Jundishapur J Microbiol. 2016 April; 9(4):e30763. doi: 10.5812/jjm.30763.

Published online 2016 April 16. **Research Article**

Phylogenetic Analysis of Prevalent Tuberculosis and Non-*Tuberculosis Mycobacteria* in Isfahan, Iran, Based on a 360 bp Sequence of the *rpoB* Gene

Bahram Nasr Esfahani,[1] Sharareh Moghim,[1] Hajieh Ghasemian Safaei,[1] Mohsen Moghoofei,[2] Mansour Sedighi,[1] and Shima Hadifar[1,*]

[1]Department of Microbiology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, IR Iran
[2]Department of Virology, Faculty of Medicine, Iran University of Medical Sciences, Tehran, IR Iran

[*]*Corresponding author*: Shima Hadifar, Department of Microbiology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, IR Iran. Tel: +98-3137922493, Fax: +98-3136688597, E-mail: aseman.shima@yahoo.com

Received 2015 July 05; **Revised** 2016 January 18; **Accepted** 2016 February 01.

Abstract

Background: Taxonomic and phylogenetic studies of *Mycobacterium* species have been based around the *16sRNA* gene for many years. However, due to the high strain similarity between species in the *Mycobacterium* genus (94.3% - 100%), defining a valid phylogenetic tree is difficult; consequently, its use in estimating the boundaries between species is limited. The sequence of the *rpoB* gene makes it an appropriate gene for phylogenetic analysis, especially in bacteria with limited variation.
Objectives: In the present study, a 360bp sequence of *rpoB* was used for precise classification of *Mycobacterium* strains isolated in Isfahan, Iran.
Materials and Methods: From February to October 2013, 57 clinical and environmental isolates were collected, subcultured, and identified by phenotypic methods. After DNA extraction, a 360bp fragment was PCR-amplified and sequenced. The phylogenetic tree was constructed based on consensus sequence data, using MEGA5 software.
Results: Slow and fast-growing groups of the *Mycobacterium* strains were clearly differentiated based on the constructed tree of 56 common *Mycobacterium* isolates. Each species with a unique title in the tree was identified; in total, 13 nods with a bootstrap value of over 50% were supported. Among the slow-growing group was *Mycobacterium kansasii*, with *M. tuberculosis* in a cluster with a bootstrap value of 98% and *M. gordonae* in another cluster with a bootstrap value of 90%. In the fast-growing group, one cluster with a bootstrap value of 89% was defined, including all fast-growing members present in this study.
Conclusions: The results suggest that only the application of the *rpoB* gene sequence is sufficient for taxonomic categorization and definition of a new *Mycobacterium* species, due to its high resolution power and proper variation in its sequence (85% - 100%); the resulting tree has high validity.

Keywords: Phylogeny, *Mycobacterium*, *rpoB* Gene

1. Background

More than 160 species of the *Mycobacterium* genus have been identified, with a third of them having the potential to cause a wide variety of human (1). These medically important species are responsible for considerable human morbidity and mortality, and mostly belong to the *Mycobacterium tuberculosis* complex (MTBC) (such as *M. tuberculosis*, *M. africanum*, *M. canettii*, and *M. bovis*) (2, 3) and the *M. avium* intracellulare complex (MAC), which include atypical *mycobacteria*. Various phylogenetic studies have been performed to gain an understanding of mycobacterial taxonomic identity, as regards the nature of pathogenic *Mycobacterium* species, the wide range of diseases they cause, and their high genetic similarity (4, 5). Phylogenetic studies have been primarily based on phenotypic characteristics and methods; however, significant progress in the genetic sciences and the development of statistical models for evolutionary analyses of nucleotide sequences have resulted in an improvement of such studies (6, 7).

For many years, phylogenetic and taxonomic studies of *Mycobacterium* species have been based on *16sRNA* analysis (8). The critical role of *16sRNA* in protein synthesis and its high information content, conserved nature, and presence in all organisms have made this gene an appropriate candidate for phylogenetic analysis, as well as for identification (6, 9). The criterion for classification based on the *16sRNA* gene is that bacterial strains differing in less than 10 - 15bp belong to the same species. However, interspecies genetic similarity in the members of the *Mycobacterium* genus is as high as 94.3% - 100% and, in some species such as the fast-growing species *Mycobacterium kansasii*, the similarity

is extremely high (> 99%) (10). The presence of more than one copy of *16sRNA* in this genus leads to complexity in interpretation of the sequence information and makes it difficult to draw a valid phylogenetic tree, thereby limiting the estimation boundary between species in phylogenetic analysis (10-12).

Accordingly, in recent decades, other genes such as *recA* (13), *hsp65* (14, 15), and *rpoB* have been considered as candidates for phylogenetic studies and diagnosis (16, 17). From these, *rpoB* encoding β subunit of RNA polymerase enzyme is thought to be the most appropriate gene for such purposes, especially in bacteria with close relativity (18-20). The first attempt to apply this gene for the differentiation of *Mycobacterial* species was carried out by Kim et al. (16, 21). Consequently, different parts of this gene were used to sequence analysis (6, 10, 14), and further studies have acknowledged the gene's adequacy for interference in phylogenetic relationships in bacterial groups with close links, such as *Mycobacteria* (22).

2. Objectives

More than 50 human pathogen species of the *Mycobacterium* genus have been identified, and interspecies similarity in the genus is high. It is important to determine the relationships between Mycobacterial species. In the present study, a 360 bp fragment of the *rpoB* gene, from *Mycobacterium* strains isolated from patients in Isfahan providence, was used to obtain a precise taxonomic classification of these isolates.

3. Materials and Methods

3.1. Mycobacteria Strains and Molecular Methods

A total of 57 *Mycobacterium* isolates were obtained between February and October 2013 at the *Mycobacterium* collection of the department of microbiology, Isfahan University of Medical Sciences and Tuberculosis center of Isfahan. Of these 57 specimens, 41 were respiratory types, including sputum, bronchoalveolar lavage, and bronchial wash specimens, 8 were non-respiratory, and 8 were environmental samples such as water which was needed to break up the mucin; each respiratory specimen was treated with an equal volume of 3.5% NaOH in a 50 mL centrifuge tube, before being vortexed for 30 seconds. Specimens that were decontaminated after being in the tubes were incubated at room temperature for 15 minutes. A sterile phosphate buffer was added to stop the digestion decontamination process; the tubes were then mixed by inversion and centrifuged at 3,000 $\times$ g for 15 minutes. The supernatant was discarded and the resultant pellet re-suspended in 3.0 mL of sterile phosphate-buffered saline (Sigma, USA). Non-respiratory specimens, after the homogenies, were processed in the same way, although sampling of water was done using the grab sampling method. Sodium thiosulfate, as an antichlor, and 0.04% cetylpyridinium chloride (Merck, Germany), as an antimicrobial agent, were added to 2l sterile Erlenmeyer flasks. Samples amounting to 500 mL were passed from 0.45 μm filters. The filters were transferred directly onto 7H10 middle brook solid media (Merck, Germany), including 15% oleic acid, albumin, dextrose, and catalase.

3.2. Primary Identification of Isolates by Conventional Methods

The *Mycobacterium* species were subcultured on LJ media and middle brook 7H9 (Merck, Germany). The isolates were identified by primary conventional methods including acid-fast staining, colony characteristics, growth at 25°C, 37°C, and 42°C, pigment production, semi-quantitative catalase test, Tween 80 hydrolysis, arylsulfatase test (3 and 14 days), heat-stable catalase (pH 7, 68°C), pyrazinamidase (4 and 7 days), urease, nitrate reduction test, and colony morphology. The reference strains used in this study were *Mycobacterium tuberculosis* H37Rv and *M. fortuitum* (ATCC 49403) (23) (Table 1).

Table 1. List of *Mycobacterium* Strains and Their Sources, as Identified by Phenotypic Tests

Mycobacterium Species	No. and Sources
***Mycobacterium tuberculosis* complex**	18 clinical isolates, H37Rv
Mycobacterium fortuitum	ATCC 49403, 28 clinical and environmental isolates
Mycobacterium avium	1 clinical isolate
Mycobacterium kansasii	1 clinical isolate
Mycobacterium smegmatis	1 environmental isolate
Mycobacterium conceptionense	1 environmental isolate
Mycobacterium gordonae	5 clinical and environmental isolates

3.3. Preparation of Genomic DNA

Genomic DNA was extracted using the cetyltrimethyl ammonium bromide (CTAB) method (24) and subjected to a high pure PCR template preparation kit (Roche Applied Science, Germany). Purified DNA was stored at -70°C for subsequent experiments.

3.4. PCR Amplification of rpoB

The reaction was performed in a thermal cycler from PCR Hybid (Omnigene). The partial

rpoB gene (360bp) was amplified using primers (RPO5') 5'-TCAAGGAGAAGCGCTACGA-3' and (RPO3') 5'-GGATGTTGATCAGGGTCTGC-3' (25). Each 25 μL PCR mixture contained 2μL of DNA supernatant (5ng genomic DNA), 1μL of each primer (10 pmoL/mL) (eurofins MWG/Operon, Ebersberg, Germany), 1.25μL of $MgCl_2$ (1.5 mM) (Fermentas-Canada), 0.5 μL of dNTP (200 mM) (Fermentas Canada), 0.25 μL of Taq polymerase (500 U) (Cinnagen, Iran), and 2.5 μL of 10x buffer. The PCR program was performed in a thermo cycler (Eppendorf) including at 94°C for 5 minutes, 35 cycles of 94°C for 1 minute each, 60°C for 1 minute, 72°C for 1 minute, and a final step at 72°C for 7 minutes. The PCR products were run on 1.5% agarose gel, which was visualized by green viewer staining (Pars Tous, Iran) and gel documentation.

3.5. Sequencing and Analysis

PCR products from amplification of the *rpoB* gene were sent to the Pishgam Company (Iran) for purification and sequencing. Chromatogram sequences, using the Bio Edite software, were applied to investigate the outset and termination of gene sequences that were improperly read; they were removed, and a blast of each sequence, using NCBI blast, was performed. Mega software (version 5.2) (26) was used to construct the phylogenetic tree, and the alignment was performed using Clustal W format. The methodology used was maximum likelihood, and the phylogram stability was evaluated by parsimony bootstrapping with 500 simulations. A *Nocardia asteroids* sequence (accession number AB219431.1) was used as the out group.

4. Results

Of 57 isolates, 19 (33.33%) were identified as *M. tuberculosis* and 38 (66.66%) were identified as NTM, using the phenotypic method. Among the 38 NTM isolates, 30 (78.94%) were clinical and 8 (21.05%) were environmental. The clinical isolates included *M. tuberculosis* complex (MTBC), *M. avium*, *M. fortuitum, M. gordonae*, and *M. kansasii*, while the environmental isolates consisted of *M. fortuitum*, *M. smegmatis*, *M. conceptionense*, and *M. gordonae*.

The constructed phylogenetic tree of the 360 bp fragment of the *rpoB* gene sequence was more or less similar to the results obtained from other studies (10, 12, 22, 27). Based on the tree for 56 *Mycobacterium* isolates that are prevalent in Isfahan, fast and slow-growing groups could be completely separated so that each species was identified in the tree by a unique title. There were 13 nodes supported that had a bootstrap value of over 50%.

The slow-growing group included *M. avium*, *M. kansasii*, *M. gordonae*, and *M. tuberculosis* complex. *Mycobacterium kansasii* and *M. tuberculosis* were in a cluster with a bootstrap value of 98%, while *M. gordonae* was in another cluster with a bootstrap value of 90%. In addition, in the fast-growing group, one cluster with a bootstrap value of 89% was identified, which included all members of the fast-growing group that were studied in the current paper (*M. fortuitum*, *M. smegmatis*, and *M. conceptionense*). In this cluster, *M. smegmatis* was in one sub-cluster while the rest of the members were in another sub-cluster with a bootstrap value of 90%.

All of these results indicate the validity of drawing a tree based on the sequence of this gene (Figure 1). Evolutionary interval matrix results show that interspecies similarity in *M. tuberculosis* was 98% - 100%, while it was 99% - 100% in *M. fortuitum* species (in most of the *M. fortuitum* isolates, intraspecies similarity was 100%, and in a few of them it was 99%) and 98% - 100% in *M. gordonae*; these results illustrate high conservation in the *rpoB* gene.

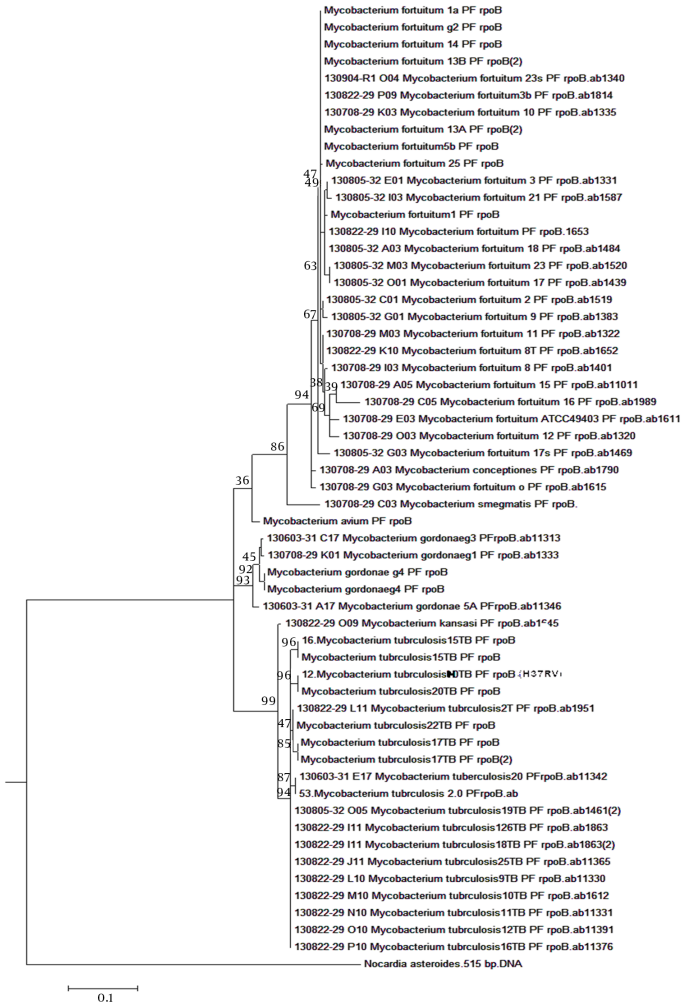
Among the 57 investigated isolates, interspecies similarity was 83/73% between *M. conceptionense* and *M. tuberculosis*, 86/25% between *Mycobacterium gordonae* and *M. fortuitum*, and 90/80% between *M. smegmatis* and *M. conceptionense* (Figure 2).

5. Discussion

The *16SrRNA* gene sequence has been used as a reference method for the detection and characterization of *Mycobacteria* and has helped to define over 45 new Mycobacterial species (19). The criteria used for the classification of *Mycobacteria*, based on the *16SrRNA* gene, is that the bacterial strains would belong to the same species if they had only 10 - 15 base pair differences with other species. The high similarity between *Mycobacterium* species (about 94.3% - 100%) (» 99%) (10), the existence of 2 copies of *16SrRNA*, and the complexity of interpretation of the resulting information made it difficult to define a valid phylogenetic tree and limited the use of it for estimating the boundaries between species in the phylogenetic analysis (10, 12, 22). Accordingly, in recent decades, other genes have been considered as candidates for phylogenetic studies, including *hsp65* (14), *recA* (13), and *rpoB* (16). Among these, a single-copy *rpoB* gene encoding a β subunit of RNA polymerase enzyme has been identified as the most appropriate gene for phylogenetic analysis (18, 19, 22, 28). This gene has variable and conserved regions and different fragments of it are used for bacterial analysis (29, 30).

Many studies have used this gene for *Mycobacterium* genus analysis (10, 16, 18, 27). For the development and completion data resulting from *16SrRNA* gene sequences to distinguish bacterial groups with close relationships, several studies have used the sequence of some housekeeping

Figure 1. A Phylogenetic Tree Based on *rpoB* Gene Sequences Shows the Relationship of the 57 *Mycobacterium* Isolates and 1 out-Group Strain.



This tree was constructed using the maximum likelihood method and the evolutionary history was inferred using the same method, based on the Jukes-Cantor model. Evolutionary analyses were conducted in MEGA5.

genes; the *rpoB* gene is one of those that is applied using the MLST method. This is supported by the resulting information of the *rpoB* gene sequence. Considering the results of tree drawing using other genes with a high bootstrap value, this gene has been deemed appropriate to infer the phylogenetic relationships of bacterial groups with close links, such as *Mycobacteria* (22, 29, 31). Accordingly, notice the frequency of *Mycobacterium* species, especially atypical *Mycobacteria*, in this geographic region, and that these bacteria are widely isolated from environmental, animal, and human resources.

Based on the above, it is necessary to determine interspecies diversity and conduct precise taxonomic classification in Isfahan. To this end, this study utilizes a portion of the gene sequence for classification of prevalent *Mycobacterium* isolates. According to the results of the phylogenetic tree in the present study, all the tested species were completely separated, so that slow and fast-growing atypical *Mycobacteria* and *M. tuberculosis* complex included separate clusters. These results are consistent with those of other studies; for example, in 2005, Devulder et al. (10) illustrated that the resulting phylogenetic tree of the *rpoB* gene can properly separate *Mycobacterium* genus members so that, based on a 396 bp sequence of this gene, fast-growing (such as *M. smegmatis*) and slow-growing (such as *Mycobacterium tuberculosis*) groups are also totally separated. Based on the results represented in this study, the bootstrap values of the phylogenetic tree for

Figure 2. Estimates of Evolutionary Divergence Between Sequences

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
4	130603-31_E17_Mycobacterium_tuberculosis20_PFrpoB.ab11342	0.08	0.18	0.15		0.04	0.04	0.04	0.05	0.05	0.04	0.04	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.02	0.04	0.02	0.02	0.01	0.04
5	130708-29_A03_Mycobacterium_conceptiones_PF_rpoB.ab1790	0.59	0.18	0.18	0.24		0.11	0.04	0.11	0.08	0.08	0.03	0.03	0.01	0.01	0.01	0.06	0.06	0.02	0.03	0.07	0.07	0.06	0.04	0.07	0.09	0.04	0.04	0.09	0.08
6	130708-29_A05_Mycobacterium_fortuitum_15_PF_rpoB.ab11011	0.70	0.21	0.21	0.29	1.06		0.10	0.04	0.08	0.08	0.03	0.03	0.01	0.01	0.01	0.06	0.06	0.02	0.03	0.07	0.09	0.06	0.04	0.08	0.10	0.04	0.04	0.12	0.09
7	130708-29_C03_Mycobacterium_smegmatis_PF_rpoB.	0.71	0.17	0.18	0.28	0.38	0.89		0.11	0.11	0.11	0.04	0.03	0.02	0.02	0.03	0.07	0.08	0.03	0.04	0.09	0.09	0.08	0.05	0.09	0.11	0.05	0.05	0.11	0.11
8	130708-29_C05_Mycobacterium_fortuitum_16_PF_rpoB.ab1989	0.76	0.24	0.23	0.33	1.10	0.48	1.00		0.08	0.08	0.04	0.03	0.01	0.01	0.02	0.06	0.07	0.02	0.03	0.08	0.09	0.07	0.05	0.09	0.10	0.05	0.05	0.12	0.10
9	130708-29_E03_Mycobacterium_fortuitum_ATCC49403_PF_rpoB.ab1611	0.65	0.20	0.20	0.29	0.76	0.72	0.95	0.77		0.06	0.03	0.03	0.01	0.01	0.01	0.05	0.06	0.01	0.03	0.08	0.07	0.06	0.04	0.08	0.08	0.04	0.04	0.10	0.08
10	130708-29_G03_Mycobacterium_fortuitum_o_PF_rpoB.ab1615	0.58	0.18	0.18	0.26	0.77	0.74	0.99	0.79	0.61		0.03	0.03	0.01	0.01	0.01	0.06	0.05	0.01	0.03	0.07	0.08	0.06	0.04	0.06	0.08	0.03	0.03	0.08	0.09
11	130708-29_I03_Mycobacterium_fortuitum_8_PF_rpoB.ab1401	0.44	0.18	0.18	0.27	0.26	0.22	0.34	0.26	0.23	0.25		0.03	0.01	0.01	0.01	0.03	0.03	0.01	0.03	0.04	0.03	0.03	0.04	0.05	0.03	0.04	0.04	0.05	0.03
12	130708-29_K01_Mycobacterium_gordonaeg1_PF_rpoB.ab1333	0.14	0.02	0.01	0.15	0.15	0.16	0.15	0.16	0.16	0.15	0.16		0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
13	130708-29_K03_Mycobacterium_fortuitum_10_PF_rpoB.ab1335	0.20	0.18	0.18	0.22	0.04	0.03	0.12	0.05	0.03	0.03	0.03	0.16		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.04	0.04	0.03	0.01
14	130708-29_M03_Mycobacterium_fortuitum_11_PF_rpoB.ab1322	0.21	0.19	0.19	0.22	0.04	0.04	0.14	0.05	0.04	0.04	0.04	0.16	0.03		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.03	0.03	0.03	0.01
15	130708-29_O03_Mycobacterium_fortuitum_12_PF_rpoB.ab1320	0.23	0.22	0.22	0.25	0.06	0.04	0.15	0.06	0.04	0.06	0.05	0.16	0.04	0.04		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.04	0.01	0.04	0.04	0.04	0.01
16	130805-32_A03_Mycobacterium_fortuitum_18_PF_rpoB.ab1484	0.53	0.18	0.18	0.27	0.53	0.48	0.66	0.52	0.48	0.47	0.23	0.16	0.01	0.04	0.06		0.03	0.01	0.02	0.03	0.03	0.04	0.03	0.05	0.04	0.04	0.04	0.05	0.04
17	130805-32_C01_Mycobacterium_fortuitum_2_PF_rpoB.ab1519	0.52	0.19	0.19	0.26	0.55	0.54	0.73	0.62	0.52	0.47	0.24	0.17	0.02	0.03	0.05	0.19		0.01	0.02	0.04	0.03	0.03	0.03	0.05	0.04	0.03	0.03	0.05	0.04
18	130805-32_E01_Mycobacterium_fortuitum_3_PF_rpoB.ab1331	0.23	0.20	0.19	0.27	0.07	0.07	0.15	0.07	0.06	0.05	0.05	0.16	0.04	0.04	0.06	0.03	0.05		0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.04	0.04	0.03	0.01
19	130805-32_G01_Mycobacterium_fortuitum_9_PF_rpoB.ab1383	0.35	0.20	0.20	0.28	0.18	0.21	0.30	0.25	0.18	0.19	0.19	0.16	0.03	0.04	0.04	0.08	0.09	0.05		0.02	0.02	0.02	0.02	0.04	0.02	0.04	0.04	0.04	0.02
20	130805-32_G03_Mycobacterium_fortuitum_17s_PF_rpoB.ab1469	0.56	0.18	0.18	0.24	0.61	0.60	0.81	0.66	0.64	0.57	0.25	0.17	0.03	0.05	0.06	0.26	0.32	0.05	0.10		0.05	0.04	0.03	0.05	0.05	0.04	0.04	0.06	0.05
21	130805-32_I03_Mycobacterium_fortuitum_21_PF_rpoB.ab1587	0.60	0.18	0.18	0.28	0.64	0.80	0.86	0.82	0.67	0.74	0.23	0.17	0.02	0.05	0.06	0.22	0.25	0.03	0.11	0.40		0.04	0.03	0.05	0.05	0.04	0.04	0.05	0.05
22	130805-32_M03_Mycobacterium_fortuitum_23_PF_rpoB.ab1520	0.57	0.16	0.17	0.27	0.54	0.51	0.66	0.60	0.48	0.52	0.23	0.16	0.03	0.05	0.06	0.30	0.29	0.04	0.11	0.37	0.29		0.02	0.05	0.05	0.04	0.04	0.05	0.05
23	130805-32_O01_Mycobacterium_fortuitum_17_PF_rpoB.ab1439	0.39	0.20	0.20	0.28	0.32	0.30	0.40	0.37	0.27	0.26	0.25	0.17	0.05	0.05	0.06	0.20	0.17	0.06	0.10	0.20	0.18	0.11		0.05	0.03	0.04	0.04	0.04	0.03
24	130805-32_O05_Mycobacterium_tubrculosis19TB_PF_rpoB.ab1461(2)	0.29	0.17	0.15	0.07	0.62	0.66	0.71	0.76	0.65	0.56	0.38	0.14	0.18	0.19	0.23	0.43	0.43	0.22	0.26	0.44	0.47	0.45	0.39		0.06	0.01	0.01	0.02	0.06
25	130822-29_I10_Mycobacterium_fortuitum_PF_rpoB.1653	0.58	0.19	0.18	0.29	0.87	0.89	1.00	0.92	0.76	0.77	0.21	0.15	0.04	0.04	0.05	0.31	0.36	0.05	0.13	0.46	0.51	0.43	0.21	0.51		0.04	0.04	0.06	0.07
26	130822-29_I11_Mycobacterium_tubrculosis126TB_PF_rpoB.ab1863	0.06	0.18	0.17	0.10	0.28	0.28	0.32	0.32	0.30	0.23	0.28	0.14	0.22	0.21	0.23	0.24	0.21	0.25	0.24	0.22	0.23	0.22	0.25	0.06	0.22		0.00	0.01	0.04
27	130822-29_I11_Mycobacterium_tubrculosis18TB_PF_rpoB.ab1863(2)	0.06	0.18	0.17	0.10	0.28	0.28	0.32	0.32	0.30	0.23	0.28	0.14	0.22	0.21	0.23	0.24	0.21	0.25	0.24	0.22	0.23	0.22	0.25	0.06	0.22	0.00		0.01	0.04
28	130822-29_J11_Mycobacterium_tubrculosis25TB_PF_rpoB.ab11365	0.33	0.14	0.14	0.07	0.83	0.92	0.91	0.92	0.85	0.71	0.41	0.14	0.18	0.20	0.24	0.41	0.45	0.22	0.27	0.49	0.47	0.44	0.34	0.19	0.56	0.02	0.02		0.07
29	130822-29_K10_Mycobacterium_fortuitum_8T_PF_rpoB.ab1652	0.61	0.22	0.22	0.29	0.77	0.81	1.00	0.90	0.82	0.82	0.23	0.16	0.02	0.02	0.04	0.34	0.40	0.06	0.13	0.47	0.54	0.43	0.26	0.52	0.71	0.27	0.27	0.60	
30	130822-29_L10_Mycobacterium_tubrculosis9TB_PF_rpoB.ab11330	0.34	0.14	0.13	0.07	1.44	1.68	1.23	1.96	1.21	1.01	0.39	0.14	0.18	0.20	0.24	0.49	0.52	0.22	0.31	0.60	0.70	0.48	0.33	0.24	0.94	0.04	0.04	0.18	0.92
31	130822-29_L11_Mycobacterium_tubrculosis2T_PF_rpoB.ab1951	0.38	0.20	0.17	0.09	1.81	2.22	1.54	2.39	1.35	1.13	0.42	0.14	0.22	0.19	0.23	0.55	0.61	0.22	0.35	0.71	0.84	0.80	0.52	0.29	0.99	0.10	0.10	0.32	1.05
32	130822-29_M10_Mycobacterium_tubrculosis10TB_PF_rpoB.ab1612	0.35	0.15	0.14	0.08	1.15	1.21	1.12	1.35	1.30	1.10	0.43	0.14	0.17	0.20	0.24	0.46	0.50	0.22	0.27	0.49	0.83	0.53	0.36	0.19	0.85	0.04	0.04	0.24	0.68
33	130822-29_N10_Mycobacterium_tubrculosis11TB_PF_rpoB.ab11331	0.05	0.17	0.15	0.06	0.24	0.24	0.28	0.28	0.24	0.21	0.24	0.14	0.21	0.20	0.23	0.20	0.18	0.21	0.21	0.19	0.20	0.20	0.24	0.05	0.20	0.05	0.05	0.03	0.24

The analysis involved 57 nucleotide sequences; all ambiguous positions were removed for each sequence pair. There were a total of 1,103 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

Mycobacterium tuberculosis complex, *M. kansasii-gastri*, and *M. smegmatis-goodii* were 100%, 84%, and over 50%, respectively. In the present study, the same results were obtained, with *M. kansasii* and *M. tuberculosis* included in the cluster with a bootstrap value of 98% and *M. fortuitum* in the cluster with a bootstrap value of 89%. In another survey from 2004, Adekambi et al. (22) demonstrated that the bootstrap value of the phylogenetic tree derived from the *rpoB* gene sequence for each cluster was more than 90%, and all tested species, according to this sequence, were completely separated. In addition, the tree resulting from this gene is more valid than those of the *16sRNA* and *recA* genes.

In the present study, the slow-growing group included *M. kansasii*, *M. tuberculosis* complex, *M. avium*, and *M. gordonae*. *M. kansasii* and *M. tuberculosis* were in a cluster with a bootstrap value of 98% and *M. gordonae* was in another cluster with a bootstrap value of 90%. Fast-growing group members in one cluster with a bootstrap value of 98% included all members of the fast-growing group in this study (*M. conceptionense* and *M. smegmatis*). *Mycobacterium smegmatis* was in one sub-cluster, while the other members were in another sub-cluster with a bootstrap value of 90%. In 2006, Adekambi et al. (12) reported that the bootstrap value of the *rpoB* gene sequence for the fast-growing *Mycobacterium* group, in both the *M. smegmatis-goodii* cluster and the *M. fortuitum-houstonense* cluster, was 100%. Lee et al. (32), in 2014, identified that NTM performed *rpoB*, *16srRNA*, and *hsp65* gene sequencing. The results showed that, based on the sequence of the *16SrRNA* gene, isolates were indistinguishable and unidentified. In this study, the slow and fast-growing groups in the phylogenetic tree were drawn based on a sequence of the *rpoB* gene that was sufficiently separated; the bootstrap value for the fast-growing *Mycobacterium* group was 95%. In this tree, *M. Conceptionense* was in a sub-cluster with a bootstrap value of 77%, while the cluster for *M. fortuitum-farcinogenes-senegalense* was 95%. *Mycobacterium tuberculosis* and *M. avium*, as members of the slow-growing group, were in their own distinct cluster.

Yamada-Noda et al. (33), in 2007, designed a study to investigate 56 *Mycobacterium* species, based on several genes such as *16SrRNA*, *dnaJ*, *hsp65*, and *rpoB*. The results of the phylogenetic tree based on the *rpoB* sequence showed that the bootstrap value for the *M. fortuitum-houstonense* cluster, as a fast-growing *Mycobacterium*, was 100%, and that this value for the *M. tuberculosis* complex, *M. avium*, and *M. kansasii-gastri* clusters, as members of the slow-growing mycobacteria group, was 100%, 100%, and 61%, respectively.

In another study, Kim et al. (16) reported, based on this gene, that the studied fast and slow-growing groups were totally separated and that the bootstrap value of the drawn tree was 100% for the *M. fortuitum* cluster, 62% for the *M. gordonae-intermedium* cluster, and 100% for the *M. avium-paratuberculosis* cluster. This slight difference could be due to variations between geographical regions and the use of different parts of the gene. Like the inter-species and intra-species similarities shown in the present study, Kim et al. intra-species and inter-species similarity was 98% - 100% and 80.05% - 99%, respectively. Adekambi et al. (29), in 2003, reported similar results whereby inter- and intra-species similarity was 83.9% - 97% and 98.3% - 100%, respec-

tively. Based on the *rpoB* gene sequence, in 2007 Simmon et al. (34) reported that inter-species similarity was 99.3% - 100%, and also introduced the *rpoB* gene as a proper goal for study, with much more distinctive power than *sodA* and *hsp65* in the study of *Mycobacteria*.

Our conclusion is that only the application of the *rpoB* gene sequence is sufficient for Mycobacterial phylogenetic study, due to its high resolution power and proper variation in its sequence (85% - 100%) for taxonomic categorization and definition of new *Mycobacterium* species; the resulting tree has high validity.

Acknowledgments

The authors are grateful to Manuchehr Homaee and Maryam Mokhtari and the staff of the Tuberculosis center, Isfahan provincial health office, Isfahan University of Medical Sciences, IR Iran.

Footnotes

Authors' Contribution: All authors contributed to the research equally.

Financial Disclosure: The vice chancellor of research at Isfahan University of Medical Sciences financially supported this study.

Funding/Support: This work was financially supported by Isfahan University of Medical Sciences (grant 191123).

References

1. de Zwaan R, van Ingen J, van Soolingen D. Utility of rpoB gene sequencing for identification of nontuberculous mycobacteria in the Netherlands. *J Clin Microbiol.* 2014;**52**(7):2544–51. doi: 10.1128/JCM.00233-14. [PubMed: 24808238].
2. Supply P, Marceau M, Mangenot S, Roche D, Rouanet C, Khanna V, et al. Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of Mycobacterium tuberculosis. *Nat Genet.* 2013;**45**(2):172–9. doi: 10.1038/ng.2517. [PubMed: 23291586].
3. de Jong BC, Antonio M, Gagneux S. Mycobacterium africanum–review of an important cause of human tuberculosis in West Africa. *PLoS Negl Trop Dis.* 2010;**4**(9):eee744. doi: 10.1371/journal.pntd.0000744. [PubMed: 20927191].
4. Bouakaze C, Keyser C, Gonzalez A, Sougakoff W, Veziris N, Dabernat H, et al. Matrix-assisted laser desorption ionization-time of flight mass spectrometry-based single nucleotide polymorphism genotyping assay using iPLEX gold technology for identification of Mycobacterium tuberculosis complex species and lineages. *J Clin Microbiol.* 2011;**49**(9):3292–9. doi: 10.1128/JCM.00744-11. [PubMed: 21734028].
5. Field SK, Cowie RL. Lung disease due to the more common nontuberculous mycobacteria. *Chest.* 2006;**129**(6):1653–72. doi: 10.1378/chest.129.6.1653. [PubMed: 16778288].
6. Tortoli E. Phylogeny of the genus Mycobacterium: many doubts, few certainties. *Infect Genet Evol.* 2012;**12**(4):827–31. doi: 10.1016/j.meegid.2011.05.025. [PubMed: 21684354].
7. Zakham F, Aouane O, Ussery D, Benjouad A, Ennaji MM. Computational genomics-proteomics and Phylogeny analysis of twenty one mycobacterial genomes (Tuberculosis & non Tuberculosis strains). *Microb Inform Exp.* 2012;**2**(1):7. doi: 10.1186/2042-5783-2-7. [PubMed: 22929624].
8. Mignard S, Flandrois JP. A seven-gene, multilocus, genus-wide approach to the phylogeny of mycobacteria using supertrees. *Int J Syst Evol Microbiol.* 2008;**58**(Pt 6):1432–41. doi: 10.1099/ijs.0.65658-0. [PubMed: 18523191].
9. Kazumi Y, Mitarai S. The evaluation of an identification algorithm for Mycobacterium species using the 16S rRNA coding gene and rpoB. *Int J Mycobacteriol.* 2012;**1**(1):21–8. doi: 10.1016/j.ijmyco.2012.01.004. [PubMed: 26786945].
10. Devulder G, Perouse de Montclos M, Flandrois JP. A multigene approach to phylogenetic analysis using the genus Mycobacterium as a model. *Int J Syst Evol Microbiol.* 2005;**55**(Pt 1):293–302. doi: 10.1099/ijs.0.63222-0. [PubMed: 15653890].
11. Rogall T, Wolters J, Flohr T, Bottger EC. Towards a phylogeny and definition of species at the molecular level within the genus Mycobacterium. *Int J Syst Bacteriol.* 1990;**40**(4):323–30. doi: 10.1099/00207713-40-4-323. [PubMed: 2275850].
12. Adekambi T, Berger P, Raoult D, Drancourt M. rpoB gene sequence-based characterization of emerging non-tuberculous mycobacteria with descriptions of Mycobacterium bolletii sp. nov., Mycobacterium phocaicum sp. nov. and Mycobacterium aubagnense sp. nov. *Int J Syst Evol Microbiol.* 2006;**56**(Pt 1):133–43. doi: 10.1099/ijs.0.63969-0. [PubMed: 16403878].
13. Blackwood KS, He C, Gunton J, Turenne CY, Wolfe J, Kabani AM. Evaluation of recA sequences for identification of Mycobacterium species. *J Clin Microbiol.* 2000;**38**(8):2846–52. [PubMed: 10921937].
14. Bubrick MP, Hitchcock CR. Renal cyst causing afferent loop obstruction and acute pancreatitis. *Am Surg.* 1975;**41**(7):440–3. [PubMed: 1147399].
15. Varma-Basil M, Garima K, Pathak R, Dwivedi SK, Narang A, Bhatnagar A, et al. Development of a novel PCR restriction analysis of the hsp65 gene as a rapid method to screen for the Mycobacterium tuberculosis complex and nontuberculous mycobacteria in high-burden countries. *J Clin Microbiol.* 2013;**51**(4):1165–70. doi: 10.1128/JCM.03067-12. [PubMed: 23363822].
16. Kim BJ, Lee SH, Lyu MA, Kim SJ, Bai GH, Chae GT, et al. Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (rpoB). *J Clin Microbiol.* 1999;**37**(6):1714–20. [PubMed: 10325313].
17. Khosravi AD, Seghatoleslami S. Genotyping and identification of mycobacteria by fingerprinting techniques. *Jundishapur J Microbiol.* 2009;**2009**(3):81–91.
18. Adekambi T, Drancourt M, Raoult D. The rpoB gene as a tool for clinical microbiologists. *Trends Microbiol.* 2009;**17**(1):37–45. doi: 10.1016/j.tim.2008.09.008. [PubMed: 19081723].
19. Pauls RJ, Turenne CY, Wolfe JN, Kabani A. A high proportion of novel mycobacteria species identified by 16S rDNA analysis among slowly growing AccuProbe-negative strains in a clinical setting. *Am J Clin Pathol.* 2003;**120**(4):560–6. doi: 10.1309/VF40-1U7H-7DHE-0FRE. [PubMed: 14560567].
20. Rajendhran J, Gunasekaran P. Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. *Microbiol Res.* 2011;**166**(2):99–110. doi: 10.1016/j.micres.2010.02.003. [PubMed: 20223646].
21. Kim BJ, Lee KH, Park BN, Kim SJ, Bai GH, Kim SJ, et al. Differentiation of mycobacterial species by PCR-restriction analysis of DNA (342 base pairs) of the RNA polymerase gene (rpoB). *J Clin Microbiol.* 2001;**39**(6):2102–9. doi: 10.1128/JCM.39.6.2102-2109.2001. [PubMed: 11376042].
22. Adekambi T, Drancourt M. Dissection of phylogenetic relationships among 19 rapidly growing Mycobacterium species by 16S rRNA, hsp65, sodA, recA and rpoB gene sequencing. *Int J Syst Evol Microbiol.* 2004;**54**(Pt 6):2095–105. doi: 10.1099/ijs.0.63094-0. [PubMed: 15545441].

23. Nasr-Esfahani B, Sarikhani E, Moghim S, et al. Molecular characterization of environmental non-Tuberculous Mycobacteria using PCR-RFLP analysis of 441 Bp heat shock protein 65 fragments. *Iran J of Public Health.* 2012;**41**(4):108–114. [PubMed: 23113172].
24. Clarke JD. Cetyltrimethyl ammonium bromide (CTAB) DNA miniprep for plant DNA isolation. *Cold Spring Harb Protoc.* 2009;**2009**(3):pdb prot5177. doi: 10.1101/pdb.prot5177. [PubMed: 20147112].
25. Lee H, Park HJ, Cho SN, Bai GH, Kim SJ. Species identification of mycobacteria by PCR-restriction fragment length polymorphism of the rpoB gene. *J Clin Microbiol.* 2000;**38**(8):2966–71. [PubMed: 10921960].
26. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 2011;**28**(10):2731–9. doi: 10.1093/molbev/msr121. [PubMed: 21546353].
27. Kim BJ, Kim CJ, Chun J, Koh YH, Lee SH, Hyun JW, et al. Phylogenetic analysis of the genera Streptomyces and Kitasatospora based on partial RNA polymerase beta-subunit gene (rpoB) sequences. *Int J Syst Evol Microbiol.* 2004;**54**(Pt 2):593–8. doi: 10.1099/ijs.0.02941-0. [PubMed: 15023980].
28. Rubin CJ, Thollesson M, Kirsebom LA, Herrmann B. Phylogenetic relationships and species differentiation of 39 Legionella species by sequence determination of the RNase P RNA gene rnpB. *Int J Syst Evol Microbiol.* 2005;**55**(Pt 5):2039–49. doi: 10.1099/ijs.0.63656-0. [PubMed: 16166707].
29. Adekambi T, Colson P, Drancourt M. rpoB-based identification of non-pigmented and late-pigmenting rapidly growing mycobacteria. *J Clin Microbiol.* 2003;**41**(12):5699–708. [PubMed: 14662964].
30. Boor KJ, Duncan ML, Price CW. Genetic and transcriptional organization of the region encoding the beta subunit of Bacillus subtilis RNA polymerase. *J Biol Chem.* 1995;**270**(35):20329–36. [PubMed: 7657605].
31. Rokas A, Williams BL, King N, Carroll SB. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature.* 2003;**425**(6960):798–804. doi: 10.1038/nature02053. [PubMed: 14574403].
32. Lee JY, Kim SH, Shin JH, Lee H-K, Lee YM, Song SA, et al. Postsurgical Wound Infection Caused by Mycobacterium conceptionense Identified by Sequencing of 16S rRNA, hsp65, and rpoB Genes in an Immunocompetent Patient. *Annals of Clinical Microbiology.* 2014;**17**(1):23–7. doi: 10.5145/ACM.2014.17.1.23.
33. Yamada-Noda M, Ohkusu K, Hata H, Shah MM, Nhung PH, Sun XS, et al. Mycobacterium species identification–a new approach via dnaJ gene sequencing. *Syst Appl Microbiol.* 2007;**30**(6):453–62. doi: 10.1016/j.syapm.2007.06.003. [PubMed: 17640840].
34. Simmon KE, Pounder JI, Greene JN, Walsh F, Anderson CM, Cohen S, et al. Identification of an emerging pathogen, Mycobacterium massiliense, by rpoB sequencing of clinical isolates collected in the United States. *J Clin Microbiol.* 2007;**45**(6):1978–80. doi: 10.1128/JCM.00563-07. [PubMed: 17409204].