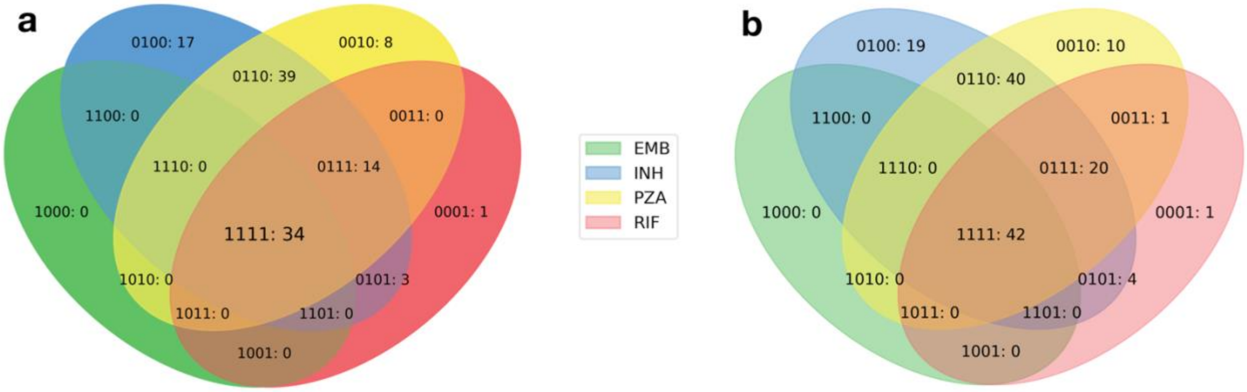

Fig. S1. Venn diagrams showing how the four sets of selected features overlap with each other. These features were selected for training 1D CNN models for the four first-line drugs. **a** Only variant features. **b** All types of features. These Venn diagrams were generated by using a python package called pyvenn (https://github.com/tctianchi/pyvenn)

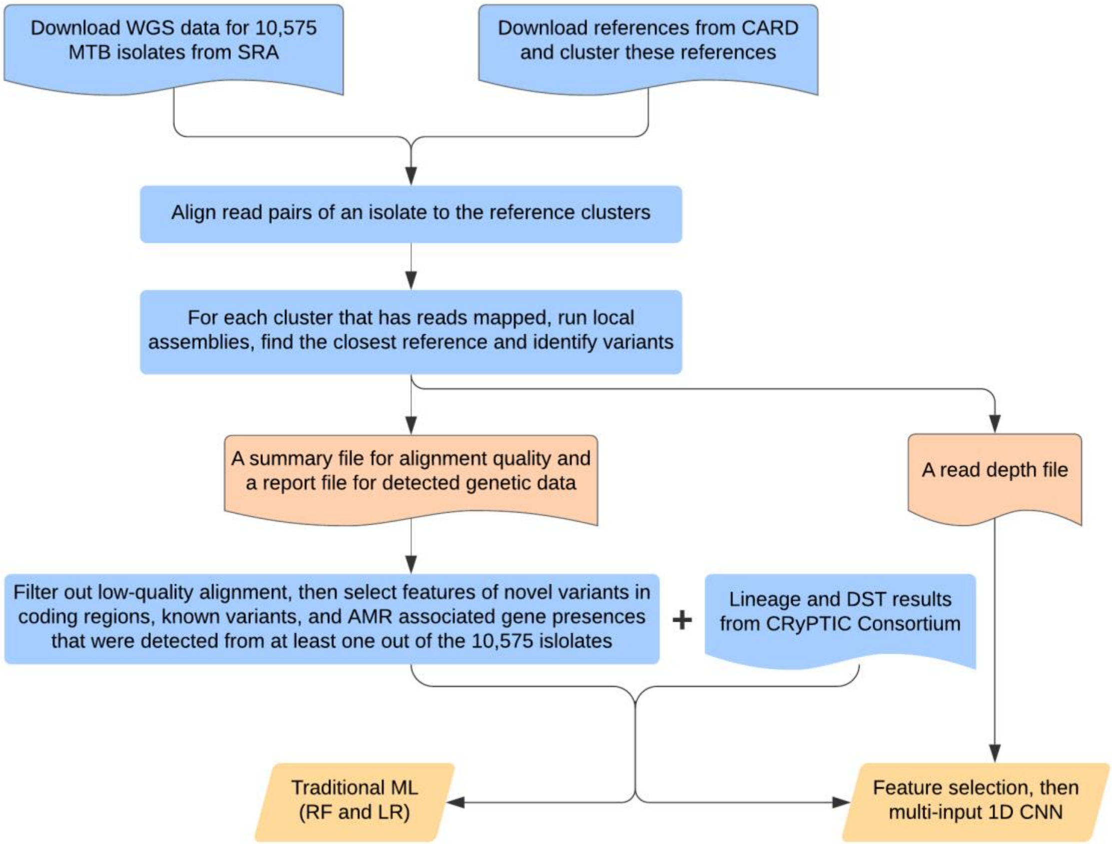

Fig. S2. Flowchart of our ML model development for MTB drug resistance classification.

Table S1. Selected variant features for 1D CNN model training of the eight (first-line and second-line) anti-TB drugs. Variants that are also in the list of AMR-associated TB mutations recently published by WHO are highlighted in bold.

Drugs	Selected variants	Covered by WHO (%)
	First-line	
RIF	**embB.D1024N** **embB.G406A** **embB.G406D** **embB.G406S** **embB.M306I** **embB.M306L** **embB.M306V** **embB.Q497K**	82.35%

	embB.Q497R **embC.R738Q** **embC.T270I** **embC.V981L** **gyrA.A90V** **gyrA.D94A** **gyrA.D94G** **gyrA.D94N** **gyrA.D94Y** **gyrA.S91P** **inhA.I194T** **inhA.I21T** **inhA.S94A** **katG.R463L** **katG.S315N** **katG.S315T** **rpoB.S450L** **rpsL.K43R** **rpsL.K88R** **Mycobacterium_tuberculosis_16S.A1401G** katG.A234G katG.G300W katG.A431V kasA.G269S Mycobacterium_avium_23S.A2274G gidB.A134E	
EMB	**embA.P913S** **embB.D1024N** **embB.D328G** **embB.D328Y** **embB.E378A** **embB.G406A** **embB.G406C** **embB.G406D** **embB.G406S** **embB.M306I** **embB.M306L** **embB.M306V** **embB.Q497K** **embB.Q497R** **embB.S297A** **embB.Y319C** **embB.Y334H** **embC.N394D** **embC.R738Q** **embC.T270I**	79.44 %

	embC.V981L embR.C110Y ethA.A381P gyrA.A90V gyrA.D89G gyrA.D89N gyrA.D94A gyrA.D94G gyrA.D94N gyrA.D94Y gyrA.G88A gyrA.G88C gyrA.S91P inhA.I194T inhA.I21T inhA.I21V inhA.S94A katG.Q127P katG.R463L katG.S140N katG.S315G katG.S315I katG.S315N katG.S315T katG.Y337C ndh.R268H ndh.V18A pncA.A146T pncA.C14R pncA.D12A pncA.D12N pncA.D49A pncA.F58L pncA.G132S pncA.G97S pncA.H82R pncA.L172P pncA.L27P pncA.L4W pncA.L85P pncA.L85R pncA.M175V pncA.P54L pncA.Q10P pncA.Q10R pncA.Q141P	

	pncA.R154G **pncA.S67P** **pncA.T135P** **pncA.T160P** **pncA.T47A** **pncA.T76P** **pncA.V139A** **pncA.V155G** **pncA.V7G** **pncA.V9G** **pncA.W68G** **pncA.W68R** **pncA.Y103H** **rpoB.S450L** **rpsL.K43R** **rpsL.K88Q** **rpsL.K88R** **Mycobacterium_tuberculosis_16S.A1401G** **tlyA.N236K** katG.A234G katG.G300W katG.A431V rpoB.L511R rpoB.D516G rpoB.H526T kasA.G312S kasA.G269S thyA.T202A gidB.S70R Mycobacterium_avium_23S.A2274G folC.A420V iniA.S501W gidB.P75S gidB.A134E folC.S150G gidB.A138E folC.I43S thyA.H207R gidB.H48Y folC.I43A katG.P131Q	
INH	**embA.P913S** **embB.D1024N** **embB.D328Y** **embB.E378A** **embB.G406A**	84.62%

	embB.G406C **embB.G406D** **embB.G406S** **embB.M306I** **embB.M306L** **embB.M306V** **embB.Q497K** **embB.Q497R** **embC.N394D** **embC.R738Q** **embC.T270I** **embC.V981L** **embR.C110Y** **ethA.A381P** **gyrA.A90V** **gyrA.D94A** **gyrA.D94G** **gyrA.D94N** **gyrA.D94Y** **gyrA.G88A** **gyrA.S91P** **inhA.I194T** **inhA.I21T** **inhA.I21V** **inhA.S94A** **katG.G279D** **katG.Q127P** **katG.R463L** **katG.S315G** **katG.S315I** **katG.S315N** **katG.S315T** **ndh.R268H** **ndh.V18A** **pncA.H82R** **rpoB.S450L** **rpsL.K43R** **rpsL.K88R** **Mycobacterium_tuberculosis_16S.A1401G** katG.A234G katG.G300W katG.A431V kasA.G269S thyA.T202A Mycobacterium_avium_23S.A2274G gidB.A134E	

	thyA.H207R	
PZA	**embA.P913S** **embB.D1024N** **embB.D328Y** **embB.E378A** **embB.G406A** **embB.G406C** **embB.G406D** **embB.G406S** **embB.M306I** **embB.M306L** **embB.M306V** **embB.Q497K** **embB.Q497R** **embB.S297A** **embB.Y319C** **embB.Y334H** **embC.N394D** **embC.R738Q** **embC.T270I** **embC.V981L** **embR.C110Y** **ethA.A381P** **ethA.D58A** **gyrA.A90V** **gyrA.D89G** **gyrA.D94A** **gyrA.D94G** **gyrA.D94N** **gyrA.D94Y** **gyrA.G88A** **gyrA.G88C** **gyrA.S91P** **inhA.I194T** **inhA.I21T** **inhA.I21V** **inhA.S94A** **katG.R463L** **katG.S315N** **katG.S315T** **ndh.R268H** **ndh.V18A** **pncA.A134V** **pncA.D12A**	82.11%

	pncA.D12N **pncA.D49A** **pncA.F58L** **pncA.G132S** **pncA.H51Q** **pncA.H57D** **pncA.H82R** **pncA.L159R** **pncA.L172P** **pncA.L27P** **pncA.L4W** **pncA.L85P** **pncA.L85R** **pncA.M175V** **pncA.P54L** **pncA.Q10P** **pncA.Q10R** **pncA.Q141P** **pncA.S67P** **pncA.T135P** **pncA.T160P** **pncA.T76P** **pncA.V139A** **pncA.V139L** **pncA.V7G** **pncA.V9G** **pncA.W119R** **pncA.W68G** **pncA.W68R** **rpoB.S450L** **rpsL.K43R** **rpsL.K88Q** **rpsL.K88R** **Mycobacterium_tuberculosis_16S.A1401G** **tlyA.N236K** katG.A234G katG.G300W katG.A431V kasA.G312S kasA.G269S thyA.T202A gidB.S70R Mycobacterium_avium_23S.A2274G folC.A420V iniA.S501W	

	folC.I43T gidB.A134E folC.S150G gidB.A138E folC.I43S thyA.H207R folC.I43A	
	Second-line	
AMK	**katG.S315T** **rpoB.S450L** **rpsL.K43R** **Mycobacterium_tuberculosis_16S.A1401G** **embB.M306V**	100.00%
CM	**Mycobacterium_tuberculosis_16S.A1401G**	100.00%
KM	**embA.P913S** **embB.D1024N** **embB.E378A** **embB.G406D** **embB.G406S** **embB.M306I** **embB.M306V** **embB.Q497R** **embC.N394D** **embC.R738Q** **embC.T270I** **embC.V981L** **embR.C110Y** **gyrA.A90V** **gyrA.D94G** **gyrA.D94N** **gyrA.D94Y** **gyrA.S91P** **katG.S315T** **ndh.V18A** **rpoB.S450L** **rpsL.K43R** **rpsL.K88R** **Mycobacterium_tuberculosis_16S.A1401G** gyrB.V340L thyA.T202A kasA.G269S folC.A420V iniA.S501W gidB.L79S	80.00%

OFX	**katG.S315T** **gyrA.A90V** **gyrA.D94G**	100.00%
Nonredundant Union	**embA.P913S** **embB.D1024N** **embB.D328G** **embB.D328Y** **embB.E378A** **embB.G406A** **embB.G406C** **embB.G406D** **embB.G406S** **embB.M306I** **embB.M306L** **embB.M306V** **embB.Q497K** **embB.Q497R** **embB.S297A** **embB.Y319C** **embB.Y334H** **embC.N394D** **embC.R738Q** **embC.T270I** **embC.V981L** **embR.C110Y** **ethA.A381P** **ethA.D58A** **gyrA.A90V** **gyrA.D89G** **gyrA.D89N** **gyrA.D94A** **gyrA.D94G** **gyrA.D94N** **gyrA.D94Y** **gyrA.G88A** **gyrA.G88C** **gyrA.S91P** **inhA.I194T** **inhA.I21T** **inhA.I21V** **inhA.S94A** **katG.G279D** **katG.Q127P** **katG.R463L** **katG.S140N** **katG.S315G**	78.81%

	katG.S315I katG.S315N katG.S315T katG.Y337C ndh.R268H ndh.V18A pncA.A134V pncA.A146T pncA.C14R pncA.D12A pncA.D12N pncA.D49A pncA.F58L pncA.G132S pncA.G97S pncA.H51Q pncA.H57D pncA.H82R pncA.L159R pncA.L172P pncA.L27P pncA.L4W pncA.L85P pncA.L85R pncA.M175V pncA.P54L pncA.Q10P pncA.Q10R pncA.Q141P pncA.R154G pncA.S67P pncA.T135P pncA.T160P pncA.T47A pncA.T76P pncA.V139A pncA.V139L pncA.V155G pncA.V7G pncA.V9G pncA.W119R pncA.W68G pncA.W68R pncA.Y103H rpoB.S450L rpsL.K43R	

	rpsL.K88Q **rpsL.K88R** **Mycobacterium_tuberculosis_16S.A1401G** **tlyA.N236K** Mycobacterium_avium_23S.A2274G folC.A420V folC.I43A folC.I43S folC.I43T folC.S150G gidB.A134E gidB.A138E gidB.H48Y gidB.L79S gidB.P75S gidB.S70R gyrB.V340L iniA.S501W kasA.G269S kasA.G312S katG.A234G katG.A431V katG.G300W katG.P131Q rpoB.D516G rpoB.H526T rpoB.L511R thyA.H207R thyA.T202A	