Original Article

JHSMR Journal of Health Science and Medical Research

Whole Genome Sequencing of A Pooled Sample of Bedaquiline and Clarithromycin Resistant *Mycolicibacterium smegmatis* Exhibits Mutations in Genes Associated with Efflux Family Protein and ATP Synthase

Muhamad Frendy Setyawan, B.Sc., M.Trop.Med[1], Amina K. Shaban, Ph.D.[2], Hayato Takihara, Ph.D.[3], Yoshitaka Tateishi, Ph.D.[2], Akihito Nishiyama, Ph.D.[2], Yuriko Ozeki, Ph.D.[2], Zakiyathun Nuha, B.Sc.[2], Ni Made Mertaniasih, M.D.[4,9], Shujiro Okuda, Ph.D.[3], Soedarsono Soedarsono, M.D.[5], Maryam Omrani, Ph.D.[6], Wayan Tunas Artama, DVM[7], Mochammad Aqilah Herdiansyah, B.Sc.[8], Sohkichi Matsumoto, M.D., Ph.D.[2,9]

[1]Doctoral Program of Medical Sciences, Faculty of Medicine, Airlangga University, Surabaya 60132, Indonesia.
[2]Department of Bacteriology, Graduate School of Medical and Dental Sciences, Niigata University, Niigata 951-8510, Japan.
[3]Medical AI Center, Graduate School of Medical and Dental Sciences, Niigata University, Niigata 951-8514, Japan.
[4]Department of Clinical Microbiology, Faculty of Medicine, Airlangga University, Surabaya 60132, Indonesia.
[5]Sub-pulmonology Department of Internal Medicine, Faculty of Medicine, Hang Tuah University, Surabaya 60111, Indonesia.
[6]Emerging Bacterial Pathogens Unit, IRCCS San Raffaele Scientific Institute, Milan 20132, Italy.
[7]Research Center for Biotechnology, Graduate School, Gadjah Mada University, Yogyakarta 55281, Indonesia.
[8]Department of Biology, Faculty of Science and Technology, Airlangga University, Surabaya 60115, Indonesia.
[9]Laboratory of Tuberculosis, Institute of Tropical Disease, Airlangga University, Surabaya 60115, Indonesia.

Received 27 January 2025 ● Revised 28 March 2025 ● Accepted 28 March 2025 ● Published online 6 August 2025

Abstract:

Objective: This study aimed to evaluate the pooled sequencing of mixed samples with different minimum inhibitory concentrations (MIC) of bedaquiline (BDQ) and clarithromycin to identify genes associated with drug resistance and explore those that could predict cross-resistance to both drugs. Additionally, it aimed to preliminarily investigate the association between single-nucleotide polymorphisms (SNPs) from pooled samples and the progression of drug resistance.

Contact: Prof. Dr. Ni Made Mertaniasih, M.D.
Department of Clinical Microbiology, Faculty of Medicine, Airlangga University, Surabaya 60132, Indonesia.
E-mail: ni-made-m@fk.unair.ac.id
Contact: Prof. Dr. Sohkichi Matsumoto, M.D.
Department of Bacteriology, Graduate School of Medical and Dental Sciences, Niigata University, Niigata 951-8510, Japan.

J Health Sci Med Res 2026;44(2):e20251245
doi: 10.31584/jhsmr.20251245
www.jhsmr.org

Material and Methods: *Mycolicibacterium smegmatis* wild type (WT) and *hup B* (MDP1) knock-out strains were obtained from the Bacteriology Laboratory, Graduate School of Medical and Dental Science, Niigata University. Drug-resistant strains were generated through stepwise passaging on drug-containing agar plates, followed by MIC determination. Whole-genome sequencing was performed using the MinION Nanopore platform, with pooled samples categorized based on MIC levels. Bioinformatics analysis, including SNP identification and annotation, was conducted using burrows wheeler aligner, Samtools, genome analysis toolkit, and SnpEff pipelines. Molecular docking and molecular dynamics simulations were employed for in silico validation of SNP-drug interactions. SNP frequency was calculated, and statistical analysis was performed using a paired t-test in Prism® software.

Results: High-impact SNPs that changed the amino acid were found in both pooled samples, one in pooled sample A (*atpA* p.Glu429Gln) and 5 in pooled sample B (MSMEG_0639, MSMEG_1867, MSMEG_4765, and MSMEG_6302).

Conclusion: Pooled samples of BDQ and clarithromycin-resistant *M. smegmatis* strains exhibited mutations in the genes associated with the efflux family and the adenosine triphosphate synthase protein. A pooled sample sequencing method can be used to identify variant sites and predict genes associated with cross-resistance.

Keywords: bedaquiline, hup, Mycolicibacterium smegmatis, single nucleotide polymorphisms, whole genome sequencing

Introduction

A major challenge in developing molecular diagnostic tools to confirm drug-resistant strains of mycobacteria, including *Mycobacterium tuberculosis,* has been the exploration of genetic markers[1–4]. The World Health Organization recommended using next-generation sequencing (NGS) to help diagnose drug-resistant tuberculosis worldwide in 2024[5,6]. The NGS-based diagnosis uses a group of specific genes that can find mutations in targeted sites[7]. Genetic markers are known as single-nucleotide polymorphisms (SNPs), including insertions and deletions (Indels) of single or multiple base nucleotides. Whole genome sequencing (WGS) offers a wider opportunity to find variants in the genome that could be associated with a decrease in susceptibility to drugs[8–10]. Currently, the molecular diagnosis of drug-resistant strains of mycobacteria depends on genotypic markers and drug susceptibility testing. As finding the correct genotypic marker for drug resistance requires further investigation, WGS serves as a preliminary tool to screen out variants and SNPs, and to identify the genes most likely related to changes in phenotypic susceptibility. *Mycolicibacterium smegmatis* MC2 155, a fast-replicating and nonpathogenic mycobacterium, has been widely used as a model to investigate the BDQ drug resistance mechanism because of: (1) its similarity with *M. tuberculosis* in the genes targeted by BDQ, (2) fast-growing bacteria, (3) nonpathogenic, and (3) biosafety level 2 (BSL 2) can be used for experiments. *M. smegmatis* also represents all the non-tuberculous mycobacteria (NTM) species, thus making it a suitable model to investigate the cross-resistance mechanism between BDQ and clarithromycin. Clarithromycin has been used to treat infections caused by NTM species, meanwhile, BDQ has not been widely explored as a potential drug for treating infections by NTM[11,12].

The development of a rapid molecular diagnostic tool for the new drug BDQ has not yet been fully established. Thus, we performed a stepwise acquisition mutation against the new drug BDQ in order to generate a mutant strain. The *Mycolicibacterium smegmatis* mutant strain resistant to BDQ has been found to harbor mutations in *atpE* after 8 passages[13,14]. In this study, we closely investigated the serial adaptation of *M. smegmatis* after 4 passages. WGS using the Oxford Nanopore Technologies nanopore platform can be used to sequence group samples. In this study, we differentiated the strain on the basis of the minimum inhibitory concentration (MIC) of each strain. We also identified the mutations that confer resistance to BDQ by sequencing based on the MIC level, with the first and second strains having a lower MIC than the third and fourth strains. Pooled sequencing was chosen in this study to provide an effective approach to using WGS in a genomic study, moreover, numerous challenges arise from single-sample sequencing in a developing country, due to the high costs and time constraints; thus, pooled-sample sequencing is a good option for overcoming this burden.

The *hup* gene (or *hlp* or *MDP1*) is essential for the growth of mycobacterial species. Thus, the strain of mycobacteria BCG with knocked down *hup B* presented increased susceptibility to isoniazid (H) and BDQ[15–17]. A previous study reported cross-resistance between BDQ and clarithromycin in *M. smegmatis* laboratory-acquired resistant strain[14,18]. This study revealed variations that could be associated with cross-resistance between BDQ and clarithromycin in *M. smegmatis*, due to the combination of pooled samples used in the sequencing process. This study is the first to analyze data from the pooled samples of low and high MIC isolates to identify the genetic markers of the level of drug resistance.

Material and Methods

Biosafety level criteria and mycobacterial strains

All procedures in this experiment were performed in a BSL 2 facility at the Graduate School of Medical and Dental Science, Niigata University, Japan. *Mycolicibacterium smegmatis* MC2 155 strain wild type and hup gene (or *hlp* or MDP1) knockout strains were obtained from the Bacteriology Laboratory, Graduate School of Medical and Dental Sciences, Niigata University, Japan[19].

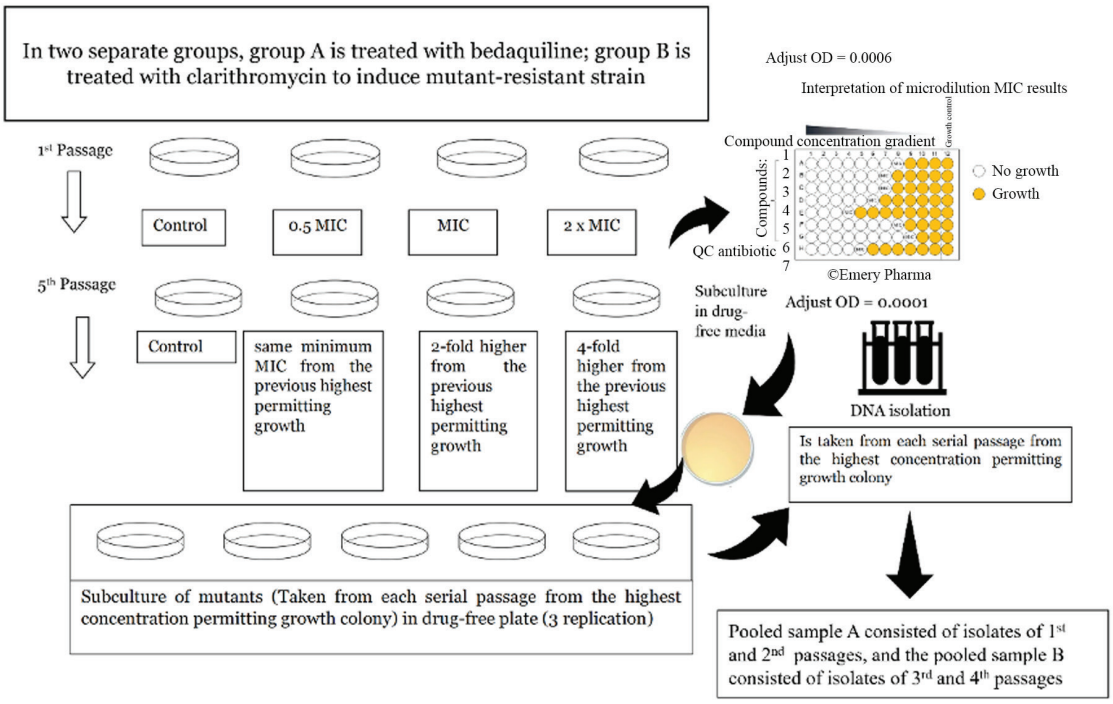
Bacterial culture

Bacteria were cultured in a liquid Middlebrook 7H9 broth (BD Difco, Burlington, USA), supplemented with 10% albumin-dextrose-catalase (BD Difco, Burlington, USA), 0.2% glycerol, and 0.05% between 80 (Sigma-Aldrich, Missouri, USA), and in Middlebrook 7H10 (BD Difco) supplemented with 10% oleic-albumin-dextrose-catalase (BD Difco, Burlington, USA) and 0.5% glycerol. The knockout strain was maintained in media supplemented with 10 mg kanamycin sulfate (Ks). The liquid culture mixture was grown for 3 days at 37 °C in a shaking incubator and subcultured continuously. The OD_{600} was maintained at OD_{600} = 0.2–0.3 during the stationary phase.

Generation and confirmation of mutant strains resistant to drugs

The bacterial cultures were adjusted to an OD_{600} of 0.0001 to be plated on agars containing drugs with twofold serial dilutions of the MIC (0.5×, 1×, 2×). Generation of the mutant strain using the stepwise passaging method was performed 5 times. The colony was harvested from the highest media permitting growth and subjected to mutant selection by MIC determination in a 96-well microtiter plate with a round bottom. The mutant strain was cultured

with another twofold increase in drug concentration from the first passage, as shown in Figure 1. Fifty microliters of bacterial suspensions (OD_{600} = 0.00025–0.0006) were transferred to each serially diluted broth media plate with drugs. The MICs for each drug were determined as the lowest concentration that inhibited 90% of bacterial growth, as observed by its turbidity. The MIC test for each group was performed in triplicate to ensure validity. The drug stocks were diluted as follows: 0.1 g of BDQ in 10 mL of UPW, 0.5 g of Ks in 10 mL of UPW, and 0.1 g of clarithromycin in 10 mL of ultrapure water. The concentration of each drug was maintained at the following dose using the stock solution above (0.015–8 µg/ml for clarithromycin; 0.003–2 µg/ml for BDQ).



MIC=minimum inhibitory concentration, DNA=Deoxyribonucleic acid

Figure 1 Stepwise acquisition mutation procedure of M. smegmatis against drugs to obtain a pooled sample of mutant strains with increasing MIC. Pooled A consisted of isolates with low MICs, whereas pooled sample B consisted of isolates with high MICs.

Deoxyribonucleic acid (DNA) extraction and quality control

Colonies from a single plate were subjected to DNA extraction. Genomic DNA (gDNA) purification was performed using the phenol/chloroform/isoamyl alcohol (25:24:1) method. The quality and quantity of all the DNA were checked via Nanodrops, and the quality of the DNA was maintained with absorbances of A260/280 and A260/230 of approximately 1.8 and 2.0, respectively, while the DNA concentration was greater than 5.5 µg/mL.

Pooled sample sequencing and bioinformatics analysis

WGS was performed using the Oxford Nanopore Technology Sequencing Kit with a MinION device and R10.4.1 flow cells (FLO-MIN114). DNA was repaired and end-prepped using NEBNext® FFPE DNA repair mix (NEB, cat # M6630). A pooled sample sequencing was prepared using a total of 18 DNA samples separated into 2 batches for sequencing: the 1st batch consisted of the 1st to 2nd passage genomic isolates. In contrast, the other batch consisted of the 3rd and 4th passage isolates, with relatively high MIC. Genomic DNA from isolates in the first and second passages was mixed to form pooled sample A, whereas gDNA from isolates in the third and fourth passages was mixed to form pooled sample B. Pooled sample B had higher MICs than pooled sample A. Both samples were subjected to whole-genome sequencing. Bioinformatics analysis was performed with the Burrows-Wheeler Aligner (BWA), samtools, genome analysis toolkit (GATK), and SnpEff pipelines. The sequence quality was checked, and to prevent bias in analysis, genomic analysis was performed using BWA and annotated with the *M. smegmatis* MC2 155 reference from downloaded Ensembl bacteria using samtools[20]. The alignment and variant search and SNPs were searched using GATK and SnpEff. To ensure average read quality, a Phred score of ≥Q20 (≥99% accuracy) was used, and the reads were mapped against the *M. smegmatis* MC2 155 reference genome using BWA[21]. We performed base call recalibration and realignment around Indels with the GATK[22]. Both Samtools and GATK were used to call SNPs and Indels, and the results were annotated via SnpEff and SnpSift[23]. In silico validations of the SNPs interaction toward drug receptors were performed using molecular docking with AutoDock Vina software and BIOVIA Discovery Studio. Further analysis of molecular dynamics was performed with CABS-flex.

SNP frequency analysis

The frequency of SNPs found in the genes in each pooled sample was calculated based on the number of times each variant occurred.

Statistical analysis

Confirmation of the genes responsible for the progression of drug resistance between pooled sample A and B was statistically analyzed using a paired t-test with Prism® software.

Results

WGS reveals dynamic changes in gene polymorphisms that are possibly associated with multidrug resistance.

The development of drug resistance and cross-resistance to both conjugated linoleic acid (CLA) and BDQ over successive passages is shown in Table 1. In both the wild-type and knockout strains, MICs for CLA and BDQ increased with successive passages, indicating adaptation. Notably, strains initially exposed to CLA (WTCLA and

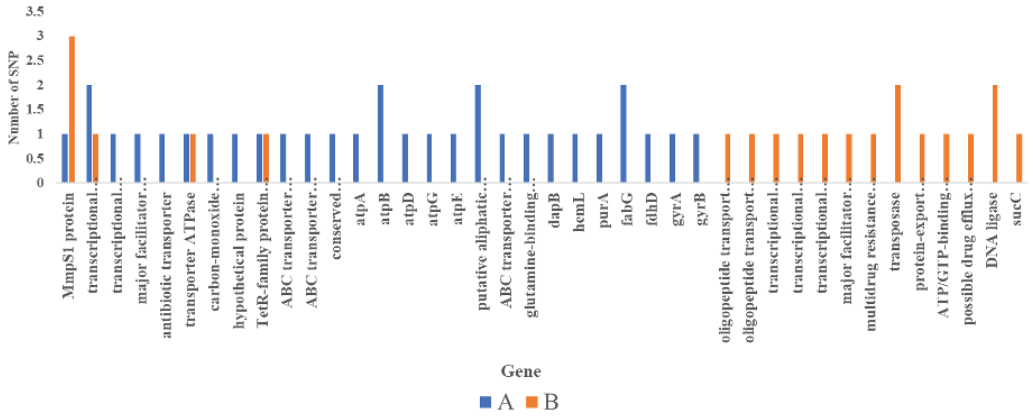
KOCLA) also showed a gradual increase in BDQ MICs, despite not being directly exposed to BDQ. Similarly, strains selected with BDQ (WTBDQ and KOBDQ) showed an increase in CLA MICs.

Table 1 A pooled sample of the sequence group based on the minimum inhibitory concentration against bedaquiline and clarithromycin

Group (s)	Pooled sample	MIC	
		CLA (μg/ml)	BDQ (μg/ml)
Control (WTCLA-0)	A	0.5	0.125
1st P (WTCLA-1)	A	64	0.125
2nd P (WTCLA-2)	A	256	0.125
3rd P (WTCLA-3)	B	256	0.125
4th P (WTCLA-4)	B	512	2
Control (KOCLA-0)	A	0.5	0.125
1st P (KOCLA-1)	A	64	0.125
2nd P (KOCLA-2)	A	512	0.125
3rd P (KOCLA-3)	B	512	2
4th P (KOCLA-4)	B	1024	2
Control (WTBDQ-0)	A	0.5	0.125
1st P (WTBDQ-1)	A	0.5	2
2nd P (WTBDQ-2)	A	0.5	32
3rd P (WTBDQ-3)	B	2	128
4th P (WTBDQ-4)	B	2	128
Control (KOBDQ-0)	A	0.5	0.5
1st P (KOBDQ-1)	A	0.5	2
2nd P (KOBDQ-2)	A	2	128
3rd P (KOBDQ-3)	B	0.5	128
4th P (KOBDQ-4)	B	2	256

WT=wild type, KO=knockout *hup gene* (or MDP1) knockout, P=passage, A=pooled sample A, B=pooled sample B, CLA=clarithromycin, BDQ=bedaquiline, MIC=minimum inhibitory concentration

The complete results of WGS of the 2 pooled samples are shown in Supplementary Table 1, and the MICs of both groups are presented in Table 1. Figure 2 revealed that most of the variants were harbored from genes associated with efflux pump protein and adenosine triphosphate (ATP) synthase. Variants from pooled sample A (low MIC isolates) consisted of 10 upstream variants, 9 downstream variants, 9 frameshift variants, 1 missense variant, and 2 synonymous variants. Moreover, from pooled sample B (high MIC isolates) we identified 5 upstream variants, 9 downstream variants, 5 missense variants, and 2 synonymous variants. We also selected some significantly high-impact SNPs that altered the amino acid sequence of the protein and changed the amino acid type, as shown in Table 2. All the SNPs identified in this study have never been reported before and are possibly new candidates for drug resistance markers; however, some studies have reported similar genes found in the BDQ-resistant strains of M. smegmatis. Statistical testing using the paired t-test resulted in p-value 0.1848 >0.005 (95% confidence interval [CI]: 0.1246), which means no significant difference between the 2 groups, suggesting that the mutations found in the 2 groups have no significant difference, as shown in Supplementary Table 2. Some gene-associated mutations remained after higher drug exposure in the high MIC isolate group. In this study, a mutation in the higher pooled isolates qualifies as a fixed or permanent polymorphism, additionally some mutations found in both samples were also considered as fixed variants, whereas in low MIC pooled samples, the variants were considered as temporary or dynamic polymorphism since they were mostly not found in the higher MIC strains, due to genomic adaptations in the bacteria.



A=pooled sample A, B=pooled sample B, SNPs=single-nucleotide polymorphisms

Figure 2 Frequency analysis of SNPs from pooled-sample sequencing of *M. smegmatis* resistance to bedaquiline and clarithromycin

Table 2 Single nucleotide polymorphisms of genes with a high impact on amino acid substitutions; pooled sample B had a higher minimum inhibitory concentrations than pooled sample A

Gene name	Sample	Gene Id	Variant effect	Gene changes post.	Ori	Alt	Amino acid	p-value (95% CI)
atpA	Pooled A	MSMEG_4938	Missense variant	c.1285G>C	G	C	p.Glu429Gln	
Oligopeptide transport ATP-binding protein AppF		MSMEG_0639	Missense variant	c.815G>A	G	A	p.Arg272His	
Transposase	Pooled B	MSMEG_1867	Missense variant	c.109G>A	G	A	p.Glu37Lys	0.1848 (95%
Transcriptional regulator, MerR family protein		MSMEG_4765	Missense variant	c.583A>G	A	G	p.Thr195Ala	CI: 0.1246)
DNA ligase		MSMEG_6302	Missense variant	c.466T>C	T	C	p.Cys156Arg	
				c.463T>C	T	C	p.Trp155Arg	

p-value (>0.005) of a paired t-test result indicates the difference is not statistically significant between the 2 groups, the 95% CI indicate confidence interval of the test, Ori=origin of replication, Alt=alternative lengthening of telomeres, DNA=deoxyribonucleic acid, CI=confidence interval

Based on the in-silico validation shown in Table 3 and Figure 3, using molecular docking analysis on ligand-receptor binding affinity, we found that the DNA ligase mutants (p.Cys156Arg) and (p.Trp155Arg) had the strongest binding affinity. This suggests that these mutations might stabilize the protein-ligand complex, and possibly influence protein function. The atpA (p.Glu429Gln) mutation had a slightly weaker binding affinity than its wild type, indicating a minimal effect on interaction strength. The oligopeptide transport ATP-binding protein AppF (p.Arg272His) mutant had a slightly stronger binding affinity than its wild type, which indicates a minor effect. The transposase mutant and the wild type both had the weakest binding affinity compared with the others. Based

on the MD simulation, the root mean square fluctuation (RMSF) analysis indicated protein flexibility, with higher values suggesting greater flexibility. The *atpA* wild type showed the highest RMSF, with a significant increase in rigidity upon mutation. The DNA ligase mutants exhibited slightly lower RMSF values compared with their wild type, suggesting minor structural rigidity upon mutation, similar to the oligopeptide transport ATP-binding protein mutant, meaning that the mutation may have interrupted enzyme activity, substrate binding, and protein binding activities. The transposase wild type and its mutant showed an increase in flexibility upon mutation, and thus could lead to conformational changes, which might affect the functional dynamics of the protein. The complete difference between the wild type and mutant protein interactions with the ligand is shown in Supplementary Figure 1 and 2.

Table 3 In-silico validation of the SNPs interaction toward bedaquiline with molecular docking and dynamic simulation

No	Compound	Sample	Complex binding affinity (kcal/mol)	RMSF value (Å)	Type of interaction	Amino acids involved (BIOVIA discovery studio)
1	BDQ	atpA_(p.Glu429Gln)	−7.3	1.53	VDW	Ser275 (B), Glu331 (B), Thr332 (B), Arg294 (B)
					HI	Ala339 (B), Lys333 (B), Phe297 (B), Pro342 (B), Ile341 (B)
					PHI	Phe340 (B) Tyr436 (B), Ser218 (B), Lys178 (B), Asp272 (B), Lys276 (B), Asp273 (B), Thr215 (B), Lys212 (B),
2		atpA_Wild type	−7.4	2.85	VDW	Gln211 (B), Lys175 (B), Thr176 (B), Gly177 (B)
					HI	Val219 (B), Ala222 (B), Ala180 (B), Arg174 (B)
					PHI	Glu331 (B) Trp96 (A), Arg37 (A), Gly35 (A), Ser51 (A),

Table 3 (continued)

No	Compound	Sample	Complex binding affinity (kcal/mol)	RMSF value (Å)	Type of interaction	Amino acids involved (BIOVIA discovery studio)
3		DNA Ligase (p.Cys156Arg)	−7.9	0.96	VDW	His312 (A),
					HI	Asn53 (A), Lys55 (A), Tyr60 (A) Leu57 (A), Arg52 (A) His312 (A), Tyr60 (A), Ser51 (A), Lys55 (A),
4		DNA Ligase (p.Trp155Arg)	−7.9	0.96	VDW	Asn53 (A),
					HI	Trp96 (A), Arg37 (A), Gly35 (A) Leu57 (A), Arg52 (A) Tyr60 (A), Ser51 (A), Trp96 (A),
5		DNA Ligase wild type	−7.4	0.981	VDW	Arg37 (A), Gly35 (A), Asn53 (A), Lys55 (A)
					HI	Ile103 (A), Leu57 (A), Arg52 (A)
6		Oligopeptide transport ATP-binding protein AppF wild type	−6.8	1.205	VDW	Val95 (A),
						Met97 (A), Gln96 (A), Arg168 (A), Leu172 (A),
7		Oligopeptide transport ATP-binding protein AppF (p.Arg272His)	−7.0	1.205		His123 (A),
					HI	Arg60 (A) Pro119 (A), Arg92 (A), Leu63 (A), Leu122 (A), Phe99 (A)

Table 3 (continued)

No	Compound	Sample	Complex binding affinity (kcal/mol)	RMSF value (Å)	Type of interaction	Amino acids involved (BIOVIA discovery studio)
					VDW	Arg60 (A), Val95 (A), Met97 (A), Gln96 (A), Arg168 (A), Leu172 (A), His123 (A)
					HI	Pro119 (A), Leu122 (A), Phe99 (A), Arg92 (A), Leu63 (A)
					VDW	Thr28 (A), Arg47 (A)
8		Tranposase wild type	−6.3	0.810	HI	Pro40 (A), Ala41 (A)
					PHI	Gly29 (A), Glu30 (A)
					VDW	Phe203 (A), Phe211 (A), Thr167 (A), Thr159 (A)
9		Transcriptional regulator, MerR family protein wild type	−7.0	1.579	HI	Val163 (A), Ala160 (A), Leu143 (A), Ala164 (A), Leu128 (A), Ala126 (A)
					VDW	Thr167 (A), Phe203 (A), Phe211 (A), Thr159 (A)
10		Transcriptional regulator, MerR family protein (p.Thr195Ala)	−7.1	1.260	HI	Val163 (A), Ala160 (A), Leu143 (A), Leu128 (A), Ala164 (A), Ala126 (A)
					PHI	Asp140 (A) Thr28 (A),
11		Transposase mutant (p.Glu37Lys)	−6.3	1.049	VDW	Arg47 (A)
					HI	Pro40 (A), Ala41 (A)
					PHI	Glu30 (A), Gly29 (A)

VDW=VanderWaals, HI=hydrophobic interaction, PHI=polar hydrogen interaction, A=acidic, B=Basic, RMSF=root mean square fluctuation, DNA= deoxyribonucleic acid

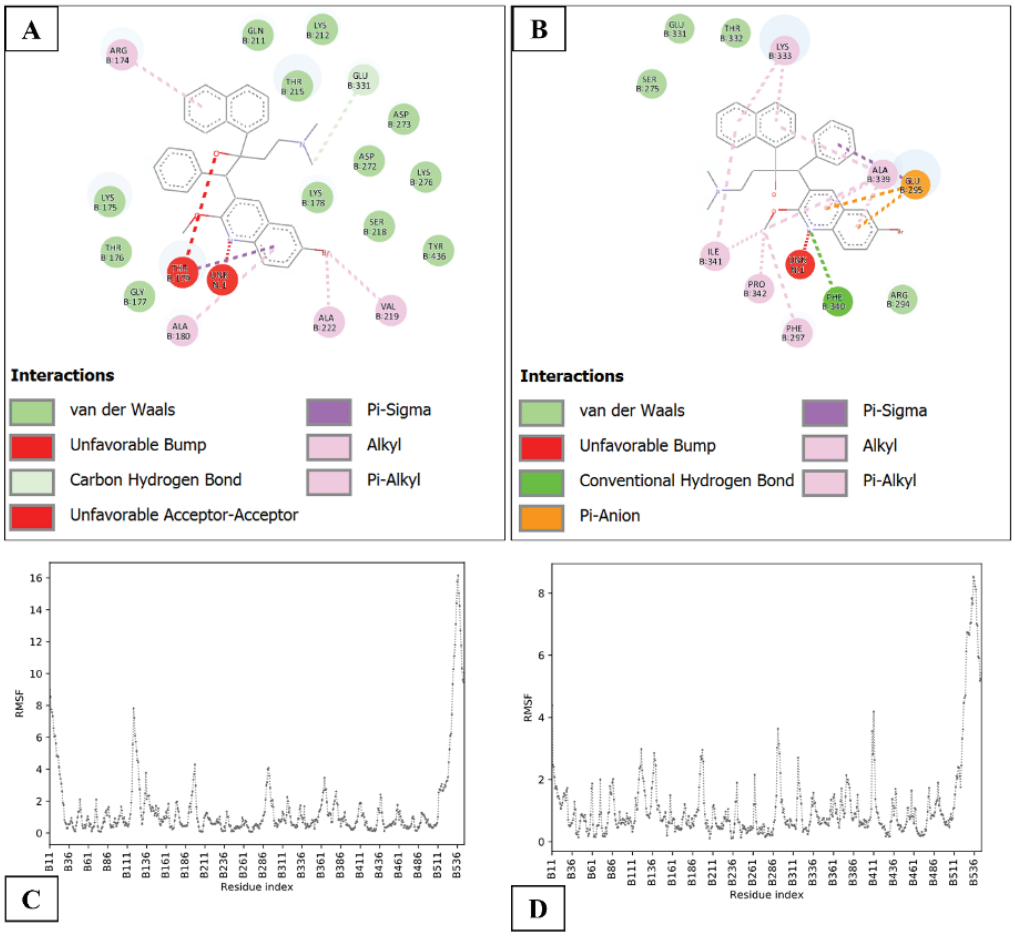
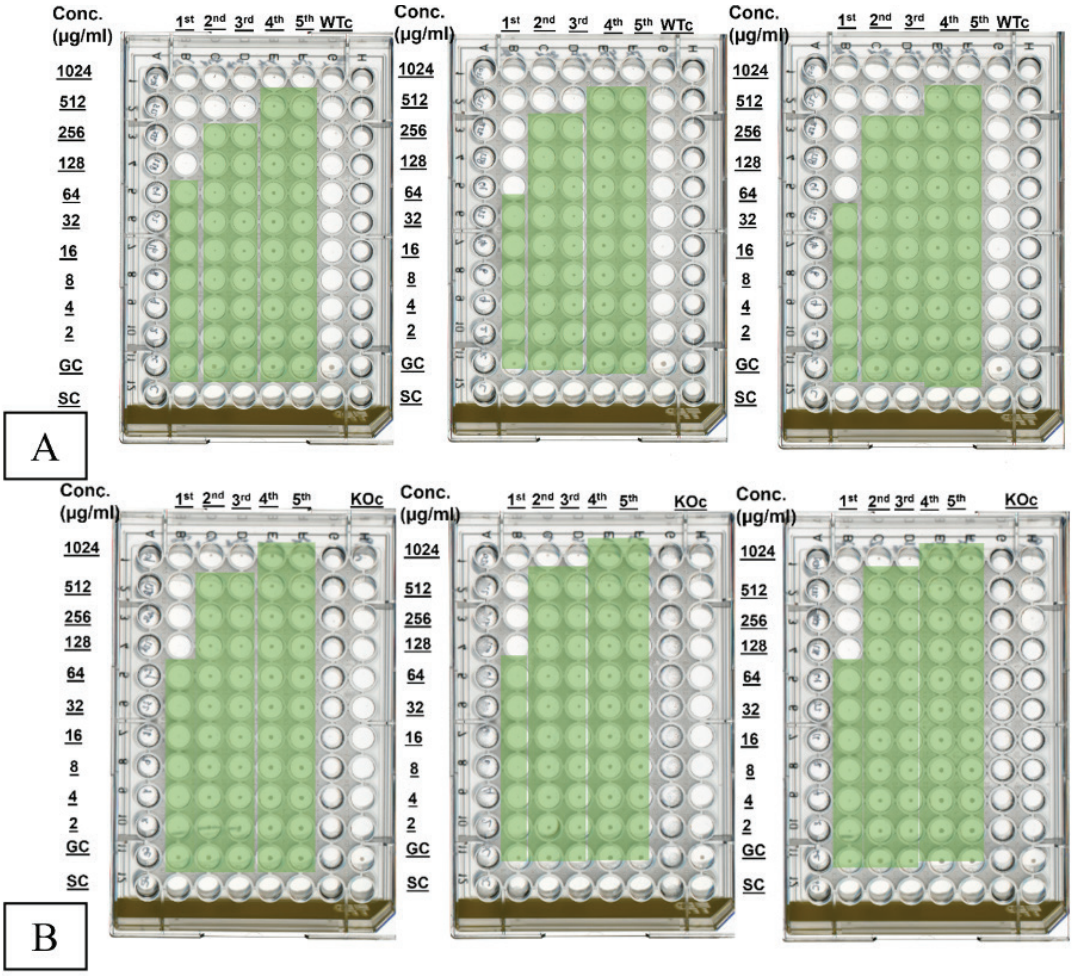


Figure 3 Comparison of the in-silico validation of the atpA wild-type and atpA mutant (p.Glu429Gln). (A) amino acid interaction with bedaquiline of the atpA WT, (B) amino acid interactions with bedaquiline of the atpA mutant using BIOVIA discovery studio, (C) Molecular dynamic (MD) simulation of the atpA WT, (D) MD of atpA mutant using Cabsflex

Discussion

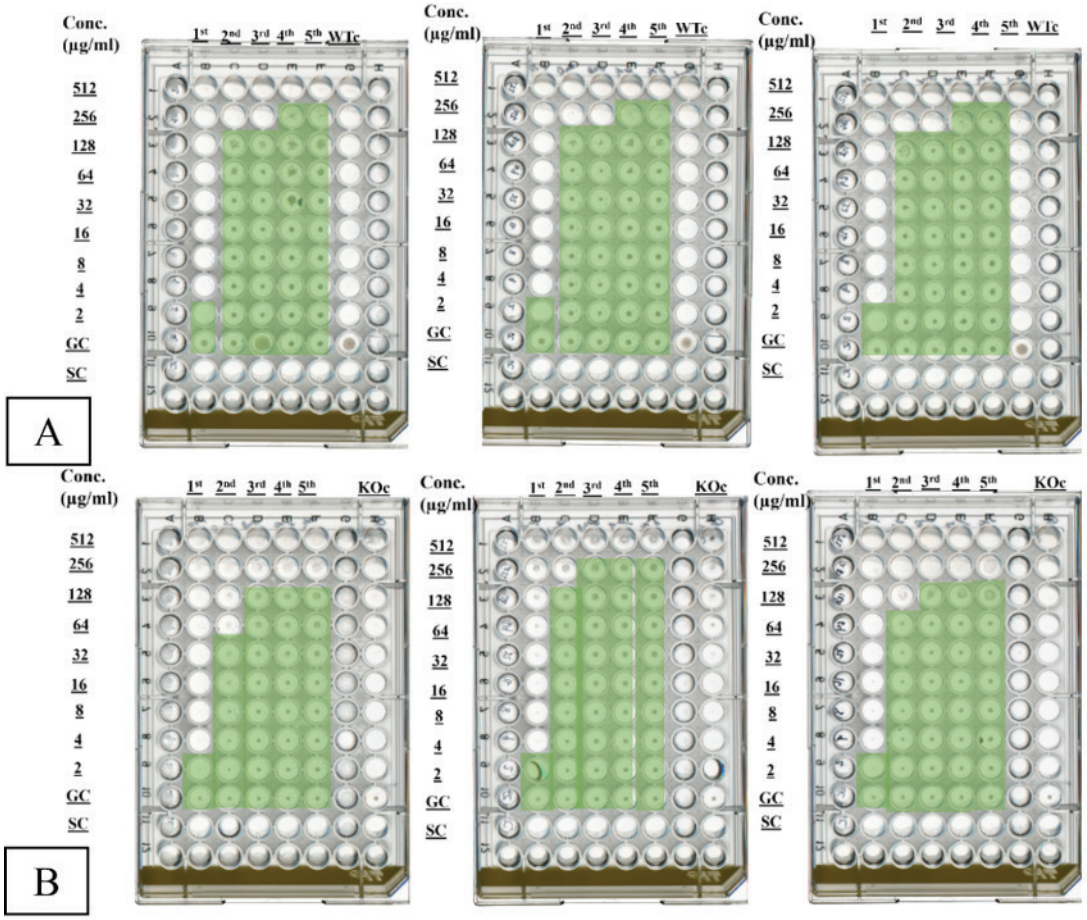
Deciphering drug resistance-associated genes that determine phenotypic traits is the most important step in developing diagnostic markers for drug resistance. WGS helps in the exploration of potential genetic traits linked to drug resistance. The discovery of the SNPs from a drug-resistant strain could help narrow potential candidates for selective markers[24,25]. Thus, we demonstrate the natural adaptive selection mechanism of both SNPs and Indels from the laboratory-induced mutant strain *M. smegmatis*. This study is the first to present a laboratory-based method for selecting fixed mutations from pooled samples with different MIC groups, as shown in Figure 1. We present an experimental hypothesis for the association between the *hup* gene and in vitro BDQ susceptibility in *M. smegmatis*. Five sequential *M. smegmatis* populations obtained through serial passaging were tested for MIC, and 4 series were sequenced using WGS. We found additional genomics data related to the dynamic changes in the gene population detected at lower and higher MICs. The phenotypic profile was consistently

resistant, progressing from low to high MIC values. No fixed *atpE* or *rv0678* mutations were found in this study; however, we detected many genes that can cause drug resistance, including genes associated with the efflux family protein, ATP synthase, and transmembrane protein. Thus, a future study is needed to verify the genetic variants. The detailed changes in the MIC for each group are shown in Figure 4 and 5. Pooled sequencing could raise economic benefits for researchers in developing countries, such as cost-saving options for genomic studies for exploring potential genes related to drug resistance mechanisms, along with comprehensive bioinformatic analysis. Furthermore, pooled-sample sequencing provides deeper genomic interpretation from a whole diverse population of genomes than a single-strain sequencing method.



GC=growth control, SC=sterile control

Figure 4 The MIC test of the WT and KO groups against BDQ was performed in triplicate and displayed in the microtiter plates. (A) Resistance of the *M. smegmatis* mutant wild-type strain to clarithromycin, (B) resistant mutant *M. smegmatis* knockout strain to clarithromycin



GC=growth control, SC=sterile control, MIC=minimum inhibitory concentrations, WT=wild type, CLA=clarithromycin

Figure 5 The MIC test of the WT and KO groups against CLA was performed in triplicate and displayed in the microtiter plates. (A) Resistance of the *M. smegmatis* wild-type strain to bedaquiline, (B) resistant mutant *M. smegmatis* hup (MDP1) knockout strain to bedaquiline

A previously reported study involving WGS of a laboratory-evolved clarithromycin-resistant mutant *M. smegmatis* strain revealed no associated gene mutations. Furthermore, mutation-conferred BDQ resistance has been reported in *atpE* and *gyrB*[14]. In the initial passages, we found many nucleotide polymorphisms that directly impact the MIC value, and many are considered dynamic. In this study, we were unable to separate the genome sequencing data from wild-type and knock-out groups; however, the interpretation was performed in general to see the dynamic gene polymorphisms in the initial and last passages. Most reported variants associated with resistance from the study are categorized as new or have never been reported before. However, we found variants in genes associated with BDQ resistance based on previously reported data. The BDQ-resistant strain of *M. tuberculosis* has variants in the *atpE* and *rv0678* genes, whereas in *M. smegmatis*, only variants in the *atpE* gene have been reported.

Previous studies have reported that mutations in *atpE* were associated with a permanent mutation, whereas dynamic mutations occur in *rv0678*, *rv1979c,* and *pepQ*. In this study, we detected mutations in ATPase family proteins encoded by the genes *atpA*, *atpB*, *atpD*, *atpG*, and *atpE* from pooled sample A (low MIC isolates), and those mutations were not detected in pooled sample B (high MIC isolates). These authors suggest that the influence of efflux family proteins is more significant than that of ATPase family proteins in causing drug resistance to BDQ and others. In a previous report, a mutation in MmpR resulted in efflux pump overexpression, making the strain resistant to isoniazid[26,27]. However, in the present study, we observed no variants in mmpR in the multidrug-resistant mutant, instead, we found mutations in the SMR superfamily and ATP binding cassette (ABC) superfamily proteins. We found more mutations in pooled sample A (low MIC colonies). We also found mutations in the possible drug efflux membrane protein (MSMEG-3815) in the high MIC isolate with orthologs in the *M. tuberculosis* gene *rv1634*. In a previous study, the overexpression of *rv1634*, an major facilitator superfamily (MFS) superfamily protein, was found to confer resistance to fluoroquinolones[28,29]. Cross-resistance between MDQ and clarithromycin has been reported to be due to variations in the efflux pump-associated genes[14], thus, this study also revealed many mutations related to the Tet-R regulatory genes, membrane protein-associated genes, and MFS superfamily proteins[30]. We predict that the drug molecules are pumped out by the exporter protein family and maintained at a low concentration inside the cells, and hence lose the ability to inactivate the bacteria.

In this study, we explain the details of how the identified mutations might contribute to resistance at a functional level. Based on the pooled sample A, we found many frameshift and nonsense mutations, those that are affected by it can result in premature amino acids and consequently produce no protein, and thus will disturb the metabolism of bacteria. Polymorphism in the *atpA* gene (p.Glu429Gln) converts Glutamic acid (E) to glutamine (Q) at position 429 in a protein. Thus, the interaction between the amino acids is due to the changes in the negatively charged and acidic amino acids to neutral and polar amino acids. This mutation removes a negatively charged residue, which may affect protein stability and impair function or interactions, particularly when *atpA* is important for binding and structural integrity. The mutation may alter hydrogen bonding patterns, which could affect the folding of a local protein or interaction site. In a pooled sample B, with high MIC, we found mutations in the ATP-binding site MSMEG_0639, which change the amino acid from Arginine (R) to Histidine (H) at position 272 in a protein. Thus, replacement changes a positively charged and basic amino acid with a weakly positively charged amino acid, thus affecting the electrostatic interactions and protein stability. Arginine is known as a highly polar amino acid, while histidine is less polar and has less hydrophilicity, and thus the differences might impact function, solubility, and alter any interactions with nearby residues. Histidine acts as a proton acceptor that might introduce new pH-dependent traits, affecting function under specific environments. Other studies support this finding by revealing the mechanism of drug resistance caused by the efflux system MmpR-MmpSL5 in mycobacteria[31,32].

Mutations in transposase (p.Glu37Lys) change glutamic acid (E) with lysine (K), switching the interactions with the nearby amino acids, disrupting key hydrogen bonds, or forming new ones. This is one of the destructive mutations due to charge reversal, which can significantly alter protein-protein interactions, enzyme activity, and binding affinity, including impaired enzyme function and destabilized local fold. A mutation in the MerR family protein (p.Thr195Ala) induces some effects, such as loss of polarity and hydrogen bonding, possible effects on DNA binding, protein disruption of regulatory function, and loss of phosphorylation potential. Some ability to activate or repress genes in response to stress could no longer be available,

and impair regulatory control mechanisms. Mutations in the DNA ligase (p.Cys156Arg and p.Trp155Arg) could possibly cause several functional impacts, including if (1) cys156 is involved in a disulfide bond, replacing it with arginine could destabilize the protein, (2) hydrophobic interaction may disrupt local interactions, (3) if Cys156 is in the active site or DNA-binding domain, this mutation could impair protein function, (4) Tryptophan has a major function in protein stability due to its bulky aromatics, while replacing it with arginine will remove these stabilizing effects, (5) a positively charged arginine in place of a neutral Trp may alter electrostatic repulsion if the region interacts with negatively charged DNA[33,34].

Mutations in the MmpS1 protein (MSMEG_0575) are also suggested to play a role in the mechanism of drug resistance. We found 3 variants in pooled sample B (high MIC colonies), which indicates a possible mechanism of survival. A similar article also revealed that the mechanism of resistance against imidazole [1,2-b] [1,2,4,5] tetrazines might affect mycobacterial iron metabolism. MmpS4-MmpL4 and MmpS5-MmpL5 are controlled by the *hup* (or *hlp*) gene and iron metabolism[31]. However, this study revealed a mutation in the MmpS1 and SMR family proteins, possibly contributing to the development of a multidrug-resistant strain. The MmpS1 protein (MSMEG_0575) functions to transport small molecules in the mycobacterial transmembrane, and this study is the first to report mutations in the multidrug-resistant strains of mycobacteria that are maintained consecutively in the mutant strain for 5 sequential passages. This study also confirmed cross-resistance between BDQ and clarithromycin in 2 ways[14]. These findings suggest that clarithromycin mutants also harbor mutations in efflux protein-associated genes. A mechanism involving the newly recommended drug BDQ has not yet been confirmed, owing to the many potential gene-associated targets that induce drug resistance. Based on the WHO resistance catalog for BDQ resistance markers, including *atpE* and *rv0678*, there are some SNPs that strongly correlate with BDQ resistance in the *atpE* of MTB, including Asp28Val, Asp28Pro, Glu61Asp, Ala63Pro, Ile66Met, Asp28Ala, Asp28Gly; meanwhile in Rv0678, there are also some significant SNPs, including Glu49fs, Asp47fs, Ile67fs, Gly121Arg, Leu117Arg, Met146Thr, Ile67Ser, Cys46Arg, Ala36Val, Asn70Asp, and Leu32Ser. Most mutations are missense and nonsense mutations that change the conformation of the amino acid and protein structure. Other studies have revealed that an overexpression of the membrane transport systems of MSMEG_1380 and MSMEG_1382 in *M. smegmatis* confers multidrug resistance. In the present study, we also found mutations in genes associated with the membrane transport system in *M. smegmatis*. In this study, based on bioinformatics analysis, we discovered that pooled-sample sequencing captured SNPs in the efflux family protein, ABC superfamily transporter, and the MFS proteins that are involved in the extrusion of drugs in mycobacteria[35].

Since the mechanism of multidrug resistance pathways is not yet understood, in the present study, we found that mutants resistant to BDQ and clarithromycin are becoming multidrug-resistant strains. Resistant strains in this study have mutations in the efflux family proteins. We also detected mutations in the transcriptional regulatory genes encoding the LysR and TetR family proteins. Previously reported mutations in the regulatory TetR family protein caused resistance to BDQ in *M. smegmatis*[36,37]. Mutations in the gene region that encode the MFS family protein and ABC superfamily proteins are maintained during the initial and final passages, suggesting a significant role in the emerging mechanism of multidrug resistance. A mutation in the SMR family protein was observed in pooled sample B (high MIC isolates). Previously reported mutations in mycobacterial MSMEG_3762/63/65 are involved in the extrusion of rifampicin and ciprofloxacin (CIP)[38]. In this study we detected no mutation in those genes; however, we discovered variants in genes associated with the TetR-regulatory genes MSMEG_0712/4033, which resulted in high MICs.

The *hup* (or *hlp* or MDP1) gene knockout mutant shows increased susceptibility to isoniazid in several studies, which provides evidence that the loss of this gene might contribute naturally as an efflux pump inhibitor that could work to address the emergence of multidrug resistance. The ATP levels are significantly important for *M. tuberculosis* to extrude drugs. This study found a similar mechanism in *M. smegmatis,* due to a variant found in the ABC superfamily protein. *M. smegmatis* has been reported to display upregulation of F_1F_0-ATP synthase upon exposure to ethambutol[39]. These data prove that antibiotic treatment inflicts physiological stresses on bacteria that increase energy production requirements. This study had limitations in its ability to perform comprehensive genomic analysis in order to validate the SNPs experimentally with an in vitro study; however, this provides future insight for research in genomics and to experimentally evaluate the reported variants and their associations with cross-resistance to drugs.

Conclusion

The sequence of the pooled sample can be used to collect variant sites and predict cross-resistance-associated genes. The WGS process confirmed that numerous mutations were found in the hotspot mutations in genes *atpA*, MSMEG_0639, MSMEG_1867, MSMEG_4765, and MSMEG_6302 that coded for the F1FO ATPase protein and the efflux family proteins, including the MFS and ABC superfamily proteins that were discovered in both pooled samples, and were strongly predicted to be associated with BDQ and clarithromycin drug resistance.

Acknowledgement

This research was fully supported by the PMDSU scholarship and the Laboratory of Bacteriology and Artificial Intelligence (AI) Department of the Graduate School of Medical and Dental Science, Niigata University, Japan.

Funding sources

This research is supported by a PhD fellowship of the "Pendidikan Magister Menuju Doktor untuk Sarjana Unggul" (PMDSU) Scholarship under the Directorate General of Higher Education, Research and Technology (DIKTI), Ministry of Education and Culture, Republic of Indonesia, Decree Number 1004/UN3/2022 and Agreement/Contract Number 0483/E5/DT.05.00/2024.

Author contributions

Conceptualization of the study was performed by MFS, AKS, and SKM; data analysis was performed by MFS, AKS, HYT, YST, ZKY, MAH, and SJO; the main manuscript was written by MFS; the manuscript was reviewed by SKM, WTA, SDS, NMM, YST, HYT, and YRO.

Ethical approval

This study's ethical approval certificate, No. 0757/KEPK/VIII/2023, was issued by the university's academic hospital research committee.

Data availability

Raw sequence data can be found on the NCBI public database under the BioProject accession numbers PRJNA1129677 and PRJNA1129801.

Conflict of interest

The authors declare no conflict of interest.

References

1. Guerra-Assunção JA, Crampin AC, Houben RM, Mzembe T, Mallard K, Coll F, et al. Large-scale whole genome sequencing of M. tuberculosis provides insights into transmission in a high prevalence area. Elife 2015;4:e05166.
2. Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, et al. Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. Nat Genet 2011;44:106–10.

3. Merker M, Kohl TA, Roetzer A, Truebe L, Richter E, Rüsch-Gerdes S, et al. Whole genome sequencing reveals complex evolution patterns of multidrug-resistant Mycobacterium tuberculosis Beijing strains in patients. PLoS One 2013;6:8.
4. Maladan Y, Krismawati H, Wahyuni T, Tanjung R, Awaludin K, Audah KA, et al. The whole-genome sequencing in predicting Mycobacterium tuberculosis drug susceptibility and resistance in Papua, Indonesia. BMC Genomics 2021;22:844.
5. World Health Organization. Guidance for the surveillance of drug resistance in tuberculosis. 6th ed. Geneva: WHO; 2020.
6. World Health Organization. The use of next-generation sequencing for the surveillance of drug-resistant tuberculosis an implementation manual [homepage on the Internet]. Geneva: WHO; 2024 [cited 2024 Oct 4]. Available from: https://iris.who.int/handle/10665/274453
7. World Health Organization. Module 3: Diagnosis. WHO consolidated guidelines on tuberculosis. Rapid diagnostics for tuberculosis detection. 3rd ed. Geneva: WHO; 2024.
8. Guo Q, Bi J, Lin Q, Ye T, Wang Z, Wang Z, et al. Whole genome sequencing identifies novel mutations associated with bedaquiline resistance in Mycobacterium tuberculosis. Front Cell Infect Microbiol 2022;12:807095.
9. Wang L, Yang J, Chen L, Wang W, Yu F, Xiong H. Whole-genome sequencing of Mycobacterium tuberculosis for prediction of drug resistance. Epidemiol Infect 2022;150:e22.
10. Ellingford JM, Ahn JW, Bagnall RD, Baralle D, Barton S, Campbell C, et al. Recommendations for clinical interpretation of variants found in non-coding regions of the genome. Genome Med 2022;14:73.
11. Sparks IL, Derbyshire KM, Jacobs WR, Morita YS. Mycobacterium smegmatis: the vanguard of Mycobacterial research. J Bacteriol 2023;205:e0033722.
12. Brown TS, Tang L, Omar SV, Joseph L, Meintjes G, Maartens G, et al. Genotype–phenotype characterization of Serial Mycobacterium tuberculosis Isolates in bedaquiline-resistant tuberculosis. Clin Infect Dis 2024;78:269–76.
13. Madadi-Goli N, Ahmadi K, Kamakoli MK, Azizi M, Khanipour S, Dizaji SP, et al. The importance of heteroresistance and efflux pumps in bedaquiline-resistant Mycobacterium tuberculosis isolates from Iran. Ann Clin Microbiol Antimicrob 2024;23:36.
14. Maeda T, Kawada M, Sakata N, Kotani H, Furusawa C. Laboratory evolution of Mycobacterium on agar plates for analysis of resistance acquisition and drug sensitivity profiles. Sci Rep 2021;11:15136.
15. Pandey SD, Choudhury M, Yousuf S, Wheeler PR, Gordon S V, Ranjan A, et al. Iron-regulated protein HupB of Mycobacterium tuberculosis positively regulates siderophore biosynthesis and is essential for growth in macrophages. J Bacteriol 2014;196:1853–65.
16. Shaban AK, Gebretsadik G, Hakamata M, Takihara H, Inouchi E, Nishiyama A, et al. Mycobacterial DNA-binding protein 1 is critical for BCG survival in stressful environments and simultaneously regulates gene expression. Sci Rep 2023;13:14157.
17. Nishiyama A, Shimizu M, Narita T, Kodera N, Ozeki Y, Yokoyama A, et al. Dynamic action of an intrinsically disordered protein in DNA compaction that induces mycobacterial dormancy. Nucleic Acids Res 2024;52:816–30.
18. Maslov DA, Shur K V, Vatlin AA, Danilenko VN. Mmps5-MmpL5 transporters provide Mycobacterium smegmatis resistance to imidazo[1,2-b][1,2,4,5] tetrazines. Pathogens 2020;9:166.
19. Niki M, Niki M, Tateishi Y, Ozeki Y, Kirikae T, Lewin A, et al. A novel mechanism of growth phase-dependent tolerance to isoniazid in mycobacteria. J Biol Chem 2012;287:27743–52.
20. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078–9.
21. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–60.
22. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, et al. From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. Curr Protoc Bioinformatics 2013;43:11.10.1-33.
23. Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, et al. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. Front Genet 2012;3:35.
24. Esteves LS, Gomes LL, Brites D, Fandinho FCO, Bhering M, Pereira MA da S, et al. Genetic Characterization and population structure of drug-resistant Mycobacterium tuberculosis Isolated from Brazilian patients using whole-genome sequencing. Antibiotics 2024;13:496.
25. Cox H, Goig GA, Salaam-Dreyer Z, Dippenaar A, Reuter A, Mohr-Holland E, et al. Whole-genome sequencing has the potential to improve treatment for rifampicin-resistant

tuberculosis in high-burden settings: a retrospective cohort study. J Clin Microbiol 2022;60:e0236221.
26. Hasan Z, Razzak SA, Kanji A, Shakoor S, Hasan R. Efflux pump gene single nucleotide variants associated with resistance in Mycobacterium tuberculosis isolates with discrepant drug genotypes. J Glob Antimicrob Resist 2024;38:128-39.
27. Sachan RSK, Mistry V, Dholaria M, Rana A, Devgon I, Ali I, et al. Overcoming Mycobacterium tuberculosis drug resistance: novel medications and repositioning strategies. ACS Omega 2023;8:32244-57.
28. De Siena B, Campolattano N, D'Abrosca G, Russo L, Cantillon D, Marasco R, et al. Characterization of the Mycobacterial MSMEG-3762/63 efflux pump in Mycobacterium smegmatis drug efflux. Front Microbiol 2020;11:575828.
29. da Silva PEA, von Groll A, Martin A, Palomino JC. Efflux as a mechanism for drug resistance in Mycobacterium tuberculosis. FEMS Immunol Med Microbiol 2011;63:1-9.
30. Gopi Reji J, K Edison L, Raghunandanan S, Pushparajan AR, Kurthkoti K, Ajay Kumar R. Rv1255c, a dormancy-related transcriptional regulator of TetR family in Mycobacterium tuberculosis, enhances isoniazid tolerance in Mycobacterium smegmatis. J Antibiot 2023;76:720-7.
31. Farnia P, Besharati S, Farina P, Ayoubi S, Marjani M, Ghanavi J, et al. The role of efflux pumps transporter in multi-drug resistant tuberculosis: Mycobacterial memberane protein(MmpL5). Int J Mycobacteriol 2024;13:7–14.
32. Shyam M, Thakur A, Velez C, Daniel C, Acevedo O, Bhakta S, et al. Mycobactin analogue interacting with siderophore efflux-pump protein: insights from molecular dynamics simulations and whole-cell assays. Front Antibiot 2024;3:1362516.
33. Molnár D, Surányi ÉV, Gálik N, Tóth J, Hirmondó R. Assessing the impact of bedaquiline, clofazimine, and linezolid on Mycobacterial genome integrity. Biomolecules 2024;14:1451.
34. Molnár D, Surányi ÉV, Trombitás T, Füzesi D, Hirmondó R, Toth J. Genetic stability of Mycobacterium smegmatis under the stress of first-line antitubercular agents. Elife 2024;13:RP96695.
35. World Health Organization. Catalogue of mutations in Mycobacterium tuberculosis complex and their association with drug resistance Second edition [homepage on the Internet]. Geneva: World Health Organization; 2023 [cited 2025 Mar 10]. Available from: https://www.who.int/publications/i/item/9789240082410
36. Colman RE, Seifert M, De la Rossa A, Georghiou SB, Hoogland C, Uplekar S, et al. Evaluating culture-free targeted next-generation sequencing for diagnosing drug-resistant tuberculosis: a multicentre clinical study of two end-to-end commercial workflows. Lancet Infect Dis 2025;25:325-34.
37. Rukminiati Y, Mesak F, Lolong D, Sudarmono P. First Indonesian report of WGS-based MTBC L3 discovery. BMC Res Notes 2024;17:176.
38. Krah A, Ragunathan P, Bond PJ, Grüber G. Variations of the Mycobacterium abscessus F-ATP synthase subunit a-c interface alter binding and potency of the anti-TB drug bedaquiline. Biochem Biophys Res Commun 2024;690:149249.
39. Vatlin AA, Shitikov EA, Shahbaaz M, Bespiatykh DA, Klimina KM, Christoffels A, et al. Transcriptomic Profile of Mycobacterium smegmatis in response to an Imidazo[1,2-b] [1,2,4,5] tetrazine reveals its possible impact on iron Metabolism. Front Microbiol 2021;12:724042.

Supplementary Table 1 List of potentially SNPs associated with bedaquiline and clarithromycin drug resistance

Pool sample A Gene name	Gene Id	Variant effect	Gene changes post.	Ori	Alt	Amino acid
MmpS1 protein	MSMEG_0575	Upstream variant	c.-1392delG	G	-	
Transcriptional regulator, TetR family protein	MSMEG_0712	Upstream variant	c.-3192delG	G	-	
Transcriptional regulator, MerR family protein	MSMEG_0713	Downstream variant	c.*3779delC	C	-	
Major facilitator superfamily protein	MSMEG_0716	Frameshift	c.969delC	G	-	p.Ile323fs
Antibiotic transporter	MSMEG_0763	Downstream variant	c.*2648_*2649delGC	GC	-	
Transporter ATPase	MSMEG_2459	Downstream variant	c.*3822delG	G	-	
Carbon-monoxide dehydrogenase	MSMEG_2462	Frameshift	c.2200delG	G	-	p.Glu734fs
Hypothetical protein	MSMEG_4031	Frameshift	c.140_141delGC	GC	-	p.Arg47fs
TetR-family protein transcriptional regulator	MSMEG_4033	Upstream variant	c.-1383_-1382delCG	CG	-	
ABC transporter ATP-binding protein	MSMEG_4761	Upstream variant	c.-2743C>T	C	T	
ABC transporter binding protein	MSMEG_4762	Upstream variant	c.-3870C>T	C	T	
Transcriptional regulator, MerR family protein	MSMEG_4765	Upstream variant	c.-3534delG	G	-	
Conserved hypothetical protein	MSMEG_4769	Frameshift	c.54delC	C	-	p.Gly19fs
atpA	MSMEG_4938	Missense variant	c.1285G>C	G	C	p.Glu429Gln
atpB	MSMEG_4942	Upstream variant	c.-2795G>C	G	C	
		Downstream variant	c.*3553G>C	G	C	
atpD	MSMEG_4936	Upstream variant	c.-1331G>C	G	C	
atpG	MSMEG_4937	Upstream variant	c.-370G>C	G	C	
atpE	MSMEG_4941	Downstream variant	c.*3191G>C	G	C	
Putative aliphatic sulfonates transport permease protein SsuC	MSMEG_5530	Synonymous variant	c.348C>G	C	G	p.Leu116Leu
		Frameshift	c.353dupG	G	GG	p.Leu119fs
ABC transporter ATP-binding protein	MSMEG_5531	Upstream variant	c.-522C>G	C	G	
Glutamine-binding periplasmic protein/glutamine transport system permease protein	MSMEG_6307	Frameshift	c.1192delC	C	-	p.Arg398fs
dapB		Frameshift	c.1001delA	A	-	p.Tyr334fs
hemL	MSMEG_0969	Synonymous variant	c.555T>C	T	C	p.Gly185Gly
purA	MSMEG_0759	Frameshift	c.765_766delGC	GC	-	p.Pro256fs
fabG	MSMEG_3150	Downstream variant	c.*4402A>G	A	G	
		Downstream variant	c.*2762delC	C	-	
fdhD	MSMEG_4669	Frameshift	c.170_171dupTC	TC	TCTC	p.Leu58fs
gyrA	MSMEG_0006	Downstream variant	c.*2434_*2435delTT	TT	-	
gyrB	MSMEG_0005	Downstream variant	c.*4992_*4993delTT	TT	-	
Pool sample B Gene name	**Gene Id**	**Variant effect**	**Gene changes post.**	**Ori**	**Alt**	**Amino acid**
MmpS1 protein	MSMEG_0575	Upstream gene variant	c.-4833C>A	C	A	-
			c.-4831C>G	C	G	
			c.-4830A>C	A	C	
Oligopeptide transport ATP-binding protein AppF	MSMEG_0639	Missense variant	c.815G>A	G	A	p.Arg272His
Oligopeptide transport ATP-binding protein OppD	MSMEG_0640	Downstream Gene variant	c.*811G>A	G	A	

Supplementary Table 1 (continued)

Pool sample B Gene name	Gene Id	Variant effect	Gene changes post.	Ori	Alt	Amino acid
Transcriptional regulator, LysR family protein	MSMEG_0708	Downstream gene variant	c.*1619A>G	A	G	
Transcriptional regulator, TetR family protein	MSMEG_0712	Downstream gene variant	c.*2689T>C	T	C	
Transcriptional regulator, MerR family protein	MSMEG_0713	Upstream gene variant	c.-1194A>G	A	G	
Major facilitator superfamily protein	MSMEG_0716	Upstream gene variant	c.-4364A>G	A	G	
TetR-family protein transcriptional regulator	MSMEG_0962	Downstream gene variant	c.*1222A>G	A	G	
Multidrug resistance protein, SMR family protein	MSMEG_1659	Downstream gene variant	c.*2383C>T	C	T	
Transposase	MSMEG_1867	Missense variant	c.109G>A	G	A	p.Glu37Lys
		Synonymous variant	c.108C>G	C	G	p.Leu36Leu
Transporter ATPase	MSMEG_2459	Downstream gene variant	c.*3942C>G	C	G	
Protein-export membrane protein SecF	MSMEG_2962	Downstream gene variant	c.*3635delC	C	-	
ATP/GTP-binding protein	MSMEG_3321	Downstream gene variant	c.*803delG	G	-	
Possible drug efflux membrane protein	MSMEG_3815	Downstream gene variant	c.*961G>T	G	T	
Transcriptional regulator, MerR family protein	MSMEG_4765	Missense variant	c.583A>G	A	G	p.Thr195Ala
DNA ligase	MSMEG_6302	Missense variant	c.466T>C	T	C	p.Cys156Arg
			c.463T>C	T	C	p.Trp155Arg
sucC	MSMEG_5525	Synonymous variant	c.864T>C	T	C	p.Gly288Gly

Ori=original allele, Alt=alternative allele

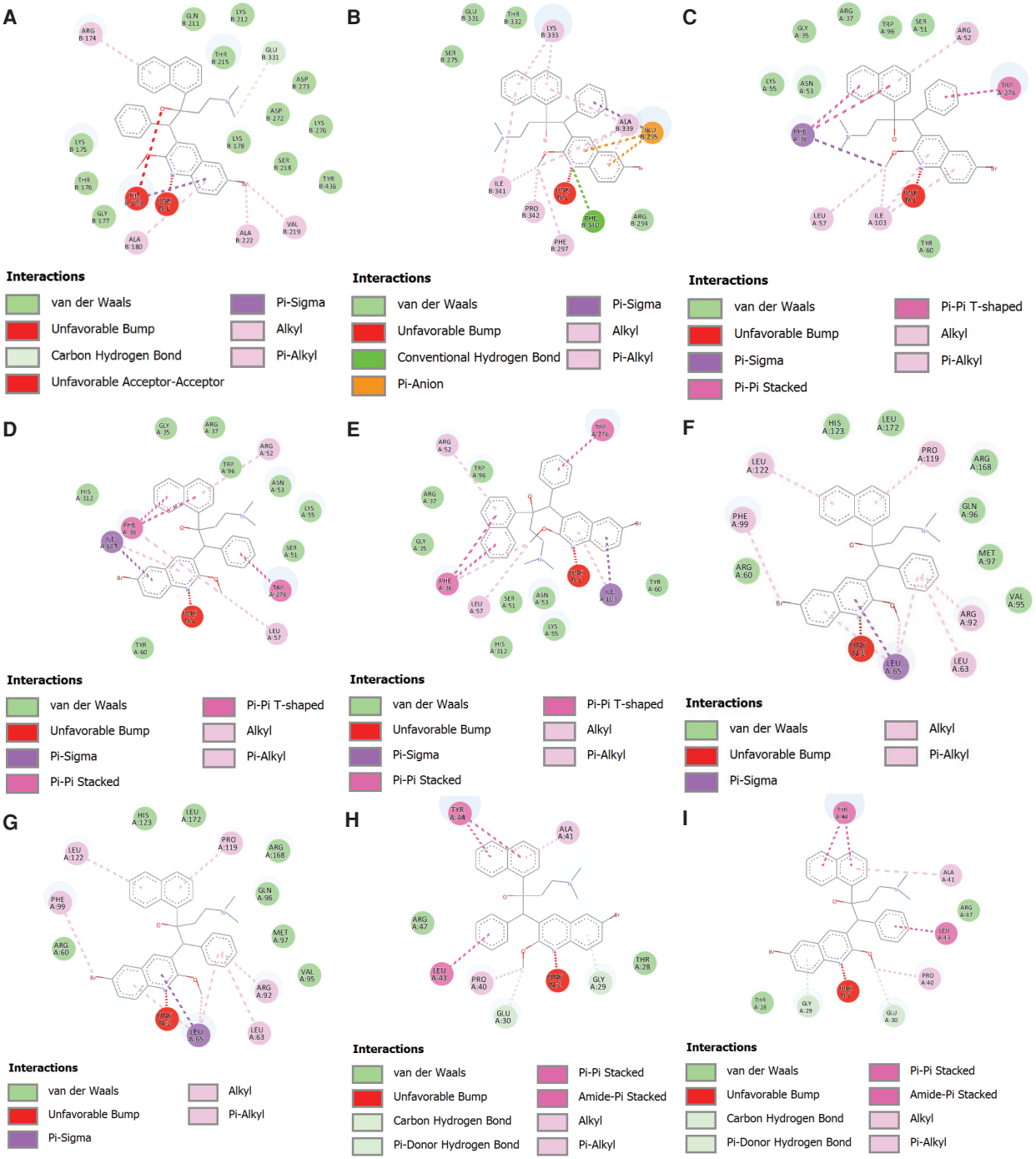
Supplementary Table 2 Descriptive statistic test result using paired T-Test

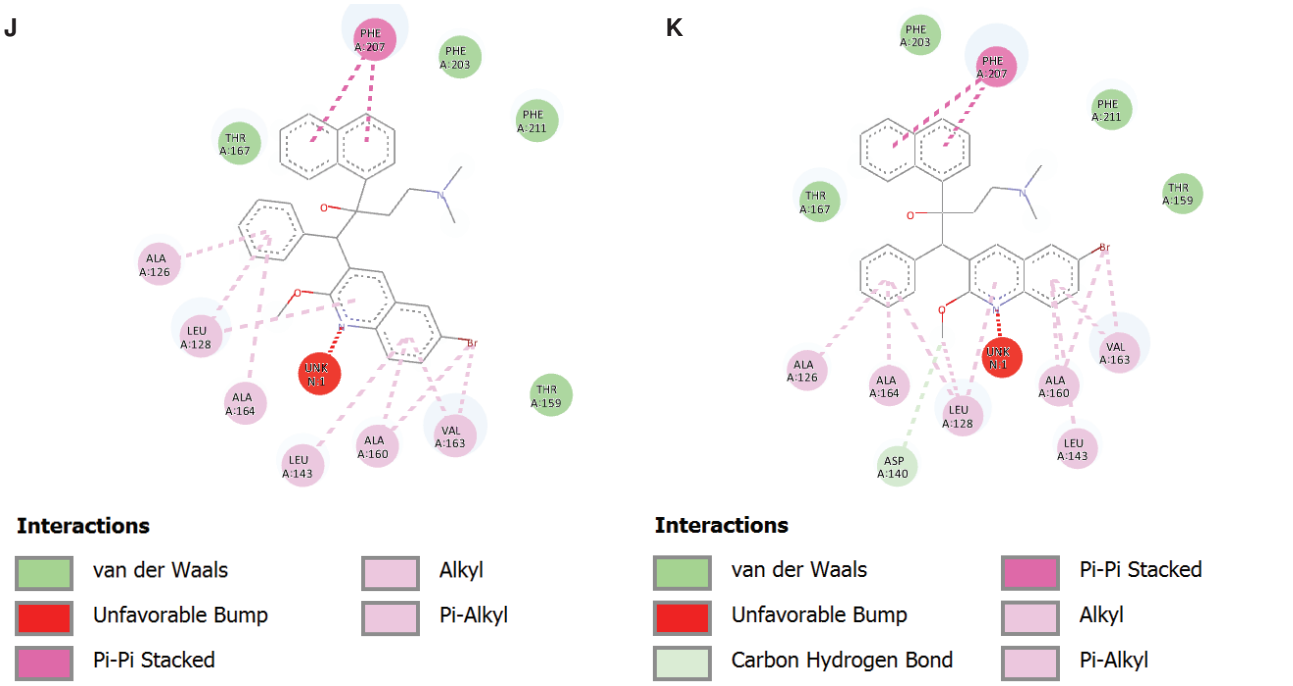
Paired t test	
Table analyzed	Data 1
Column B	B
vs.	vs.
Column A	A
Paired t test	
p-value	0.1848
p-value summary	ns
Significantly different (p-value<0.05) ?	No
One- or two-tailed p-value ?	Two-tailed
t, df	t=1.350, df=39
Number of pairs	40

Supplementary Table 2 (continued)

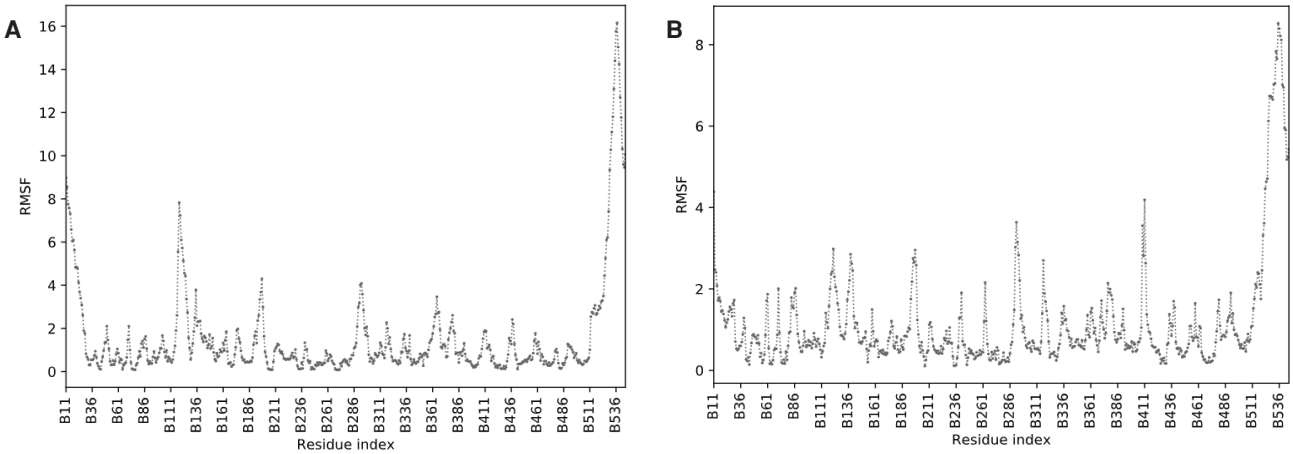
Paired t test	
How big is the difference ?	
Mean of differences (B - A)	-0.2500
S.D. of differences	1.171
SEM of differences	0.1852
95% confidence interval	-0.6246 to 0.1246
R squared (partial eta squared)	0.04464
How effective was the pairing?	
Correlation coefficient (r)	-0.5363
p-value (one tailed)	0.0002
p-value summary	***
Was the pairing significantly effective?	Yes

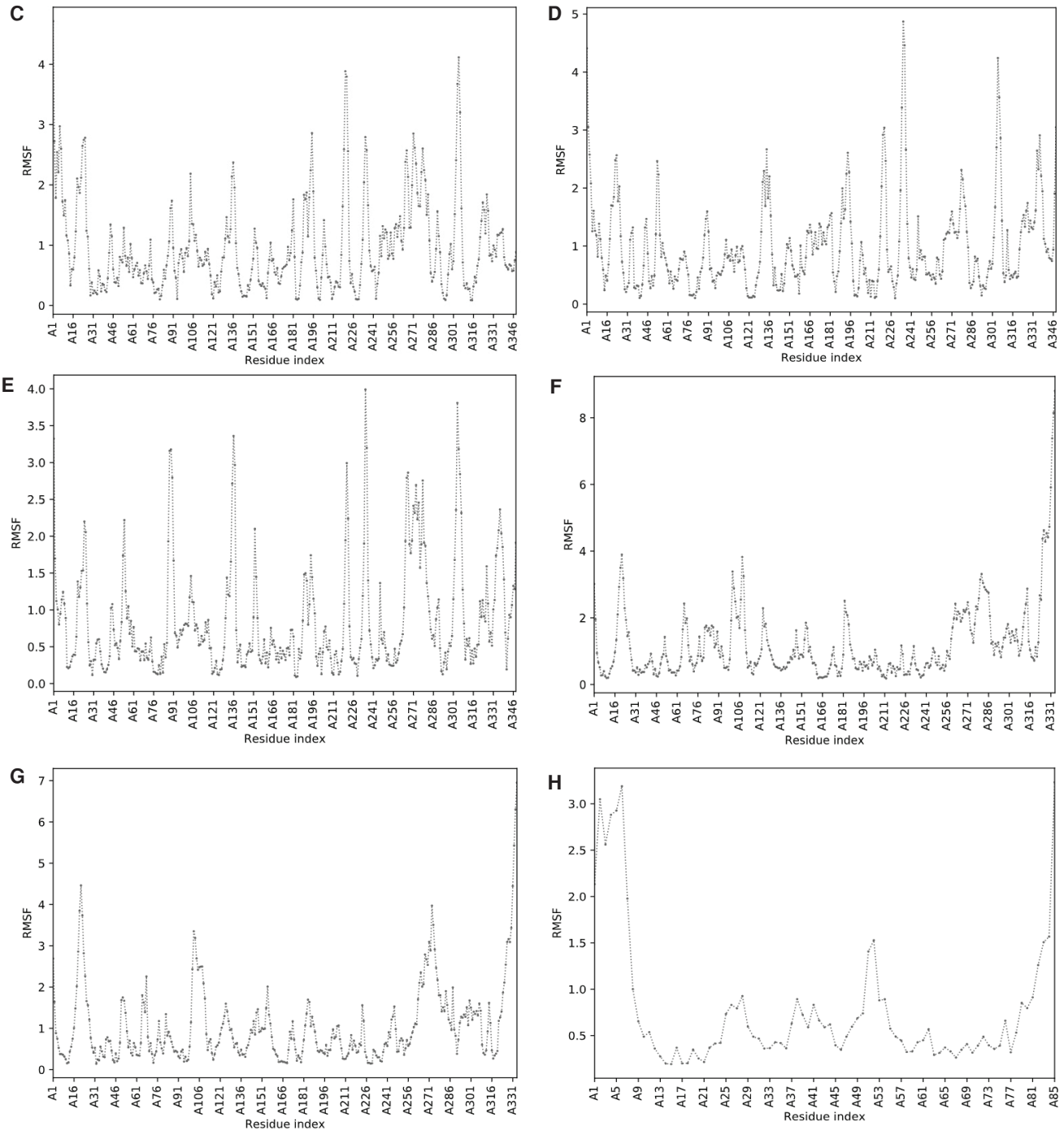
S.D.=standard deviation, SEM=standard error of the mean, df=degree of freedom, t=t-value

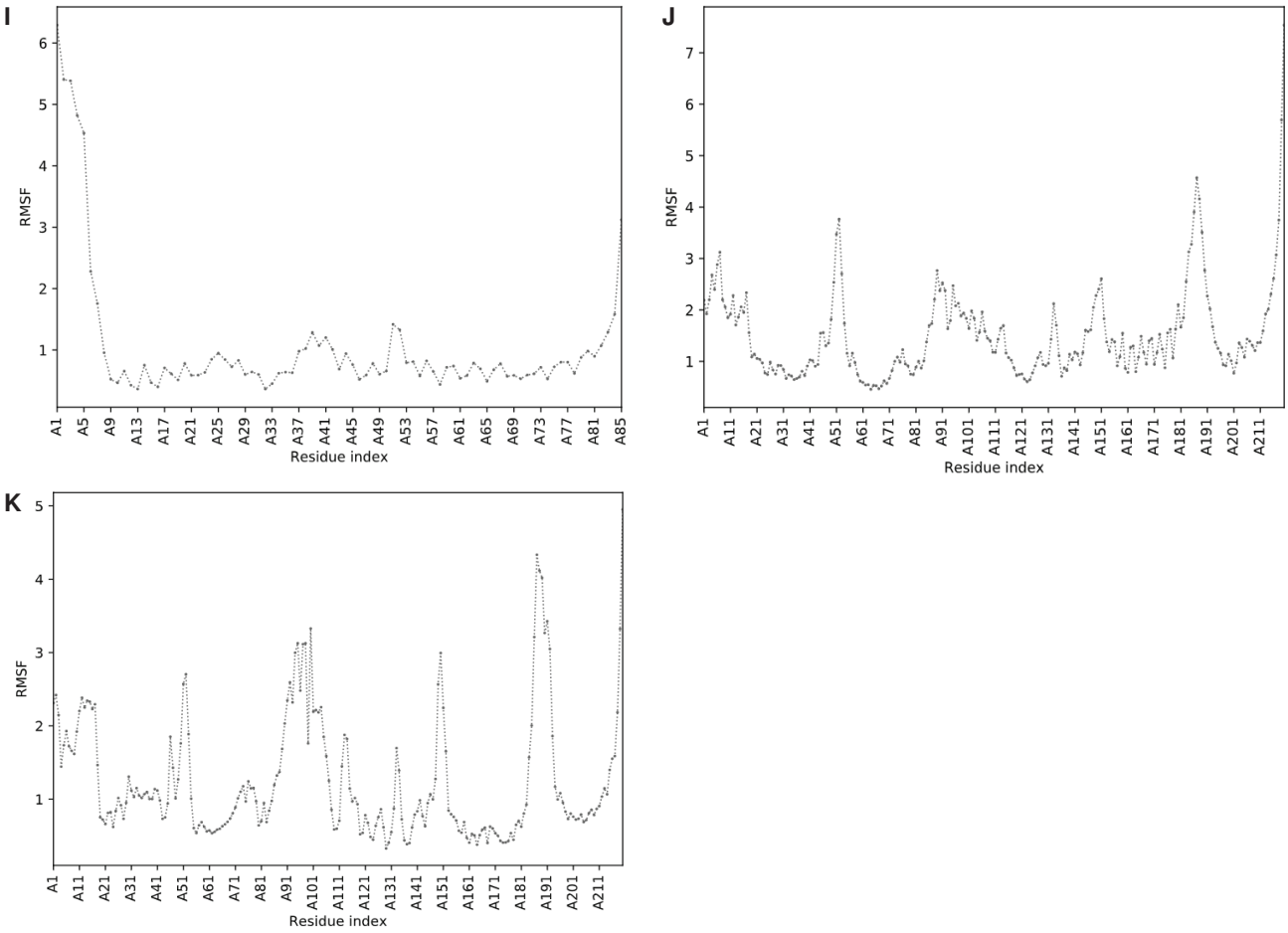

A
Interactions
van der Waals
Unfavorable Bump
Carbon Hydrogen Bond
Unfavorable Acceptor-Acceptor
Pi-Sigma
Alkyl
Pi-Alkyl
B
Interactions
van der Waals
Unfavorable Bump
Conventional Hydrogen Bond
Pi-Anion
Pi-Sigma
Alkyl
Pi-Alkyl
C
Interactions
van der Waals
Unfavorable Bump
Pi-Sigma
Pi-Pi Stacked
Pi-Pi T-shaped
Alkyl
Pi-Alkyl
D
Interactions
van der Waals
Unfavorable Bump
Pi-Sigma
Pi-Pi Stacked
Pi-Pi T-shaped
Alkyl
Pi-Alkyl
E
Interactions
van der Waals
Unfavorable Bump
Pi-Sigma
Pi-Pi Stacked
Pi-Pi T-shaped
Alkyl
Pi-Alkyl
F
Interactions
van der Waals
Unfavorable Bump
Pi-Sigma
Alkyl
Pi-Alkyl
G
Interactions
van der Waals
Unfavorable Bump
Pi-Sigma
Alkyl
Pi-Alkyl
H
Interactions
van der Waals
Unfavorable Bump
Carbon Hydrogen Bond
Pi-Donor Hydrogen Bond
Pi-Pi Stacked
Amide-Pi Stacked
Alkyl
Pi-Alkyl
I
Interactions
van der Waals
Unfavorable Bump
Carbon Hydrogen Bond
Pi-Donor Hydrogen Bond
Pi-Pi Stacked
Amide-Pi Stacked
Alkyl
Pi-Alkyl



Supplementary Figure 1 Molecular docking and molecular dynamic simulation of the wild type and mutant protein with bedaquiline using BIOVIA and Cabsflex. (A) atpA WT, (B) atpA_(p.Glu429Gln), (C) DNA ligase WT, (D) DNA Ligase (p.Cys156Arg), (E) DNA Ligase (p.Trp155Arg), (F) oligopeptide transport ATP-binding protein AppF Wild Type, (G) oligopeptide transport ATP-binding protein AppF (p.Arg272His), (H) Tranposase wild type, (I) Transposase mutant (p.Glu37Lys), (J) transcriptional regulator, MerR family protein Wild Type, (K) transcriptional regulator, MerR family protein (p.Thr195Ala)




C
D
E
F
G
H
RMSF
Residue index



(A) atpA WT, (B) atpA_(p.Glu429Gln), (C) DNA ligase WT, (D) DNA Ligase (p.Cys156Arg), (E) DNA Ligase (p.Trp155Arg), (F) oligopeptide transport ATP-binding protein AppF wild type, (G) oligopeptide transport ATP-binding protein AppF(p.Arg272His), (H) Tranposase wild type, (I) Transposase mutant (p.Glu37Lys), (J) transcriptional regulator, MerR family protein wild type, (K) transcriptional regulator, MerR family protein (p.Thr195Ala)

Supplementary Figure 2 Root mean square fluctuation graph