Detection and Characterization of Mutations of MultiDrug-Resistant Tuberculosis Isolates of the Philippine General Hospital

Franco Antonio C. Catangui,[1] Ann Camille Q. Yuga,[1] Sharie Keanne C. Ganchua,[2] Concepcion F. Ang,[3] Myrna T. Mendoza,[3] Marilen M. Parungao-Balolong,[1,2] Raul V. Destura[2] and Francisco M. Heralde III[4]

[1] *Department of Biology, College of Arts and Sciences, University of the Philippines Manila*
[2]*Institute of Biotechnology and Molecular Biology, National Institutes of Health, University of the Philippines Manila*
[3]*Section of Infectious Diseases, Department of Medicine, College of Medicine and Philippine General Hospital, University of the Philippines Manila*
[4]*Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines Manila*

ABSTRACT

Background. Emergence of multidrug-resistant tuberculosis (MDR-TB) poses a major challenge to prevailing disease management. MDR-TB arises from mutations in several genes comprising the resistance determining regions, including rpoB, katG and gyrA.

Objective. To detect and characterize mutations in rpoB, katG and gyrA.

Methods. Thirty selected Mycobacterium tuberculosis isolates from the IDS-PGH were subjected to PCR amplification and sequencing. Sequences were compared to the wild type strain H37Rv.

Results. Mutations were detected in codons 512, 513, 516, 522, 526, 531 and 533 of *rpoB*, codons 280, 281, 315 and 333 of *katG*, and codons 90 and 94 of *gyrA* sequences. The most frequently mutating codons for rpoB, katG and gyrA were 531, 315 and 94, respectively. A clustering analysis of the sequences showed occurrence of seven, four and three clusters for the genes rpoB, katG and gyrA, respectively. The eight clusters obtained from the concatenated sequences of the three genes represent the eight potential genotypes of local strains. One cluster represents the wild type strain genotype, another cluster represents the XDR strain genotype, and six clusters represent the MDR strain genotypes.

Conclusion. These findings indicate the utility of multiple RDR sequence analysis in both identifying specific drug resistance mutation and genotyping of various M. tuberculosis isolates.

Key Words: Multi-drug Resistance Tuberculosis, rpoB, katG

Introduction

For years, effective treatments have been utilized to control tuberculosis (TB). However, the emergence of multidrug-resistant tuberculosis (MDR-TB) made the disease management more challenging than ever. Thus, it has become the second-leading cause of mortality among infectious diseases worldwide.[1] MDR-TB occurs when Mycobacterium tuberculosis, the causative agent for TB becomes resistant to the first line drugs, rifampicin (RIF) and isoniazid (INH).[2] Drug resistance in M. tuberculosis has been associated with mutations in several genes such as rpoB, katG, and gyrA. These genes encode for RNA polymerase β-subunit, catalase-peroxidase, and the A subunit of DNA gyrase, respectively.[2] This study focused on these resistance determining genes for rifampicin, isoniazid, and ciprofloxacin, as these drugs are among those primarily included in the standard treatment regimen for TB in the Philippines. The first two drugs are the first-line drugs against TB and ciprofloxacin, a fluoroquinolone drug, is prescribed for MDR cases.[3] The development of a nucleic acid-based assay capable of rapid MDR-TB detection is one important approach in the management of TB. This approach, however, will not progress without a priori characterizing the target population genome to amplify. Implicit is the requirement for sequence analysis of candidate resistance genes that would reveal mutations prevalent among the local isolates. This information would constitute the initial database for various downstream strategies for TB management namely disease diagnosis, epidemiology and therapy.

Multi-loci sequence alignment (MLSA) has found several applications in molecular phylogenetic analysis of various organisms ranging from bacteria to higher animals. This arises from the apparent weakness of single genes in resolving phylogenetic ambiguities. In this study, detection

Presented at the Young Scientists' Forum, 35th Annual Philippine Society for Biochemistry and Molecular Biology (PSBMB), December 2, 2008, National Institute for Science and Mathematics Education Development (NISMED) auditorium, University of the Philippines, Diliman, Quezon City.

Corresponding author: Marilen M. Parungao-Balolong, MSc
Department of Biology
College of Arts and Sciences
University of the Philippines Manila
Institute of Molecular Biology and Biotechnology
National Institutes of Health
University of the Philippines Manila
625 Pedro Gil Street, Ermita, Manila 1000 Philippines
Telephone: +632 5265861
Email: lhenparungao@yahoo.com / mmparungao@post.upm.edu.ph

and characterization of mutations is done for the rpoB, katG and gyrA genes of multidrug-resistant M. tuberculosis samples of the IDS-PGH. MLSA is utilized to establish the genotype of these different MTB samples.

Methods

Thirty (30) heat-deactivated M. tuberculosis isolates were provided by the Infectious Disease Section of the Philippine General Hospital (IDS-PGH) (Figure 1). Three to five μL aliquots of deoxyribonucleic acid (DNA) isolates were used as template for the polymerase chain reaction (PCR). The primers used in the study were those previously validated for the genes rpoB, katG and gyrA.[5] The rpoB forward and reverse primers were 5′-GGGAGCGGATGACCACCCAGG-3′ and 5′-CAGCGAGCCGATCAGACCGATG-3′, respectively. For katG, the forward and reverse primers were 5′-CTCGGCGATGAGCGTTACAGCG-3′ and 5′-CCCGCAGCGAGAGGTCAGTGG-3′, respectively. The forward and reverse primers for gyrA gene were 5′-GCAGCTACATCGACTATGCGATG-3′ and 5′-CGGGCTTCGGTGTACCTCATCG-3′, respectively.

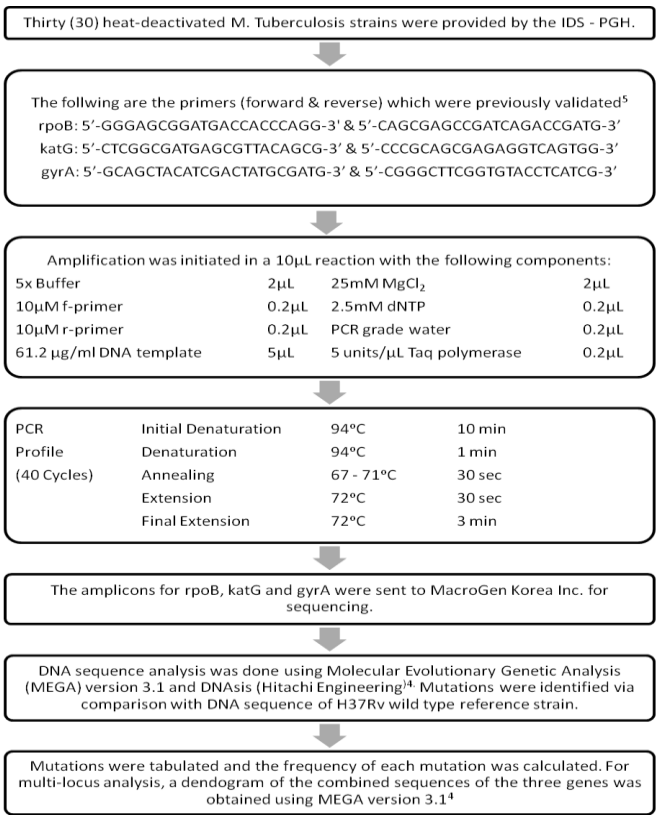

Figure 1. Flowchart of the methodology

Amplification was initiated in a 10μL reaction containing 2μL of 5x buffer, 0.2μL of 10 μM of each primers, 0.2μL of 2.5mM dNTP, 2μL of 25mM Magnesium Chloride, 0.2μL of 5 units/μL Taq Polymerase, 5μL of 61.2 μg/mL DNA template and 0.2μL of PCR-grade water. The PCR profile included an initial denaturation of 10 min at 94°C, followed by 40 cycles of denaturation for 1 min at 94°C, annealing for 30 sec at 67-71°C depending on the primers, and extension for 30 sec at 72°C; and terminated with a final extension of 3 min at 72°C.

The amplicons for rpoB, katG and gyrA were sent to MacroGen Korea Inc. for sequencing.[4] DNA sequence analysis was done using Molecular Evolutionary Genetic Analysis (MEGA) version 3.1 and DNAsis (Hitachi Engineering). Mutations were identified by comparing the obtained sequences to the DNA sequence of H37Rv wild type reference strain. The mutations of each gene were tabulated and the frequency of each mutation was calculated. For multi-locus analysis, a dendogram of the combined sequences of the three genes was obtained using MEGA version 3.1.[4]

Results

rpoB

Six codons of the rpoB sequence of the thirty isolates exhibited single-point mutations (Table 1). Codon 513 changed from CAA to CCA (Glutamine to Proline); Codon 516 mutated into two types of amino acid: GTC replaced GAC (Aspartate to Valine) in the first type and in the second type, the codon changed to TTC (Aspartate to Phenylalanine). Mutation in codon 522 replaced TCG to TTG (Serine to Leucine). Codon 526 mutated into three types of amino acid similar to codon 516: CAC (Histidine) changed to TAC (Tyrosine), TGC (Cysteine) or CGC (Arginine). Each of these mutations comprises 3.33% of the total isolates. Several isolates mutated in codon 531; where a change occurred from TCG to TTG (Serine to Leucine). On the other hand, codon 533 changed from CTG to CCG (Leucine to Proline) occurring in 6.67% of the isolates.

A double-point mutation was observed in 6.67% of the samples. Codon 512 mutated from AGC to ACC (Serine to Threonine); while codon 516 changed from GAC to TAC (Aspartate to Tyrosine). Thirty percent of the samples did not exhibit any mutation and are considered to be wild type.

The clustering of the samples according to their mutations in the rpoB gene is shown in Figure 2. The dendogram showed seven clusters, which included the RIF-resistant strains (i.e., with accession numbers). The first cluster, which included the non-mutated sample, is represented by sample 4. The samples were grouped together with the H37Rv rpoB sequence. The second cluster, represented by sample 1, consisted of ten samples. The third cluster, consisted of four samples, was represented by sample 14. The fourth and fifth cluster each has a single sample. The sixth cluster consisted of three samples, which is represented by sample 20. The last cluster had two samples represented by sample 19.

Table 1. Characteristics of *M. tuberculosis* samples from the Infectious Disease Section-Philippine General Hospital

Sample No.	Phenotypic Results[a*]	Mutations[δ]								
		rpoB			*katG*			*gyrA*		
		Codon	Nucleotide	Amino Acid	Codon	Nucleotide	Amino Acid	Codon	Nucleotide	Amino Acid
1	R^r	531	TCG to TTG	Ser to Leu						
4	I^r, C^r				315	AGC to ACC	Ser to Thr		NM	
6	I^r				315	AGC to AAC	Ser to Asn			
9	I^r				315					
11	I^r									
13	I^r, C^r				315	AGC to AAC	Ser to Asn	94	GAC to TTC	Asp to Phe
14	R^r, I^r	512 516	AGC to ACC GAC to TAC	Ser to Thr Asp to Tyr	315	AGC to ACC	Ser to Thr			
15	R^r, I^r	531	TCG to TTG	Ser to Leu		NM				
16	R^r, I^r	531	TCG to TTG	Ser to Leu	280 315	TCG to CCG AGC to ACC	Ser to Pro Ser to Thr			
17	R^r, I^r	512 516	AGC to ACC GAC to TAC	Ser to Thr Asp to Tyr		NM				
18	R^r, I^r	531	TCG to TTG	Ser to Leu	333	CTC to CTG	Leu to Leu			
19	R^r, I^r	533	CTG to CCG	Leu to Pro		NM		94	GAC to GGC	Asp to Gly
20	R^r, I^r, C^r	526	CAC to TAC	His to Tyr	281 315	GCC to GTC AGC to AGA	Ala to Val Ser to Arg	90	GCG to GTG	Ala to Val
21	R^r, I^r	531	TCG to TTG	Ser to Leu	315	AGC to ACC	Ser to Thr			
22	R^r, I^r	531	TCG to TTG	Ser to Leu	315	AGC to ACC	Ser to Thr			
23	R^r, I^r		NM		315	AGC to ACC	Ser to Thr			
24	R^r, I^r	531	TCG to TTG	Ser to Leu	315	AGC to ACC	Ser to Thr			
25	R^r, I^r	531	TCG to TTG	Ser to Leu	315	AGC to ACC	Ser to Thr			
30	R^r, I^r, C^r	526	CAC to TGC	His to Cys		NM			NM	
31	R^r, I^r	516	GAC to GTC	Asp to Val	315	AGC to ACC	Ser to Thr			
35	R^r, I^r	531	TCG to TTG	Ser to Leu	315	AGC to ACC	Ser to Thr			
37	R^r, I^r	516	GAC to TTC	Asp to Phe	315	AGC to ACC	Ser to Thr			
39	R^r, I^r, C^r	526	CAC to CGC	His to Arg		NM			NM	
40	R^r, I^r	513	CAA to CGA	Gln to Pro		NM				
41	R^r, I^r	531	TCG to TTG	Ser to Leu	315	AGC to ACC	Ser to Thr			
42	R^r, I^r	522	TCG to TTG	Ser to Leu	315	AGC to ACC	Ser to Thr			
43										
44										
45	I^r	533	CTG to CCG	Leu to Pro	315	AGC to ACC	Ser to Thr			
46	R^r		NM							

[a] *R^r, rifampicin resistant; I^r, isoniazid resistant; C^r, ciprofloxacin resistant*

[*] *The phenotypic results were obtained from Mrs. Conception F. Ang of the Infectious Diseases Section (IDS) of the Microbiology Research Laboratory of the Philippine General Hospital.*

[δ] *NM, no mutation.*

katG

The katG sequences exhibited two single-point and two double-point mutations. Codon 315 changed into two types of amino acids (Table 1) wherein AGC was changed to ACC (Serine to Threonine) and to AAC (Serine to Asparagine). On the other hand, CTC of codon 333 was replaced by CTG (Leucine to Leucine) in 3.33%. The two double-point mutations were observed in 3.33% of the isolates. Initial grouping of the katG mutations would show that mutations in codons 280 TCG to CCG (Serine to Proline) and 315 AGC to ACC (Serine to Threonine) comprised the first type of katG isolates. The second type includes those with mutations in codons 281 GCC to GTC (Alanine to Valine) and 315 AGC to AGA (Serine to Arginine). Those not exhibiting any mutation comprise 36.67% of the isolates and are considered to fall under the wild type group.

Figure 3 shows the dendogram of the samples according to their mutations in the katG gene. The first cluster consisted of non-mutated strains. This cluster had eleven samples and was represented by sample 1. They were shown to group accordingly with H37Rv. The second and third cluster each has one sample, sample 20 and sample 18, respectively. The fourth cluster had the most number of samples. This group consisted of seventeen samples and was represented by sample 4.

gyrA

Ninety percent (90%) of the isolates did not exhibit any mutation in the gyrA region (Table 1). Each of the three single-point mutations was observed in 3.33% of the isolates. Codon 90 mutated from GCG to GTG (Alanine to Valine). Codon 94 changed into two types of amino acid: GAC (Alanine) was changed into TTC (Phenylalanine) or GGC (Glycine).

Figure 4 shows three clusters according to the mutations in the gyrA gene. The first cluster was composed of twenty-

seven samples. This was represented by sample 1 and was grouped with H37Rv. The second cluster was composed of a single sample, sample 20. On the other hand, two samples belonged to the last cluster and were represented by sample 13.

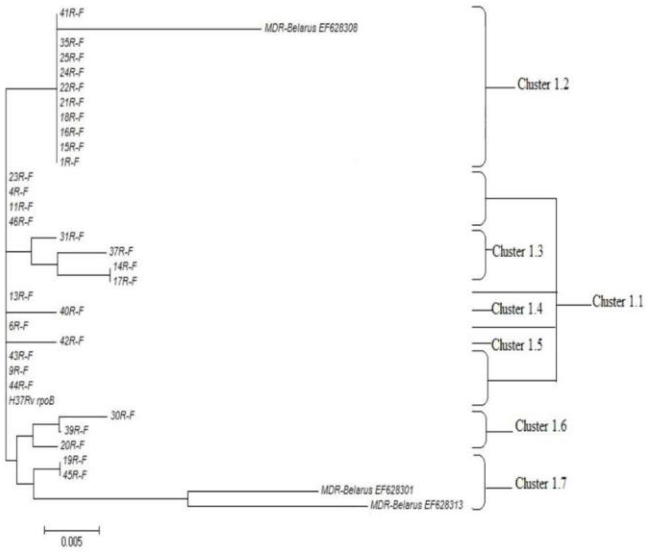

Figure 2. Dendogram of 297-nt partial rpoB gene sequence of thirty local isolates

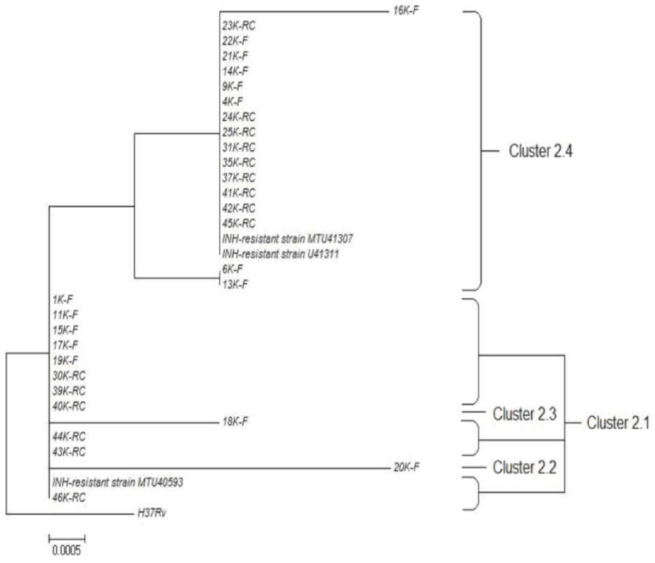

Figure 3. Dendogram of 543-nt partial katG gene sequence of thirty local isolates

Multi-loci Sequence Analysis

The clustering of samples according to their mutations in all three genes was shown in Figure 5. There were eight clusters and the first cluster, represented by sample 11, includes the H37Rv wild type strain. Sample 40 was the lone isolate for the second cluster. Meanwhile, there were two samples for the third cluster as represented by sample 20. Four samples were placed in the fourth cluster, represented by sample 6. The fifth cluster, represented by sample 39, had five samples. Six samples belonged to the sixth cluster with sample 14 as representative. Three samples comprised the seventh cluster and the representative was sample 1. The last cluster, represented by sample 21, had seven samples.

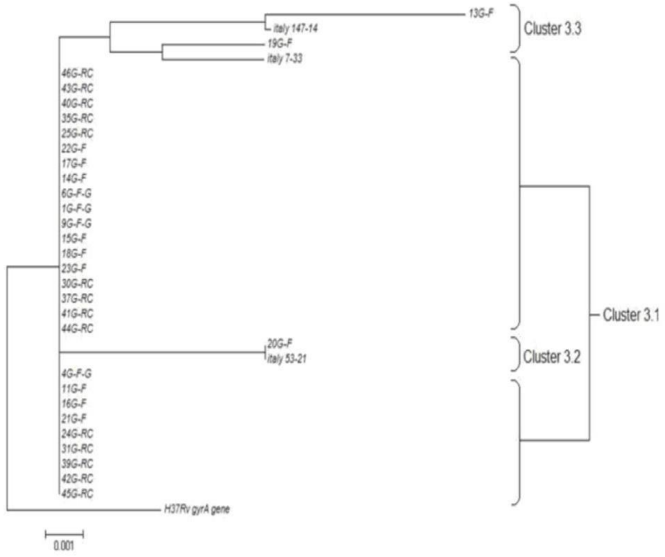

Figure 4. Dendogram of 320-nt partial gyrA gene sequence of thirty local isolates

Discussion

rpoB

The mutated samples were found to have point mutations in codons 513, 516, 522, 526, 531 and 533 in this gene. Interestingly, 33.33% of the isolates showed mutations in codon 531. This implies that codon 531 is the most frequently mutating codon with a high tendency to undergo mutation. More so, two samples showed mutations in codon 516 while three samples had mutations in codon 526. As shown in Table 1, codons 516 ad 526 had more than one type of mutation resulting to a change from Aspartate to Tyrosine, Valine and Phenylalanine, and from Histidine to Tyrosine, Cysteine and Arginine, respectively. Consistent with the findings of other investigators, the codons 516, 526 and 531 were among the most commonly mutating codons.[2,6]

Figure 2 showed seven clusters which included strains resistant to rifampicin as reported in the Genbank. The first cluster is composed of H37Rv rpoB sequence and the nine non-mutated samples. The samples that belonged to the second cluster had mutations at codon 531, one of the most frequent and most commonly mutating codons.[2,6] The samples in the third cluster had mutations at codon 516, which also falls among the commonly mutating codons. It is

important to note that samples 14 and 17 had double mutations involving two codons, 512 and 516 as shown in Table 1. The occurrence of this double mutation in two samples may suggest a mutational mechanism wherein one codon mutation would likely predict a secondary mutation. However, this mechanism merits further investigation. The fourth cluster, which included sample 40, showed a mutation in codon 513; while the fifth cluster, which included sample 42 had a mutation in codon 522. The members of the sixth cluster all showed mutations in codon 526. The last cluster, with only two samples had mutations in codon 533. The Genbank reported RIF-resistant strains, the multidrug-resistant strains common in other countries, belong to clusters second and seven. Apparently, the TB samples from the IDS-PGH showed more varied mutations than those reported in the Genbank. Whether this reflects on the nature of the local TB therapeutic management and should necessitate a call for alarm, remains to be further verified.

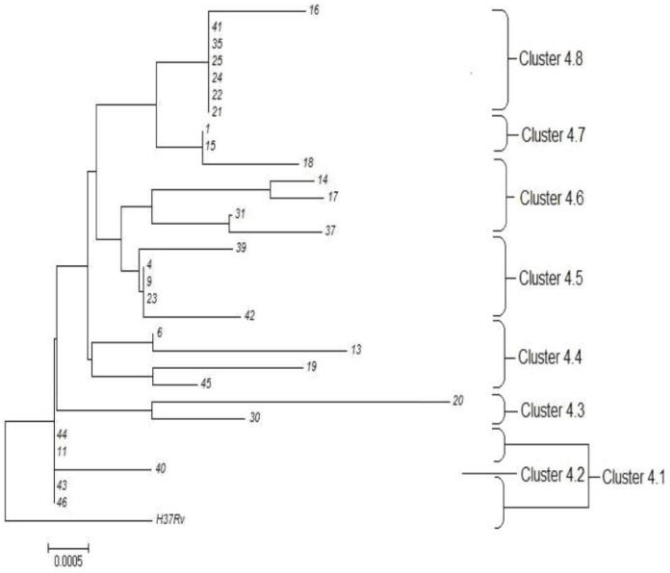

Figure 5. Dendogram of concatenated rpoB, katG, and gyrA sequences from thirty local isolates

katG

More than half of the samples had mutations in codons 280, 281, 315, and 333 in the katG region and almost 50% of the mutated samples were found to have single mutation in codon 315. Likewise, this codon showed a myriad of nucleotide substitutions. The type of mutation in codon 315 where AGC was changed to ACC (Serine to Threonine), was the most frequently observed, occurring in 46.67% of the samples.[1,2] The same mutation was found most prevalent elsewhere.[3] This high percentage shows its importance in development of resistance to isoniazid.[2] This observation finds support in that residues encoded by those including codon 315 participate in binding the heme group of katG enzyme. Thus, mutations in this region have been proposed to cause the loss of the enzymatic function of katG enzyme.[3]

In Figure 3, the first cluster comprised the non-mutated strains together with H37Rv. Interestingly, one of the Genbank reported INH-resistant strains belonged to this cluster. This implies that not all non-mutated samples for katG were of the wild type strains. Some may show no mutation within the target region but may be resistant to isoniazid.[2] Possibly, mutations outside the target region such as in genes such as inhA, kasA and ahpC, may have contributed to the resistance to isoniazid.[6] Sample 20 of the second cluster mutated in two codons, codon 281 and codon 315. In codon 281, Alanine was changed to Valine while in codon 315, Arginine was originally Serine. Sample 18 of the third cluster had a mutation in codon 333 CTC to CTG (Leucine to Leucine), which may be considered a silent mutation since there was no amino acid change. All the members of the fourth cluster had mutations in codon 315. Threonine replaced Serine in majority of the samples. Sample 16 was located at some distance from the other samples since aside from having mutated codon 315, it also had mutation in codon 280. The two INH-resistant strains (i.e. with accession number), which are Genbank reported M. tuberculosis multidrug-resistant strains, belong to the last cluster.

gyrA

Among the thirty isolates, only about 10% was found to have point mutations in codons 90 and 94 (Table 1). Mutations in codon 94 were among the common mutations occurring in the gyrA gene. Among the isolates, mutation in codon 95 had not been observed.[3] Mutation in codon 95 is considered the most common mutation in isolates resistant to ciprofloxacin. However, it was reported that it is not associated in the development of resistance to fluoroquinolones such as ciprofloxacin since it also occurs in susceptible isolates.[3,6]

The first cluster includes the H37Rv strain and the samples wherein no mutation was identified. The second cluster consisted of samples that had mutations in codon 90. The two samples that belonged to the last cluster both had mutations in codon 94. Aspartate was changed to Glycine and to Phenylalanine in sample 19 and sample 13, respectively.

Multi-locus Sequence Analysis

The Multi-loci sequence alignment (MLSA) analysis showed the occurrence of eight clusters as presented in the results. The samples in the first cluster had no mutation and were grouped with H37Rv. Sample 40, the lone sample for the second cluster, had a mutation in its rpoB region. Meanwhile, the two samples for the third cluster had mutations in the same codon of the rpoB region although of

different resulting amino acid. Sample 20 was placed farther from the point of the wild type strains since it also had mutations in katG and gyrA.

Samples from the fourth cluster had similar mutation in one gene while different in another. Samples 6 and 13 had the same mutation in the katG region. The latter had an additional mutation in the gyrA region placing it at some distance from the former. On the other hand, samples 19 and 45 had the same mutation in the rpoB region. Sample 45 had an additional mutation in the katG region in the same codon with that in sample 6 and 13. Likewise, sample 19 also had a mutation in gyrA in the same codon with that in sample 13.

The fifth cluster was the group with odd composition. Sample 39 should have been placed with sample 30 since they both had mutation in the same codon in rpoB. The remaining samples, however, had the same mutations in the katG region. Sample 42 had another mutation in the rpoB region making it different from the other samples in the cluster.

The samples in cluster six had similar mutations for one codon of the rpoB region. Samples 14 and 17 were separated from the other two since both have double-point mutations for the rpoB region. Samples 31 and 37 had different nucleotide substitution in the same codon for the rpoB region but had the same mutation in the katG region. The samples in cluster seven all had the same mutations in the rpoB region. Only, sample 18 was placed farther since it had an additional mutation in the katG gene. The last cluster was composed of samples with the same mutations in both rpoB and katG regions. They were all reported in the Genbank to be multidrug-resistant tuberculosis. Sample 16, however, had an additional mutation in the katG region.

The MLSA clustering showed all samples to be consistent in their clustering except for sample 39. Based on previous dendograms (Figures 2-4), rpoB showed seven clusters. The katG gene showed four clusters while gyrA showed three clusters. The eight clusters presented by MLSA suggested a possibility to genotype the isolates based on the three genes. The first cluster would comprise the wild type strains due to the absence of mutations while the third cluster would comprise the extensively drug resistant (XDR) strains due to the presence of mutations in all three genes (as exemplified by sample 20). The remaining six clusters would comprise the MDR strains. This suggests the existence of six genotypes for the MDR isolates and one genotype for XDR. The use of multiple loci (rpoB, katG and gyrA) provided a better genotyping option and identification of MDR-TB rather than the independent genes alone as criteria. However, the initially identified six clusters as putative MDR-TB genotypes need further validation in terms of agreement of their IS6110, spoligotype or MIRU-VNTR profiles (i.e., gold standards for TB genotyping).The putative genotypes also should be taken into account when developing an appropriate detection kit.

Acknowledgment

The authors would like to thank the Endowment Fund of the University of the Philippines -Manila for funding the study.

References

1. Marin M, Garcia de Viedma D, Ruiz-Serrano MJ, Bouza E. Rapid direct detection of multiple rifampin and isoniazid resistance mutations in Mycobacterium tuberculosis in respiratory samples by real-time PCR. Antimicrob Agents Chemother. 2004; 48(11):4293-300.
2. Tracevska T, Jansone I, Broka L, Marga O, Baumanis V. Mutations in the rpoB and katG genes leading to drug resistance in Mycobacterium tuberculosis in Latvia. J Clin Microbiol. 2002; 40(10):3789-92.
3. Sekiguchi J, Miyoshi-Akiyama T, Augustynowicz-Kopec E, et al. Detection of multidrug resistance in Mycobacterium tuberculosis. J Clin Microbiol. 2007; 45(1):179- 92.
4. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform. 2004; 5(2):150-63.
5. Heralde F, et al. Development of Nucleic Acid-Based Diagnostics for Multidrug-Resistant Tuberculosis Project, Project Proposal under the Biotechnology Alliance for Health and Infection (BAHaI) Research Program. National Institutes of Health. 2008.
6. Siddiqi N, Shamin M, Hussain S, et al. Molecular characterization of multidrug-resistant isolates of Mycobacterium tuberculosis from patients in North India. Antimicrob Agents Chemother. 2002; 46(2):443–50.