Supplement A

Table 2. A list of 23 genes. § denotes genes suspected to confer resistance to four first-line anti-TB drugs. Starred genes contain specific loci previously documented in the literature as being associated with drug resistance.

Genes	Function	Relevant drug
§ahpC*	Oxidative stress	INH
eis*	Cell surface involvement	Aminoglycosides
embA, §embB*,embC,embR	Cell wall bisynthesis	EMB
§fabG1*	Fatty acid biosynthesis	INH
gidB	Glucose-inhibited division protein B	Streptomycin
gyrA*,gyrB	Enzyme for DNA coiling	Fluoroquinolones
§inhA	Fatty acid biosynthesis	INH
§iniA,§iniC	Likely transmembrane protein	EMB, INH
§katG	Multifunctional enzyme	INH
§manB	GDP-mannose biosynthesis	EMB
ndh	Transfer of electrons	INH
§pncA*	Intermediary metabolism	PZA
rmlD	Sugar biosynthesis	EMB
§rpoB*	Transcriptional enzyme	RIF
rpsA	Binds mRNA	
rpsL*	Translation initiation step	Streptomycin
rrs	Ribosomal RNA 16S	Aminoglycosides
tlyA	Virulence; methylation	Aminoglycosides

Supplement B

Table 3. "-L" SNP Library for Direct Association: 108 resistance-determinants were reported in T. M. Walker et al, 2015. There were 108 resistance-determinants for the interested eight drugs considered in analyses (isolates resistant to AK, CAP, KAN were too few for analysis). *** stands for omitted long sequence of amino acids that is inserted or deleted. * stands for a stop.

SNP	Drug	SNP	Drug	SNP	Drug
ahpC_C-57T	INH	embB_D354A	EMB	embB_G406A	EMB
embB_G406D	EMB	embB_G406S	EMB	embB_H1002R	EMB
embB_M306I	EMB	embB_M306V	EMB	embB_Q497K	EMB
embB_Q497R	EMB	fabG1_C-15T	INH	fabG1_G-17T	INH
fabG1_T-8C	INH	gidB_40_delG	SM	gidB_101_delC	SM
gidB_101_delC	SM	gidB_A134E	SM	gidB_A138T	SM
gidB_A138V	SM	gidB_A200E	SM	gidB_A80P	SM
gidB_G69D	SM	gidB_H48N	SM	gidB_L91P	SM
gidB_P75L	SM	gidB_R137W	SM	gidB_S70N	SM
gidB_V65G	SM	gidB_V88A	SM	gyrA_A74S	CIP
gyrA_A90V	MOX,OFX	gyrA_D94A	MOX, OFX	gyrA_D94G	CIP,MOX, OFX
gyrA_D94N	MOX, OFX	gyrA_S91P	CIP, MOX, OFX	inhA_I194T	INH
inhA_I21T	INH	inhA_S94A	INH	katG_1349_delG***	INH
katG_1809_delA***	INH	katG_370_delC	INH	katG_L159P	INH
katG_S315N	INH	katG_S315T	INH	katG_T180K	INH
katG_V633A	INH	katG_W191R	INH	katG_W300C	INH
katG_W328L	INH	katG_W90R	INH	pncA_76_delG	PZA
pncA_191_insT	PZA	pncA_308_delGTAC	PZA	pncA_390_insCC	PZA
pncA_393_del***	PZA	pncA_427_ins***	PZA	pncA_469_insT	PZA
pncA_A-11G	PZA	pncA_C138R	PZA	pncA_C14R	PZA
pncA_D12A	PZA	pncA_D136N	PZA	pncA_D49N	PZA
pncA_D8G	PZA	pncA_D8N	PZA	pncA_G132D	PZA
pncA_G162D	PZA	pncA_G78C	PZA	pncA_G97D	PZA
pncA_H57D	PZA	pncA_H57R	PZA	pncA_K96T	PZA
pncA_L172P	PZA	pncA_L27P	PZA	pncA_L4S	PZA
pncA_Q10*	PZA	pncA_Q141*	PZA	pncA_S104R	PZA
pncA_T-12C	PZA	pncA_V125G	PZA	pncA_V139L	PZA
pncA_V180F	PZA	pncA_V7L	PZA	pncA_W68C	pZA
rpoB_1296_insTTC	RIF	rpoB_1326+10_TGGCCCC	RIF	rpoB_D435F	RIF
rpoB_D435V	RIF	rpoB_H445D	RIF	rpoB_H445N	RIF
rpoB_H445R	RIF	rpoB_H445Y	RIF	rpoB_I491F	RIF
rpoB_L452P	RIF	rpoB_Q432K	RIF	rpoB_S431G	RIF
rpoB_S450F	RIF	rpoB_S450L	RIF	rpoB_S450W	RIF
rpoB_V170G	RIF	rpoB_V262A	RIF	rpoB_V359A	RIF
rpsA_A440T	PZA	rpsL_K43R	SM	rpsL_K88R	SM
rrs_A514C	SM	rrs_C513T	SM	rrs_C517T	SM

Note: katG_1809_delA*** denotes katG_1809_delACGGGTT; katG_1349_delG*** is katG_1349_delGACGAGGTCGTG;pncA_393_del*** is pncA_393_delCGACCACAT; pncA_427_ins*** is pncA_427_insGCCGTCTGGC

Supplement C

Table 4. "-D" SNP Library for Direct Association: Characterisation of mutations previously identified in the literature as resistance-determinants, which was reported in T. M. Walker et al, 2015. The 'literature' was defined as any mutation listed in the Dream TB database project. There were 68 resistance-determinants for the interested eight drugs considered. * stands for a stop.

SNP	Drug	SNP	Drug	SNP	Drug
embB_D354A	EMB	embB_G406A	EMB	embB_G406D	EMB
embB_G406S	EMB	embB_M306I	EMB	embB_M306V	EMB
embB_Q497R	EMB	embB_Q497K	EMB	fabG1_C-15T	INH
fabG1_G-17T	INH	fabG1_T-8C	INH	gidB_A134E	SM
gidB_A200E	SM	gidB_A80P	SM	gidB_R137W	SM
gyrA_A74S	CIP	gyrA_A90V	MOX, OFX	gyrA_D94A	MOX, OFX
gyrA_D94G	CIP, MOX, OFX	gyrA_D94N	MOX, OFX	gyrA_S91P	CIP, MOX, OFX
inhA_I194T	INH	inhA_I21T	INH	inhA_S94A	INH
katG_S315N	INH	katG_S315T	INH	katG_T180K	INH
katG_W191R	INH	katG_W328L	INH	pncA_A-11*	PZA
pncA_C138R	PZA	pncA_C14R	PZA	pncA_D12A	PZA
pncA_D136N	PZA	pncA_D8N	PZA	pncA_G132D	PZA
pncA_G162D	PZA	pncA_G97D	PZA	pncA_H57D	PZA
pncA_K96T	PZA	pncA_L172P	PZA	pncA_L27P	PZA
pncA_L4S	PZA	pncA_Q10*	PZA	pncA_Q141*	PZA
pncA_S104R	PZA	pncA_V125G	PZA	pncA_V139L	PZA
pncA_V180F	PZA	pncA_W68C	PZA	rpoB_D435*	RIF
rpoB_H445D	RIF	rpoB_H445N	RIF	rpoB_H445R	RIF
rpoB_H445Y	RIF	rpoB_I491F	RIF	rpoB_L452P	RIF
rpoB_Q432K	RIF	rpoB_S431G	RIF	rpoB_S450D	RIF
rpoB_S450F	RIF	rpoB_S450L	RIF	rpoB_S450W	RIF
rpsL_K43R	SM	rpsL_K88R	SM	rrs_A514C	SM
rrs_C517T	SM	pncA_G97D	PZA		

Supplement D

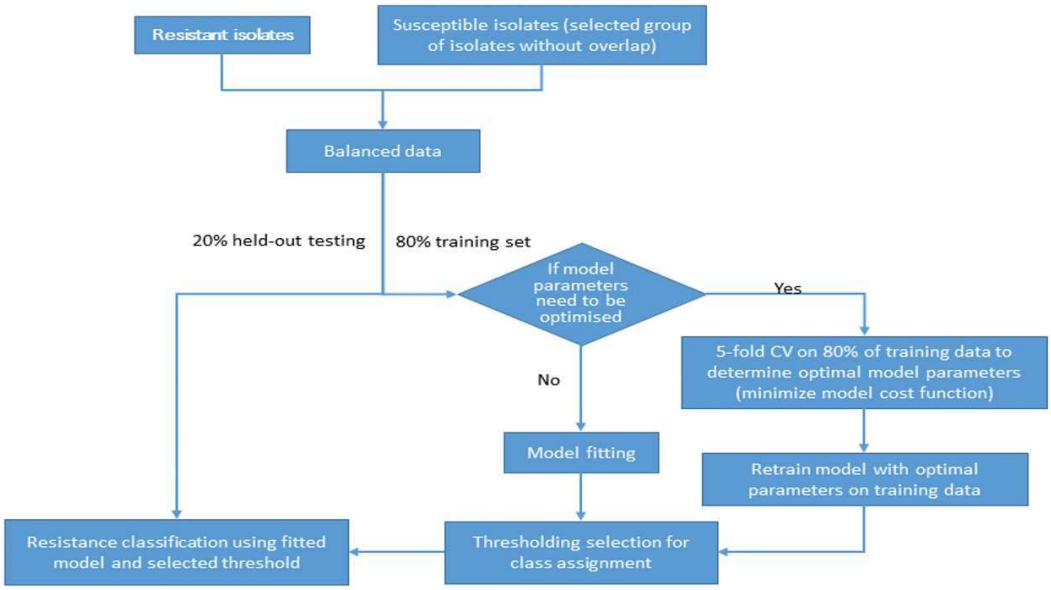

Fig. 4: Flowchart for examined classifiers. There are three main steps shown in the figure: 1) Assembling balanced data sets; 2) Training a classifier; 3) Testing a classifier. In step 1, we randomly divided all susceptible isolates into a series of groups, the number of which equals to the number of resistant isolates, to avoid bias in the classifier. The number of experiments equals to the number of the groups of susceptible isolates with respect to every drug. In step 2, the hyperparameters of the supervised models were determined based on internal five-fold cross-validation on 80% of the training data. The optimised parameters were then used to train a final model using all the training data. The decision threshold was determined as the point on the ROC curve that maximised accuracy. In step 3, the trained model and decision threshold were used for classification on the "held-out" 20% of data in the test set. All performance values were averaged over iterations determined by number ratio of susceptible over resistant isolates.

Supplement E

Table 5. Classification methods

Method	Details
LR and SVM	LR and SVM-L2 predictive modelling was performed using the LIBLINEAR library Version 2.1. SVM-RBF was performed using the SVM toolbox in Matlab.
RF	A committee of 40-400 weak"base learners", or trees, was built using a random selection of half of the features, which has been found to be a suitable means of initialising the various parameters for the problems involving genomic loci as features. It was performed using the TreeBagger toolbox in Matlab.
PM	This model is based on the assumption of independent conditional probability between the input variables. In addition, it uses a Dirichlet and Beta prior on the probability of each class and the probability of features in each class, respectively. A Beta(0.5,0.5) prior was used for every SNP, except for the resistance determinants. For the resistance-determinants, a Beta(1,0.25) prior for the resistant class, and a Beta(0.25,1) prior for the susceptible class, was used.
CBMM	This models each class with a multivariate Bernoulli mixture model and uses Bayes rule to classify. The number of the mixture components in each class was learned in the cross-validation iterations within the training stage using grid search. The new examples were assigned to the class with highest posterior probability.

Supplement F

Table 6. Confusion table for direct association (DA). -D and -L correspond to the two libraries described in Methods; R and S correspond to resistant and susceptible populations, respectively.

INH		Phenotype R	Phenotype S
DA-D	R	245	9
	S	21	1536
DA-L	R	248	11
	S	18	1534

RIF		Phenotype R	Phenotype S
DA-D	R	90	28
	S	7	1600
DA-L	R	91	28
	S	6	1600

EMB		Phenotype R	Phenotype S
DA-D	R	45	50
	S	2	1647
DA-L	R	45	52
	S	2	1645

CIP		Phenotype R	Phenotype S
DA-D	R	23	3
	S	3	267
DA-L	R	23	3
	S	3	267

MOX		Phenotype R	Phenotype S
DA-D	R	16	4
	S	4	65
DA-L	R	16	4
	S	4	65

OFX		Phenotype R	Phenotype S
DA-D	R	15	3
	S	4	65
DA-L	R	15	3
	S	4	65

PZA		Phenotype R	Phenotype S
DA-D	R	36	1
	S	23	1665
DA-L	R	41	2
	S	18	1664

SM		Phenotype R	Phenotype S
DA-D	R	22	1
	S	19	325
DA-L	R	26	4
	S	15	322

MDR		Phenotype R	Phenotype S
DA-D	R	71	6
	S	10	1621
DA-L	R	73	7
	S	8	1620

Supplement G

Table 7. Comparison of predictive performance for INH resistance using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	92 † ± 0.3	100 ± 0.1	96 † ± 0.0	92 ± 0.3	99 ± 0.1	96 † ± 0.0	92 ± 0.4	99 ± 0.1	96 † ± 0.0
DA-L	93 ± 0.3	99 ± 0.1	96 ± 0.0	93 ± 0.3	99 ± 0.1	96 ± 0.0	93 ± 0.4	99 ± 0.1	96 ± 0.0
LR-L1	96 † ± 0.3	97 † ± 0.3	99 † ± 0.1	93 ± 0.3	99 ± 0.2	96 ± 0.1	91 † ± 0.4	99 † ± 0.2	96 † ± 0.0
LR-L2	95 † ± 0.3	96 † ± 0.3	98 † ± 0.1	95 † ± 0.3	97 † ± 0.2	97 † ± 0.0	90 † ± 0.5	99 † ± 0.2	95 † ± 0.0
SVM-L2	94 ± 0.4	97 † ± 0.3	98 † ± 0.1	95 † ± 0.3	97 † ± 0.2	97 † ± 0.0	90 † ± 0.5	99 † ± 0.2	95 † ± 0.0
SVM-RBF	96 † ± 0.4	95 † ± 0.4	99 † ± 0.0	96 † ± 0.2	97 † ± 0.3	97 † ± 0.1	91 † ± 0.4	98 † ± 0.3	98 † ± 0.0
RF	**97** † ± **0.3**	**94** † ± **0.4**	**99** † ± **0.0**	95 † ± 0.3	97 † ± 0.2	97 † ± 0.0	92 ± 0.4	97 † ± 0.3	96 † ± 0.1
PM	96 † ± 0.4	94 † ± 0.4	99 † ± 0.0	96 † ± 0.3	97 † ± 0.2	97 † ± 0.0	94 ± 0.4	97 † ± 0.3	97 † ± 0.1
CBMM	93 ± 0.4	93 † ± 0.4	97 † ± 0.1	95 † ± 0.3	96 † ± 0.3	98 † ± 0.0	91 † ± 0.5	95 † ± 0.4	97 † ± 0.0

Table 8. Comparison of predictive performance for RIF resistance using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01).The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	93 ± 0.5	98 ± 0.3	96 † ± 0.1	93 ± 0.6	99 ± 0.3	96 † ± 0.1	94 ± 0.5	98 ± 0.3	96 † ± 0.1
DA-L	94 ± 0.5	98 ± 0.3	96 ± 0.1	94 ± 0.5	99 ± 0.3	96 ± 0.1	95 ± 0.5	98 ± 0.3	96 ± 0.1
LR-L1	96 † ± 0.5	94 † ± 0.6	97 † ± 0.1	94 ± 0.5	97 † ± 0.5	97 † ± 0.1	95 ± 0.5	98 ± 0.3	98 † ± 0.1
LR-L2	91 ± 0.9	93 † ± 0.7	96 ± 0.2	96 † ± 0.6	94 † ± 0.6	97 † ± 0.1	92 † ± 0.7	98 ± 0.3	95 † ± 0.2
SVM-L2	90 ± 1.0	92 † ± 0.8	97 † ± 0.2	96 ± 0.5	94 † ± 0.6	97 † ± 0.1	93 ± 0.7	98 ± 0.3	96 ± 0.2
SVM-RBF	97 † ± 0.4	92 † ± 0.7	98 † ± 0.1	97 † ± 0.4	96 † ± 0.5	98 † ± 0.1	98 † ± 0.3	95 † ± 0.5	98 † ± 0.1
RF	90 † ± 0.9	91 † ± 0.8	97 † ± 0.1	94 ± 0.5	97 ± 0.5	98 † ± 0.1	95 ± 0.5	98 ± 0.3	93 † ± 0.4
PM	93 ± 0.7	95 † ± 0.6	98 † ± 0.1	95 ± 0.4	97 ± 0.4	99 † ± 0.1	97 † ± 0.4	98 ± 0.3	98 † ± 0.1
CBMM	90 † ± 0.8	91 † ± 0.9	96 ± 0.1	95 ± 0.5	97 † ± 0.4	99 † ± 0.1	**97** † ± **0.4**	**97** ± **0.4**	**99** † ± **0.1**

Table 9. Comparison of predictive performance for EMB resistance using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	95 ± 0.7	97 ± 0.6	96 ± 0.1	95 ± 0.7	98 ± 0.4	97 ± 0.1	96 ± 0.6	97 ± 0.6	96 ± 0.1
DA-L	95 ± 0.7	97 ± 0.6	96 ± 0.1	95 ± 0.7	98 ± 0.4	97 ± 0.1	96 ± 0.6	97 ± 0.6	96 ± 0.1
LR-L1	91 † ± 1.0	96 ± 0.6	97 ± 0.2	94 ± 0.9	96 † ± 0.7	97 ± 0.2	93 ± 0.8	98 ± 0.5	97 † ± 0.2
LR-L2	88 † ± 1.3	94 † ± 1.0	96 ± 0.2	96 ± 1.0	97 ± 0.6	98 † ± 0.2	95 ± 0.7	97 ± 0.7	97 † ± 0.1
SVM-L2	87 † ± 1.5	94 ± 0.9	97 ± 0.2	96 ± 0.9	97 ± 0.6	98 † ± 0.2	94 ± 0.9	97 ± 0.5	98 † ± 0.2
SVM-RBF	94 ± 0.9	92 † ± 0.9	98 † ± 0.1	97 † ± 0.7	94 † ± 0.8	98 † ± 0.1	94 ± 1.1	96 ± 0.7	96 ± 0.1
RF	89 † ± 1.2	93 † ± 0.9	97 † ± 0.1	95 ± 0.9	97 ± 0.5	99 † ± 0.1	90 † ± 1.3	96 ± 0.8	84 † ± 0.9
PM	89 † ± 1.3	95 ± 0.8	98 † ± 0.1	96 † ± 1.1	97 ± 0.6	99 † ± 0.1	91 † ± 1.1	98 ± 0.6	97 † ± 0.2
CBMM	88 † ± 1.4	93 † ± 0.9	97 † ± 0.1	**97** † ± **1.0**	**96** † ± **0.6**	**99** † ± **0.1**	91 † ± 1.0	97 ± 0.6	97 ± 0.2

Table 10. Comparison of predictive performance for PZA resistance using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	57 †± 1.5	100 ± 0.0	79 †± 0.0	59 †± 1.4	100 ± 0.0	79 †± 0.0	57 †± 1.4	100 ± 0.0	79 †± 0.0
DA-L	69 ± 1.4	100 ± 0.0	85 ± 0.0	70 ± 1.4	100 ± 0.0	85 ± 0.0	69 ± 1.3	100 ± 0.0	85 ± 0.0
LR-L1	82 †± 1.2	88 †± 1.1	90 †± 0.2	81 †± 1.3	93 †± 0.8	88 †± 0.3	50 †± 1.4	100 ± 0.0	75 †± 0.2
LR-L2	80 †± 1.3	87 †± 1.2	89 †± 0.3	88 †± 1.0	89 †± 1.0	90 †± 0.2	50 †± 1.4	100 ± 0.0	75 †± 0.2
SVM-L2	80 †± 1.6	88 †± 1.0	90 †± 0.3	87 †± 1.1	89 †± 1.0	90 †± 0.2	50 †± 1.4	100 ± 0.0	75 †± 0.2
SVM-RBF	89 †± 1.1	85 †± 1.3	94 †± 0.2	89 †± 1.0	88 †± 1.1	92 †± 0.2	75 †± 1.5	98 †± 0.4	88 †± 0.2
RF	81 †± 1.2	87 †± 1.2	92 †± 0.2	85 †± 1.1	90 †± 0.9	89 †± 0.2	50 †± 1.4	100 ± 0.0	71 †± 0.3
PM	**84** †± **1.2**	**90** †± **1.1**	**95** †± **0.2**	90 †± 1.1	90 †± 0.9	93 †± 0.2	72 ± 1.3	100 ± 0.1	87 †± 0.1
CBMM	82 †± 1.2	87 †± 1.1	93 †± 0.2	88 †± 1.1	90 †± 0.9	93 †± 0.2	75 †± 1.5	98 †± 0.4	87 †± 0.1

Table 11. Comparison of predictive performance for CIP resistance using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	87 ± 1.0	99 ± 0.4	94 ± 0.1	88 ± 1.1	99 ± 0.3	94 ± 0.1	88 ± 1.0	99 ± 0.3	94 ± 0.1
DA-L	87 ± 1.0	99 ± 0.4	94 ± 0.1	88 ± 1.1	99 ± 0.3	94 ± 0.1	88 ± 1.0	99 ± 0.3	94 ± 0.1
LR-L1	88 ± 1.4	94 †± 0.8	96 †± 0.3	89 ± 1.3	95 †± 0.9	93 †± 0.3	86 ± 1.4	94 †± 1.0	96 †± 0.3
LR-L2	87 ± 1.5	98 ± 0.5	94 †± 0.5	92 ± 1.2	96 †± 0.8	94 ± 0.3	87 ± 1.5	93 †± 1.0	95 †± 0.5
SVM-L2	85 ± 1.4	98 ± 0.7	96 †± 0.4	92 ± 1.1	96 †± 0.7	94 ± 0.4	87 ± 1.5	94 †± 1.0	95 †± 0.4
SVM-RBF	86 ± 1.5	99 ± 0.5	97 †± 0.3	92 †± 1.6	91 †± 1.3	98 †± 0.2	85 ± 1.6	91 †± 1.6	93 ± 0.5
RF	87 ± 1.4	92 †± 1.2	88 †± 1.0	95 †± 1.1	99 ± 0.5	96 †± 0.4	86 ± 1.4	93 †± 1.1	93 ± 0.5
PM	88 ± 1.4	96 ± 1.1	97 †± 0.3	**96** †± **0.9**	**98** ± **0.4**	**98** †± **0.3**	89 ± 1.5	95 †± 1.0	97 †± 0.3
CBMM	86 ± 1.6	96 ± 1.1	95 †± 0.3	95 †± 1.1	95 †± 0.8	98 †± 0.1	86 ± 1.4	94 †± 1.0	96 †± 0.3

Table 12. Comparison of predictive performance for MOX resistance using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	83 ± 1.4	93 ± 0.8	87 ± 0.1	80 ± 1.5	93 ± 0.9	87 ± 0.1	79 ± 1.6	96 ± 0.7	87 ± 0.2
DA-L	83 ± 1.4	93 ± 0.8	87 ± 0.1	80 ± 1.5	93 ± 0.9	87 ± 0.1	79 ± 1.6	96 ± 0.7	87 ± 0.2
LR-L1	70 †± 1.8	91 ± 1.1	87 ± 0.6	73 †± 1.9	88 ± 1.7	79 †± 0.7	88 †± 1.7	90 †± 1.1	91 †± 0.6
LR-L2	69 †± 1.9	84 †± 1.7	80 †± 0.8	78 ± 1.8	85 †± 1.8	80 †± 0.6	90 †± 1.6	90 †± 1.2	87 ± 0.9
SVM-L2	69 †± 1.8	86 †± 1.4	83 †± 0.6	76 ± 2.0	85 †± 1.9	79 †± 0.6	91 †± 1.4	91 †± 1.2	90 †± 0.7
SVM-RBF	64 †± 2.3	96 †± 0.8	85 ± 0.7	62 †± 2.7	94 ± 1.0	88 †± 0.6	89 †± 2.1	84 †± 1.6	93 †± 0.5
RF	71 †± 2.1	84 †± 1.7	77 †± 1.0	80 ± 1.8	90 ± 1.4	85 ± 0.5	90 †± 1.6	90 †± 1.3	88 ± 0.8
PM	74 †± 2.1	87 ± 1.8	89 †± 0.4	85 ± 1.7	88 ± 1.6	88 †± 0.4	**95** †± **1.4**	**93** ± **1.0**	**95** †± **0.4**
CBMM	71 †± 2.1	87 ± 1.8	87 ± 0.5	79 ± 1.9	85 †± 1.8	85 ± 0.6	88 †± 1.8	90 †± 1.3	92 †± 0.5

Table 13. Comparison of predictive performance for OFX resistance using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	81 ± 1.5	95 ± 0.9	87 ± 0.3	77 ± 1.4	96 ± 0.9	87 ± 0.3	78 ± 1.5	95 ± 1.1	87 ± 0.3
DA-L	81 ± 1.5	95 ± 0.9	87 ± 0.3	77 ± 1.4	96 ± 0.9	87 ± 0.3	78 ± 1.5	95 ± 1.1	87 ± 0.3
LR-L1	67 †± 2.2	90 ± 1.6	83 †± 0.7	75 ± 1.8	86 †± 2.0	80 †± 0.7	85 †± 2.3	91 †± 1.2	91 †± 0.6
LR-L2	64 †± 2.2	78 †± 2.2	75 †± 0.8	79 ± 1.7	86 †± 2.0	80 †± 0.6	86 †± 2.0	92 ± 1.2	82 ± 1.5
SVM-L2	62 †± 2.2	82 †± 2.1	77 †± 0.8	80 ± 1.8	84 †± 2.2	80 †± 0.7	87 †± 2.1	93 ± 1.2	85 ± 1.1
SVM-RBF	63 †± 2.3	87 ± 2.5	78 †± 0.6	75 ± 2.3	89 †± 2.0	86 ± 0.9	92 †± 2.1	78 †± 2.7	93 †± 0.7
RF	65 †± 2.3	76 †± 2.3	69 †± 1.3	81 ± 1.7	88 †± 1.8	84 †± 0.5	88 †± 2.0	90 †± 1.3	88 ± 0.9
PM	63 †± 2.2	89 †± 1.6	83 †± 0.9	84 †± 1.5	89 †± 1.9	86 ± 0.6	**96** †± **1.4**	**92** ± **1.3**	**95** †± **0.5**
CBMM	62 †± 2.3	85 †± 2.3	79 †± 0.7	79 ± 1.9	81 †± 2.5	85 ± 0.7	85 †± 2.0	89 †± 1.5	92 †± 0.6

Table 14. Comparison of predictive performance for SM resistance using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	52 †± 2.0	99 ± 0.5	77 †± 0.0	56 †± 2.0	100 ± 0.2	77 †± 0.0	52 †± 1.6	100 †± 0.1	77 †± 0.0
DA-L	63 ± 1.8	98 ± 0.6	81 ± 0.1	65 ± 1.8	99 ± 0.4	81 ± 0.1	62 ± 1.5	99 ± 0.4	81 ± 0.1
LR-L1	83 †± 1.2	88 †± 1.1	89 †± 0.4	76 †± 1.6	94 †± 0.8	86 †± 0.3	59 ± 1.7	93 †± 1.1	72 †± 0.3
LR-L2	85 †± 1.2	89 †± 1.1	88 †± 0.3	81 †± 1.5	92 †± 1.0	88 †± 0.1	61 ± 1.6	89 †± 1.2	74 †± 0.5
SVM-L2	84 †± 1.3	88 †± 1.2	88 †± 0.3	81 †± 1.6	92 †± 1.0	88 †± 0.2	60 ± 1.7	90 †± 1.1	73 †± 0.4
SVM-RBF	80 †± 1.3	86 †± 1.5	91 †± 0.3	**87** †± **1.5**	**90** †± **1.0**	**91** †± **0.3**	73 †± 1.9	80 †± 1.5	81 ± 0.3
RF	85 †± 1.2	89 †± 1.2	87 †± 0.3	81 †± 1.5	92 †± 0.9	86 †± 0.2	60 ± 1.8	86 †± 1.3	66 †± 0.7
PM	86 †± 1.1	87 †± 1.3	89 †± 0.2	84 †± 1.5	92 †± 0.9	91 †± 0.2	76 †± 1.4	80 †± 1.4	84 †± 0.3
CBMM	83 †± 1.6	83 †± 1.7	89 †± 0.3	85 †± 1.4	91 †± 1.0	91 †± 0.3	66 ± 1.8	78 †± 1.8	77 †± 0.3

Table 15. Comparison of predictive performance for MDR using all methods with all three feature sets. Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Methods	F1			F2			F3		
	Sens	Spec	AUC	Sens	Spec	AUC	Sens	Spec	AUC
DA-D	88 ± 0.8	100 ± 0.2	94 †± 0.0	87 ± 0.9	100 ± 0.2	94 †± 0.0	87 ± 0.8	100 ± 0.1	94 †± 0.0
DA-L	90 ± 0.7	100 ± 0.2	95 ± 0.0	90 ± 0.7	99 ± 0.2	95 ± 0.0	90 ± 0.7	100 ± 0.2	95 ± 0.0
LR-L1	96 †± 0.6	96 †± 0.5	98 †± 0.1	96 †± 0.7	95 †± 0.7	98 †± 0.1	96 †± 0.7	96 †± 0.6	97 †± 0.1
LR-L2	91 ± 0.9	95 †± 0.5	97 †± 0.1	98 †± 0.4	94 †± 0.6	97 †± 0.1	96 †± 0.7	96 †± 0.6	97 †± 0.1
SVM-L2	90 ± 1.1	95 †± 0.6	97 †± 0.1	97 †± 0.6	94 †± 0.7	98 †± 0.1	95 †± 0.7	96 †± 0.5	97 †± 0.1
SVM-RBF	97 †± 0.6	92 †± 0.7	99 †± 0.1	97 †± 0.6	93 †± 0.7	99 †± 0.1	98 †± 0.5	94 †± 0.7	99 †± 0.0
RF	92 ± 0.8	91 †± 0.7	97 †± 0.1	96 †± 0.8	95 †± 0.6	99 †± 0.1	94 †± 0.8	96 †± 0.6	97 †± 0.2
PM	94 †± 0.9	94 †± 0.7	98 †± 0.1	94 †± 0.7	96 †± 0.6	99 †± 0.1	**96** †± **0.6**	**98** †± **0.5**	**100** †± **0.1**
CBMM	90 ± 1.1	92 †± 0.7	96 †± 0.1	96 †± 0.8	94 †± 0.6	99 †± 0.0	95 †± 0.7	96 †± 0.5	99 †± 0.1

Supplement H

Table 16. Resistance-conferring mutation candidates, with the number of false negative isolates (the susceptible isolates that were classified to be resistant by all machine learning classifiers) harboring the mutation, n_{fn}, followed by n_{res} in parenthesis, which indicates that in addition to the noted number of false negative isolates with the mutation, the number resistant isolate containing an established mutation also shared the mutation. These candidates are defined as those (a) found in at least two isolates resistant to the given drug, (b) found in at least one isolate lacking any alternative established resistance-conferring mutation, and (c) having a positive predictive value of 1.0 (i.e., every isolate with the SNP was classified as drug-resistant by all machine learning classifiers).

Drug	SNPs	$n_{fn}(n_{res})$	Drug	SNPs	$n_{fn}(n_{res})$
INH	katG_1899_insG	2(2)	RIF	rpoB_1298_ins***	3(3)
PZA	ahpC_T-42C	2(5)	CIP,MOX,OFX	katG_V473L	3(14)
	gidB_A27B	2(2)			
	gyrA_S91P	3(4)			
	rpoB_L452P	2(5)			

Note: rpoB_1298_ins*** denotes rpoB_1298_insCTTCATGGACCAGAAC

Supplement I

Table 17. Comparing performance between best classifier and DA-L for resistance prediction with 8 drugs and MDR-TB within subclade C1 (Beijing, EuroAmer, LAM, Tur and Uganda.) Sensitivity (sens) and specificity (spec) are shown with AUC, where results are reported as mean and standard error. †represents that p-value is lower than 0.01 (p<0.01). The p-value of performance measurement of the examined classifier compared to the DA-L was obtained by Wilcoxon signed-rank test.

Drug	Clades	DA			Best classifier			
		Sens	Spec	AUC	(Feature set)	Sens	Spec	AUC
INH	C1	93 ± 0.5	100 ± 0.2	96 ± 0.0	PM(F1)	96 †± 0.6	94 †± 0.6	99 †± 0.0
	Delhi_CAS	96 ± 0.4	100 ± 0.1	98 ± 0.0	PM(F1)	97 †± 0.5	97 †± 0.6	100 †± 0.0
	EAI	98 ± 0.4	99 ± 0.3	98 ± 0.1	LR-L1(F1)	96 ± 0.9	97 ± 0.8	99 †± 0.1
EMB	C1	92 ± 1.0	98 ± 0.9	95 ± 0.1	PM(F2)	94 †± 1.5	97 ± 0.7	99 †± 0.1
	Beijing	90 ± 1.4	97 ± 0.7	93 ± 0.1	RF(F2)	92 ± 1.8	97 ± 1.1	99 †± 0.2
RIF	C1	88 ± 0.9	99 ± 0.3	94 ± 0.1	PM(F2)	95 †± 0.7	97 †± 0.6	100 †± 0.0
	Beijing	84 ± 1.3	100 ± 0.0	61 ± 4.6	PM(F2)	96 †± 0.8	98 †± 0.6	67 †± 5.0
PZA	C1	78 ± 1.4	100 ± 0.0	90 ± 0.0	PM(F2)	86 ± 2.0	95 †± 1.2	96 †± 0.2
MDR	C1	85 ± 1.0	99 ± 0.3	91 ± 0.1	PM(F2)	98 †± 0.6	93 †± 0.8	99 †± 0.1
	Beijing	83 ± 1.4	100 ± 0.2	91 ± 0.1	SVM-RBF(F3)	96 †± 1.7	91 †± 1.9	99 †± 0.1

Supplement J

Table 18. The SNPs with high posterior probability using product of marginals for PZA and SM. The SNPs were selected with higher posterior in resistant class and lower posterior in susceptible class (For SM, the thresholds were set to 0.1 for both resistant and susceptible classes, respectively; For PZA, the thresholds were 0.05 for susceptible class and 0.08 for resistant class). We reported the SNPs pulled out from training model in descend order associated with the number of training samples that have the SNPs, n_{tot}, the number of training samples resistant to PZA that have the SNPs, n_{res}, and the posterior log probability in resistant class, p_{res}. The highlighted SNPs are known resistance-determinants for PZA and SM, respectively. We selected the model with the best test results to pull out the interested SNP, and reported the corresponding test performance. At this stage, no epistasis effect is considered. We note that these SNPs are statistically relevant for resistance classification instead of being associated to resistance genetically for the given drug.

PZA PM(F1) Test [sen=100%, spec=100%,auc=100%]		SM PM(F1) Test [Sen = 94%, Spec = 90%, AUC= 100%]	
SNPs	$n_{tot}(n_{res})$	SNPs	$n_{tot}(n_{res})$
pncA_H57D	9(9)	gyrA_E21Q	31(29)
embB_G406S	7(7)	gyrA_G668D	28(26)
gidB_E92D	9(8)	gyrA_S95T	27(25)
rpsA_A440T	7(7)	katG_S315T	22(22)
embB_N13S	7(7)	rpoB_S450L	16(16)
iniA_N88S	7(7)	rpsL_K43R	11(11)
manB_-51_delAGTGAACTGCGC	7(7)	embB_Q497R	5(5)
gyrA_D94G	7(6)	gyrA_D94G	6(5)
ndh_R284W	6(6)	embB_G406S	4(4)
rpsL_K43R	6(6)	**rpsL_K88**	3(3)
katG_V473L	5(5)	manB_T-56N	5(5)
rpoB_ I491F	5(5)		
ahpC_T-42C	4(4)		
rpoB_L452P	4(4)		
pncA_L4S	3(3)		
pncA_V125G	3(3)		

Supplement K

Table 19. The selected SNPs based on variable importance measures in random forest given different drugs (we only list the selected SNPs within suspect genes for given drugs). The SNPs were selected using random forest with best classification on testing set. The highlighted SNPs are established resistance-determinant of given drugs.

INH	RIF	EMB	PZA
katG_S315T	**rpoB_S450L**	**embB_M306V**	**pncA_H57D**
fabG1_C-15T	**rpoB_H445Y**	**embB_M306I**	**rpsL_A440T**
katG_1899_insG	**rpoB_H445D**	**embB_G406S**	**pncA_H57R**
ahpC_T-42C	**rpoB_H445N**	**embB_Q497R**	**pncA_A-11G**
fabG1_G-47C	rpoB_C-61T		**pncA_T-12C**
ndh_G-70T	**rpoB_I491F**		**pncA_L4S**
ahpC_G-88A	**rpoB_S450W**		
ahpC_G32D	**rpoB_L452P**		
fabG1_T-8C	**rpoB_I480V**		
katG_G-76A	**rpoB_H445R**		
	rpoB_S450F		

MOX,OFX	CIP	SM	MDR
gyrA_D94G	**gyrA_D94G**	**rpsL_K43R**	**rpoB_S450L**
gyrA_S91P	gyrA_A384V	gidB_E92D	**katG_S315T**
gyrA_A90V	gyrA_D641E	**rpsL_K88R**	**rpoB_H445Y**
		gidB_S100F	ahpC_T-42
		rrs_C1257T	rpoB_1299insC
			rpoB_S450F
			rpoB_L452P
			fabG1_C-15T
			rpoB_C-61T
			rpoB_D435V
			rpoB_I480V
			katG_V473L