Six species of non-tuberculosis mycobacteria carry non-identical 16S rRNA gene copies

Keita Takeda,^{a,b,c#} Kinuyo Chikamatsu,^a Yuriko Igarashi,^a Yuta Morishige,^a Yoshiro Murase,^a
Akio Aono,^a Hiroyuki Yamada,^a Akiko Takaki,^a Satoshi Mitarai,^{a,c}

^aDepartment of Mycobacterium Reference and Research, the Research Institute of
Tuberculosis, Tokyo, Japan

^bCenter for Pulmonary Diseases, National Hospital Organization Tokyo National Hospital,
Tokyo, Japan

^cDepartment of Basic Mycobacteriology, Graduate School of Biomedical Science, Nagasaki
University, Nagasaki, Japan

Running Head: 16S rRNA copies of non-tuberculosis mycobacteria

#Address correspondence to Keita Takeda, takedak@tokyo-hosp.jp

Abstract

Non-tuberculosis mycobacteria (NTM) can carry two or more 16S rRNA gene copies that are, in some instances, non-identical. In this study, we used a combined cloning and sequencing approach to analyze the 16S rRNA gene sequences of six NTM species, *Mycobacterium cosmeticum*, *M. pallens*, *M. hodleri*, *M. crocinum*, *M. flavescens*, and *M. xenopi*. The approach facilitated the identification of two distinct gene copies in each species. The two *M. cosmeticum* genes had a single nucleotide difference, whereas two nucleotide polymorphisms were identified in *M. hodleri*, *M. flavescens*, and *M. xenopi*. *M. pallens* had a difference in four nucleotides and *M. crocinum* in 23. Hence, we showed that the six NTM species possess at least two non-identical 16S rRNA gene copies.

Importance

The presence of multiple 16S rRNA gene copies with nucleotide polymorphisms represents a challenge for species identification using 16S rRNA as the target sequence. Our analysis was focused on six NTM species, *M. cosmeticum*, *M. pallens*, *M. hodleri*, *M. crocinum*, *M. flavescens*, and *M. xenopi*. As a result, we generated the full-length sequences of two non-identical 16S rRNA copies for each NTM species. The data will be helpful for the sequence analysis of specimens or other samples.

Keywords: 16S rRNA, gene copy, *Mycobacterium*, non-tuberculous mycobacteria

Introduction

Non-tuberculosis mycobacteria (NTM) comprise over 190 species of the genus *Mycobacterium* on the “List of prokaryotic names with standing in nomenclature” (<http://www.bacterio.net/mycobacterium.html>; accessed on April 13, 2018). NTM species are classified based on sequence comparisons of select housekeeping genes, such as 16S rRNA, *rpoB*, and *hsp65*, and the 16S-23S rRNA internal transcribed spacer region (1–10). Several commercial tests are available that use these genomic sequences for species identification (11–15). However, some NTM species cannot be identified using these tests because of insufficient genome sequence information.

Specifically, Chikamatsu et al. reported failed attempts for the identification of several NTM species using the PyroMark Q24 test kit (QIAGEN, Tokyo, Japan), which is based on pyrosequencing (16). Ambiguous bases were found within the 16S rRNA gene in six NTM species, *Mycobacterium cosmeticum*, *M. pallens*, *M. hodleri*, *M. crocinum*, *M. flavescens*, and *M. xenopi* (16). Direct sequencing using the Sanger method suggested that these species carry two non-identical 16S rRNA gene copies. In general, rapid-growing mycobacteria carry two copies of the 16S rRNA gene (17), whereas, with a few exceptions, slow-growing mycobacteria possess only one copy (18). For example, it was reported earlier that isolates of

the slow-growing *M. terrae* complex (19) and *M. celatum* (20) harbor two non-identical copies of the 16S rRNA gene. Here, we applied a combined cloning and sequencing approach to unequivocally determine the copy numbers and complete sequences of all 16S rRNA genes of the six NTM species investigated earlier (16), *M. cosmeticum*, *M. pallens*, *M. hodleri*, *M. crocinum*, *M. flavescens*, and *M. xenopi*.

Materials and Methods

Bacterial strains. *M. cosmeticum* JCM14739, *M. pallens* JCM16370, *M. hodleri* JCM12141, and *M. crocinum* JCM16369 were obtained from the Japan Collection of Microorganisms (JCM, Ibaraki, Japan). *M. flavescens* ATCC14474 and *M. xenopi* ATCC19250 were acquired from the American Type Culture Collection (ATCC, Manassas, VA). All strains were initially grown on 7H10 agar and then cloned from single colonies. The isolates were sub-cultured in 2% Ogawa medium at 37 °C.

DNA extraction. Bacterial DNA was extracted using the Isoplant Kit (Nippon Gene Co., Ltd, Toyama, Japan). Briefly, one inoculation loop (approximately 10 µl) of fresh colonies grown on Ogawa medium were suspended in 300 µl of extraction buffer and again suspended in 150 µl of lysis buffer for a 15-min incubation at 50 °C. Genomic DNA was extracted with sodium acetate (pH 5.2) on ice for 15 min. After centrifugation (12,000×g, 15 min at 4 °C), the upper phase was transferred to a new tube and the genomic DNA was precipitated with 70% ethanol.

The DNA pellet was dissolved in 50 μ L TE buffer (10 mM Tris-HCl, 1 mM EDTA).

Cloning. The target 16S rRNA genes from each bacterial DNA preparation were amplified with primers 285 (5' GAG AGT TTG ATC CTG GCT CAG 3') and rp2 (5' ACG GCT ACC TTG TTA CGA CTT 3') yielding the almost complete 16S rRNA gene (1, 21). In brief, 25 μ L of a mixture containing ExTaq HS (TaKaRa Bio Inc., Shiga, Japan), 2.5 mM dNTP mixture, 10 μ M of each primer, and 5 μ L template DNA was used for PCR. Amplification was performed in a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA) using 30 cycles of 30 s at 94 °C, 30 s at 60 °C, and 90 s at 72 °C. Then, the PCR products were purified and cloned using a TOPO TA Cloning Kit (Invitrogen, USA). In brief, the PCR products, salt solution, water, and TOPO[®] vector using vaccinia topoisomerase I were mixed at room temperature (22–23 °C) and incubated for 30 min. The recombinant TA cloning mixes were incubated with *E. coli* competent cells (DH5 alpha) for 30 min on ice to perform the transformation. The process was stopped by incubating the samples at 42 °C for 60 s (heat shock), immediately followed by an incubation on ice. Super optimal broth with catabolite repression (SOC) was added to the samples and incubated at 37 °C for one hour.

The competent cells were cultured on Luria-Bertani (LB) agar supplemented with 2 mg of X-gal. Ten white colonies of each transformation were picked from the LB agar and individually cultured in LB broth. Plasmid DNA was isolated and purified using the miniprep Flexiprep Kit (Amersham Biosciences, Buckinghamshire, England) and a

column method with the FastGene Gel/PCR Extraction Kit (Nippon Genetics Co., Ltd, Tokyo, Japan).

Sequence analysis. The sequencing of each 16S rRNA clone was performed using the primers M13 Forward (5′ GTA AAA CGA CGG CCA GT 3′), M13 Reverse (5′ CAG GAA ACA GCT ATG AC 3′) and 264 (5′ TGC ACA CAG GCC ACA AGG GA 3′) with the BigDye Terminator Cycle sequencing kit ver. 3.1 (Applied Biosystems) in an ABI 3500 Genetic Analyzer (Applied Biosystems). Finally, the sequences (approximately 1,500 bp each) of the 10 clones of each species (approximately 1,500 bp each) were aligned and further analyzed using the software package Molecular Evolutionary Genetics Analysis Ver. 7 (22).

Results

The sequence alignments led to the identification of 2 non-identical 16S rRNA copies for each of the six NTM species. The results have been deposited with GenBank under accession number as follow: *M. cosmeticum* (MH169224 and MH169226), *M. pallens* (MH169208 and MH169209), *M. hodleri* (MH169216 and MH169217), *M. crocinum* (MH169218 and MH169219), *M. flavescens* (MH169220 and MH169222), and *M. xenopi* (MH169221 and MH169241). The nucleotide polymorphisms are shown in Figure 1. *M. cosmeticum* had a single nucleotide difference between the two sequences. Two-nucleotide differences were found in *M. hodleri*, *M. flavescens*, and *M. xenopi*. *M. pallens* had a difference in four

nucleotides and *M. crocinum* in 23.

Discussion

It is well documented that some species of the genus *Mycobacterium* harbor multiple 16S rRNA gene copies with distinct sequences (16, 19, 20, 23–26). In this study, the cloning experiments targeting 16S rRNA genes facilitated the identification two distinct copies in six NTM species, *M. cosmeticum*, *M. pallens*, *M. hodleri*, *M. crocinum*, *M. flavescens*, and *M. xenopi*. Each strain was re-isolated from a single colony, and each of the two 16S rRNA gene copies was reproducibly obtained from multiple clones derived from these isolates. In addition, the nucleotide polymorphisms observed for the species-specific gene copies were supported by earlier findings obtained by direct sequencing, e.g., the nucleotide position 185 of the 16S rRNA copies in *M. cosmeticum* was A or G by our cloning-sequencing experiment (Figure 1), but a mixture of A and G at this position had been indicated earlier by direct Sanger sequencing (16). Our new data were confirmed by multiple clones per strain to minimize the possible impact of technical sequencing errors.

Chilia et al. reported earlier that sequences obtained from clones may be more definitive than sequence data obtained from direct sequencing (23). Indeed, our sequence data obtained from the cloning experiment unequivocally established the existence of two non-identical gene copies per species whereas the earlier direct sequencing only suggested a polymorphism

based on sequence ambiguities. Hence, it is recommended to avoid direct sequencing for species identification if there are non-identical genomic copies of the target sequence.

However, our study could not reveal whether the NTM species had more than two 16S rRNA gene copies per genome. It is possible that one genome carries several 16S rRNA copies with identical sequences. This problem might be resolved by whole-genome sequencing (WGS). A database search revealed that *M. cosmeticum* DSM 44829 has two 16S rRNA genes (GenBank accession: NZ_CCBB010000003.1), *M. flavescens* strain M6 has three (GenBank accession: NZ_MIHA000000000.1), and *M. xenopi* has one (Strain DSM 43995, GenBank accession: LQQB01000023.1) or two (Strain RIVM700367, GenBank accession: NZ_AJFI01000116.1). These WGS data were obtained by shotgun sequencing, which also has limitations regarding the identification of identical or almost identical gene copies (27–29). Hence, cloning along with sequencing is still required, but an improvement in WGS data accuracy will be obtained in near future by implementing long-read sequencing using the single molecular, real-time sequencing technology (28, 29). Currently, the exact 16S rRNA gene copy number is not yet known for *M. pallens*, *M. hodleri*, and *M. crocinum*, requiring further analysis.

In this study, we established the existence of two 16S rRNA gene copies for each of the six NTM species examined. However, the detected nucleotide polymorphism can be a challenge for species identification using 16S rRNA sequencing. The identification of the two

non-identical 16S rRNA copies in NTM species will be helpful for sequence analyses of specimens or other samples and sequencing efforts.

References

- 1) Kirschner P, Springer B, Vogel U, Meier A, Wrede A, Kiekenbeck M, Bange FC, Böttger EC. 1993. Genotypic identification of mycobacteria by nucleic acid sequence determination: report of a 2-year experience in a clinical laboratory. J Clin Microbiol 31:2882–2889.
- 2) Kim SY, Park HY, Jeong BH, Jeon K, Huh HJ, Ki CS, Lee NY, Han SJ, Shin SJ, Koh WJ. 2015. Molecular analysis of clinical isolates previously diagnosed as *Mycobacterium intracellulare* reveals incidental findings of "*Mycobacterium indicus pranii*" genotypes in human lung infection. BMC Infect Dis 15:406.
- 3) Tortoli E, Fedrizzi T, Meehan CJ, Trovato A, Grottola A, Giacobazzi E, Serpini GF, Tagliazucchi S, Fabio A, Bettua C, Bertorelli R, Frascaro F, De Sanctis V, Pecorari M, Jousson O, Segata N, Cirillo DM. 2017. The new phylogeny of the genus *Mycobacterium*: The old and the news. Infect Genet Evol 56:19–25.
- 4) Telenti A, Marchesi F, Balz M, Bally F, Böttger EC, Bodmer T. 1993. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. J Clin Microbiol 31:175–178.

- 5) Domenech P, Menendez MC, Garcia MJ. 1994. Restriction fragment length polymorphisms of 16S rRNA genes in the differentiation of fast-growing mycobacterial species. FEMS Microbiol Lett 116:19–24.
- 6) Roth A, Reischl U, Streubel A, Naumann L, Kroppenstedt RM, Habicht M, Fischer M, Mauch H. 2000. Novel diagnostic algorithm for identification of mycobacteria using genus-specific amplification of the 16S-23S rRNA gene spacer and restriction endonucleases. J Clin Microbiol 38:1094–1104.
- 7) Lee H, Park HJ, Cho SN, Bai GH, Kim SJ. 2000. Species identification of mycobacteria by PCR-restriction fragment length polymorphism of the *rpoB* gene. J Clin Microbiol 38:2966–2971.
- 8) Rogall T, Wolters J, Flohr T, Böttger EC. 1990. Towards a phylogeny and definition of species at the molecular level within the genus *Mycobacterium*. Int J Syst Bacteriol 40:323–330.
- 9) Stahl DA, Urbance JW. 1990. The division between fast- and slow-growing species corresponds to natural relationships among the mycobacteria. J Bacteriol 172:116–124.
- 10) Tortoli E. 2012. Phylogeny of the genus *Mycobacterium*: many doubts, few certainties. Infect Genet Evol 12:827–831.
- 11) Huh HJ, Kwon HJ, Ki CS, Lee NY. 2015. Comparison of the genedia MTB detection kit and the cobas TaqMan MTB assay for detection of *Mycobacterium tuberculosis* in

respiratory specimens. J Clin Microbiol 53:1012–1014.

12) Bergmann JS, Woods GL. 1996. Clinical evaluation of the Roche AMPLICOR PCR *Mycobacterium tuberculosis* test for detection of *M. tuberculosis* in respiratory specimens. J Clin Microbiol 34:1083–1085.

13) Tanaka H, Hirose H, Kato Y, Kida S, Miyajima E. 2010. Clinical evaluation of TRCRapid M.TB for detection of *Mycobacterium tuberculosis* complex in respiratory and nonrespiratory specimens. J Clin Microbiol 48:1536–1541.

14) Helb D, Jones M, Story E, Boehme C, Wallace E, Ho K, Kop J, Owens MR, Rodgers R, Banada P, Safi H, Blakemore R, Lan NT, Jones-López EC, Levi M, Burday M, Ayakaka I, Mugerwa RD, McMillan B, Winn-Deen E, Christel L, Dailey P, Perkins MD, Persing DH, Alland D. 2010. Rapid detection of *Mycobacterium tuberculosis* and rifampin resistance by use of on-demand, near-patient technology. J Clin Microbiol 48:229–237.

15) Lee AS, Jelfs P, Sintchenko V, Gilbert GL. 2009. Identification of non-tuberculous mycobacteria: utility of the GenoType Mycobacterium CM/AS assay compared with HPLC and 16S rRNA gene sequencing. J Med Microbiol 58:900–904.

16) Chikamatsu K, Aono A, Hata H, Igarashi Y, Takaki A, Yamada H, Sakashita K, Mitarai S. 2018. Evaluation of PyroMark Q24 pyrosequencing as a method for the identification of mycobacteria. Diagn Microbiol Infect Dis 90:35–39.

17) Domenech P, Menendez MC, Garcia MJ. 1994. Restriction fragment length polymorphism

of 16S rRNA genes in the differentiation of fast-growing mycobacterial species. FEMS Microbiol Lett 116:19–24.

18) Ji YE, Colston MJ, Cox RA. 1994. The ribosomal RNA (*rrn*) operons of fast-growing mycobacteria: primary and secondary structure and their relation to *rrn* operons of pathogenic slow-growers. Microbiol 140:2829–2840.

19) Ninet B, Monod M, Emler S, Pawlowski J, Metral C, Rocner P, Auckenthaler R, Hirschel B. 1996. Two different 16S rRNA genes in a mycobacterial strain. J Clin Microbiol 34:2531–2536.

20) Reischl U, Feldmann K, Naumann L, Gaugler BJ, Ninet B, Hirschel B, Emler S. 1998. 16S rRNA sequence diversity in *Mycobacterium celatum* strains caused by presence of two different copies of 16S rRNA gene. J Clin Microbiol 36:1761–1764.

21) Adekambi T, Drancourt M. 2004. 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 54:2095–2105.

22) Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870–1874.

23) Cilia V, Lafay B, Christen R. 1996. Sequence heterogeneities among 16 ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. Mol Biol Evol 13:451–461.

- 229 24) Menendez MC, Garcia MJ, Navarro MC, Gonzalez-y-Merchand JA, Rivera-Gutierrez S,
230 Garcia-Sanchez L, Cox RA. 2002. Characterization of an rRNA operon (*rrnB*) of
231 *Mycobacterium fortuitum* and other mycobacterial species: implications for the
232 classification of mycobacteria. J Bacteriol 184:1078–1088.
- 233 25) Viezens J, Arvand M. 2008. Simultaneous presence of two different copies of the 16S
234 rRNA gene in *Bartonella henselae*. Microbiol 154:2881–2886.
- 235 26) Conville PS, Witebsky FG. 2005. Multiple copies of the 16S rRNA gene in *Nocardia nova*
236 isolates and implications for sequence-based identification procedures. J Clin Microbiol
237 43:2881–2885.
- 238 27) Goodwin S, McPherson JD, McCombie WR. 2016. Coming of age: ten years of
239 next-generation sequencing technologies. Nat Rev Genet 17:333–351.
- 240 28) Schadt EE, Turner S, Kasarskis A. 2010. A window into third-generation sequencing.
241 Hum Mol Genet 19:227–240.
- 242 29) Nakano K, Shiroma A, Shimoji M, Tamotsu H, Ashimine N, Ohki S, Shinzato M, Minami
243 M, Nakanishi T, Teruya K, Satou K, Hirano T. 2017. Advantages of genome sequencing
244 by long-read sequencer using SMRT technology in medical area. Hum Cell 30:149–161.

Figure 1

Nucleotide polymorphism between two non-identical 16S rRNA sequences of six NTM species. Nucleotide positions were derived from an alignment with the 16S rRNA gene of *M.tuberculosis* H37Rv ATCC272 (GenBank accession: NC_000962)

M. cosmeticum

185

Variant 1: ATAGGACTCCAGCCTTCATGG

Variant 2:G.....

M. pallens

185–187

470

Variant 1: ATAGGACCACATGCTTCATGGTG

GACGGTACCTATAGAAGAAGC

Variant 2:GGC.....

.....G.....

M. hoderi

197198

Variant 1: GATGCATGTCTCTTGGTGGA

Variant 2:CT.....

M. crocinum

73

77

80

87

90

93

183

185–187

196197

Variant 1: AACGGTAAG-GCCCTTCGGGGT-ACACGAGT

GGACCACGGCCTTCATGGGTGTGG

Variant 2:A...T...T.....A...G...T.....

.....G...ATG.....TG.....

441–445

467–471

1245

Variant 1: TTTCAGGTAGGGACGAAGCGCAAGTGACGGTACCTATAGAAG

CCGGTACAAAAGGCTGCGATG

Variant 2:CCCAC.....GTGGG.....

.....G.....

M. flavescens

184185

Variant 1: AATATTCCCTATTGGTCGCATG

Variant 2:GC.....

M. xenopi

210

Variant 1: TGGTGGAAAGTGTTTGGTAGC

Variant 2:C.....

434

GTTGTAAACCTCTTTCAGCCT

.....C.....