Evaluation of *Mycobacterium tuberculosis* rpoB for detecting rifampicin resistance

Maharani Pertiwi Koentjoro[1] and Endry Nugroho Prasetyo[2]*
1. Department of Medical Laboratory Technology, Universitas Nahdlatul Ulama Surabaya, Surabaya, INDONESIA
2. Department of Biology, Institut Teknologi Sepuluh Nopember, Surabaya, INDONESIA
*endry@bio.its.ac.id

Abstract
The resistance mechanism of Mycobacterium tuberculosis to rifampicin (RIF) involves point mutation in the 81 bp region in the rpoB, namely, rifampicin resistance–determining region (RRDR). Mutations in these regions did not fully cause resistance to RIF. Therefore, a molecular rapid test would be less effective in diagnosing multidrug–resistant TB (MDR–TB). Here, we evaluated the rpoB of M. tuberculosis to obtain the variation and differentiation of its nucleotide sequences leading to MDR–TB. A total of 6,972 M. tuberculosis rpoB were identified in the NCBI, but only 200 data were considered eligible for analysis. Inclusion criteria were nucleotide bases determined through sequencing methods and isolated from human.

We found that the polymorphic base of M. tuberculosis mutation–causing rpoB involved codons 413, 435, 451, 490, 511, 513, 516, 521, 522, 526, 530, 531 and 533. Phylogenetic tree analysis showed that mutation in the RRDR was a result of evolution in different geographical regions. This study could be used as a basis for rationally designing molecular tests to rapidly screen RIF resistance–related mutations that might contribute to MDR–TB control in Indonesia.

Keywords*: Mycobacterium tuberculosis*, Rifampicin resistance determining region, Tuberculosis, rpoB, Rifampicin.

Introduction

Tuberculosis (TB) is a lung disease caused by *Mycobacterium tuberculosis* infection and one of the causes of death worldwide[21]. In 2016, 10.4 million people were affected by TB, and 1.7 million people died; furthermore, 24% of patients with TB had HIV. According to the WHO, the incidence of TB is 64 % which is recorded from low– and middle–income or developing countries such as India, Indonesia, China, the Philippines, Pakistan, Nigeria, and South Africa[21].

The incidence of TB is exacerbated by the transition of multidrug–resistant TB (MDR–TB)[16]. The WHO estimated that 600,000 new cases have resistance to rifampicin (RIF) and isoniazid which are the most effective first–line drugs whereas 490,000 cases are MDR–TB. In Indonesia, 2 % of TB cases are expected to occur and 12 % of TB cases are estimated to be MDR–TB[21]. Furthermore, 55 % of MDR–TB cases have not been diagnosed and yet to receive correct treatment. MDR–TB cases are suspected to be those with failed treatment or strains transmitted by patients with MDR–TB[8].

Since 1957, RIF has been a widely used anti–TB drug and component of TB treatment in Indonesia[18]. RIF inhibits bacterial growth by strongly binding to DNA–dependent RNA polymerase (RNA polymerase β–subunit–bacterial DNA). It also inhibits bacterial DNA–dependent RNA synthesis, thus blocking RNA transcription by inhibiting RNA chain initiation and elongation[2,11]. However, susceptible bacteria from selective pressure may have favored the survival of resistant strains[4]. Therefore, molecular technology is needed to overcome the drawbacks of MDR–TB treatment.

RIF resistance in *M. tuberculosis* is largely due to changes in the nucleotide composition of *rpoB* encoding the RNA polymerase subunit. These alterations can cause changes in the expression of the structure and formation of *rpoB* with other RNA polymerase β–subunits. Moreover, the activity of target drugs is dysfunctional because of changes in the structure of *rpoB*[3,8].

RIF resistance inside the 81 bp core of *rpoB* is associated with alterations in the nucleotides between codons 431, 507, 531 and 533. Therefore, this region is called RIF resistance–determining region (RRDR)[17,20]. However, numerous studies have assessed that mutations in RRDR cannot be overestimated to cause RIF resistance[19]. As such, molecular rapid tests have become less effective in MDR–TB diagnosis.

Methods for rapidly diagnosing drug resistance, especially RIF resistance in TB, should be developed. In this study, as an initial step, *M. tuberculosis rpoB*, as a gene that supports resistance to RIF, was explored through meta–analysis. Further information was needed to assess the dynamic evolution of RIF resistance. Hence, meta–analysis as an analytical method of pre–existing analysis or preliminary studies is necessary to elucidate the changes in the molecular characteristics of *M. tuberculosis* rpo*B*[1].

Material and Methods

Search strategy: This research was a secondary study or meta–analysis of *M. tuberculosis* rpoB found in the DNA database. This study was designed with a retrospective observational study on *M. tuberculosis* resistance to RIF. The following diagrams were used to collect information

(Figure 1). Previous studies on selected *M. tuberculosis* isolates in human were analyzed. rpoB sequences were retrieved from the rpoB or genome sequences of *M. tuberculosis* deposited to the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) from 1990 to 2018. The keywords used were rpoB and *M. tuberculosis*. Titles, abstracts, and keywords were screened using the following search string: (rpoB) AND "Mycobacterium tuberculosis" [porgn:_txid1773]. The bibliographies of each paper that met our inclusion criteria were manually searched. The search returned 244 nucleotide data.

Inclusion and selection criteria: The suitability of the publications was assessed by scanning their titles and abstracts. References from the selected studies were also assessed to ensure that no relevant studies were omitted. The inclusion criteria or requirements applied to rpoB in this meta–analysis was as follows: (a) studies involving strain identification; (b) studies on *M. tuberculosis* with nongenetically modified rpoB (genetically modified organisms); and (c) data from nucleotide sequencing. Nucleotide data that met the next requirements were grouped on the basis of the year of study (10–20 data per year) and mutation analysis (sorted from articles or publications in the database). All potentially relevant studies had 200 data. ID data from the NCBI were written on the analyzed data.

Data analysis and interpretation: The following data were extracted manually by using Excel spreadsheets: year of publication, country, source of sample and isolates found to be phenotypically resistant to RIF. Each of the sequence of RIF–resistant bacteria was imported into Fasta format which consisted of approximately 130 nucleotide lengths.

The results were compared with *M. tuberculosis* H37Rv strain in the NCBI database with the accession number ID 888164. The RRDR nucleotide sequence was edited and processed using BLAST, MEGA7, and Phylogeny.fr (https://www.phylogeny.fr) to know the nucleotide variation[1]. The Phylogeny.fr platform was processed with the following steps: (1) Sequences were aligned with MUSCLE (v3.8.31) configured for the highest accuracy (MUSCLE with default settings). (2) After alignment, ambiguous regions (i.e. containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b) by using the following parameters: a) the minimum length of a block after gap cleaning was 10; b) no gap positions were allowed in the final alignment; c) all the segments with contiguous nonconserved positions larger than 8 were rejected and d) the minimum number of sequences for a flank position was 85 %.

(3) A phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.1/3.0 aLRT). An HKY85 substitution model was selected by assuming an estimated proportion of invariant sites of 0.403 and four gamma–distributed rate categories to account for the heterogeneity rate across sites. The gamma shape parameter was estimated directly from the data (gamma = 0.862). The reliability of the internal branch was assessed using an aLRT test (SH–like). (4) The phylogenetic tree was graphically represented and edited using TreeDyn (v198.3).

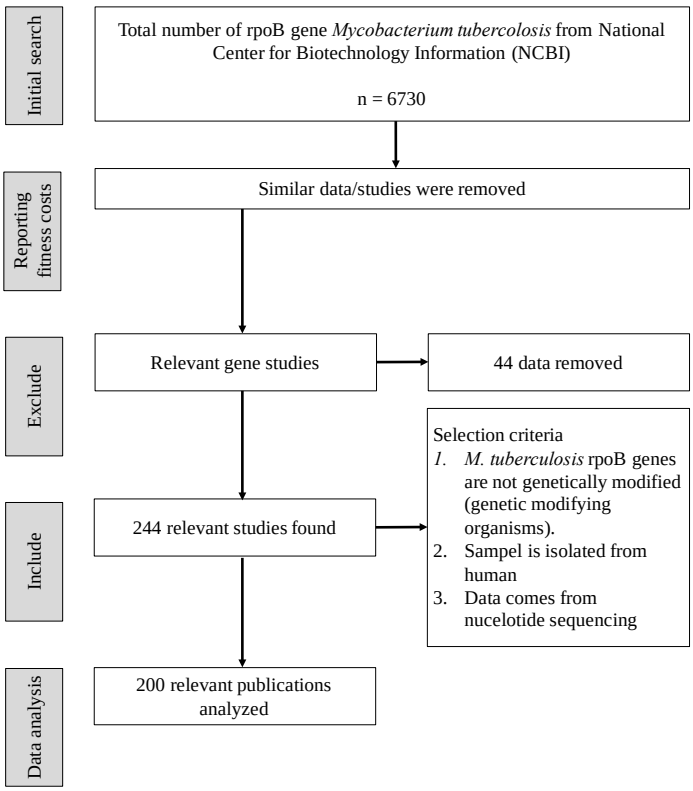

Figure 1: Selection of studies

Results

A total of 6,730 nucleotides were identified from the NCBI database. All the obtained nucleotide base sequences were checked on the basis of title and source of data. The application of the exclusion criteria is shown in figure 1. A total of 244 studies that appeared relevant were identified. However, 44 studies (18 %) did not meet our inclusion criteria.

Most of the excluded studies were eliminated for one of the following reasons: studies involving a nucleotide sequence without a report or an article, studies that did not specify *M. tuberculosis* strains, studies that were methodologically weak for our purposes, review articles/editorials, duplicate publications of the same study, studies on genetically modified organisms, and studies available only in an abstract form or had no full text. Articles for review were selected in three stages: looking at nucleotide sequences in titles, reviewing abstracts, and reading full–text articles.

Several nucleotide sequences from databases were excluded because out of scope of sequence rules and sequence shorter than 200 nucleotides.

The nucleotide sequences encoding the amino acid of the RRDR are highly conserved in *M. tuberculosis*. Figure 2 shows that *rpoB* consists of 1,173 codons (3519 bp). The sequence was divided into five fragments to facilitate analysis: NI, NII, RRDR, CII, and CIII. The RRDR was then used to identify the variation mutation.

Table 1 shows that diverse mutations occurred in *M. tuberculosis* isolates. Numerous nucleotide changes had mutation at codon 526. Single–nucleotide substitutions caused amino acid changes in all the mutations. Remarkably, two substitutions at codon 516 were found in the Asp/Val amino acid change.

RIF resistance is strongly correlated with mutations in a small stretch of DNA in *rpoB* which encodes the RNA polymerase β–subunit. These mutations are responsible for a RIF resistance phenotype[6,22]. The tree obtained in our study was better resolved than a sequence–based tree in terms of confidence values. The phylogenetic tree was arranged to determine the rearrangement of the RRDR of *M. tuberculosis* isolates worldwide (Figure 3). It was drawn to scale and its branch lengths were measured in terms of the number of mutations per site. All positions with gaps and missing data were eliminated.

The sequences are the closest distance indicated by the highest similarity or level of kinship. Four subgroups were identified in the tree. This result suggested that various strains from different origins underwent gene rearrangements. For example, JN819069, JQ414013, KP658716, AY544974, AF057454, and JX303328 came from the Netherlands, China, India, France, South Korea, and Brazil respectively. The six isolates have the same base nucleotide arrangement and are at the highest level of kinship. JQ414013 and KP658716 showed a separation level which indicated that these isolates have the lowest kinship level among the isolates.

Single–gene sequence alteration is based on geographical location. Thus, gene order method was used to determine the distinct phenotypic characteristics of *M. tuberculosis* isolates compared in this study.

The sequence of the molecular data responsible for the RIF resistance of *M. tuberculosis* was analyzed[15]. Our initial analysis of the entire sample of studies on *rpoB* revealed that nucleotide bases in *rpoB* changed and consequently caused an amino acid change compared with that in *M. tuberculosis* H37Rv *rpoB* as a wild type. The results of the extraction of *rpoB* from the NCBI bank gene showed that *M. tuberculosis* isolates that experienced resistance to RIF antibiotics came from various geographic regions. The analysis of gene bank data revealed that *rpoB* mutation varied in *M. tuberculosis* isolates from patients who had TB and were resistant to RIF.

Mutations in *rpoB* are found in different geographical regions including continental Europe, the United States of America, and Asia. This database was needed to determine the influence of geography on the conservative regions of mutations that occur in *M. tuberculosis*[7,10,12-14].

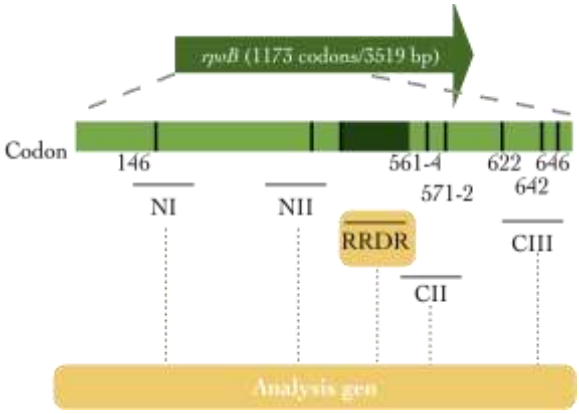

Figure 2: An overview of the rpoB gene region that has mutations causing resistance to RIF. Fragments of NI, NII, RRDR, CII and CIII were made to facilitate primary manufacturing.

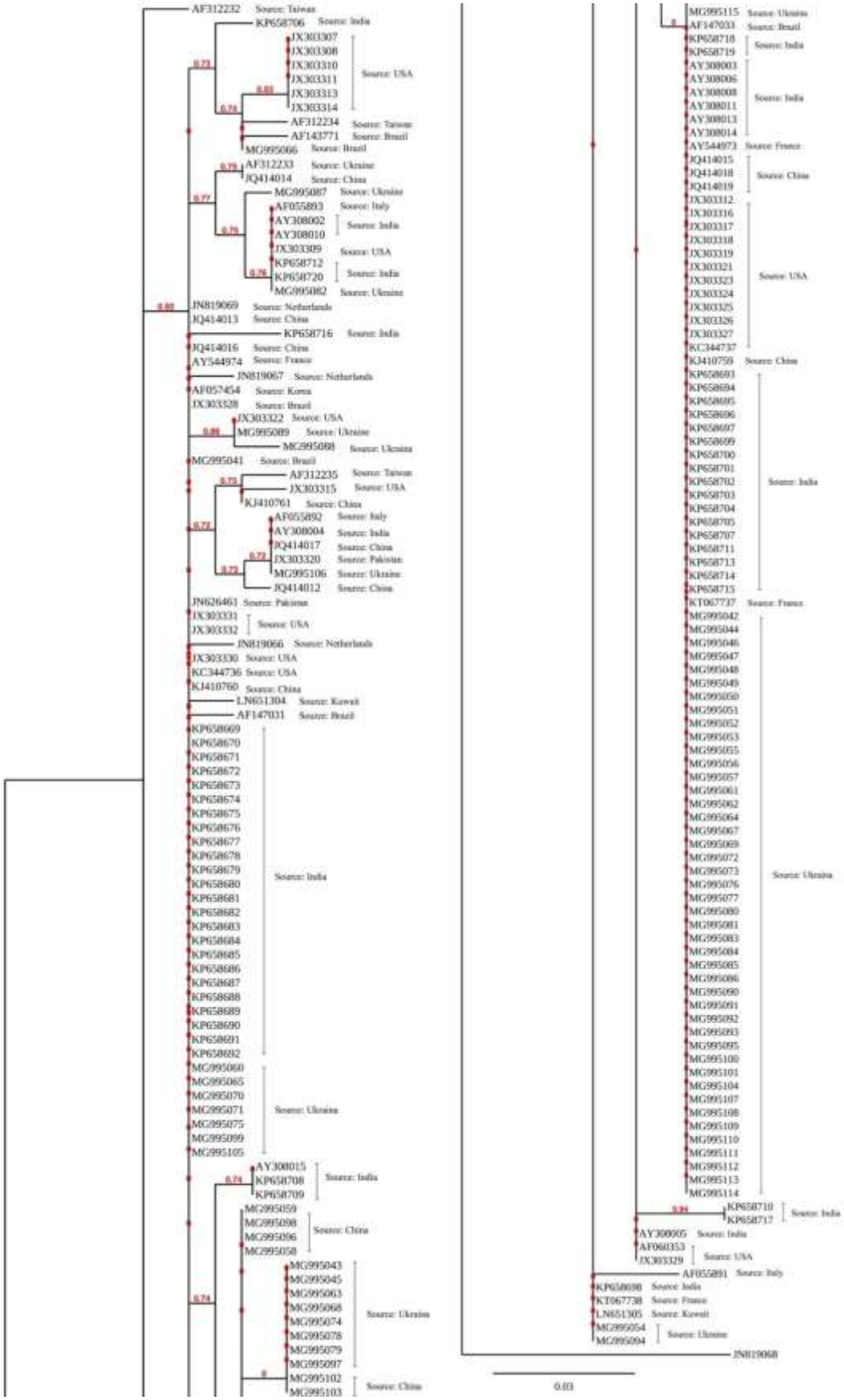

Fig. 3: Phylogenetic tree mutations in the RRDR of *M. tuberculosis rpoB*. A total of 200 published sequences (NCBI accession number) are reported from different countries. The scale bar represents amino acid changes per site.

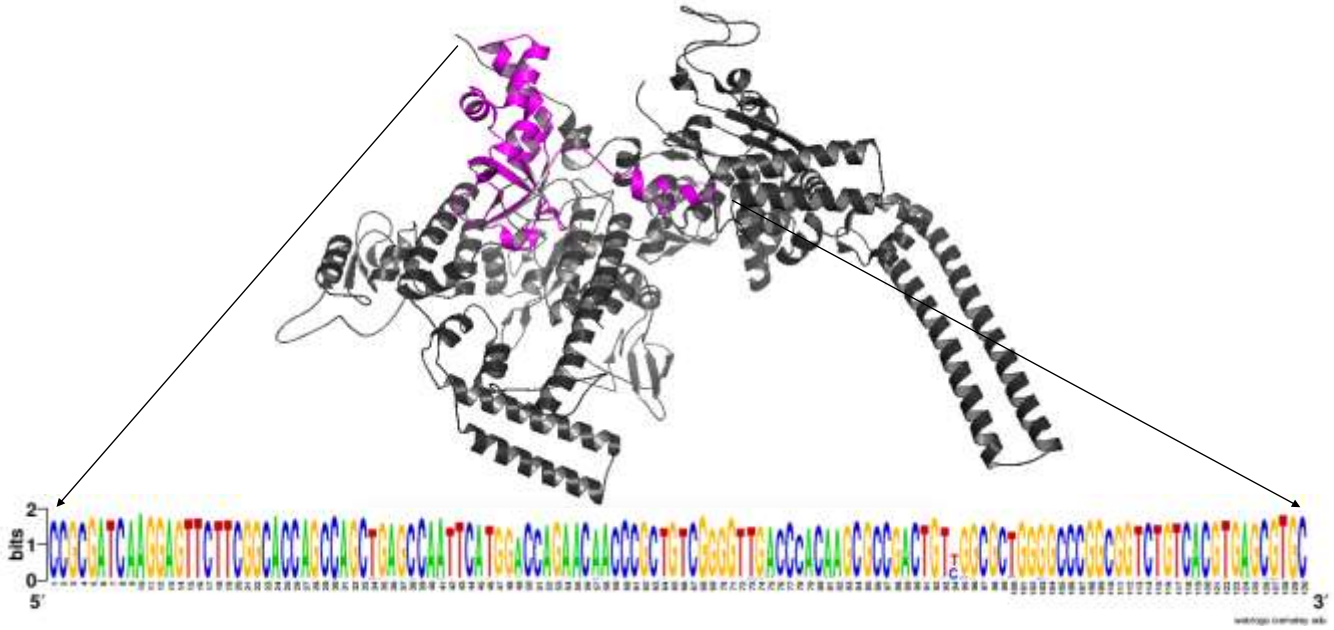

Figure 4: Cartoon diagram of the superposition of the RRDR of DNA–directed RNA polymerase subunit alpha (PDB 5UHB) HBV. Consensus nucleotide base sequence appears at the under of the cartoon diagram. Arrow and magenta color indicate nucleotide base sequence shows high mutations. Figures were regenerated by (PS)2v2: Protein Structure Prediction Server and http://weblogo.threeplusone.com

Table 1
Mutations observed in *M. tuberculosis* rpob gene

Codon No	Nucleotide change Wild type → Mutant	Amino Acid Change
413	AAC → CAC	Asn → His
435	GAC → GAG	Asp → Glu
451	GCA → GAC	Ala → Asp
490	CAG → CAT	Gln → His
	CAG → CGG	Gln → Arg
511	CGC → TGC	Arg → Cys
513	GTC → GAC	Val → Asp
516	GAC → GTG	Asp → Val
	GAC → GTC	Asp → Val
521	GAG → GAC	Glu → Asp
522	TCG → TTG	Ser → Leu
526	CAC → CTC	His → Leu
	CAC → TAC	His → Tyr
	CAC → GAC	His → Asp
	CAC → CGC	His → Arg
	CAC → AGC	His → Ser
	CAC → ACC	His → thr
530	CTG → ATG	Leu → Met
531	TCG → TTG	Ser → Leu
	TCG → TGG	Ser → Trp
533	CTG → CCG	Leu → Pro

Our data would be needed as the first step in designing a region–specific rapid method to detect most strains resistant to RIF in each region.

The nucleotide base change in *rpoB* that made the RNA polymerase β–subunit less favorable to the binding of other RNA polymerase subunits was determined (Figure 4). The nucleotide base changes in RRDR associated with the mutation in *rpoB* are evident in a biochemical assay[12].

Discussion

Our nucleotide base analysis revealed that mutations in the form of nucleotide base replacement at nucleotide bases caused changes in the results of *rpoB* expression namely

rpoB. Mutations in codon 531 (serine–leucine) in the yellow region induced the changes in the nature and structure of these amino acids. The gel shift mobility assay test showed a decrease in the *rpoB* binding activity which resulted in the RIF resistance phenotype. This change influenced the effectiveness of drugs.

In sequence alignment analysis, 200 sequences of *M. tuberculosis* isolates were compared with the *M. tuberculosis* H37Rv sequence as a control, and the nucleotide bases with mutations in the RRDR were traced. Data in the NCBI indicated that these 200 sequences encoded the RIF resistance trait in *M. tuberculosis* isolates. Table 1 shows the codon number, nucleotide changes, and amino acid changes. Similar to other previous studies from various regions of the world, our study revealed that Ser531Leu and His526Cys were the most common point mutations. These mutations showed a variety of entrained mutations in *RRDR* in *M. tuberculosis* BPB.

Information on the genetic variation and differentiation of *rpoB* from *M. tuberculosis* against RIF resistance from RRDR caused genetic changes including point mutation, which is the replacement of one nucleotide base, resulting in changes in gene expression. The analysis of primary data indicated that the RRDR in *rpoB* was the result of evolution or RIF consumption in a certain period. The mutation data obtained in this research could be useful for the rational design of molecular tests to rapidly screen RIF resistance–related mutations as one way of controlling MDR–TB. Therefore, the meta–analysis could be used to monitor the changes in RIF–resistant strains and TB distribution.

Conclusion

Our meta–analysis reveals that the resistance of *M. tuberculosis* to RIF has reached alarming levels which greatly affect efficacy treatment. These mutations are caused by the polymorphism of *rpoB* in various nucelotide bases. Mutations in *rpoB* involved codons 413, 435, 451, 490, 511, 513, 516, 521, 522, 526, 530, 531, and 533. Point mutations in one nucleotide base led to changes in the *rpoB* structure probably through evolution and antibiotic use.

Variations in nucleotide base mutations occurred in different geographical regions. The analysis of a set of studies generated several predictors that were useful for rationally designing molecular tests to rapidly screen RIF resistance–related mutations that might contribute to MDR–TB control in Indonesia.

Acknowledgement

This work was supported by Universitas Nahdlatul Ulama Surabaya and Institut Teknologi Sepuluh Nopember.

References

1. Bell B.G., Schellevis F., Stobberingh E., Goossens H. and Pringle M., A systematic review and meta–analysis of the effects of antibiotic consumption on antibiotic resistance, *BMC Infectious Diseases*, **9**, 13 **(2014)**

2. Campbell E.A. et al, Structural mechanism for rifampicin inhibition of bacterial RNA Polymerase, *Cell Press*, **104**, 901–912 **(2001)**

3. Comas I. et al, Whole–genome sequencing of rifampicin–resistant *M. tuberculosis* strains identifies compensatory mutations in RNA Polymerase, *Nature Genetic*, **44**, 106–110 **(2012)**

4. Davies J. and Davies D., Origins and evolution of antibiotic resistance, *Microbiology and Molecular Biology Reviews*, **74**, 417–33 **(2016)**

5. Dereeper A., Audic S., Claverie J.M. and Blanc G., BLAST–EXPLORER helps you building datasets for phylogenetic analysis, *BMC Evolutionary Biology*, **10**, 1471–2148 **(2010)**

6. Khan M.M., Khan M. H., Alves M.S., Sherwani S.K., Baig M.A. and Kamal M., Structure based assessment of rpoB gene from multiple–drug resistant *Mycobacterium tuberculosis* clinical isolate, *International Journal of Advanced Research*, **1**, 567–575 **(2013)**

7. Kit G. et al, Mutations outside the rifampicin resistance–determining region associated with rifampicin resistance in *Mycobacterium tuberculosis*, *Journal of Antimicrobial Chemotherapy*, **66**, 730–733 **(2011)**

8. Lahti J.L., Tang G.W., Capriotti E., Liu T. and Altman R.B., Bioinformatics and variability in drug response: a protein structural perspective, *Journal of The Royal Society Interface*, **9**, 1409–1437 **(2012)**

9. Lechartier B., Rybniker J., Zumla A. and Cole S.T., Tuberculosis Drug Discovery in the Post–Post–Genomic Era, *EMBO Molecular Medicine*, **6**, 158–168 **(2014)**

10. Lawrence K.A., Stefans M., Yates Christopher M., Mark W.N. and Michael J.E.S., The phyre 2 web portal for protein modeling, prediction and analysis, *Nature Protocols*, **10**, 845–858 **(2015)**

11. Ma C., Yang X. and Lewis P.J., Bacterial transcription as a target for antibacterial drug development, *Microbiology and Molecular Biology Reviews*, **80**, 139–160 **(2016)**

12. Meftahi N., Namouchi A., Mhenni B., Brandis G., Hugher D. and Mardassi H., Evidence for the critical role of a secondary site rpoB mutation in the compensatory evolution and successful transmission of an MDR tuberculosis outbreak strain, *Journal of Antimicrobial Chemotherapy*, **71**, 324–332 **(2016)**

13. Molodtsov V., Scharf N.T., Stefan M.A., Garcia G.A. and Murakami K.S., Structural basis for rifamycin resistance of bacterial RNA polymerase by the three most clinically important RpoB mutations found in *Mycobacterium tuberculosis*, *Molecular Microbiology*, **103**, 1034–1045 **(2017)**

14. Noura M.A.M., Ahmad S., Mokaddas E., Eldeen H.S. and Joseph S., Occurrence of disputed rpoB mutations among *Mycobacterium tuberculosis* isolates phenotypically susceptible to rifampicin in a country with a low incidence of multidrug-resistant tuberculosis, *BMC Infectious Diseases*, **19**, 3 **(2019)**

15. Palomino J.C. and Martin A., Drug Resistance Mechanisms in *Mycobacterium tuberculosis*, *Antibiotics*, **3**, 317–40 **(2014)**

16. Rajendra P., Gupta N. and Banka A., Multidrug–Resistant Tuberculosis/Rifampicin–Resistant Tuberculosis: Principles of Management, *Lung India*, **35**, 78–81 **(2018)**

17. Riccardi G., Pasca M.R. and Buroni S., *Mycobacterium tuberculosis*: drug resistance and future perspectives, *Future Microbiology*, **4**, 597–614 **(2009)**

18. Soeroto A.Y., Lestari B.W., Santoso P., Chaidir L., Andriyoko B., Alisjahbana B., van Crevel R. and Hill P.C., Evaluation of Xpert MTB–RIF Guided Diagnosis and Treatment of Rifampicin–Resistant Tuberculosis in Indonesia: A Retrospective Cohort Study, *PLoS ONE*, **14**, e0213017 **(2019)**

19. Takawira F.T., Mandishora R.S.D., Dhlamini Z., Munemo E. and Stray–Pedersen B., Mutations in rpoB and katG genes of multidrug resistant *Mycobacterium tuberculosis* undetectable using genotyping diagnotic methods, *The Pan African Medical Journal*, **27**, 145 **(2017)**

20. Ullah I.A.A. et al, Rifampicin resistance mutations in the 81 bp RRDR of rpoB gene in *Mycobacterium tuberculosis* clinical isolates using Xpert MTB/RIF in Khyber Pakhtunkhwa, Pakistan: a retrospective study, *BMC Infectious Diseases*, **16**, 413 **(2016)**

21. World Health Organization, Global Tuberculosis Report 2018, WHO Press, France, 1–4 **(2018)**

22. Zheng H. et al, Genetic basis of virulence attenuation revealed by comparative genomic analysis of *Mycobacterium tuberculosis* strain H37Ra versus H37Rv, *PLoS One*, **3**, e2375 **(2011)**.

(Received 10th August 2020, accepted 05th October 2020)