TÉCNICO LISBOA

A genome-wide perspective on local and global genetic determinants of drug resistant tuberculosis

(Article)

Pedro Gomes Duarte

Supervisors: Dr. João Ruben Lucas Mota Perdigão and Professor Isabel Maria de Sá Correia Leite de Almeida

Abstract

Tuberculosis (TB) is the leading infectious disease caused by a single agent, and multidrug-resistant (MDR) TB poses a serious health hazard, given its low treatment success rate. The advent of whole genome sequencing (WGS) has widened the analytical spectrum of the *Mycobacterium tuberculosis* genome, enabling a broader characterization of genetic determinants of drug-resistant TB. A dedicated WGS pipeline was employed to uncover the relationship between drug resistance-associated genotypes of *M. tuberculosis* clinical isolates from Portugal and resistance levels of first- and second-line drugs. Differing resistance levels were found among same-class drugs harbouring identical drug resistance-associated genotypes. Novel *gid* Ala167Asp, *rrs* 1076insT, *ethA* Met1Leu, *alr* Met343Thr and Phe4Leu mutations were characterized as associated with drug resistance, while the previously reported association between *embB* Asp354Ala and ethambutol resistance was discredited. A global analysis of the diversity of drug resistance associated mutations and their association with the human-adapted *M. tuberculosis* complex lineages has demonstrated a significant association between prevalent *fabG1* C-15T, *rpoB* Ser450Leu, *embB* Met306Val, *rpsL* Lys43Arg and *rrs* A1401G mutations and lineages 2 and 4, elucidating these lineages proclivity to develop MDR.

Introduction

Tuberculosis (TB) is caused by the bacterial pathogen *M. tuberculosis*, and it is estimated that nearly 2 billion people worldwide have a latent TB infection (Zumla et al., 2013). Since 2007, TB has been the leading cause of death from a single infectious agent, responsible for 1.2 million deaths globally, in 2018. Multidrug-resistant (MDR) TB, which is defined as having resistance to at least isoniazid (INH) and rifampin (RIF), presents a global concern, with a worldwide prevalence average of 18 %, reaching proportions >50 % in countries of the former Soviet Union. Considering the treatment success rates of MDR and extensive drug-resistant (XDR) TB – defined as MDR TB plus resistance to a fluoroquinolone (FLQ) and at least one of the following second-line injectable drugs (SLIDs): amikacin (AMK), kanamycin (KAN) or capreomycin (CAP) – are at a low 56 and 39 %, respectively, urgent development of new antituberculosis drugs and improvement of current antituberculosis chemotherapy regimens is required (World Health Organization, 2019).

Understanding the molecular mechanisms behind drug resistance acquisition is crucial for helping devise better therapeutic regimens and develop improved molecular diagnostic tools to monitor TB outbreaks a detect drug-resistant TB. This study aims at attempting to establish links between whole genome sequencing (WGS)-derived drug resistance-associated mutations from a local dataset of *M. tuberculosis* clinical strains with drug resistance levels. Moreover, the association between the globally-obtained drug resistance-associated mutations and the genetic background of the MTBC will be discussed to better understand if different *M. tuberculosis* lineages are more prone to develop drug resistance.

Materials and Methods

Local analysis

A total of 207 *M. tuberculosis* clinical isolates, sourced nationwide from 15 districts in Portugal, including the Madeira and Azores archipelagos, were selected for the present study. The study sample is composed of 191 isolates retrospectively selected between 2008-2016 as part of the TB National Laboratory Surveillance VigLab program, across the National Institute for Hygiene and Tropical Medicine of the NOVA University of Lisbon and Faculty of Pharmacy of the University of Lisbon. This sample includes 71.1 % of the MDR TB cases notified by national public health authorities in the same period and comprises and excess of 31 XDR TB isolates to those notified in the same period. An additional 16 historical strains isolated between 1995-2007 were also included.

All clinical isolates had been previously subjected to WGS using an Illumina NGS sequencing platform producing 100/150-bp paired-end reads. Raw sequence data was initially subjected to quality control (QC) using Trimmomatic with default settings which removes and trims low-quality reads. QC-passed reads were subsequently mapped to the reference genome

of the *M. tuberculosis* H37Rv (NCBI reference sequence: NC_000962.3), using the Burrows-Wheeler aligner tool with the BWA-MEM algorithm. Variant calling of genome-wide SNPs and insertions and deletions (indels) was performed using both Samtools and Genome Analysis ToolKit, and only concordant variants between both tools were retained for downstream analysis.

Genotypic profiles of drug resistance were inferred through a revised mutation database comprised of 1,290 validated drug resistance-associated mutations, related to 37 different drug resistance-associated genes (Table A.1), collected from recent WGS and molecular studies (Coll et al., 2018; Coll et al., 2015; Perdigão et al., 2010). Briefly, this in-house pipeline screens variant call format (VCF) files obtained upon mapping and variant calling for variants falling within pre-specified target regions comprising drug resistance associated genes and, when applicable, respective promoter regions. The initial set of variants detected in this way is afterwards compared with the abovementioned mutation database, and isolates harbouring mutations present in this database were classed as genotypically resistant to the respective drugs to which the mutations are associated, while retaining information regarding other mutations present in all drug resistance associated genes for each drug which enables the characterization of the complete allelic configuration associated with drug resistance. This same pipeline was used for the global analysis as well.

All clinical isolates had been previously subjected to phenotypic DST as part of routine TB laboratory diagnosis DST against first-line drugs was performed using the standardized procedure of fluorometric BACTEC MGIT 960 system (Becton Dickinson, Sparks, MD, USA) for INH, RIF, streptomycin (STR), ethambutol (EMB), pyrazinamide (PZA). For clinical isolates dating before 2000, DST had been carried out through the radiometric BACTEC 460 (Becton Dickinson, Sparks, MD, USA) for susceptibility to all first-line drugs except PZA which was not determined.

A subset of 40 *M. tuberculosis* clinical isolates, available at the Faculty of Pharmacy of the University of Lisbon biobank was subjected to Minimum Inhibitory Concentration (MIC) determination for 12 antituberculosis drugs. All clinical isolates were previously grown in Löwenstein-Jensen medium during 2–3 weeks at 37 °C, after which a bacterial suspension was prepared by homogenization in a glass tube with glass beads, and resuspended in sterile bidistilled water. Mycobacterial cell suspension turbidity was adjusted to 0.5 MacFarland units using a densitometer (Grant Instruments, Cambridge, UK). Mycobacterial cell suspensions were then diluted to 1:100 in Middlebrook 7H9 broth (BD Difco, Berkshire, UK) supplemented with 10 % Middlebrook OADC. 96-well microplates with prepared with serial dilutions of each 12 drugs in a scheme identical to the Sensititre *M. tuberculosis* MYCOTB (Thermo Scientific, Waltham, MD, USA) MIC plates. The following drugs at the specified concentration intervals were included in each plate: INH (0.03–4.00 mg/L), RIF (0.12–16.00 mg/L), rifabutin (RFB) (0.12–16.00 mg/L), STR (0.25–32.00 mg/L), AMK (0.12–16.00 mg/L), KAN (0.6–40.0 mg/L), ofloxacin (OFX) (0.25–32.00 mg/L), moxifloxacin (MFX) (0.06–8.00 mg/L), ethionamide (ETH) (0.3–40.0 mg/L), para-aminosalicylic acid (PAS) (0.5–64.0 mg/L) and cycloserine (DCS) (2–256 mg/L). Each well was inoculated with 0.1 mL of the prepared bacterial suspension and incubated at 37 °C with growth

assessments at 10 and 14 days post inoculation by microscopic observation in an inverted microscope. MICs were determined as the lowest drug concentration preventing visible growth.

Global analysis

The study includes an initial set of 28,385 M. tuberculosis isolates for which publicly available raw sequence data was retrieved from the European Nucleotide Archive. This dataset includes publicly available sequence data made available until the 31st of July, 2018, composed of whole genome sequencing libraries generated by an Illumina platform under a paired-end sequencing mode.

Upon raw read trimming and filtering, using Trimmomatic, initial mapping and variant calling was carried out using the Snippy pipeline and the reference genome of M. tuberculosis H37Rv (NCBI reference sequence: NC_000962.3). Snippy implements read mapping using the Burrows-Wheeler Aligner tool with the BWA-MEM algorithm and, variant calling using Freebayes.

The lineage profiles were determined for the global dataset of *M. tuberculosis* clinical isolates, using a dedicated SNP barcode (Coll et al., 2014). A χ^2 test was employed to determine the association between the genetic background, particularly the L1, L2, L3 and L4, and genomic variants, normalized for drug-resistant isolates. Only mutations with expected frequencies, among lineage-specific associated *M. tuberculosis* isolates of 5 or more were considered suitable for the aforementioned statistical analysis. Mutations significantly associated with the MTBC genetic background were subjected to a second set of four χ^2 tests to determine the single-association with each of the abovementioned lineages. The latter analyses were performed using the Yates correction term. The level of significance was 0.05 (Jones, 2002).

Results and Discussion

The INH MIC analysis has revealed 14 unique genotypes with varying resistance levels (Figure 1). Isolates with single *fabG1* C-15T promoter mutations show a minor resistance level (0.50 mg/L), but isolates with double *inhA* mutations, in particular the *fabG1* C-15T promoter mutations complemented with *inhA* Ser94Ala mutations show a high-level resistance to INH (3.75 mg/L). These results are concordant with previous findings on INH-resistant *M. tuberculosis* clinical isolates from Portugal where double *inhA* mutations appear to confer higher levels of INH resistance as opposed to single *fabG1* C-15T promoter mutations (Machado et al., 2013). The *katG* Ser315Thr is associated with high-level INH resistance (4.0 mg/L), and the *katG* Arg463Leu mutation does not seem associated with INH resistance, as it was exclusively found among INH-susceptible isolates in its single allelic configuration, as this is a phylogenetic polymorphism that can be found on isolates belonging to the Principal Genetic Group 1 (Sreevatsan et al., 1997a).

Regarding rifamycin associated genotypes, high-level RIF and RFB resistance can be attributed to mutations in the *rpoB* gene, despite being variations in the degree of resistance level

between the two drugs. A noteworthy difference seems to be the *rpoB* Asp435Val mutation, which in the case of RIF is associated with high-level resistance (16.0 mg/L) but in the case of RFB its associated MIC (0.12 mg/L) does not exceed the critical concentration (CC) for RFB. Genotypes harbouring *rpoC* mutations like the Gly594Glu substitution, is not associated with RIF or RFB resistance, but rather as compensatory mutations (Comas et al., 2012).

The EMB MIC analysis (Figure 1) has revealed that high-level resistance to EMB is dependent on multigenic mutations, particularly the combination of *embB* and *embA* promoter mutations. The C-16T *embA* promoter mutation in combination with the Met306Val and Met423Thr *embB* mutations have shown to confer high-level EMB resistance (16.0 mg/L), while double Met306Val and Met423Thr *embB* mutations only confer a low level of EMB resistance (6.0 mg/L). In fact, the latter MIC can be solely attributed to the *embB* Met306Val, considering single *embB* Met423Thr mutations are not associated with EMB resistance, which discredits its previously reported association with microevolution events towards EMB resistance by *M. tuberculosis* strains from the Q1, from Portugal (Perdigão et al., 2014b). The *embB* Asp354Ala, which is present in the current database, does not seem to be associated with EMB resistance and therefore should not be included for future genomic analysis of EMB resistance.

Concerning isolates with double *gid* mutations that have the *gid* Leu16Arg mutation and are STR-resistant, their acquired resistance can be solely attributed to the remaining *gid* mutation, since the *gid* Leu16Arg mutation is a natural polymorphism that is not associated with STR resistance (Spies et al., 2011). The *gid* Ala180Pro was found to be associated with STR resistance, and has been previously demonstrated as a phylogenetic marker for the Q1 clade (Perdigão et al., 2014a). The *rpsL* Lys43Arg, which has been found as a highly frequent among *M. tuberculosis* clinical isolates from Portugal, is associated with high-level STR resistance (32 mg/L) (Perdigão et al., 2008). The analysis revealed a novel mutation, the *gid* Ala167Asp substitution to be associated with low-level STR resistance (2 mg/L) (Figure 1), and should be incorporated into the present mutation database to improve genome-based detection of STR-resistant isolates.

Both the AMK and KAN MIC analysis attributed the *rrs* A1401G mutation with high-level resistance (16.0 and 40.0 mg/L, respectively) (Figure 1). The G-10A *eis* promoter mutation is associated low-level AMK resistance, as it only exhibited a MIC of 2 mg/L, slightly above the CC for AMK. In contrast, the same mutation contributes to intermediate-level KAN resistance (20 mg/L), with a higher increase relative to the CC for KAN. These results are directly correlated with KAN increased affinity to the Eis protein when compared to AMK (Cohen et al., 2014). The *rrs* 1076insT is a novel mutation is associated with both high-level resistance to AMK and KAN, for which it should be added to the present database to improve detection of AMK- and KAN-resistant isolates.

Ignoring the *gyrA* Glu21Gln Gly668Asp Ser95Thr genotype, whose associated mutations are only natural occurring polymorphisms exclusively associated with the L4 genetic background, and not associated with FLQ resistance (Lau et al., 2011), both OFX and MFX MIC analysis reveal

that mutations in the 91 and 94 codons of the *gyrA* gene are associated with high-level FLQ resistance (Figure 1). Albeit, by comparing the relative increases of MICs to their respective CCs, the two FLQs show significant differences, with the abovementioned mutations contributing to higher levels of resistance to MFX than to OFX, which contradict previous findings (Sirgel et al., 2012).

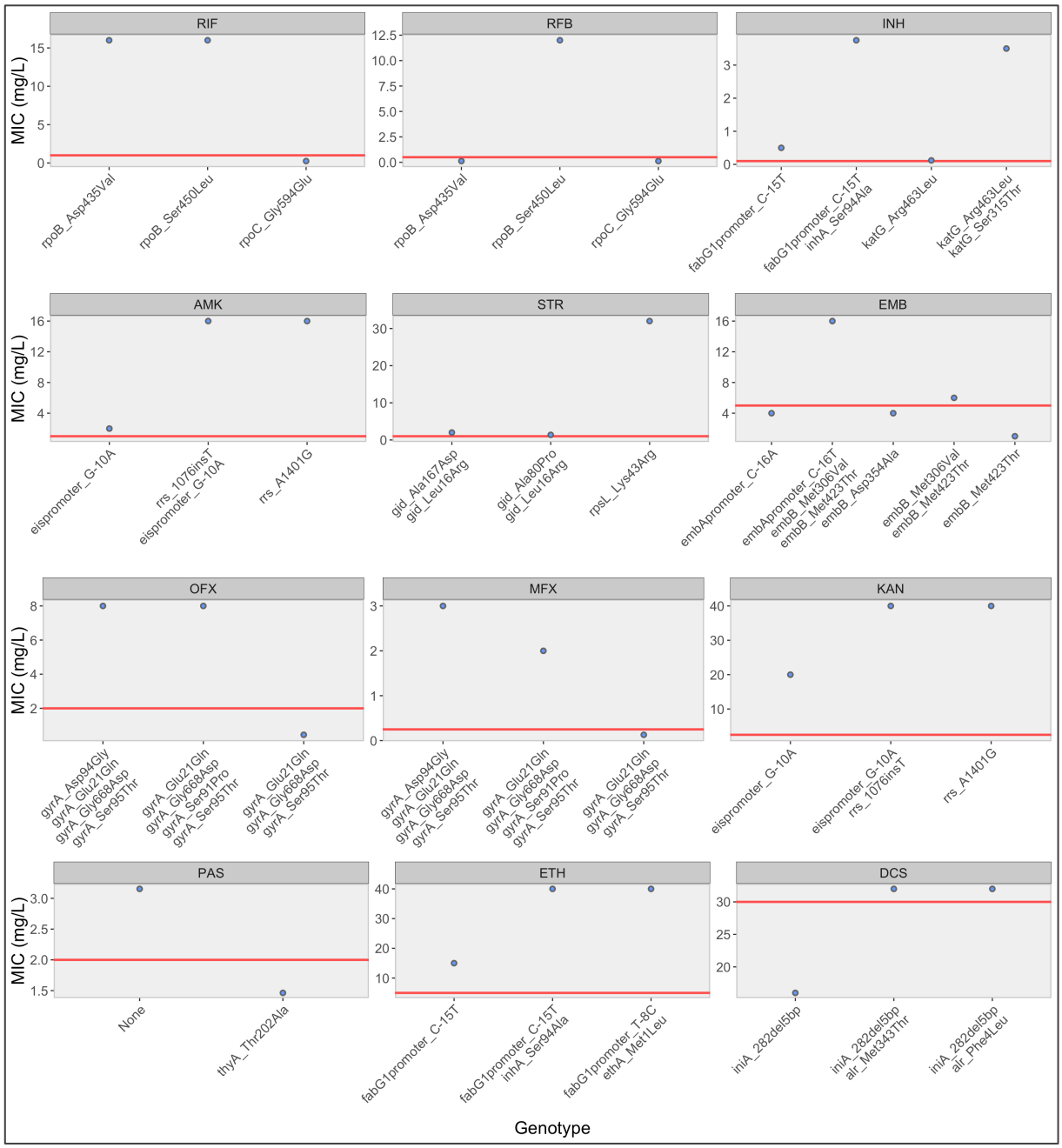

Figure 1. Mean MICs of 12 first- and second-line antituberculosis drugs relative to a selection of respective drug associated genotypes. Genotypes pertain to a subset of 40 *M. tuberculosis* clinical isolates from Portugal. Red lines indicate the CC for each drug (INH: 0.1 mg/L, RIF: 1.0 mg/L, RFB: 0.5 mg/L, EMB: 5.0 mg/L, STR: 1.0 mg/L, AMK: 1.0 mg/L, KAN: 2.5 mg/L, OFX: 2.0 mg/L, MFX: 0.25 mg/L, ETH: 5.0 mg/L, PAS: 2.0 mg/L, DCS: 30.0 mg/L).

The majority of INH resistance-associated mutations also contribute to ETH cross-resistance. This is the case with *inhA* mutations, and much like with INH, double *fabG1-inhA* promoter mutations contribute to higher levels of ETH resistance (40 mg/L) than single *fabG1* promoter mutations (15 mg/L) (Machado et al., 2013). The T-8C *fabG1* promoter *ethA* Met1Leu double mutation also displayed a high MIC to ETH (40 mg/L). Following the reasoning that *fabG1* promoter mutations only confer intermediate-level resistance to ETH, there is plausible evidence that the *ethA* Met1Leu mutation alone is associated with ETH resistance.

The PAS MIC analysis showed a lower variety of associated genotypes. The *thyA* Thr202Ala mutation was the only genotype with an associated MIC (1.5 mg/L), despite being below the CC, which is attributed to its sole association with L4 of the MTBC (Feuerriegel et al., 2010). Isolates without an associated genotype were found to display a MIC of 3.2 mg/L, which suggest other genes might be involve in PAS resistance acquisition.

Regarding the DCS MIC analysis, two new mutations, the *alr* Met343Thr and Phe4Leu substitutions were found to be associated with DCS resistance, for which their inclusion in future molecular analyses could improve detection of DCS-resistant isolates.

The global analysis of *M. tuberculosis* clinical isolates has revealed a total of 8,235 unique mutations, of which 624 (7.6 %) were known drug resistance-associated mutations. The association between the identified mutations and the MTBC genetic background was conducted, by examining their distribution across different lineages. Within the total identified mutations, only 31 showed a significant association with any of the of the 4 *M. tuberculosis sensu stricto* lineages: L1, L2, L3 or L4. Within these 31 mutations, only 15 were associated with drug resistance, and of these 9 will be discussed in this study (Table 2).

Regarding INH resistance-associated mutations, the C-15T *fabG1* promoter mutation and the *kasA* Gly269Ser mutations were found to be significantly associated with L4 (Table 2). Contradicting evidence suggests the C-15T *fabG1* promoter mutation is primarily linked with L1 strains. Despite this, the majority of isolates with low-level INH resistance that harboured the C-15T *fabG1* promoter mutation belonged to L4 (Fenner et al., 2012). Similarly, with INH-resistant isolates, the C-15T *fabG1* promoter mutation, which confers cross-resistance to ETH, was also found to be associated with the L4 among ETH-resistant isolates.

Two RIF resistance-associated mutations were found to be associated with the MTBC genetic background. The Asp435Val substitution in the *rpoB* gene was found to be associated exclusively with L4 while the Ser450Leu substitution is associated with L2. The *rpoB* Asp435Val mutation has been previously associated with L4-specific WGS analyses (Mortimer et al., 2018), suggesting that it is positively selected under this phylogenetic subgroup of the MTBC. There is ambiguity on which lineage the prevalent *rpoB* Ser450Leu mutation is associated to. A study has shown that the Ser450Leu mutation is less likely to be associated with the Beijing genotype – the main L2 sub-lineage – than with non-Beijing strains (Park et al., 2005), while another study revealed the opposite (Dymova et al., 2011). The data herein obtained demonstrates its association with the L2 lineage probably owing to the highly dynamic spread of the Beijing family

(L2) across Asia towards Europe. Specific MDR TB subclades of the Europe-Russia W148 and Central Asia Outbreak clades presently comprise MDR TB superclusters spanning mainly across Europe and Asia with RIF resistance mainly driven by the *rpoB* Ser450Leu mutation (Merker et al., 2015).

EMB resistance is attributed to wide variety of mutations, from multiple loci. The lineage-association analysis has revealed 3 EMB resistance-associated mutations with significant association with the MTBC genetic background (Table 2). The *embB* Glu378Ala and the *embC* Thr270Ile mutations were found to be exclusively associated with L1 strains. Concordant evidence confirms these findings while also revealing their added association with the remaining ancestral L5 and L6 (Brossier et al., 2015). While Mortimer and co-workers (2018) have found that the *embB* Met306Val mutation is more associated with L4 strains, the present study contradicts this, suggesting that the abovementioned mutation to be significantly more associated with L2.

The *rpsL* Lys43Arg mutation, which confers high-level STR resistance was found to be associated primarily with L2 (Table 2). This mutation has been previously described having a significant association with Beijing strains, when compared with non-Beijing strains (Smittipat et al., 2016). The *rrs* A1401G mutation was found to be significantly associated with L2 among AMK, KAN and CAP resistant *M. tuberculosis* isolates. Altogether, *rrs* mutations have been previously described to be significantly more associated with the Beijing genotype than with non-Beijing strains, among SLID-resistant strains (Miotto et al., 2012).

Table 2. Drug resistance-associated mutations with a significant association with one of the 4 main human-adapted MTBC lineages, normalized for their respective drug-resistant *M. tuberculosis* clinical isolates. P-values <0.05 imply statistical significance.

Mutation	Drug	P-value				Relative frequency (%)			
		L1	L2	L3	L4	L1	L2	L3	L4
fabG1	INH	0.021	<0.001	<0.001	<0.001	11.3	32.8	3.2	52.6[1]
promoter_C-15T	ETH	0.001	<0.001	0.174	<0.001	11.3	32.8	3.2	52.6[1]
*kasA*_Gly269Ser	INH	<0.001	<0.001	<0.001	<0.001	0.1	0.0	0.0	99.9[1]
*rpoB*_Asp435Val	RIF	0.016	<0.001	0.001	<0.001	0.8	33.9	9.2	56.2[1]
*rpoB*_Ser450Leu	RIF	<0.001	<0.001	<0.001	<0.001	1.8	53.9[1]	3.6	40.7
*embB*_Glu378Ala	EMB	<0.001	<0.001	<0.001	<0.001	99.8[1]	0.0	0.0	0.2
*embB*_Met306Val	EMB	<0.001	<0.001	0.208	<0.001	2.9	63.5[1]	3.7	29.9
*embC*_Thr270Ile	EMB	<0.001	<0.001	<0.001	<0.001	99.8[1]	0.0	0.0	0.2
*rpsL*_Lys43Arg	STR	<0.001	<0.001	0.182	<0.001	1.9	80.6[1]	3.4	14.1
*rrs*_A1401G	AMK	0.002	0.043	0.398	<0.001	2.3	57.3[1]	3.1	37.3
	KAN	0.003	0.038	0.394	<0.001	2.3	57.3[1]	3.1	37.3
	CAP	<0.001	0.009	0.018	<0.001	2.3	57.3[1]	3.1	37.3

[1] Proportions >50 % imply a significant association with the lineage in question.

Conclusion

The MIC analysis of genomic variant profiles revealed novel resistance level dynamics with known drug resistance-associated, while also characterizing novel mutations drug resistance, like the *gid* Ala167Asp, *rrs* 1076insT, *ethA* Met1Leu, *alr* Met343Thr and Phe4Leu mutations. Their subsequent introduction in resistance-predicting WGS pipelines may improve detection of genomic resistance profiles to second-line drugs. Contrarily, the *embB* Asp354Ala, which is currently considered to be associated with EMB resistance, did not exhibit an associated MIC above the CC for EMB resistance, and it should be disregarded in future analyses.

The global genome-wide analysis revealed a higher number of resistance-associated mutations to be associated with L2 and L4. Prevalent drug resistance-associated mutations, like the *rpoB* Ser450Leu and *embB* Met306Val, which are known to confer high- and intermediate-levels of RIF and EMB, respectively, were more associated with L2 strains, a characteristic which is likely responsible for their prolific worldwide dissemination and for their overrepresentation in MDR TB populations.

References

Brossier, F., Sougakoff, W., Bernard, C., Petrou, M., Adeyema, K., Pham, A., … Veziris, N. (2015). Molecular analysis of the *embCAB* locus and *embR* gene involved in ethambutol resistance in clinical isolates of *Mycobacterium tuberculosis* in France. *Antimicrobial Agents and Chemotherapy*, *59*(8), 4800–4808. doi: 10.1128/AAC.00150-15

Cohen, K. A., Bishai, W. R., & Pym, A. S. (2014). Molecular Basis of Drug Resistance in Mycobacterium tuberculosis. *Molecular Genetics of Mycobacteria*, *2*(3), 413–429. doi: 10.1128/microbiolspec.mgm2-0036-2013

Coll, F., McNerney, R., Guerra-Assunção, J. A., Glynn, J. R., Perdigão, J., Viveiros, M., … Clark, T. G. (2014). A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nature Communications*, *5*, 4–8. doi: 10.1038/ncomms5812

Coll, F., McNerney, R., Preston, M. D., Guerra-Assunção, J. A., Warry, A., Hill-Cawthorne, G., … Clark, T. G. (2015). Rapid determination of anti-tuberculosis drug resistance from whole genome sequences. *Genome Medicine*, *7*(1), 1–10. doi: 10.1186/s13073-015-0164-0

Coll, F., Phelan, J., Hill-Cawthorne, G. A., Nair, M. B., Mallard, K., Ali, S., … Matsumoto, T. (2018). Genome-wide analysis of multi- and extensively drug-resistant. *Nature Genetics*, *50*, 307–316 doi: 10.1038/s41588-017-0029-0

Comas, I., Borrell, S., Roetzer, A., Rose, G., Malla, B., Kato-Maeda, M., … Gagneux, S. (2012). Whole genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. *Nature Genetics*, *44*(1), 106–110. doi: 10.1038/ng.1038

Dymova, M. A., Kinsht, V. N., Cherednichenko, A. G., Khrapov, E. A., Svistelnik, A. V., & Filipenko, M. L. (2011). Highest prevalence of the *Mycobacterium tuberculosis* Beijing genotype isolates in patients newly diagnosed with tuberculosis in the Novosibirsk oblast, Russian Federation. *Journal of Medical Microbiology*, *60*(7), 1003–1009. doi: 10.1099/jmm.0.027995-0

Fenner, L., Egger, M., Bodmer, T., Altpeter, E., Zwahlen, M., Jaton, K., … Gagneux, S. (2012). Effect of mutation and genetic background on drug resistance in *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*, *56*(6), 3047–3053. doi: 10.1128/AAC.06460-11

Feuerriegel, S., Köser, C., Trübe, L., Archer, J., Gerdes, S. R., Richter, E., & Niemann, S. (2010). Thr202Ala in thyA is a marker for the Latin American mediterranean lineage of the *Mycobacterium tuberculosis* complex rather than Para-aminosalicylic acid resistance. *Antimicrobial Agents and Chemotherapy*, *54*(11), 4794–4798. doi: 10.1128/AAC.00738-10

Jones, D. (2002). *Pharmaceutical Statistics*. London, United Kingdom: Pharmaceutical Press

Lau, R. W. T., Ho, P. L., Kao, R. Y. T., Yew, W. W., Lau, T. C. K., Cheng, V. C. C., … Yam, W. C. (2011). Molecular characterization of fluoroquinolone resistance in *Mycobacterium tuberculosis*: Functional analysis of *gyrA* mutation at position 74. *Antimicrobial Agents and Chemotherapy*, *55*(2), 608–614. doi: 10.1128/AAC.00920-10

Machado, D., Perdigão, J., Ramos, J., Couto, I., Portugal, I., Ritter, C., … Viveiros, M. (2013). High-level resistance to isoniazid and ethionamide in multidrug-resistant *Mycobacterium tuberculosis* of the Lisboa family is associated with *inhA* double mutations. *Journal of Antimicrobial Chemotherapy*, *68*(8), 1728–1732. doi: 10.1093/jac/dkt090

Merker, M., Blin, C., Mona, S., Duforet-frebourg, N., Lecher, S., Willery, E., … Ballif, M. (2015). Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nature Genetics*, *47*(3), 242–249. doi: 10.1038/ng.3195

Miotto, P., Cabibbe, A. M., Mantegani, P., Borroni, E., Fattorini, L., Tortoli, E., … Cirillo, D. M. (2012). GenoType MTBDRsl performance on clinical samples with diverse genetic background. *European Respiratory Journal*, *40*(3), 690–698. doi: 10.1183/09031936.00164111

Mortimer, T. D., Weber, A. M., & Pepperell, C. S. (2018). Signatures of Selection at Drug Resistance Loci in *Mycobacterium tuberculosis*. *mSystems*, *3*(1), 1–16. doi: 10.1128/mSystems.00108-17

Park, Y. K., Shin, S., Ryu, S., Sang, N. C., Koh, W. J., O, J. K., … Gill, H. B. (2005). Comparison of drug resistance genotypes between Beijing and non-Beijing family strains of *Mycobacterium tuberculosis* in Korea. *Journal of Microbiological Methods*, *63*(2), 165–172. doi: 10.1016/j.mimet.2005.03.002

Perdigão, J., Macedo, R., Fernandes, E., Brum, L. & Portugal, I. (2008). Multidrug-Resistant Tuberculosis in Lisbon, Portugal: A Molecular Epidemiological Perspective *Microbial Drug Resistance*, *14*(2), 133–143. doi: 10.1089/mdr.2008.0798

Perdigão, J., Macedo, R., Malaquias, A., Ferreira, A., Brum, L., & Portugal, I. (2010). Genetic analysis of extensively drug-resistant Mycobacterium tuberculosis strains in Lisbon, Portugal. *Journal of Antimicrobial Chemotherapy*, *65*(2), 224–227. doi: 10.1093/jac/dkp452

Perdigão, J., Macedo, R., Machado, D., Silva, C., Jordão, L., Couto, I., … Portugal, I. (2014). GidB mutation as a phylogenetic marker for Q1 cluster *Mycobacterium tuberculosis* isolates and intermediate-level streptomycin resistance determinant in Lisbon, Portugal. *Clinical Microbiology and Infection*, *20*(5), 278–284. doi: 10.1111/1469-0691.12392

Perdigão, J., Silva, H., Machado, D., Macedo, R., Maltez, F., Silva, C., … Viveiros, M. (2014). Unraveling *Mycobacterium tuberculosis* genomic diversity and evolution in Lisbon, Portugal, a highly drug resistant setting. *BMC Genomics*, *15*, 1–20. doi: 10.1186/1471-2164-15-991

Smittipat, N., Juthayothin, T., Billamas, P., Jaitrong, S., Rukseree, K., Dokladda, K., … Palittapongarnpim, P. (2016). Mutations in *rrs*, *rpsL* and *gidB* in streptomycin-resistant *Mycobacterium tuberculosis* isolates from Thailand. *Journal of Global Antimicrobial Resistance*, *4*, 5–10. doi: 10.1016/j.jgar.2015.11.009

Spies, F. S., Ribeiro, A. W., Ramos, D. F., Ribeiro, M. O., Martin, A., Palomino, J. C., … Zaha, A. (2011). Streptomycin resistance and lineage-specific polymorphisms in *Mycobacterium tuberculosis gidB* gene. *Journal of Clinical Microbiology*, *49*(7), 2625–2630. doi: 10.1128/JCM.00168-11

Sreevatsan, S., Pan, X., Stockbauer, K. E., Connell, N. D., Kreiswirth, B. N., Whittam, T. S., & Musser, J. M. (1997). Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(18), 9869–9874. doi: 10.1073/pnas.94.18.9869

World Health Organization. (2019). Global tuberculosis report 2019.

Zumla, A., Raviglione, M., Hafner, R., & Fordham von Reyn, C. (2013). Tuberculosis. *New England Journal of Medicine*, *368*(8), 745–755. doi: 10.1056/NEJMra1200894