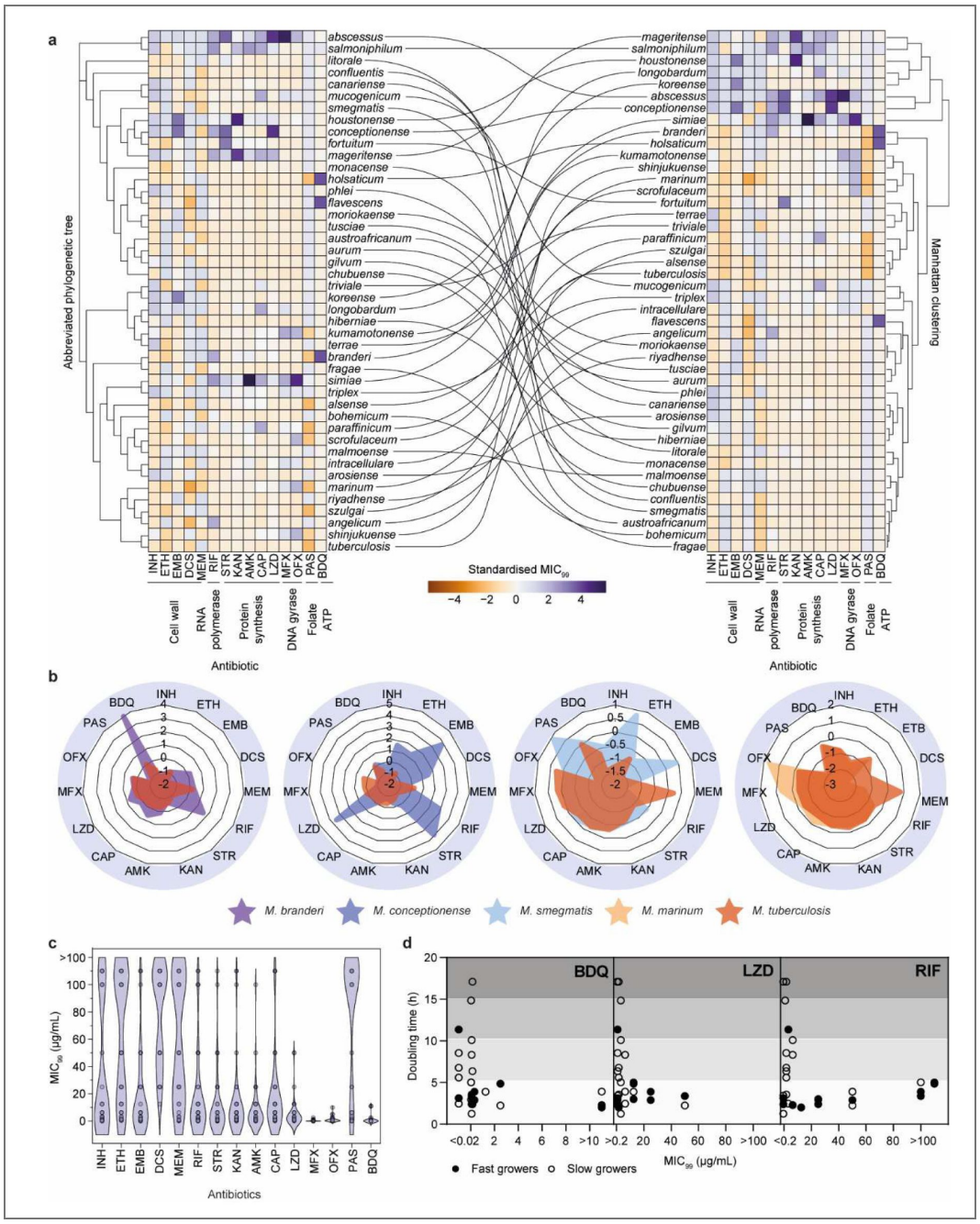


Fig. 2. Antibiotic sensitivity mapping reveals complex patterns.

a, Heatmaps of overall MIC_{99} values. In the X axis, antibiotics are organized based on their mechanism of action; in the Y axis, mycobacterial species are organized phylogenetically in the left heatmap, and based on their response to the set of antibiotics tested (Manhattan clustering) in the right heatmap. Colors represent the standardized MIC_{99} (mean/SD and centered scaled). Lower MIC_{99} values are in brown/orange and higher MIC_{99} values in lilac/purple. The details of the data can be found in Supplementary Figure 2. b, Radar plots displaying the standardized MIC_{99} for all antibiotics tested. MIC_{99} - values are normalized to be plotted in radar plots. All radar plots display the results for *M. tuberculosis* in orange. *M. branderi* is displayed in purple, *M. conceptionense* in dark blue and *M. smegmatis* in light blue. c, Violin plots showing the distribution of MIC_{99} values. In the X axis, the set of antibiotics tested; in the Y axis the MIC_{99} values in µg/mL. d, Relationship between mycobacterial doubling time (X axis) and MIC_{99} for the antibiotics BDQ, LZD and RIF (Y axis). Antibiotics targeting the cell wall are: isoniazid (INH), ethionamide (ETH), ethambutol (EMB), D-cycloserine (DCS), and meropenem (MEM); RNA/protein synthesis: rifampicin (RIF), streptomycin (STR), kanamycin (KAN), amikacin (AMK) capreomycin (CAP) and linezolid (LZD); DNA gyrase: moxifloxacin (MFX) and ofloxacin (OFX); folate metabolism: para-aminosalicylic acid (PAS); and ATP synthase: bedaquiline (BDQ).

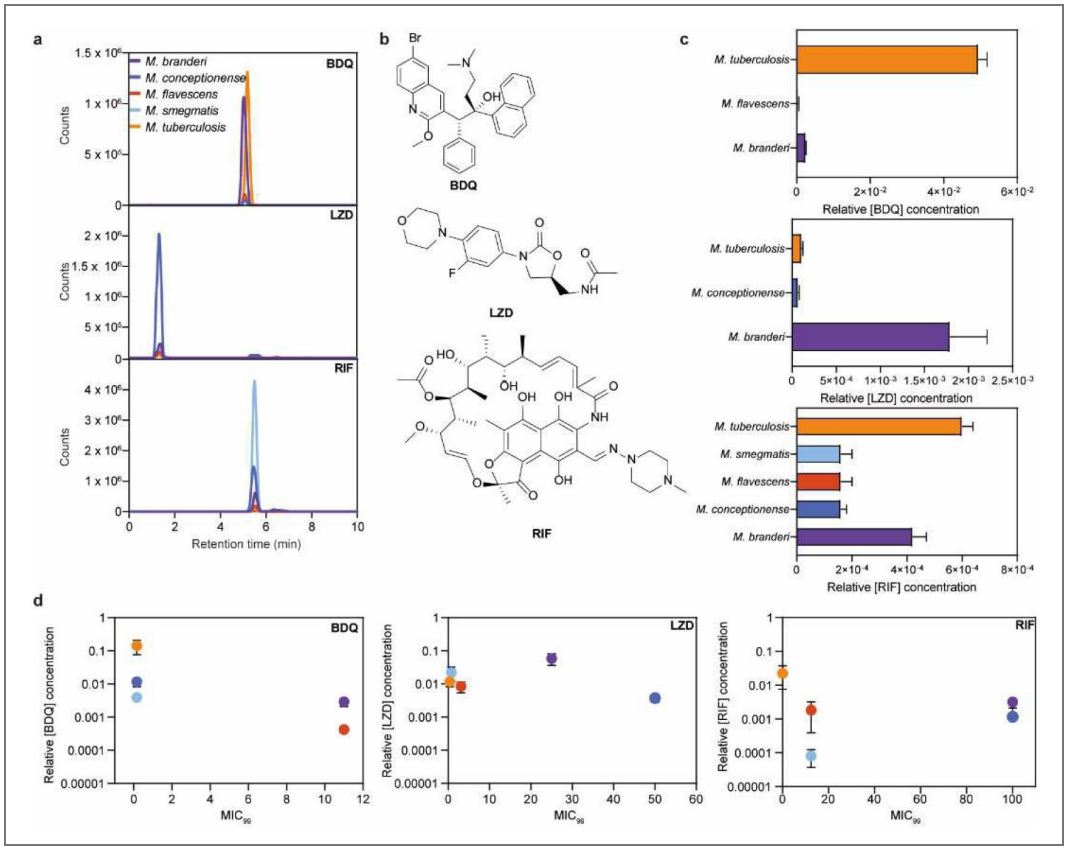


Fig. 3. Intra-bacterial antibiotic concentration does not correlate with potency.

a, Positive mode extracted ion chromatograms (EIC) of whole-cell extracts of mycobacteria treated with selected antibiotics. BDQ (m/z 555.1642), LZD (m/z 338.1511) and RIF (m/z 823.4124). b, Chemical structure of BDQ, LZD and RIF. c, Relative intracellular antibiotic concentration, obtained by comparing the peak height of the samples with and injection of 10 μM of antibiotic, then normalized by the concentration of antibiotic used to treat the cells. d, Relative intracellular antibiotic concentrations in relation to the MIC_{99}. Data in c and d come from independent experiments.

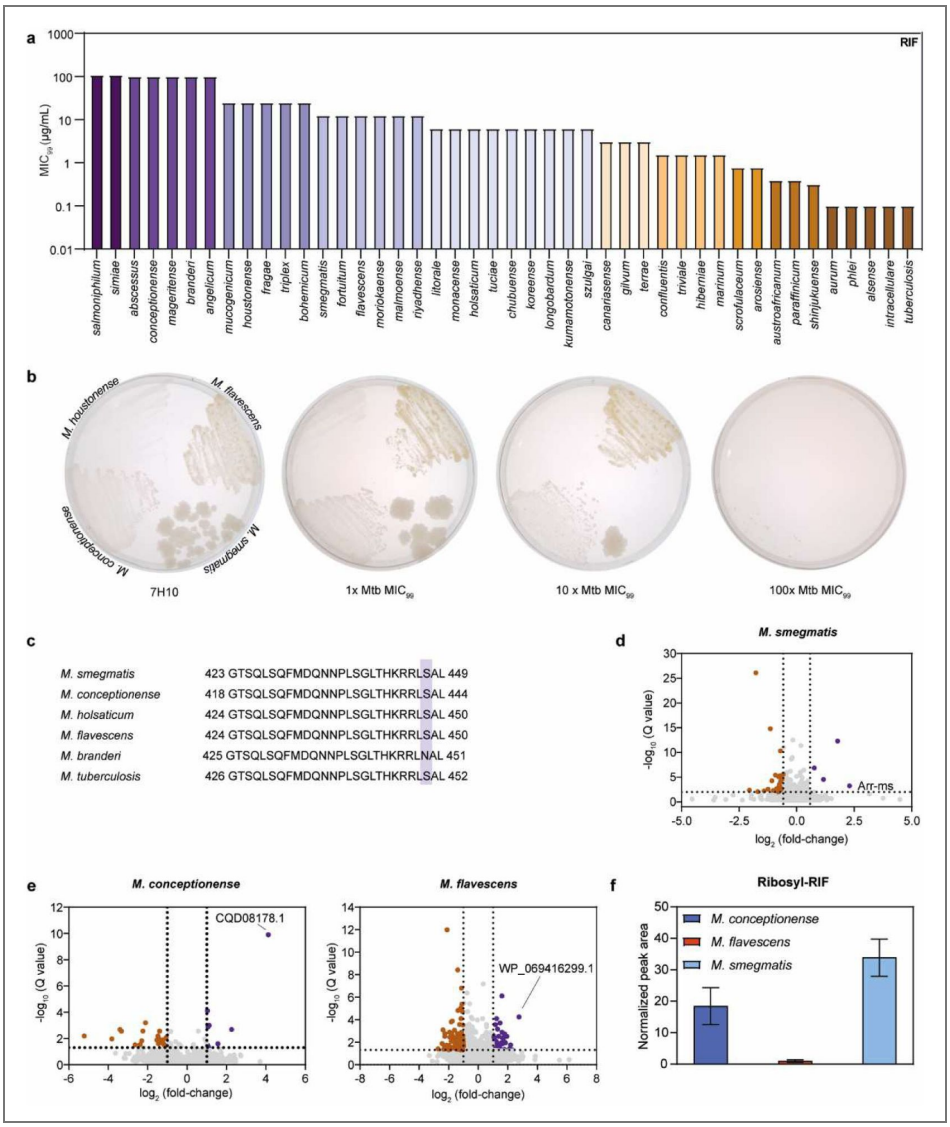


Fig. 4. High-level rifampicin resistance is caused by rifamycin modification in selected mycobacteria.

a, RIF MIC_{99} values for the mycobacterial species in our library organized in decreasing MIC_{99} value order. b, Cultures of selected species on solid medium (7H10) containing RIF at different concentrations, starting at 1× *M. tuberculosis* MIC. c, Comparison of the amino acid residue sequence of the rifampicin resistance-determining region (RRDR) of RpoB in selected mycobacterial species; the only residue that differs is Ser 450 in *M. branderi*. d and e, Volcano plots showing the differential protein expression in whole-cells with and without RIF revealing inducible expression of RIF ADP-ribosyltransferase 1 (Arr-1) in *M. smegmatis*, *M. conceptionense* and *M. flavescens*. f, Detection and quantification of ribosyl-RIF in whole-cell extracts by LC-MS.

In *M. tuberculosis*, the fixation of mutations that decrease the affinity of RIF to the RNA polymerase β subunit (RpoB) represents the dominant cause of RIF resistance (7, 33), and therefore target modification is an obvious starting point to explore probable mechanisms of resistance to the rifamycin class of antibiotics in mycobacteria. Fig. 4c shows the Rifampicin Resistance Determining Region (RRDR), the segment of RpoB where most mutations conferring resistance to RIF are found. Except for *M. branderi*, no amino acid variations are found in our species of interest. The rifampicin (RIF) binding pocket is generally conserved, but *Mycobacterium branderi* has an S45ON mutation in the RRDR region. While this specific mutation has not been found in clinical isolates (34), it is located at the rifamycin binding site and may therefore confer resistance. While S450N mutations have not yet been observed in *M. tuberculosis* related mutations like S450Q have been linked to resistance. In contrast, *M. conceptionense*, *M. flavescens*, and *M. smegmatis* show no target sequence differences that explain their resistance. This observation suggests that in contrast to *M. tuberculosis*, most mycobacteria are not resistant to RIF due to variations in the RIF binding region of RpoB.

As (RIF)IB or RpoB variation cannot account for the observed resistance to RIF, we evaluated the remaining major mechanism of resistance to rifamycins, drug modification. RIF modification is widely found in nature and is carried out by various enzyme types, including phosphotransferases, glycosyltransferases, ADP-ribosyltransferases (ARTs) and monooxygenases (35–39). Importantly, a RIF-ART homolog has been characterized in *M. smegmatis* (39); it is encoded by the gene MSMEG_1221, also known as *arr*-ms. The arr-ms encoded RIF ADP-ribosyltrasferase (RIF-ART) has been shown to be the sole determinant of RIF resistance in *M. smegmatis* by chemical and genetic methods (41, 42). We employed proteomics to first check whether Arr-ms is expressed in the absence of RIF and if it is differentially expressed in the presence of RIF. Fig. 4d shows that expression of Arr-ms is stimulated (5.6-fold) in the presence of RIF at 6×MIC_{99}, and therefore, proteomics can assist on the identification of RIF modifying enzymes in mycobacteria. Next, we evaluated whether the annotated Arr homologous proteins in *M. conceptionense* (SAMEA3305051) and in *M. flavescens* (SAMN05729960) were also induced in the presence of RIF (Fig. 4e), this was indeed the case (4.12- and 2.75-fold change respectively). To confirm that these putative RIF-ARTs are inactivating RIF, we employed LC-MS, to identify ribosyl-RIF (m/z 955.4601), a fragment of the larger ADP-ribosyl-RIF product, which fragments under our LC-MS conditions. Fig. 4f illustrates that ribosyl-RIF was observed in *M. smegmatis*, *M. conceptionense* and *M. flavescens* treated with RIF. Additionally, other mycobacteria with annotated putative *arr* genes also displayed high levels of RIF ADP-ribose (Supplementary Fig. 5a and 5b), indicating that RIF modification, and precisely ADP-ribosylation, is the dominant mechanism of resistance to RIF in mycobacteria.

A novel group of rifamycin ADP-ribosyltransferases

In order to have a comprehensive understanding of the distribution of Arrs in mycobacteria we mined for *arr* sequences in reference genomes and built a phylogenetic tree. Arr proteins were found to be widespread in both fast- and slow-growing mycobacteria, but in a dispersed pattern suggesting that both local vertical inheritance and gene losses and acquisitions have taken place. Surprisingly, mycobacterial Arrs form two orthologous groups (Fig. 5a; Supplementary Table 2). One of the groups, previously known, which we designated Arr-1, corresponds to sequences closely related to Arr-ms (median sequence identity of 80%). Arr-1 group members are predominantly Actinomycetota of the orders Geodermatophilales, Propionibacteriales, Micrococcales and Mycobacteriales. The second group, discovered here, which we have named Arr-X, is taxonomically more broadly distributed, including members from Actinomycetota, Bacillota, Pseudomonadota and Bacteroidota. Within mycobacteria, more species have an *arr*-1 gene than *arr*-X and interestingly, a few species have both, for example *M. conceptionense* and *M. flavescens* (Supplementary Fig. 5a). *M. conceptionense* Arr-1 (Uniprot A0A0U1D6J3) and Arr-X (Uniprot A0A0U1DL14) share 50% identity and 63% similarity (BLOSUM62). The equivalent of the three residues showed by Baysarowich and collaborators to be necessary for enzymatic activity in Arr-ms (Asp84, His19 and Tyr49) are conserved in all mycobacterial Arr-1 and Arr-Xs, suggesting

that they are all active ADP ribosyltransferases (38). However, the hydrophobic nature of the RIF binding cleft of Arr-ms is not completely preserved in the Arr-X group (Supplementary Table 3) hinting at probable differences in substrate binding preference.

To assess that Arr-X enzymes are indeed rifamycin-ARTs and to understand why some species have two *arr* genes, we cloned, overexpressed, purified, and tested the enzymatic activity of Arr-ms (as a control) and Arr-1 and Arr-X from both *M. conceptionense* and *M. flavescens*. Fig. 5c displays the catalytic activity (V^{app}, as a surrogate of the specificity constant *V/K*) of the different proteins with six rifamycins. All Arr-1 enzymes had similar activity and substrate preference, but Arr-X enzymes were much superior at inactivating rifamycins. For example, *M. flavescens* Arr-X is 4-fold faster with rifapentine than Arr-ms. Surprisingly, *M. conceptionense* Arr-X is 30-fold faster to inactivate rifabutin, compared to Arr-ms. These results therefore demonstrate that Arr-Xs are not only *bona fide* rifamycin inactivating enzymes, but also that they are significantly more efficient than Arr-1 s. We also determined the MIC_{99} for different rifamycins in selected species (Supplementary Table 4). Interestingly, these species are considerably resistant to all rifamycins except for rifabutin, for which the MIC is at a lower concentration. To probe whether Arr-X is active *in bacterio*, we used CRISPR interference to reduce the transcription of *arr*-1, *arr*-X and both genes in *M. conceptionense*. *M. conceptionense* continued to be resistant to rifabutin upon *arr*-1 silencing but became more sensitive when *arr*-X was knocked down (Fig. 5d). Thus, Arr-X is an active “rifabutinase” that confers rifamycin resistance in *M. conceptionense*.

Discussion

Discovery of unknown mechanisms for high-level antibiotic resistance in nature is essential for efficient antibiotic discovery and development and for the continuous treatment of patients. Our ability to discovery and improve antibiotics making them bypassing and evading resistance mechanisms is illustrated by the discovery of clavulanic acid and penems, and the rational development of glycylcyclines and more recently ethionamide booster compounds, which completely sensitize ethionamideresistant *M. tuberculosis* to ethionamide (70). Here we propose a powerful approach for the discovery of novel antibiotic resistance determinants. Our strategy consists of mapping and comparing the antibiotic response profile of a library of tractable species representative of the entire genus *Mycobacterium* (macroevolution) instead of strains of the same species (microevolution). This approach is applicable for the study of many biological traits and to other bacterial genera.

Using our mycobacterial library, we identified for the first time high- and ultra-high-level intrinsic resistance (3) to many of the antibiotics tested. Of note, the resistant phenotype is naturally occurring and not a result of mutations due to exposure to the antibiotic in the clinic – which is the more traditional approach for probing mechanisms of antibiotic resistance. Our observations revealed that resistance profiles are highly variable across the genus and do not follow phylogeny, implicating HGT as the key mechanism for acquisition of resistance determinants and evolution of antibiotic resistance in mycobacteria (43). While *M. tuberculosis* is famously regarded as “genetically locked”, which has not carried out HGT since it arose and underwent clonal expansion 35,000 years ago (71), the majority of species in the genus clearly exchanges DNA via plasmids, phages and other means. Our study revealed that resistance levels to BDQ, LZD and RIF were particularly divergent across the genus and often could not be explained by our current knowledge of antibiotic resistance mechanisms (6, 7, 31). We found that resistance to these antibiotics in mycobacteria do not correlate with by uptake/efflux mechanisms in the species tested and it does not correlate with growth rate. Identification of mycobacterial species highly resistant to BDQ and LZD is worrisome as most of this species, if not all, have never been exposed to these drugs.

We illustrated the power of this comparative method by characterizing a previously unrecognized group of rifamycin-inactivating enzymes (Arr-X) that is present in a wide range of bacteria including Actinomycetes, Bacilli and Gammaproteobacteria. We found that several mycobacterial species have the gene coding for the RIF inactivating enzyme RIF-ART (Arr-1), the founding member of this group of enzymes (38) and revealed that some species also code for a homologous

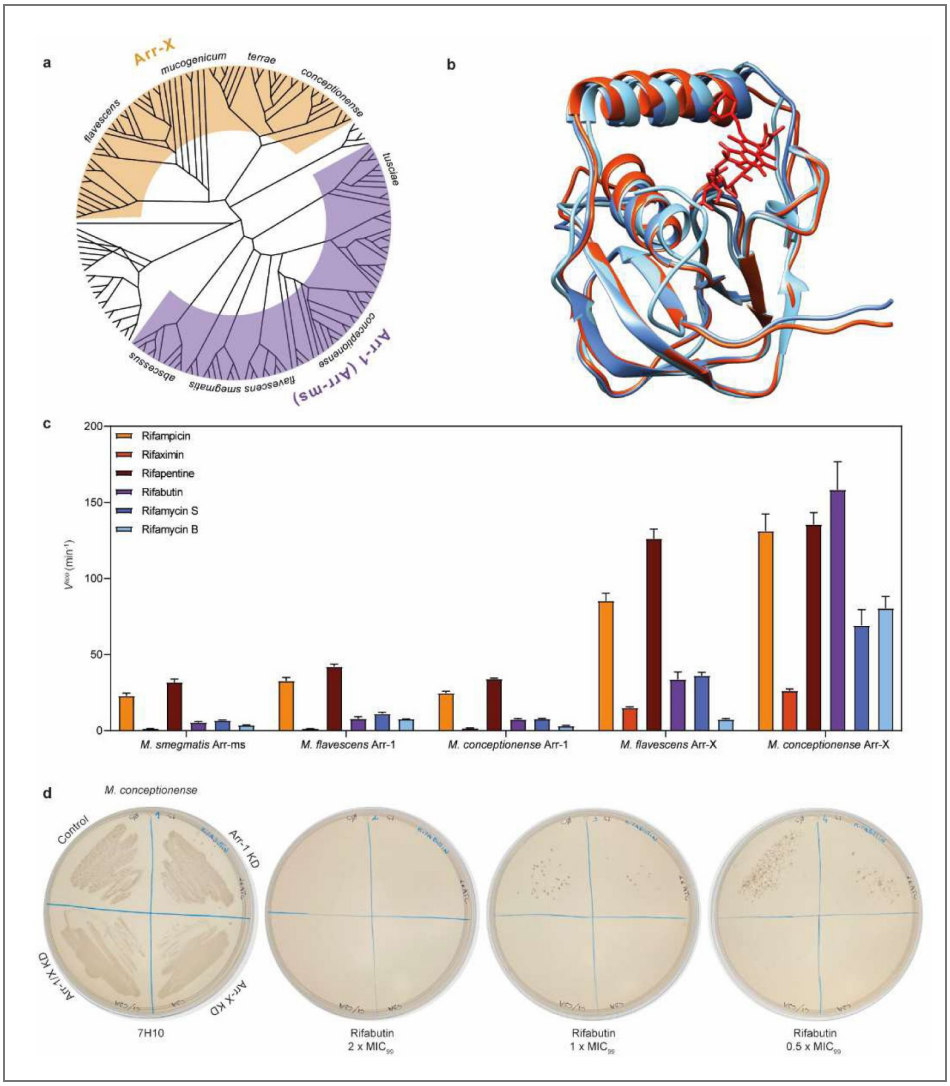


Fig. 5. Characterization of a novel rifabutin-ADP ribosyltransferase in mycobacteria.

a, Phylogenetic tree of mycobacterial RIF ADP-ribosyltransferases (RIF-ARTs) and related PFAM family PF12120 sequences. Many mycobacterial species encode the equivalent of *M. smegmatis* RIF-ART (MSMEG_1221; Arr-1/ms in purple), and some mycobacterial species encode a previously unidentified sister group we have named Arr-X (dark orange). See Supplementary Table 3 for detailed information of the sequences in the tree. b, Ribbon representation of the crystal structure of *M. smegmatis* RIF-ART (PDB code 2HW2; light blue) with RIF bound (red) overlaid with the AlphaFold2 models of *M. conceptionense* Arr-X (dark blue) and *M. flavescens* Arr-X (dark orange). c, Apparent velocity of reaction of each of the enzymes (X axis) with different rifamycins as substrate. d, *M. conceptionense* single and double knockdown (KD) arr strains in the presence of rifabutin.

protein Arr-X. Arr-X enzymes are not only faster RIF-ARTs but also display higher specificity to rifabutin. The existence of a superior rifabutin-inactivating enzyme in several mycobacterial species might jeopardize the use of rifabutin, currently an antibiotic of choice to treat *M. avium* complex-caused infections and other infections. Novel inhibitors of these two distinct Arr enzymes (41, 42) might become essential to re-sensitize mycobacteria against rifamycins and against rifabutin.

Materials and Methods

Mycobacterial species and cultures

Mycobacterial species were acquired from the German Collection of Microorganisms and Cell Cultures GmbH – DSMZ (Braunschweig, Germany). The species comprised in our library were selected based on (i) broad genus coverage; (ii) diversity with respect to niche/pathogenicity; (iii) availability of genome sequence; (iv) availability of the type or laboratory strain; and (v) ability to grow on Middlebrook 7H9 culture medium. Upon arrival, long-term (-80 °C), short-term (-20 °C) and agar plate stocks were prepared according to DSMZ's protocols.

Mycobacterial diversity

Genotypic and phenotypic diversity is an essential feature required for our approach. Genome size and guanine-cytosine content (GC%) values were obtained from NCBI Microbial Genomes (43) and ribosomal copy number was annotated using rrnDB (44) or BLAST. When more than one hit was obtained, the genomic sequence and context were analyzed to confirm duplication. The growth rate of selected species was determined by turbidity measurements (OD600) taken in equal intervals. Species were selected based on experiments carried out throughout the study (e.g., correlating growth rate with antibiotic levels or requiring growth rate data for biomass matching in metabolomics experiments). We inoculated 100 mL Middlebrook 7H9 broth complete in roller bottles. Middlebrook 7H9 broth complete contained 10% Album in-Dextrose-Catalase (ADC) supplement, 0.05% Glycerol and 0.05 % Tyloxapol. For all species, cultures were incubated at 37°C, except for *M. marinum*, *M. abscessus*, *M. angelicum*, *M. conceptionense*, *M. fallax*, *M. salmoniphilum* and *M. tusciae* which were grown at 30°C. Growth rate was calculated using the specific growth rate formula (45). Finally, encoded genes were classified based on their gene ontology using OmicsBox's functional analysis tool, with default parameters (46, 47).

Minimal inhibitory concentration measurements

Our mid-throughput minimal inhibitory concentration assay (MIC_{99}) was adapted from previously described protocols (48). Briefly, in sterile Eppendorf tubes, each antibiotic was diluted in the recommended solvent to 4 mg/mL. Fluoroquinolones and BDQ were diluted to 0.4 mg/mL. Subsequently, 200 µL of each drug was added to column 10 of a 96-well plate. In the same plate, 100 µL of DMSO or water were added to columns 1-9 and 11. The drug titration was performed using a multi-channel pipette, using the volume 100 µL for 1:2 dilutions. The remaining 100 µL left over from column 1 was added to column 12, which is the negative control (contamination control). A secondary replicate plate was prepared in the same manner, and then combined for a final volume of 200 µL in each well. These plates were the master plates which were then copied into 36 new plates by transferring 5 µL of each well using Biomek FX^P Liquid Handling Automation (Beckman Coulter, California, USA). Once antibiotics were in the wells, 195 µL of Middlebrook 7H10 containing 10% Oleic Acid Albumin Dextrose Catalase (OADC) supplement was added each well and homogenized using Multidrop™ Combi Reagent Dispenser (Thermo Fischer Scientific, Massachusetts, USA). All procedures were performed in a biosafety cabinet.

Bacterial cultures were grown in Middlebrook 7H9 broth complete at their preferred temperature, shaking at 180 rpm. Once cultures reached approximately OD600 of 1, they were aliquoted in sterile micro tubes, and frozen at -20 °C for future use. For the MIC_{99} determination, stock cultures were diluted OD600 of 0.006 in Middlebrook 7H9 broth complete and 2 µL of the dilution were spotted into columns 1-11 of the 96-well plates. Plates were then incubated at the appropriate

temperature and analyzed after confluent growth was observed in the growth control wells (column 11). Pictures of the plates were taken, and visual analysis was carried out to record the MIC_{99}. At least three independent experiments were recorded for each species-antibiotic pair, for 15 antibiotics and 44 species, totaling 1,980 individual MIC determinations.

Sample preparation for metabolomics and proteomics

Sample preparation for LC-MS and proteomics was performed as previously described (49). Mycobacterial species were grown in roller bottles in a volume of 100 mL of Middlebrook 7H9 broth complete to an OD600 of 1. Cultures were filtered using MF-Millipore Membrane Filter – 0.22 µm pore size to concentrate cell amount. For each species, 24 bacterial-laden filters were prepared and 3 were placed in petri-dish plates of Middlebrook 7H10 containing 10% OADC supplement and incubated at appropriate temperature (37 or 30 °C, depending on the species) for 5 doubling times, to expand the bacterial biomass. Subsequently, filters were transferred to fresh 7H10 containing 10 % OADC supplement plates with vehicle or antibiotic at the concentration of 6 × MIC_{99} and incubated for one doubling time, at appropriate temperature. Cells were then scraped into screw-cap tubes containing either 1 mL of ACN:MeOH:Water (2:2:1, v/v/v) and glass beads (150 µm) for LC-MS, or washed twice with 1 mL of PBS and then placed in 4% SDS/100 mM HEPES/50 mM DTT lysis buffer and glass beads (150-212 µm), for proteomics pilot experiment and 1% SDC/100 mM HEPES/50 mM DTT lysis buffer for remaining proteomics experiments. All samples were lysed by bead beating. At this stage, LC-MS samples were centrifuged, and the supernatant was collected and filtered using Corning® Costar® Spin-X® Plastic Centrifuge 0.22 µm tube filters. Proteomics samples were heat killed and subjected to acetone precipitation, only when extracted with SDS, and peptide digested with LysC and Trypsin in 100 mM HEPES pH 8, Guanidine HCl 1 M.

LC-MS metabolomics

LC-MS metabolomic samples were analyzed using a previously described method (50). Briefly, aqueous normal phase liquid chromatography was performed using an Agilent 1200 LC system at where the samples were kept at controlled temperature (4 °C), A column Cogent Diamond Hydride Type C column linked to an Agilent Accurate Mass 6220 Time of Flight (TOF) spectrometer and coupled with an Agilent 1200 LC system was used for data acquisition. Flow rate of 0.4 ml min–1 was used and the elution of polar compounds was performed using a gradient of two solvents over a 24-minute separation, A (MS-grade water and 0.1% of formic acid) and B (acetonitrile and 0.1% of formic acid) in positive mode. The gradient of solvent B was defined as 0-2 minutes 85%, 3-5 minutes 80%, 6-7 minutes 70%, 8-9 minutes 70%, 10-11 minutes 50%, 11-14 minutes 20%, 1424 minutes 5% followed by a re-equilibration of 10 minutes at 85%. Using an isocratic pump, a reference mass solution was infused with the run to allow for simultaneous mass axis calibration. The data were analyzed by MassHunter Qualitative Analysis B07.00 or XCMS (51). To verify the intracellular drug concentration for all samples, the molecular formulae of each drug was searched against the raw spectra and the integrated Extracted Ion Chromatogram (EIC) was used to extract quantitative information. The peak height of both the antibiotic standard and the samples was used to calculate the relative drug concentration in each sample. Although area would generally be the variable of choice for performing quantifications, the peak shape was not consistent for LZD, therefore the peak height would correlate better with the amount of drug inside the cell. Further, the amount of drug added to the cells was considered and used to perform normalization, given that each species was exposed to a concentration of drug proportional to their MIC_{99}. For identifying and quantifying Ribosyl-RIF, XCMS was used to perform peak picking and alignment. Metaboanalyst was used to perform batch effect corrections and statistical analysis (52). A feature with m/z 955.4601 was identified at 5.8 minutes and confirmed to be only present in RIF-treated samples. The feature abundance was normalized by the abundance in the pooled biological quality control (PBQC) samples (shown in Supplementary Fig. 3). This accurate mass was searched using MassHunter Qualitative Analysis B07.00 with an acceptable error of ±10 ppm and the obtained peaks were integrated to extract the peak area, which was normalized by the peak area of the PBQC (shown in Fig. 4). Finally, the average and the standard deviation between normalized values from six biological replicates were calculated and plotted.

Proteomics

Proteomics analyses were performed at the Proteomics Scientific Technology Platform (STP) at The Francis Crick Institute. Data Dependent Acquisition (DDA) was used to build a peptide library and Data Independent Acquisition (DIA) was used to analyze the experimental samples. For both DDA and DIA, Evosep LC system (Evosep) was employed with the standard gradient for a total LC runtime of 44 min, using their supplied 15 cm column (53). Each sample was loaded from the peptide digests at a minimum volume of 10 µL (samples diluted for the optimum load for final loading volume of 10 µL to ensure all liquid enters the tip). An aliquot of the recommended amount of iRT peptides (Biognosys AG, Switzerland) was added to each sample at the sample loading stage. The protocol supplied with the Evotips was followed for conditioning/equilibrating/loading and washing the tips.

The outlet of the analytical column was connected directly to an adapter that allowed the EasySpray nano-source to be employed on the Orbitrap Fusion Lumos (Thermo Fischer Scientific, USA) using a stainless-steel emitter. The spray voltage was set to 2.2kV. The default charge state was set to 2+. For the DDA runs, MSI data were acquired in profile mode at a resolution of 60000 (FWHM), with an AGC target of 1E6 ions and a maximum injection time of 50ms. The ion funnel RF was set at 30%. Quadrupole isolation was employed over the MS1 mass range of 375-1200 m/z. The monoisotopic precursor selection (MIPS) was set to “Peptide” and an intensity threshold of 5E4 was applied. Charge states from 2+ to 6+ were considered for MS/MS and dynamic exclusion was set to 15 seconds/10 ppm including exclusion of isotopes. Cycle time for the Data Dependent MS/MS Acquisition was set to 1 second. For the MS/MS, quadrupole isolation was set to 1.4 Da and HCD collision energy was employed at 32%. Data were acquired in the Orbitrap at a resolution of 15000 (FWHM) in centroid, with a fixed first mass of 120 m/z. The AGC was set at 1E6 and maximum injection time of 22 ms.

For the DIA data acquisition, the following parameters were adjusted. The default charge state was set to 4+, and MS1 data were acquired in profile mode at a resolution of 120,000 (FWHM) with an AGC setting of 1E6 and maximum injection time of 20ms. The MS1 scan range was set from 393-907 m/z to allow enough data points per peak. 27 DIA windows (20 Da / 1 Da overlap) were employed over this range for DIA MS2 acquisition. These data were acquired at 30000 resolution (FWHM) in the Orbitrap in centroid mode. HCD collision energy was the same as for the DDA runs. MS2 data were acquired over the mass range 200-2000 m/z with an AGC setting of 1E6 and a maximum injection time of 54 ms. Ions were injected for all available parallelizable time.

Gene knockdown using CRISPRi

Gene silencing of *M. conceptionense arr-1* and *arr*-X and both genes was performed following the protocols for *M. smegmatis* described in Wong and Rock (54) with the primers listed in Supplementary Table 4.

Electrocompetent *M. conceptionense* cells were transformed with one of the three silencing constructs and with the empty vector pLJR962 as negative control.

Transformants were selected by plating cells into Middlebrook 7H9 broth complete and 20 µg/mL of kanamycin.

Individual colonies were picked and inoculated into 10 mL of Middlebrook 7H9 broth complete and 20 µg/mL of kanamycin and glycerol stocks were prepared for subsequent experiments. To test for antibiotic sensitivity of knockdown constructs, glycerol stocks were used to inoculate 5 mL of Middlebrook 7H9 broth complete with 20 µg/mL of kanamycin and incubated at 37°C with shaking to an OD600 of ca. 1. These saturated cultures were then used to inoculate fresh 5 mL aliquots of medium with kanamycin to an OD600 of 0.05 and the cultures were incubated at 37 °C with shaking to an OD600 of 0.4-0.8. The cultures were then diluted to an OD 0.1 in Middlebrook 7H9 broth complete with 20 µg/mL of kanamycin and 100 µg/mL of anhydrotetracycline (ATc) and grown at 37°C with shaking to an OD600 of 0.4-0.8. This step was repeated a second time. Cultures were then streaked into Middlebrook 7H10 plates containing 10% OADC, 0.05% glycerol, 20 µg/mL of kanamycin, 200 µg/mL of ATc and rifabutin at 0.5 ×, 1× and 2 × the MIC99.

Cloning, expression and purification of Arr enzymes

The RIF ADP-ribosyl transferase (arr) genes from *M. smegmatis*, *M. flavescens* and *M. conceptionense* were cloned from genomic DNA by PCR and inserted via isothermal assembly into the pNIC-CTHF expression vector (a gift from Opher Gileadi; Addgene plasmid #26105) (55). The resulting plasmids were Sanger sequenced to confirm the correct insertion of the genes and then transformed into *E. coli* BL21(DE3) Gold competent cells (Agilent Technologies).

The transformed cells with each of the arr genes were grown at 37 °C and 200 rpm in 1 L of lysogeny broth (LB) supplemented with 50 µg/mL kanamycin to an OD600 of 0.6, at this point the temperature was dropped to 16 °C. Protein expression was induced by addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM, and cells allowed to grow for 20 hours. Cells were harvested by centrifugation at 4000 *g* for 30 min and stored at −80 °C.

The purification was performed according to the previous reported protocols (37). In summary, cells were thawed on ice for 30 min before being resuspended in buffer A (50 mM HEPES (pH 7.5), 1 mM EDTA) containing a tablet of cOmplete® EDTA-free protease inhibitor cocktail (Roche), 2 µL Benzonase nuclease (Millipore), and 6 mM MgCl2. Samples were further lysed by probe sonication (amplitude 35%, on 10 s, off 50 s, 2 min total on time per cycle, 2–3 × cycles) and centrifuged at 48000 g for 45 min to separate the cell debris. The supernatant was filtered through a 0.45-µm membrane and loaded onto a 20 mL HiPrep Q Sepharose Column (GE Healthcare), pre-equilibrated with buffer A. The column was washed with 5 column volumes (CV) of 5 % buffer B (50 mM HEPES (pH 7.5), 1 mM EDTA, 1 M NaCl), and the adsorbed proteins were eluted with 5 CV of a linear gradient from 10 to 30 % buffer B. Fractions containing the desired Arr protein were pooled together and brought to 1.25 M (NH4)2SO4 by dropwise addition of the ammonium sulfate solution while stirring. After 30 min, the sample was filtered through a 0.45-µm membrane and loaded onto a 1 mL HiTrap Phenyl Sepharose column (GE Healthcare) pre-equilibrated with buffer C (50 mM sodium phosphate (pH 7.0), 1.25 M (NH4)2SO4). The column was washed with 15 column volumes (CV) of 10 % buffer D (50 mM sodium phosphate (pH 7)), and the adsorbed proteins were eluted with 20 CV of a linear gradient from 10 to 90 % buffer D. Fractions were analyzed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (NuPAGE Bis-Tris 4-12% Precast gels, Thermo Fisher Scientific), pooled, dialyzed against 2 × 2 L of 20 mM HEPES (pH 8.0), concentrated using 5000-molecular-weight-cut-off (MWCO) centrifugal ultrafiltration membranes (Millipore), aliquoted, and stored at -80 °C. The concentration was determined spectrophotometrically (NanoDrop, Thermo Fisher Scientific) at 280 nm using a theoretical extinction coefficient of 16960 M-1 cm-1. (ExPASy's ProtParam, (56)

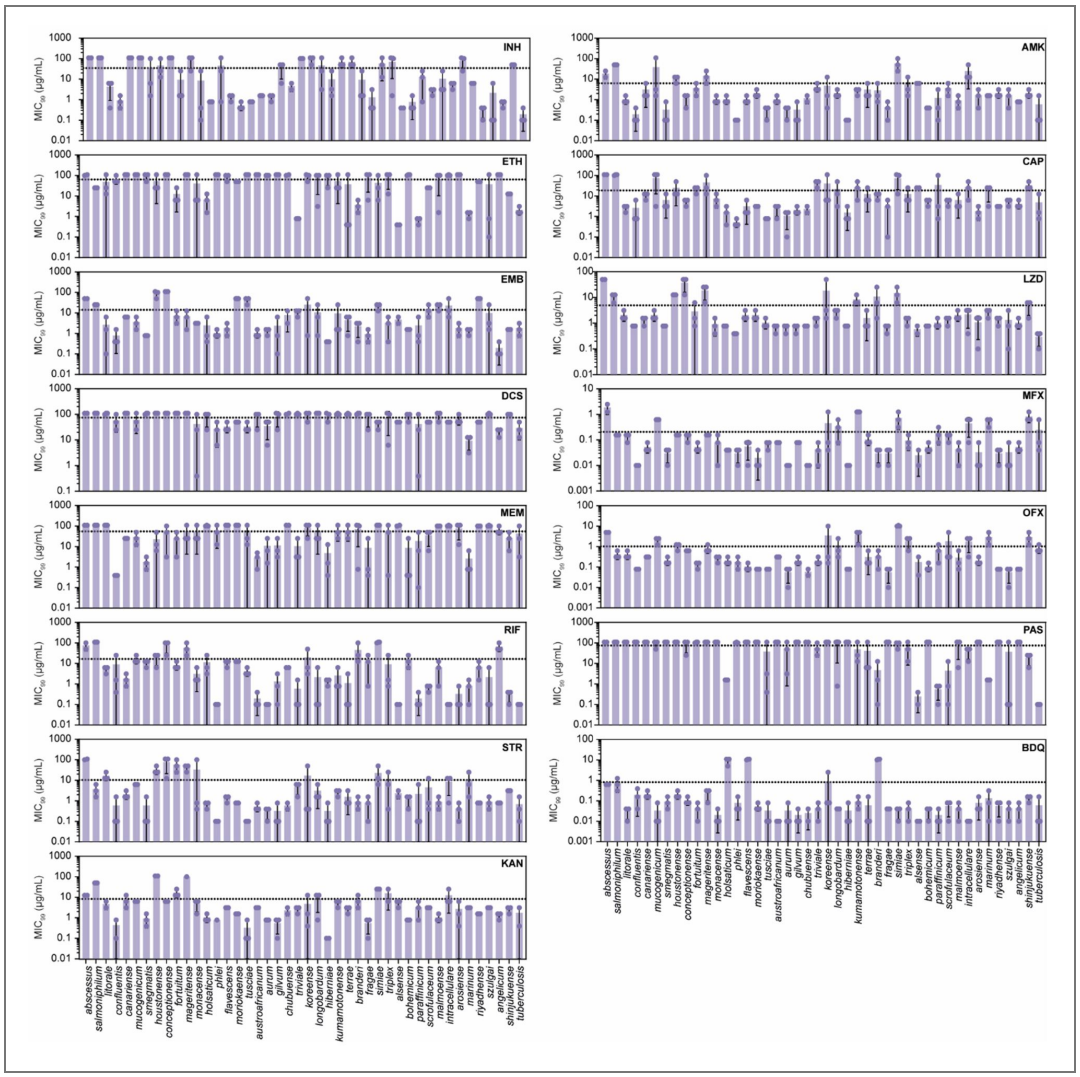
In vitro activity assay of Arr enzymes

Assays were carried out based on the previous reported methods (37). Briefly, 150 µM enzyme (except for *M. conceptionense* Arr-X where 50 µM was used) in 50 mM HEPES buffer (pH 7.5) was mixed with 150 pM of rifampicin and 2 mM NAD+. Time points (50 µL) where quenched by the addition of methanol (200 µL). The samples were then analyzed by injecting 20 µL onto the HPLC column Poroshell 120 Å, EC-C18, 3.0 × 150 mm, 2.7 µm (Agilent Technologies) and monitoring the consumption of rifampicin and formation of ADP-ribosylated rifampicin product. The same protocol was followed for each of the rifamycins in study: rifabutin, rifapentine, rifaximin, rifamycin B and rifamycin S. A standard curve of each rifamycin were measured independently.

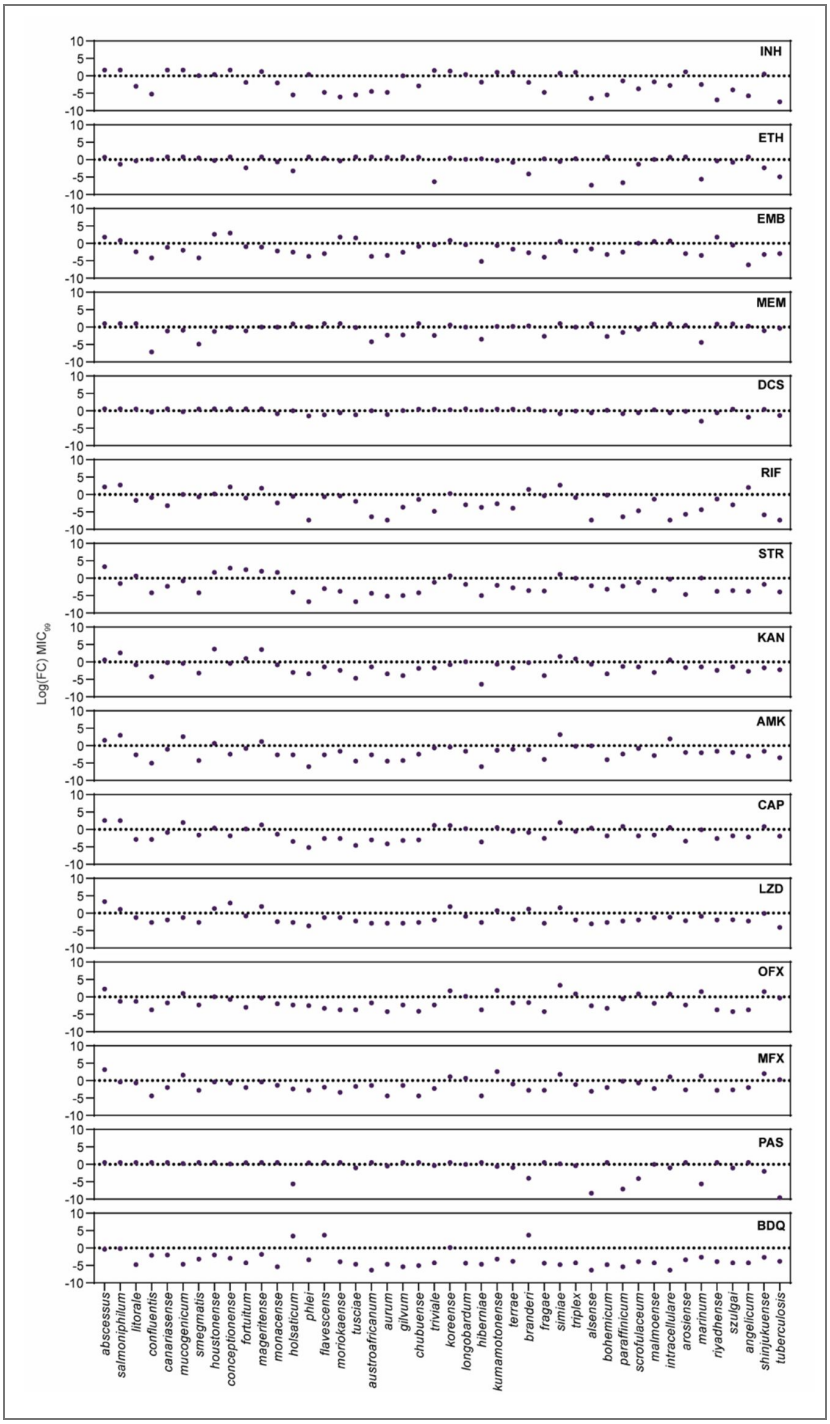
Arr distribution analysis

In order to build a phylogenetic tree of mycobacterial Arr enzymes, Arr-ms (Uniprot A0QRS5) was used to query a local BLAST (57) database of genomic coding sequences of the mycobacterial reference genomes available in NCBI (43). The matching sequences were combined with those obtained by searching UniProt (58) entries matching the RIF-ART protein family (Pfam PF12120 (59)). CD-Hit (56, 60) and Jalview (61) were used to reduce redundancy and a multiple protein sequence alignment was calculated using MUSCLE (62). Trees were generated with IQ-Tree 1.6.11 (63, 64) with 1000 ultrafast bootstrap replicates.

Supplementary Figures

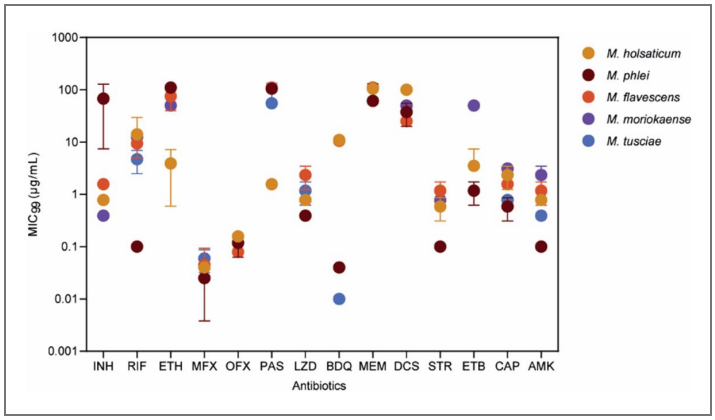


Supplementary Figure 1. Detailed overall MIC_{99} quantifications against the mycobacterial species in our library. X axes represent the species tested in phylogenetic order, and Y axes are the MIC_{99} measurements in log_{10} scale. Dotted lines represent the mean MIC_{99} for each drug.

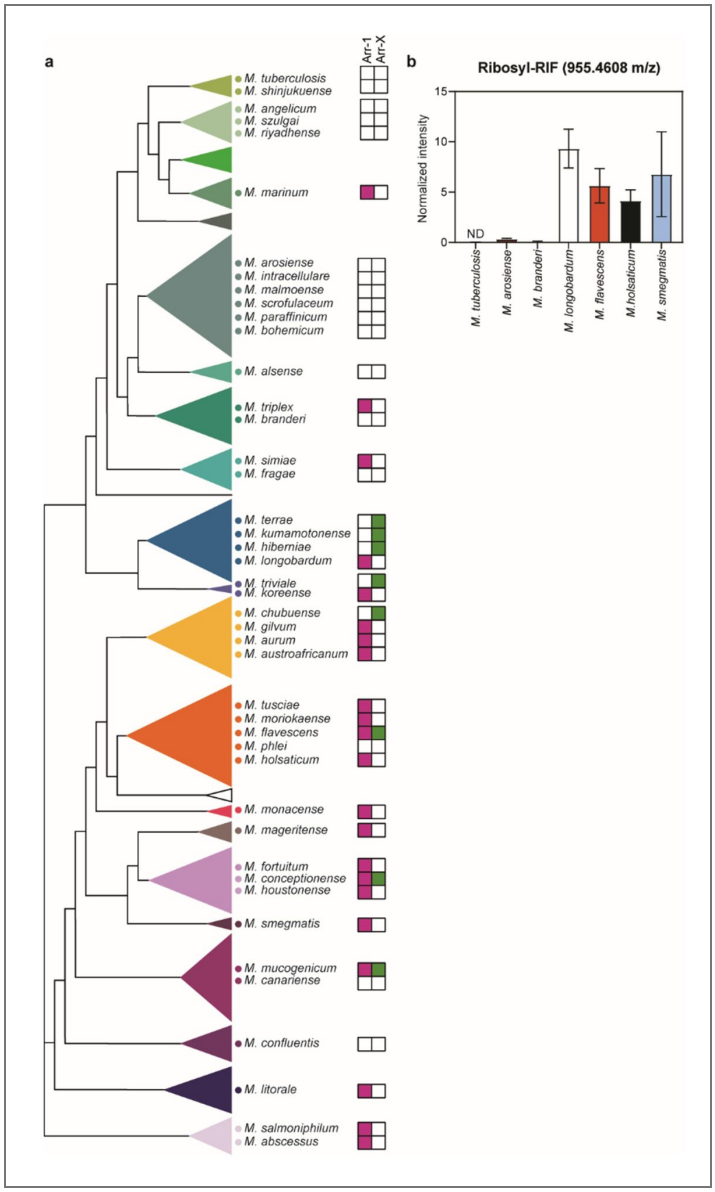


Supplementary Figure 2. **Log2(FC) MIC_{99} (Y axis) compared to the mean MIC_{99} for each antibiotic.**

Species organized by phylogeny (X axis). Dotted line represents the mean MIC_{99} for each antibiotic.



Supplementary Figure 3. **Detailed view of the *Mycobacterium holsaticum* group MIC_{99} values**



Supplementary Figure 4.

a, Phylogenetic distribution of RIF ADP-ribosyl transferases present in the mycobacterial species in our library. **b**, Quantification of ribosyl-RIF in some mycobacterial species that encode Arr-1 and/or Arr-X. Not detected (ND).

Data availability

Microbiological and biochemical data is available in the manuscript and supplementary information. Metabolomics and proteomics data are available upon request.

Additional information

Funding

This work was funded by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001060), the UK Medical Research Council (FC001060), and the Wellcome Trust (FC001060). This work was partially funded by a Wellcome Trust New Investigator Award (104785/B/14/Z) to LPSC.

Author contributions

Conceptualization: FTS, AGG, LPSC

Methodology: FTS, TFGM, AGG, JMK

Investigation: FTS, TFGM, AGG, HD, JMK

Supervision: MS, AGG, LPSC

Writing—FTS, TFGM, HD, JMK, AGG, LPSC

Funding

Funder	Grant reference number	Author
The Francis Crick Institute	FC001060	Holly Douglas Teresa FG Machado Joanna M Kirkpatrick Acely Garza-Garcia Fernanda T Subtil Mark Skehel Luiz Pedro S de Carvalho
Wellcome	https://doi.org/10.35802/104785	Luiz Pedro S de Carvalho Holly Douglas

Author ORCID iDs

Acely Garza-Garcia: https://orcid.org/0000-0003-0307-0138
Luiz Pedro S de Carvalho: https://orcid.org/0000-0003-2875-4552

Additional files

Supplementary Table 1

Supplementary Table 2

Supplementary Table 3

Supplementary Table 4

References

1. **WHO** (2015) Global action plan on antimicrobial resistance. Geneva, Switzerland: WHO.
2. **Ventola C.L.** (2015) The antibiotic resistance crisis: part 1: causes and threats. *P T* **40**:277-83
3. **Cox G.,** Wright G.D. (2013) Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *Int J Med Microbiol* **303**:287-92
4. **Dantas G.,** Sommer M.O. (2012) Context matters - the complex interplay between resistome genotypes and resistance phenotypes. *Curr Opin Microbiol* **15**:577-82
5. **Finley R.L.,** et al. (2013) The scourge of antibiotic resistance: the important role of the environment. *Clin Infect Dis* **57**:704-10
6. **Andries K.,** et al. (2014) Acquired resistance of Mycobacterium tuberculosis to bedaquiline. *PLoS One* **9**:e102135
7. **Farhat M.R.,** et al. (2019) Rifampicin and rifabutin resistance in 1003 Mycobacterium tuberculosis clinical isolates. *J Antimicrob Chemother* **74**:1477-1483
8. **Mathys V.,** et al. (2009) Molecular genetics of para-aminosalicylic acid resistance in clinical isolates and spontaneous mutants of Mycobacterium tuberculosis. *Antimicrob Agents Chemother* **53**:2100-9
9. **D'Costa V.M.,** et al. (2011) Antibiotic resistance is ancient. *Nature* **477**:457-61
10. **Djouadi L.N.,** et al. (2016) Mycobacterium icosiumassiliensis sp. nov., a New Member in the Mycobacterium terrae Complex Isolated from Surface Water in Algeria. *Curr Microbiol* **73**:255-64
11. **Hennessee C.T.,** et al. (2009) Polycyclic aromatic hydrocarbon-degrading species isolated from Hawaiian soils: Mycobacterium crocinum sp. nov., Mycobacterium pallens sp. nov., Mycobacterium rutilum sp. nov., Mycobacterium rufum sp. nov. and Mycobacterium aromaticivorans sp. nov. *Int J Syst Evol Microbiol* **59**:378-87
12. **Bouam A.,** et al. (2018) Mycobacterium terramassiliense, Mycobacterium rhizamassiliense and Mycobacterium numidiamassiliense sp. nov., three new Mycobacterium simiae complex species cultured from plant roots. *Sci Rep* **8**:9309
13. **Kazda J.,** et al. (1993) Mycobacterium hiberniae sp. nov. *Int J Syst Bacteriol* **43**:352-7
14. **Bergey D.,** et al. (1923) *Bergey's Manual of Determinative Bacteriology* (1st Edition) Baltimore: The Williams & Wilkins Co..
15. **Schwabacher H.** (1959) A strain of Mycobacterium isolated from skin lesions of a cold-blooded animal, Xenopus laevis, and its relation to atypical acid-fast bacilli occurring in man. *J Hyg (Lond)* **57**:57-67
16. **Pourahmad F.,** et al. (2015) Mycobacterium angelicum sp. nov., a non-chromogenic, slow-growing species isolated from fish and related to Mycobacterium szulgai. *Int J Syst Evol Microbiol* **65**:4724-4729
17. **Bojalil L.F.,** Cerbon J., Trujillo A. (1962) Adansonian classification of mycobacteria. *J Gen Microbiol* **28**:333-46
18. **Magee J.G.,** Ward A.C. (2012) Family III. Mycobacteriaceae. In: Goodfellow M., et al. (Eds). *Bergey's Manual of Systematic Bacteriology* New York: Springer. pp. 312-375
19. **Devulder G.,** de Montclos M.P., Flandrois J.P. (2005) A multigene approach to phylogenetic analysis using the genus Mycobacterium as a model. *International Journal of Systematic and Evolutionary Microbiology* **55**:293-302
20. **Griffith D.E.,** et al. (2007) An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. *Am J Respir Crit Care Med* **175**:367-416
21. **Johansen M.D.,** Herrmann J.L., Kremer L. (2020) Non-tuberculous mycobacteria and the rise of Mycobacterium abscessus. *Nat Rev Microbiol* **18**:392-407

22. **Abdulkadir N.,** et al. (2024) Genome-centric analyses of 165 metagenomes show that mobile genetic elements are crucial for the transmission of antimicrobial resistance genes to pathogens in activated sludge and wastewater. *Microbiol Spectr* e0291823
23. **Zhang J.,** et al. (2022) Metagenomics insights into the profiles of antibiotic resistome in combined sewage overflows from reads to metagenome assembly genomes. *J Hazard Mater* **429**:128277
24. **Valdivia-Anistro J.A.,** et al. (2015) Variability of rRNA Operon Copy Number and Growth Rate Dynamics of Bacillus Isolated from an Extremely Oligotrophic Aquatic Ecosystem. *Front Microbiol* **6**:1486
25. **Klappenbach J.A.,** Dunbar J.M., Schmidt T.M. (2000) rRNA operon copy number reflects ecological strategies of bacteria. *Appl Environ Microbiol* **66**:132833
26. **Bachmann N.L.,** et al. (2020) Key Transitions in the Evolution of Rapid and Slow Growing Mycobacteria Identified by Comparative Genomics. *Frontiers in Microbiology* **10**
27. **Luthra S.,** Rominski A., Sander P. (2018) The Role of Antibiotic-Target-Modifying and Antibiotic-Modifying Enzymes in Mycobacterium abscessus Drug Resistance. *Front Microbiol* **9**:2179
28. **Jayasingam S.D.,** Zin T., Ngeow Y.F. (2017) Antibiotic resistance in Mycobacterium Abscessus and Mycobacterium Fortuitum isolates from Malaysian patients. *Int J Mycobacteriol* **6**:387-390
29. **Vesenbeckh S.,** et al. (2017) Bedaquiline as a potential agent in the treatment of Mycobacterium abscessus infections. *Eur Respir J* **49**
30. **Yu J.F.,** et al. (2022) Decreased Methylenetetrahydrofolate Reductase Activity Leads to Increased Sensitivity to para-Aminosalicylic Acid in Mycobacterium tuberculosis. *Antimicrob Agents Chemother* **66**:e0146521
31. **Zampieri M.,** et al. (2018) High-throughput metabolomic analysis predicts mode of action of uncharacterized antimicrobial compounds. *Sci Transl Med* **10**
32. **Takaki K.,** et al. (2012) An in vivo platform for rapid high-throughput antitubercular drug discovery. *Cell Rep* **2**:175-84
33. **Lawn S.D.,** Nicol M.P. (2011) Xpert(R) MTB/RIF assay: development, evaluation and implementation of a new rapid molecular diagnostic for tuberculosis and rifampicin resistance. *Future Microbiol* **6**:1067-82
34. **Fahat M.R,** et al. (2019) Rifampicin and rifabutin resistance in 1003 Mycobacterium tuberculosis clinical isolates. *J Antimicrob Chemother* **74**:1477-83
35. **Spanogiannopoulos P.,** et al. (2014) A rifamycin inactivating phosphotransferase family shared by environmental and pathogenic bacteria. *Proc Natl Acad Sci U S A* **111**:7102-7
36. **Spanogiannopoulos P.,** et al. (2012) Characterization of a rifampin-inactivating glycosyltransferase from a screen of environmental actinomycetes. *Antimicrob Agents Chemother* **56**:5061-9
37. **Koteva K.,** et al. (2018) Rox, a Rifamycin Resistance Enzyme with an Unprecedented Mechanism of Action. *Cell Chem Biol* **25**:403-412 e5
38. **Baysarowich J.,** et al. (2008) Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr. *Proc Natl Acad Sci U S A* **105**:4886-91
39. **Imai T.,** et al. (1999) Identification and characterization of a new intermediate in the ribosylative inactivation pathway of rifampin by Mycobacterium smegmatis. *Microb Drug Resist* **5**:259-64
40. **Dabbs E.R.,** et al. (1995) Ribosylation by mycobacterial strains as a new mechanism of rifampin inactivation. *Antimicrob Agents Chemother* **39**:1007-9
41. **Zheng M.,** Lupoli T.J. (2021) Modulation of a Mycobacterial ADP-Ribosyltransferase to Augment Rifamycin Antibiotic Resistance. *ACS Infect Dis* **7**:2604-2611
42. **Ganapathy Uday S.,** et al. (2023) Blocking ADP-ribosylation expands the anti-mycobacterial spectrum of rifamycins. *Microbiology Spectrum* **11**:e01900-23
43. **Durão P.,** Balbontín R., Gordo I. (2018) Evolutionary Mechanisms Shaping the Maintenance of Antibiotic Resistance. *Trends Microbiol* **26**:677-691

44. **(MD) B.** (no date) *Genome (Internet) 2004 - 2023, National Library of Medicine (US)* National Center for Biotechnology Information.
45. **Stoddard S.F.,** et al. (2015) rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res* **43**:D593-8
46. **Maier R.M.,** Pepper J.L., Gerba C.P. (2000) *Environmental microbiology* San Diego, Calif. ; London: Academic.
47. (2019) OmicsBox - Bioinformatics Made Easy. BioBam Bioinformatics.
48. **Gotz S.,** et al. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* **36**:3420-35
49. **Guzman J.D.,** et al. (2013) Antitubercular specific activity of ibuprofen and the other 2-arylpropanoic acids using the HT-SPOTi whole-cell phenotypic assay. *BMJ Open* **3**
50. **Nandakumar M.,** et al. (2015) Metabolomics of Mycobacterium tuberculosis. *Methods Mol Biol* **1285**:105-15
51. **Serafini A.,** et al. (2019) Mycobacterium tuberculosis requires glyoxylate shunt and reverse methylcitrate cycle for lactate and pyruvate metabolism. *Mol Microbiol* **112**:1284-1307
52. **Smith C.A.,** et al. (2006) XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Analytical Chemistry* **78**:779-787
53. **Pang Z.,** et al. (2021) MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res* **49**:W388-W396
54. **Krieger J.R.,** et al. (2019) Evosep One Enables Robust Deep Proteome Coverage Using Tandem Mass Tags while Significantly Reducing Instrument Time. *J Proteome Res* **18**:2346-2353
55. **Wong A.I.,** Rock J.M. (2021) CRISPR Interference (CRISPRi) for Targeted Gene Silencing in Mycobacteria. *Methods Mol Biol* **2314**:343-364
56. **Savitsky P.,** et al. (2010) High-throughput production of human proteins for crystallization: the SGC experience. *J Struct Biol* **172**:3-13
57. **Gasteiger E.,** et al. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: Walker J.M. (Ed). *The Proteomics Protocols Handbook* Totowa, NJ: Humana Press. pp. 571-607
58. **Camacho C.,** et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**:421
59. **Consortium T.U.** (2022) UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* **51**:D523-D531
60. **Paysan-Lafosse T.,** et al. (2023) InterPro in 2022. *Nucleic acids research* **51**:D418-D427
61. **Li W.,** Godzik A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**:1658-1659
62. **Waterhouse A.M.,** et al. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**:1189-91
63. **Edgar R.C.** (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:1792-1797
64. **Kalyaanamoorthy S.,** et al. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**:587-589
65. **Nguyen L.-T.,** et al. (2014) IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution* **32**:268-274
66. **Ankenbrand M.J.,** Keller A. (2016) bcgTree: automatized phylogenetic tree building from bacterial core genomes. *Genome* **59**:783-791
67. **Eddy S.R.** (2009) A new generation of homology search tools based on probabilistic inference. *Genome Inform* **23**:205-211
68. **Castresana J.** (2000) Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution* **17**:540-552

69. **Stamatakis A.** (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**:1312-1313

70. **Willand N. Dirié,** Carette B., Bifani X., Singhal P., Desroses A., Leroux M., Willery F., Mathys E., Déprez-Poulain V., Delcroix R., *et al.* (2009) A.R. Synthetic EthR inhibitors boost antituberculous activity of ethionamide. *Nature Medicine* **15**:537-44

71. **Gutierrez M.C. Brisse,** Brosch S., Fabre R., Omaïs M., Marmiesse B., Supply P M., Vincent V. (2005) Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. *PLoS Pathogens* **1**:e5

72. **Jumper J.,** et al. (2021) Highly accurate protein structure Suprediction with AlphaFold. *Nature* **596**:583-589

Peer reviews

Reviewer #1 (Public review):

This work analyzes innate resistance to drugs in mycobacteria by comparing minimum inhibitory concentrations (MICs) across a diverse panel of mycobacterial species. The results show that MICs are poorly correlated with growth rate while phylogeny associated with horizontal gene transfer underlies the observed differences in MIC, an important demonstration. A further investigation into the driver for the vast differences in susceptibility profiles shows that for three drugs the MIC is not correlated with intrabacterial drug concentrations where intrabacterial drug concentration is comprised of cytosolic and cell wall associated drug. This is a striking observation. The authors delve into the mechanisms that drive resistance to rifamycins and confirm that resistance is driven by ADP-ribosyltransferases of which two variant groups exist, one of which is kinetically faster and apparently is superior at modifying more hydrophobic rifamycins. The relative role of the two ADP-ribosyltransferases in conferring resistance especially in the species with both orthologs is not fully understood since the modified drug can possibly be further modified and transcriptional downregulation experiments performed in this work do not provide genetic evidence of perturbation of mRNA levels of the respective open reading frames.

Comments on revisions:

Demonstration of the level of transcriptional downregulation of the two Arr orthologs would have been a nice demonstration of (1) the utility of CRISPRi in other mycobacteria, (2) that the difference in rifabutin susceptibility during knockdown of Arr-1 vs Arr-X can fully be ascribed to the role of Arr-X in modifying the drug.

https://doi.org/10.7554/eLife.101940.2.sa2

Reviewer #3 (Public review):

This manuscript presents a macroevolutionary approach to identification of novel high-level antibiotic resistance determinants that takes advantage of the natural genetic diversity within a genus (mycobacteria, in this case) by comparing antibiotic resistance profiles across related bacterial species and then using computational, molecular, and cellular approaches to identify and characterize the distinguishing mechanisms of resistance. The approach is contrasted with "microevolutionary" approaches based on comparing resistant and susceptible strains of the same species and approaches based on ecological sampling that may not include clinically relevant pathogens or related species. The potential for new discoveries with the macroevolution-inspired approach is evident in the diversity of drug susceptibility profiles revealed amongst the selected mycobacterial species and the identification and characterization of a new group of rifamycin-modifying ADP-ribosyltransferase (Arr) orthologs of previously described mycobacterial Arr enzymes.

Additional findings that intra-bacterial antibiotic accumulation does not always predict potency within this genus, that M. marinum is a better proxy for M. tuberculosis drug susceptibility than the commonly used saprophyte M. smegmatis, and that susceptibility to semi-synthetic antibiotic classes is generally less variable than susceptibility to antibiotics more directly derived from natural products strengthen the claim that the macroevolutionary lens is valuable for elucidating general principles of susceptibility within a genus.

There are some limitations to the work. The argument for the novelty of the approach could be better articulated. While the opportunities for new discoveries presented by identification of discrepant susceptibility results between related species is evident, it is less clear how the macroevolutionary approach is further leveraged for the discovery of truly novel resistance mechanisms. The example of the discovery of Arr-X enzymes presented here relied upon foundational knowledge of previously characterized Arr orthologs. There is less clarity about what the pipeline would look like for discovery of previously unknown determinants when one is agnostic to putative mechanisms. From the point at which interspecies differences in susceptibility are noted, does the framework still remain distinct from other discovery frameworks and approaches?

While the experimentation and analyses performed are generally well designed and rigorous, there are a few instances in which broad claims are based on inferences from sample sets or data sets that are, at present, too limited to provide robust support. For example, the claim that rifampicin modification, and precisely ADP-ribosylation, is the dominant mechanism of resistance to rifampicin in mycobacteria is still a bit premature or at least an over-generalization, as other enzymatic modification mechanisms and other mechanisms such as helR-mediated dissociation of rifampicin-stalled RNA polymerases, efflux, etc were not examined. CRISPR interference was used in a demonstrative example to support this assertion, but would need to be applied more systematically to be more conclusive. The general claim that intra-bacterial antibiotic accumulation does not predict potency in mycobacteria may be another over-generalization based on the limited set of drugs and species studied.

Comments on revisions:

Discussion, lines 321-323: "We found that resistance to these antibiotics in mycobacteria do not correlate with by uptake/efflux mechanisms in the species tested..." is an over-generalization and conflicts with the following statement on lines 199-201: "for BDQ we could observe some correlation between antibiotic potency and [BDQ]IB which could be indicative of efflux playing a role in antibiotic efficacy. Given that the current statement in the Discussion only applies to 2 of 3 drugs tested, a more specific or nuanced interpretation seems warranted.

https://doi.org/10.7554/eLife.101940.2.sa1

Author response:

The following is the authors' response to the original reviews

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***
>
> *This work shows that resistance profiles to a variety of drugs are variable between different mycobacterial species and are not correlated with growth rate or intrabacterial compound concentration (at least for linezolid, bedaquiline, and Rifampicin). Note that intrabacterial compound concentration does not distinguish between cytosolic and*

periplasmic/cell wall-associated drugs. The susceptibility profiles for a wide range of mycobacteria tested under the same conditions against 15 commonly used antimycobacterial drugs provide the first recorded cross-species comparison which will be a valuable resource for the scientific community. To understand the reasons for the high Rifampicin resistance seen in many mycobacteria, the authors confirm the presence of the arr gene known to encode a Rif ribosyltransferase involved in Rif resistance in M. smegmatis in the resistant mycobacteria after confirming the absence of on-target mutations in the RpoB RRDR. Metabolomic analyses confirm the presence of ribosylated Rif in some of the naturally resistant mycobacteria which may not be entirely surprising but an important confirmation. Presumably M. branderi is highly resistant despite lacking the arr homolog due to the rpoB S45N mutation. M. flavescens has an MIC similar to that of M. smegmatis, despite having both Arr-1 and Arr-X. Various Arr-1 and Arr-X proteins are expressed and characterized for catalytic activity which shows that Arr-X is a faster enzyme,, especially with respect to more hydrophobic rifamycins. M. flavescens has similar MIC values to Rifapentine and Rifabutin to M. smegmatis. Thus, the Arr-1 versus Arr-X comparison does not provide a complete explanation for the underlying reasons driving natural Rif resistance in mycobacteria. Downregulation of Arr-X expression in M. conceptionense confers increased sensitivity to Rifabutin confirming its role as a rifamycin-inactivating enzyme.

Overall, the comparison of cross-species susceptibility profiles is novel; the demonstration that MIC is not correlated with intracellular drug concentration is important but not sufficiently interrogated, the demonstration that Arr-X is also a Rif ADP-ribosyltransferase is a good confirmation and shows that it is more efficient than Arr-1 on hydrophobic rifamycins is interesting but maybe not entirely surprising. The manuscript seems to have two parts that are related, but the rifamycin modification aspect of the work is not strongly linked to the first part since it interrogates the modification of one drug but not the common cause of natural resistance for other drugs.

Reviewer #2 (Public review):

Summary:

The authors use a variety of methods to investigate the mechanisms of innate drug resistance in mycobacteria. They end up focusing on two primary determinants - drug accumulation, which correlates rather poorly with resistance for many species, and, for the rifamycins, ADP-ribosyltransferases. The latter enzymes do appear to account for a good deal of resistance, though it is difficult to extrapolate quantitatively what their relative contributions are.

Overall, they make excellent use of biochemical methods to support their conclusions. Though they set out to draw very broad lessons, much of the focus ends up being on rifamycins. This is still a very interesting set of conclusions.

Strengths:

(1) A very interesting approach and set of questions.

(2) Outstanding technical approaches to measuring intracellular drug concentrations and chemical modification of rifamycins.

(3) Excellent characterization of variant rifamycin ADP-ribosyltransferases

Weaknesses:

> *(1) Figure 3c/d: These panels show the same experiment done twice, yet they display substantially different results in certain cases. For instance, M. smegmatis appears to show an order of magnitude lower RIF accumulation in panel d compared to M. flavescens, despite them displaying equal accumulation in panel c. The authors should provide justification for this variation, particularly as quantitative intra-species comparisons are central to the conclusions of this figure.*

The data in panels 3c and 3d are from different sets of experiments. The reviewer is correct with regards to *M. smegmatis*. The data indeed is ~ 1 order of magnitude different. However, the data for other species is very similar. The reviewer may also have noticed that the error bars are also larger in 3d, compared to 3c, indicating a greater variation between independent experiments use in 3d. We do not have a good explanation for this, other than the experiments shown in 3d were associated with greater biological variability.

> *(2) There are several technical concerns with Figure 3 that affect how to interpret the work. According to the methods, the authors did not appear to normalize to an internal standard, only to an external antibiotic standard (which may account for some of the technical variation alluded to above).*

We agree that using a labeled drug as an internal standard (IS) would be ideal. However, the experiment initially followed an untargeted metabolomics approach, which later shifted to relative drug quantification. At that stage, normalizing with IS was impractical because proper implementation would require multiple IS across the chromatographic range. Therefore, we opted for total ion current (TIC) normalization, which accounts for variability in overall metabolite abundance—even though the experimental setup was already adjusted for each bacterial species' growth rate. Additionally, we prepared external standard curves for each drug to enable quantification, and the amount of drug added to each plate was considered when reporting these values.

> *Second, the authors used different concentrations of drug for each species to try to match the species' MICs. I appreciate the authors' thinking on this, but I think for an uptake experiment it would be more appropriate to treat with the same concentration of drug since uptake is likely saturable at higher drug concentrations. In the current setup, for the species with higher MIC, they have to be able to uptake substantially more antibiotics than the species with low MIC in order to end up with the same normalized uptake value in Figure 3d. It would be helpful to repeat this experiment with a single drug concentration in the media for all species and test whether that gives the same results seen here.*

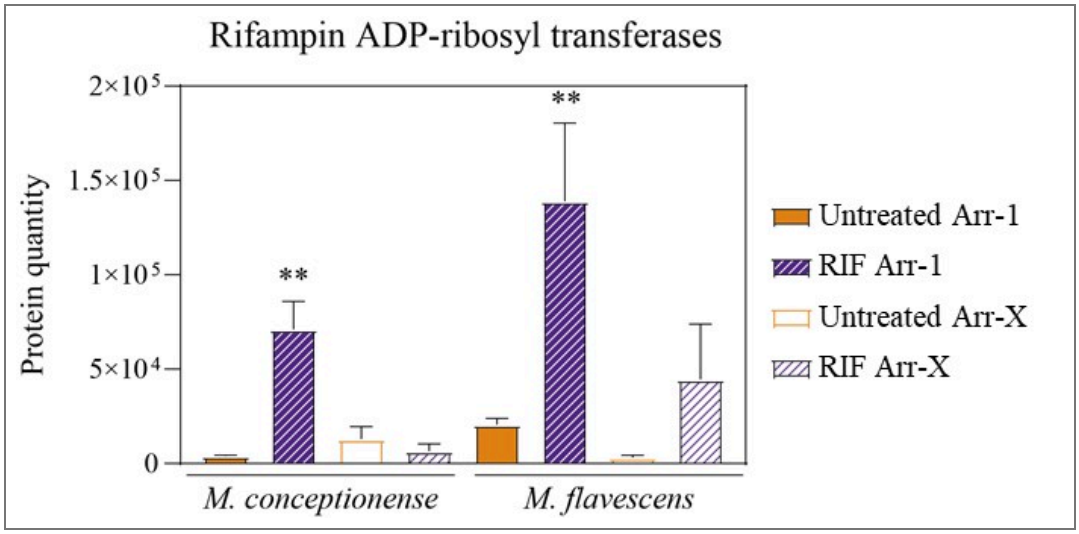
We respectfully disagree with the reviewer. Experiments such as the one proposed by the review work well when MIC values are a few fold apart, for strains of the same species, but have not been tested when MIC values are 100-1000-fold apart, with different species. Furthermore, what would be the interpretation of compound uptake at 1000-fold the MIC for one species and MIC level for another? By using antibiotic concentrations at the respective MIC for each species we are at least under conditions where we know the biological effect of the antibiotic across species is the same, based on its potency.

> *(3) Figure 4f: This panel seems to argue against the idea that the efficacy of RIF ribosylation is what's driving drug susceptibility. M. flavescens is similarly resistant to RIF as M. smegmatis, yet M. flavescens has dramatically lower riboslyation of RIF. This is perhaps not surprising, as the authors appropriately highlight the number of different rif-modifying enzymes that have been identified that likely also contribute to drug resistance. However, I do think this means that the authors can't make the claim that the resistance they observe is caused by rifamycin modification, so those claims in the text and figure legend should be altered unless the authors can provide further evidence to*

> *support them. This experiment also has results that are inconsistent with what appears to be an identical experiment performed in Supplemental Figure 5b. The authors should provide context for why these results differ.*

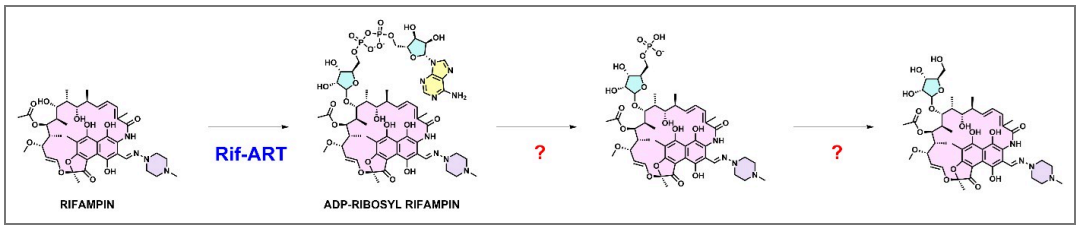
In regard to enzyme efficiency, the apparent rate of all Arr-1 is relatively similar in converting RIF into ADP-Ribosyl-Rif between species. However, Arr-X is much more efficient when compared to Arr-1 in both *M. flavescents* and *M. conceptionense*. This is indicated by the apparent rate measured and displayed on figure 5c.

Proteomics data shows that there is upregulation of Arr-1 and Arr-X upon rifampicin treatment in *M. flavescens* and *M. conceptionense*. However, the same experiment was not performed in Arr-1 KD. Therefore, we can't verify through this approach if the activity observed in vivo directly correlates with a higher expression of Arr-X alone. Of note, likely both enzymes contribute to resistance to rifamycins, as per our results with the Arr-X KD and sensitization of *M. conceptionense* to RIF.



Author response image 1.

It is also worth mentioning that there are other enzymes in the pathway of RIF ribosylation and their efficiency is unknown (Author response image 2). Therefore ADP-Ribosyl-RIF It is not an "end-metabolite" and maybe not the sole determinant of RIF resistance via ADP-ribosylation. Downstream enzymes can also account for the difference observed between *M. flavescens* and *M. smegmatis*.



Author response image 2.

It is correct that the Rifampicin MIC for *M. flavescens* is the same as *M. smegmatis*.

> *(4) Fig 4f/5c: M. flavescens has both Arr-1 and Arr-X, yet it appears to not have ribosylated RIF. This result seems to undermine the authors' reliance on the enzyme*

> *assay shown in Fig 5c - in that assay, M. flavescens Arr-X is very capable of modifying rifampicin, yet that doesn't appear to translate to the in vivo setting. This is of importance because the authors use this enzyme assay to argue that Arr-X is a fundamentally more powerful RIF resistance mechanism than Arr-1 and that it has specificity for rifabutin. However, the result in Figure 4f would argue that the enzyme assay results cannot be directly translated to in vivo contexts. For the authors to claim that Arr-X is most potent at modifying rifabutin, they could test their CRISPRi knockdowns of Arr-X and Arr-1 under treatment with each of the rifamycins they use in the enzyme assay. The authors mentioned that they didn't do this because all the strains are resistant to those compounds; however, if Arr-X is important for drug resistance, it would be reasonable to expect to see sensitization of the bacteria to those compounds upon knockdown.*

The reviewer is reading Fig. 4f incorrectly, probably because it is plotted in a linear scale instead of logarithmic scale. Ribosylated Rif is present in *M. flavescens*, just at lower levels than *M. conceptionense* and M. smegmatis. In species where there is no Arr-1 or Arr-3, ribosylated RIF is not detected at all (e.g. *M. tuberculosis*), i.e., concentration is zero. Therefore, any detection of ribosylated RIF can be considered significant. In addition, as mentioned before, ADP-ribosylation of RIF is not the final product of the reaction and further studies need to be undertaken to understand subsequent reactions.

> *(5) Figure 5d: The authors use this CRISRPi experiment to claim that ArrX from M. conceptionanse is more potent at inactivating rifabutin than Arr-1. This claim depends on there being equal degrees of knockdown of Arr-1 and Arr-X, so the authors should validate the degree of knockdown they get. This is particularly important because, to my knowledge, nobody has used this system in M. conceptionanse before.*

We agree with the reviewer that a qPCR should have been performed to define the extent of interference in the strain. generated Unfortunately, at this time a qPCR was not performed in the strains tested to confirm the extent of down regulation. Although it is the best practice to validate the strain KD, there is no indication that the effect observed is due to unspecific downregulation. The genetic environment in which Arr-X is positioned is different from Arr-1 and the targeting oligonucleotides are specific and would not promiscuously bind to Arr-1. Said that, this is indeed a fault in our setup.

> *(6) The authors' arguments about Arr-X and Arr-1 would be strengthened by showing by LC/MS that Arr-X knockdown in M. conceptionense results in more loss of ribosyl-rifabutin than knockdown of Arr-1.*

We agree with the reviewer that performing the LC-MS analysis of the Arr-x knockdown would have strengthened the argument of our paper. Unfortunately, this experiment was not performed.

> ***Reviewer #3 (Public review):***
>
> *This manuscript presents a macroevolutionary approach to the identification of novel high-level antibiotic resistance determinants that takes advantage of the natural genetic diversity within a genus (mycobacteria, in this case) by comparing antibiotic resistance profiles across related bacterial species and then using computational, molecular, and cellular approaches to identify and characterize the distinguishing mechanisms of resistance. The approach is contrasted with "microevolutionary" approaches based on comparing resistant and susceptible strains of the same species and approaches based on ecological sampling that may not include clinically relevant pathogens or related species. The potential for new discoveries with the macroevolution-inspired approach is evident in the diversity of drug susceptibility profiles revealed amongst the selected mycobacterial species and the identification and characterization of a new group of*

> *rifamycin-modifying ADP-ribosyltransferase (Arr) orthologs of previously described mycobacterial Arr enzymes. Additional findings that intra-bacterial antibiotic accumulation does not always predict potency within this genus, that M. marinum is a better proxy for M. tuberculosis drug susceptibility than the commonly used saprophyte M. smegmatis, and that susceptibility to semi-synthetic antibiotic classes is generally less variable than susceptibility to antibiotics more directly derived from natural products strengthen the claim that the macroevolutionary lens is valuable for elucidating general principles of susceptibility within a genus.*
>
> *There are some limitations to the work. The argument for the novelty of the approach could be better articulated. While the opportunities for new discoveries presented by the identification of discrepant susceptibility results between related species are evident, it is less clear how the macroevolutionary approach is further leveraged for the discovery of truly novel resistance determinants. The example of the discovery of Arr-X enzymes presented here relied upon foundational knowledge of previously characterized Arr orthologs. There is little clarity on what the pipeline for identifying more novel resistance determinants would look like. In other words, what does the macroevolutionary perspective contribute to discovery from the point of finding interspecies differences in susceptibility? Does the framework still remain distinct from other discovery frameworks and approaches? If so, how?*

Thanks for pointing this out, as this is a critical feature of our study and method. Our approach relies on inter-species comparative genomics and phenotypes, and therefore, it is distinct from inter-strains comparison. This difference is dramatic, and it becomes clearer when we are comparing the core genome of *M. tuberculosis* (one species) 92% with the core genome of the genus, circa of 1%. While we focus on rifamycin in this manuscript, future manuscripts will investigate many of the other dozens of "inconsistencies" observed between the genetic makeup of different mycobacterial species and there actual performance in the presence of different antibiotics.

> *While the experimentation and analyses performed appear well-designed and rigorous, there are a few instances in which broad claims are based on inferences from sample sets or data sets that are too limited to provide robust support. For example, the claim that rifampicin modification, and precisely ADP-ribosylation, is the dominant mechanism of resistance to rifampicin in mycobacteria may be a bit premature or an over-generalization, as other enzymatic modification mechanisms and other mechanisms such as helR-mediated dissociation of rifampicin-stalled RNA polymerases, efflux, etc were not examined nor were CRISPRi knockdown experiments conducted beyond an experiment to tease out the role of Arr-X and Arr-1 in one strain. The general claim that intra-bacterial antibiotic accumulation does not predict potency in mycobacteria may be another over-generalization based on the limited number of drugs and species studied, but perhaps the intended assertion was that antibiotic accumulation ALONE does not predict potency.*
>
> ***Recommendations for the authors:***
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *Major comments*
>
> *(1) The metabolomics is done using mycobacteria grown on filters. Initially, mycobacterial cells are grown on the filters for 5 doublings before being transferred to drug-containing (or free) agar for one doubling. Is this based on calculated doubling time in liquid culture or a true determination of the fact that the biomass increases to what would amount to 5 doublings?*

The doubling time used is the one determined in liquid media. Although it is possible that the growth kinetics in solid media is slightly different from liquid (±10%), this experimental design is well established for *M. tuberculosis* (since Proc Natl Acad Sci U S A. 2010 May 25;107(21):9819-24.) and *M. smegmatis* (unpublished). Therefore, we used the growth rate as a proxy for having the same biomass of cells for each species tested. A maximum difference of 10% was observed between *M. tuberculosis* growth in liquid and in solid media, however, cells grow exponentially for much longer in filters. This makes filter-based experiments more reliable, as few growth phase-derived differences are present.

> *(2) The demonstration that intrabacterial drug concentrations vary between mycobacterial species in a manner not related to MIC for at least LZD and RIF, is an important finding. However, intrabacterial does not mean cytoplasmic since a considerable fraction could be present in the periplasmic/cell wall layers. Ideally, this would need to be determined but would of course be a massive undertaking since the method needs validation & optimization for each mycobacterial species. Nevertheless, this has to be mentioned. In addition, three drugs are limiting. Measuring additional drug concentrations in these 5 mycobacteria would at least establish some confirmation about the extent of this lack of correlation. Thus, could the authors measure concentrations of additional drugs with intracellular targets?*

Testing additional drugs can be beneficial and would be an expansion of our paper, which will definitely be on future plans for further studies focusing on other antibiotics described here. It would also provide new insights into other possible mechanisms of resistance in mycobacterial species. However, in this study we aimed to first determine the antibiotic response profile in different mycobacterial species, and once we identified interesting resistance phenotypes that could not be readily explained by known mechanisms of resistance, we narrowed it down to certain drugs and species that would potentially provide insights into new mechanisms of antibiotic resistance. Finally, exploring drug concentration across multiple bacterial compartments is a dauting task and it has not been done extensively with any species, not to mention with multiple species, many of which are still lacking any study of their actual cell envelope.

> *(3) CRISPRi was used to reduce transcription in M. conceptionense. What was the level of gene downregulation?*

As mentioned previously, a setback from our setup is that the level of KD was not measured at this instance.

> *Minor comments:*
>
> *(1) The introduction mentions the fast and slow-growing mycobacteria which are classified based on the time that it takes to observe colonies on solid agar. However, in liquid medium, there is less correlation between the reported growth on agar and doubling time in liquid (Figure 1b, Figure 2d). This could be mentioned in the results section. In Figure 2d, the filled circles represent fast-growers but this does not hold well for liquid culture and it might make more sense to not distinguish between fast- and slow-growers in these graphs. A small complication would also be the fact that the doubling time represents growth in a liquid medium with Tyloxapol as a detergent whereas the MIC and metabolomics are done on solid agar with no detergent. The metabolomics is done after a doubling but for those where agar growth and liquid growth have large discrepancies in growth rate, there could be some differences.*

Apologies for this misunderstanding. Fast- and slow-growth phenotypes are determined in Lowenstein-Jensen (LJ) agar, not in 7H10 agar (used in our study and most studies of mycobacteria). Furthermore, this is a qualitative definition, not a quantitative one. Therefore, our measurements do not need to correlate with fast- and slow-growth phenotypes, unless we

had used that one specific medium. Furthermore, in liquid medium, we determined growth rate directly, which is never done with LJ medium.

In addition to adding the same amount of cells to each filter, we also perform TIC normalization, which should account for how rich the samples were – and therefore how much material we had. Therefore, we do not observe discrepancies due to differences in growth rate and the presence/absence of detergent in the media.

It is also worth mentioning that this experimental set up has been well established in many *M. tuberculosis* labs that study metabolism. Importantly, the use of detergent drastically affects mass spectrometry, and therefore cannot be used.

> *(2) Figure 1g in the text should be Figure 1f.*

Apologies, it has been fixed.

> *(3) Figure S1 would be ideal to have in (supplementary) table format.*

This data is now being provided in a table format.

(4) Table S1 - ethambutol misspelt.

Spelling has been corrected.

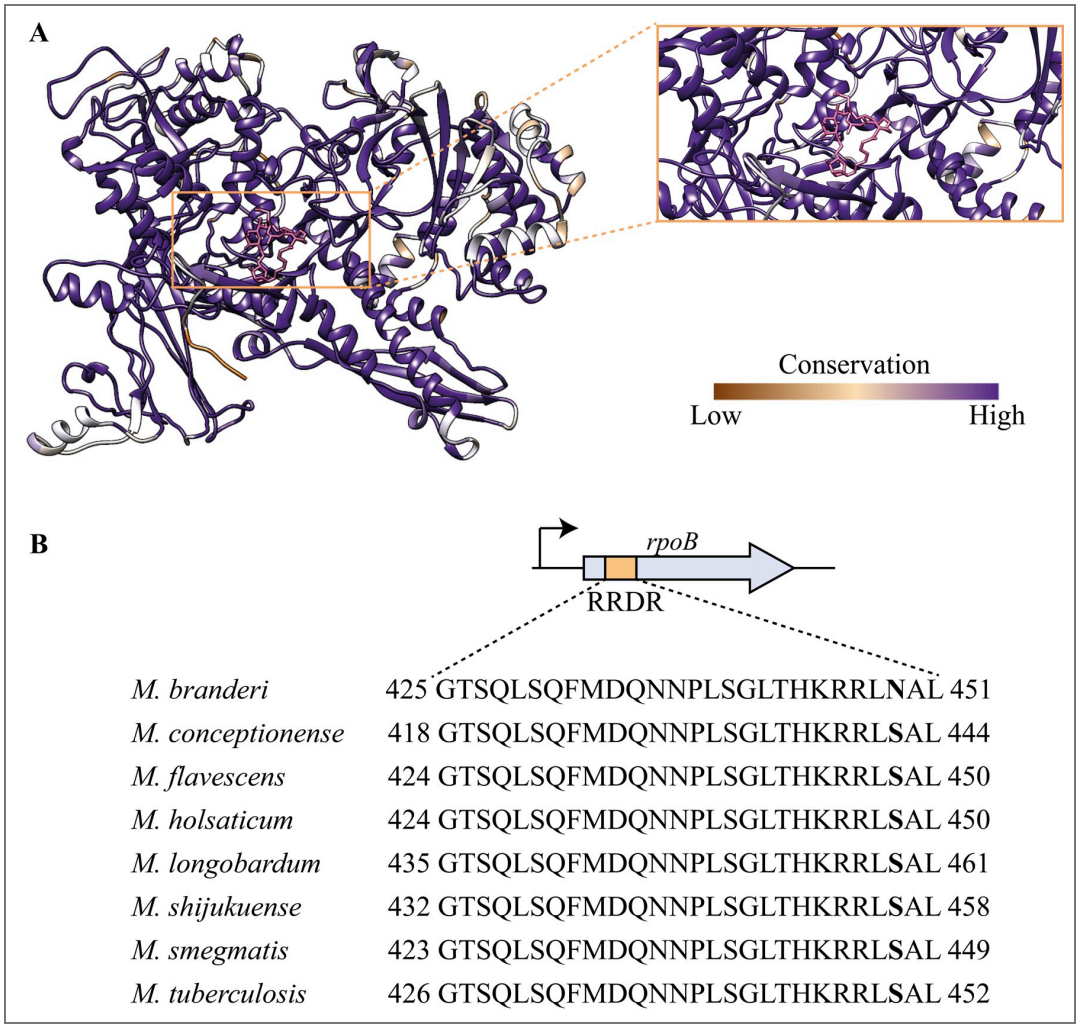
> *(5) MIC for species such as M. abscessus could depend on medium (7H9-based medium can give different MIC values than CAMH).*

Indeed, different media can significantly change MIC values, and this is true for many bacterial species, if not all. For this study we used only species that could be grown in 7H9 broth containing 10 % ADC, 0.05% glycerol 0.05% tyloxapol and 7H10 plates containing 10% OADC and 0.05% glycerol. MIC_{99} was determined in the latter as we found more efficient and robust to do our tests it in solid media. The goal of our experiment was not to the determined the “true” MIC for the antibiotics tested, as this value does not exist. It was to find lack of correlations between relative values and the presence of genes that can account for it.

> *(6) The statement "the experiment was performed at a concentration of antibiotic equal to its MIC" initially seems confusing. It was not equal to the MIC but performed at 6-fold the respective MIC of the species in question. Maybe re-phrasing this would help.*

Apologies for this oversight. It has been corrected.

> *(7) Note that some mutations outside the RRDR (eg. V170F and I491F) can also cause Rif resistance.*



Author response image 3. A Rainbow diagram of RpoB X-Ray structure coloured according to sequence conservation. Dark purple indicates high conservation, whereas dark orange indicates low conservation. RIF (showed in magenta) is bound to RpoB. Zoomed view displays that the RIF-binding pocket is considerably conserved. B RpoB protein sequence has an 81bp region called Rifampicin Resistance Determining Region (RRDR) that is known to be important for RIF binding and is where most mutations occur in drug-resistant TB. Sequence alignment displays that the RRDR region is conserved with the exception of *M. branderi*, which has an Asn instead of a Ser residue in position 456 (numbering is related to the *M. tuberculosis* sequence), highlighted in bold.

Attached we have a structural alignment of RpoB of the species highlighted on this paper. Although there is variability within the sequences, which is also displayed in Author response image 3 with the conservation analysis, the residues that have been implicated with resistance (including V170 and I491) are conserved. Alignment sent on .fasta file that can be opened in jalview.

> *(8) Discuss how the RpoB S450N mutation in M. branderi confers the observed level of resistance.*

That's a great point, thank you. Now it reads as:

"The rifampicin (RIF) binding pocket is generally conserved, but *Mycobacterium branderi* has an S450N mutation in the RRDR region. While this specific mutation hasn't been found in clinical isolates, it's located at the binding site and may confer resistance (273). Although both

serine (S) and asparagine (N) have similar side chains, related mutations like S450Q have been linked to resistance (156). Thus, M. branderi may be RIF-resistant due to this mutation. In contrast, *M. conceptionense*, *M. flavescens*, and M. smegmatis show no target sequence differences that explain their resistance"

> *(9) The statement that the three tested NTM are sensitive to rifabutin ("resistant to all rifamycins except for rifabutin") needs to be interpreted considering what sensitivity means. The MIC is still high (1.6-3.1 ug/mL) when compared to that of Mtb. The 2-fold differences in MIC between M. smegmatis and M. conceptionense do not really prove or disprove the role of Arr-X in rifabutin resistance.*

We fixed the sentence to be more careful with the language on the text. We agree, but it is worth mentioning that generally with bacteria there is a regulation by the CLSI. Each bacterial species has a range that is considered sensitive or resistant, but these are not available for the species used in this study. In general, bacteria with MIC values above 8 µg/mL are considered resistant to rifampin (J Antibiot 2014 67:625).

> *(10) Figure 1d: It's hard to quantify the sensitivity of the plates. Can this be done by MIC? Was only rifabutin tested or also rifampicin?*

The initial experiments described on the paper were all performed using Rifampicin only. Then, the MIC for the remaining rifamycins was determined for *M. smegmatis*, *M. flavescens* and *M. conceptionense*, and can be perused on "Supplementary table 4". Figure 5d is to illustrate the effect of the KD in *M. conceptionense* sensitivity to rifabutin.

> *(11) Is there data to show the ADP-ribosylation of rifabutin in M. conceptionense and the CRISPRi strains?*

Unfortunately, we did not perform LC-MS analysis on *M. conceptionense* CRISPRi strains exposed to rifabutin to measure potential ADP-ribosylation.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *(1) It would be useful if the authors would complete Figure 1A by determining growth rates for the remaining 18 strains that they currently omitted.*

These growth rates were obtained using roller bottles and in at least 3 independent experiments, unfortunately the throughput is far ideal. The goal of the experiment was to highlight difference in growth rate, beyond fast- and slow-growth, which we did. Adding the remaining values would not change this conclusion. Growth rate variation in 7H9 is significant and the point is made in our figure.

> *(2) The authors should justify their choice of species used in Figures 3-4. It would be useful to know, for instance, if the authors chose these species in an unbiased fashion, or if they were chosen because the authors had already determined that they possess rifamycin-modifying enzymes of interest. In that case, they wouldn't necessarily be a representative sample to use for the correlation analysis of antibiotic uptake and potency in Figure 3.*

They were chosen because of their resistance profile for BDQ, LZD and RIF. This has been addressed in the text, which now reads "Given the antibiotic response profiles observed, we selected BDQ, LZD and RIF to explore the molecular causes of these dramatic changes in antibiotic potency observed across the Mycobacterium genus."

> *(3) Figure 4b: The data in this panel appear inconsistent - for instance, M. houstonense appears to grow at 10X Mtb MIC, but fails to grow at 1X Mtb MIC. Repeating this experiment would better establish the validity of the authors' claims about the relative susceptibility of these strains to RIF.*

The figures got rotated when exported from illustrator. Corrected figure is uploaded, and original plate photos are also uploaded for clarity.

> *(4) Figure 4e: Does Arr-X get upregulated in these proteomic datasets? The authors' argument that proteomic upregulation correlates with important drug resistance genes would imply that it might be, so that would be useful information to provide.*

Arr-X is slightly upregulated, but not statistically significant – this could be due to the native expression of Arr-1. Data is displayed in a previous answer.

> *(5) I wasn't able to find the supplementary tables that the authors allude to - not sure if that was a file mixup, but those tables would be useful for interpreting the manuscript.*

We are sorry that you couldn't access the table. It must be a file corruption issues, as the other reviewers were able to. We will make sure that all tables are available and accessible.

> *(6) For LC/MS, the authors use peak height instead of peak area, which they argue correlates better with the amount of drug in cells because of the poor peak shape they observed for linezolid. This is not standard practice, so the authors should provide evidence to support this claim by running an LC/MS standard curve, then showing the correlation between peak height and amount of compound added as well as the correlation between peak area and compound.*

Thank you for pointing that out, accuracy calculated and displayed. Both peak area and height can be used, but indeed area is standard practice.

> *(7) The authors should provide methods information about the LC column and the gradient settings used for LC-MS, as well as the settings of the MS.*

The full method has been added to the paper.

> ***Reviewer #3 (Recommendations for the authors):***
>
> *I have only minor comments aside from the information in the Public Review:*
>
> *(1) Results, section on Intra-bacterial antibiotic accumulation, line 8: "experiment was performed at a concentration of antibiotic PROPORTIONAL to its MIC" would be more accurate?*

Agreed and adjusted according to Reviewer's suggestion.

> *(2) Results, section on A minor role for pre-existing target modification, last sentence: the mere presence of RIF-ribosylating enzymes does not, in and of itself indicate that "RIF modification, and precisely ADP-ribosylation, is the dominant mechanism of resistance to RIF in mycobacteria", as other mechanisms and other forms of modifying enzymes are known to confer rifamycin resistance, with redundancy (e.g., other rifampicin-modifying enzymes, or helR-mediated dissociation of rifampicin-stalled RNA polymerases from DNA). It would be more appropriate to suggest the results presented to this point indicate RIF modification is common among mycobacteria. The evidence from the CRISPRi knockdown of Arrs shown in Fig 5d is the kind of evidence that suggests ribosylation as a dominant mechanism, at least against rifabutin in this particular species.*

Absolutely, there are other possible modifying enzymes that could be encoded by these mycobacterial species. There is a possibility that *M. flavescens* and *M. smegmatis* encode for a putative *helR* (attached alignment) but further experiments would need to be carried out to confirm its ability to displace RIF in the RNAP. Interestingly, the presence of both Arr and

HelR has been studied in *M. abscessus* and those mechanisms of resistance are independent from each other (Molecular Cell 2022 82(17):3166-3177.e5).

> *(3) Discussion, 2nd sentence needs grammatical editing.*

Rephrased and it reads “Using our mycobacterial library, we identified for the first time high- and ultra-high-level intrinsic resistance (3) to many of the antibiotics tested. Of note, the resistant phenotype is naturally occurring and not a result of mutations due to exposure to the antibiotic in the clinic – which is the more traditional approach for probing mechanisms of antibiotic resistance. Our observations revealed that resistance profiles are highly variable across the genus and do not follow phylogeny, implicating HGT as the key mechanism for acquisition of resistance determinants and evolution of antibiotic resistance in mycobacteria (42).”

> *(4) Discussion, page 7, first line: the inclusion of LZD and BDQ in this statement seems at odds with Figure 2c and the statements in the first paragraph of page 5 highlighting these as examples of drugs to which most mycobacteria are susceptible.*

Indeed, many of the species are susceptible, however the MIC_{99} levels observed have never been reported before, and therefore we found it to be an interesting finding to highlight. From a treatment perspective, knowing which species are sensitive to which drugs is of course the most useful outcome of our study.

> *(5) The next sentence..."We found that resistance to these antibiotics in mycobacteria cannot be explained by uptake/efflux mechanisms..." is a bit of an over-generalization and conflicts with the evidence presented earlier that efflux could be playing a role in BDQ resistance and the published evidence establishing a clinically significant role for efflux-mediated BDQ resistance in M. tuberculosis, M. avium complex and M. abscessus complex.*

We rephrased it to make it more specific to our findings. It reads “We found that resistance to these antibiotics in mycobacteria do not correlate with by uptake/efflux mechanisms in the species tested and it does not correlate with growth rate. Identification of mycobacterial species highly resistant to BDQ and LZD is worrisome as most of this species, if not all, have never been exposed to these drugs.”

> *(6) Methods, section on In vitro activity assay of Arr enzymes, line 1: reference(s) should be provided for previously reported methods.*

Reference now added.

> *(7) Figure 2d: the low end of the susceptibility range is not well defined.*

In this figure the susceptibility is not defined as the lowest area of the graph, but the lower concentrations are indeed harder to be defined. Hopefully supplementary figure 1 and the additional table containing the MIC can be informative to address this comment.

> *(8) Figures 3c,d: the presentation of the relative antibiotic concentrations could be harmonized between the graphs in 3c and those in 3d to enable a more ready comparison.*

We disagree. The goal of these different panels is exactly to illustrate two distinct points. C gives the relative concentration of antibiotic, while D correlates relative concentration with MIC99. The use of log scale in D further clarifies that there is no correlation between intracellular antibiotic concentration and potency (MIC). This information is not present in C.

> *(9) Figure 4f and Supplementary Figure 5b: it is difficult to understand the limited amount of ribsosyl-RIF in M. flavescens in Fig 4f relative to Supplementary Figure 5b (esp.*

> *when considering M. smeg as a common comparator); and, further, to understand the seeming lack of correlation between RIF susceptibility, ribosylation and Arr number and catalytic efficiency for these two strains without considering additional resistance mechanisms.*

In reality the difference between figure 4f and Supplementary figure 5b is mainly due to M. smegmatis – that has an apparent lower production of ribosyl-RIF in the experiment described in the supplementary figure. The values for *M. flavescens* are relatively similar. In addition, the ADP-Ribosyl-RIF is not the final metabolite of the pathway.

In regards of having the entire picture, it is true that we were unable to completely unravel and correlate MIC value, expression of Arr-1, expression of Arr-3, efficiency of each enzyme, production of ADP-Ribosyl-RIF and the presence of other possible mechanisms of resistance and this is indeed a setback in our study, and of most studies ever published, which usually focus on one resistant determinant.

https://doi.org/10.7554/eLife.101940.2.sa0