

18 **Number of words in Abstract:** 96

19 **Number of words in Text:** 999

20 **Running title:** NTM spectrum: conventional vs. 16S sequencing

21 **Keywords:** *Mycobacterium gordonae*, *Mycobacterium scrofulaceum*, South Africa

22

DRAFT

23 **SUMMARY**

24 Non-tuberculous mycobacterial isolates from gold miners were speciated using standard
25 biochemical testing (SBT) and 16s rDNA sequencing. Of 237 isolates tested, SBT identified
26 126 compared with all 237 identified by sequencing. Of 111 isolates unspciated by SBT but
27 identified by sequencing, 38 (34.2%) were identified as *Mycobacterium gordonae* and 8
28 (7.2%) were new species. Of 126 isolates speciated by both methods, 37 were discordant,
29 with 14/17 *M. gordonae* isolates incorrectly identified as *M. scrofulaceum* using SBT. The
30 majority of these were the potentially pathogenic strain D *M gordonae*: sequencing is
31 preferable where available to guide treatment.

32

33

DRAFT

34 Identifying non-tuberculous mycobacteria (NTM) is important, especially where HIV is
35 prevalent; to distinguish potential pathogens. In South African gold mines, use of liquid
36 mycobacterial culture media has increased both the yield of positive cultures and the
37 proportion of NTM isolated [1].

38

39 Conventionally, NTM are speciated using standard biochemical testing (SBT). 16S ribosomal
40 ribonucleic acid (rRNA) gene sequence determination (16S rDNA sequencing) provides
41 faster, accurate speciation and can identify new species [2].

42

43 We compared the spectrum of NTM identified by SBT versus sequencing in a gold-mining
44 population and linked a subgroup of isolates to clinical data. This is the one of the larger
45 clinical studies of NTM reported.

46

47 **METHODS**

48 This work was part of a sub-study [1] of “Thibela TB”, a cluster-randomised trial of
49 community-wide isoniazid preventive therapy (IPT). At pre-IPT screening and follow-up
50 visits, [1] and at routine mine health facilities (restricted to those without prior tuberculosis),
51 we recruited individuals with suspected tuberculosis, between July 2006 and December 2007.
52 Participants gave one sputum specimen; all isolates with results from both SBT and
53 sequencing were included.

54

55 Following decontamination, specimens were cultured using both BACTEC MGIT 960
56 system (BD Diagnostics, Sparks MD) and Löwenstein-Jensen media [1]. *Mycobacterium*
57 *tuberculosis* complex was distinguished from NTM by detection of MPB64 antigen (Capilia
58 TB, Japan). Phenotypic identification was based on growth rate at 25⁰C, 37⁰C, 42⁰C, 45⁰C

59 and with p-nitrobenzoic acid; pigmentation and colony morphology in light and dark
60 conditions at 37⁰C. SBT included Tween hydrolysis; nitrate reduction and the catalase test.
61 For sequencing, heat-killed culture lysates were subjected to 5'-16s rDNA amplification;
62 sequenced [2] and referred to the RIDOM and NCBI GenBank sequence databases for
63 identification [3,4]. GyrB genes of heat killed lysates were sequenced to confirm the absence
64 of *M. tuberculosis*, identify other members of MTB complex and confirm *M. kansasii*
65 identification.

66
67 Where SBT and sequencing results were discrepant, sequencing was repeated. SBT was
68 repeated for isolates identified as *M. scrofulaceum* using SBT and *M. gordonae* using
69 sequencing. For a subgroup with discrepant identification, because of uncertainty regarding
70 pathogenicity, routine clinical data were collected retrospectively using a standardised case
71 report form.

73 RESULTS

74 237 isolates were included. Dominant species identified using SBT included *M. kansasii*
75 (51 isolates), *M. avium* complex (47) and *M. scrofulaceum* (17); and using sequencing,
76 *M. gordonae* (62), *M. kansasii/M. gastri* (53), *M. avium complex* (38) and *M.*
77 *parascrofulaceum* (20). 28/237 isolates (11.8%) contained mixed NTM species on
78 sequencing but none had mixed NTM/MTB. 111 isolates were not identifiable by SBT, but
79 sequenced as follows: *M. gordonae* (38 isolates), *M. fortuitum* (17), *M. parascrofulaceum*
80 (10), *M. avium complex* (7), *M. kansasii/M. gastri* (5), other NTM species (22), new
81 mycobacterial species (8) or non-mycobacterial species (4).

82 Among 126 isolates successfully speciated by both methods, 38 (30%) were discordant on
83 initial testing (table 1). Among 17 isolates identified as *M. scrofulaceum* using SBT, most
84 (14/17) were identified as *M. gordonae* by sequencing.

85

86 Figure 1 shows a portion of the 16S rRNA sequence of *M. scrofulaceum* and *M. gordonae*
87 strains (positions 392 to 446), indicating one of the few major differences between these
88 species. Differences are visible at positions 411 to 427, including a three base-pair
89 insertion/deletion. Only two minor variations within *M. gordonae* strains are observed
90 among these 10 isolates at position 412 (TC or CC) and position 426 (GC or AT). The
91 sequences of 14 isolates, biochemically identified as *M. scrofulaceum*, are identifiable as *M.*
92 *gordonae* strains by sequencing. In 9/13 strains (one was not re-sequenced fully), TC
93 replaces CC in position 412.

94

95 On repeat SBT, successful for 11/14 isolates originally identified as *M. scrofulaceum*, 10
96 were *M. gordonae* and one retained the initial identification of *M. scrofulaceum*. Among 38
97 isolates for which SBT and sequencing were discordant, repeat sequencing produced the
98 same result for 28; one isolate initially identified as *M. szulgai* was identified as *M.*
99 *parascrofulaceum* on repeat sequencing; the remaining nine isolates had poor and
100 uninterpretable results.

101

102 Clinical data were available for 8/10 *M. gordonae* strain D isolates, identified by SBT as *M.*
103 *scrofulaceum*. Six individuals were recruited at Thibela TB study sites and two at routine
104 health services. 3/8 had a history of previous tuberculosis; all were smear negative; 2/8
105 reported cough, with one additionally reporting weight loss. An HIV test result was recorded
106 for 1/8, who was HIV negative. 3/8 had cavitation on chest radiograph, only one of whom

107 had previous tuberculosis. 1/8 (HIV negative with chest cavitation and no prior tuberculosis)
108 was given standard tuberculosis treatment.

109

110 **DISCUSSION**

111 *M. gordonae* identified in sputum is generally considered to be non-pathogenic and has
112 frequently been isolated from tap water, whereas *M. scrofulaceum* is considered to cause
113 disease [5]. In nine of our *M. gordonae* isolates, a polymorphism (TC replacing CC at
114 position 412) was shown that is associated with *M. gordonae rpoB* cluster D, which may be
115 more pathogenic than other strains [6]. We note that *M. gordonae* can be pathogenic in the
116 immunocompromised [7, 8] and may be causing disease in some individuals in this
117 population, although relatively low numbers make it difficult to be certain. Accurate
118 distinction between species is therefore important in populations with high HIV prevalence,
119 such as this.

120

121 The dominant NTM species were *M. kansasii*, *M. gordonae*, *M. parascrofulaceum* and
122 members of *M. avium* and *M. fortuitum* complexes. *M. kansasii* is known to be prevalent
123 among miners [9]. In previous studies of NTM in miners, SBT was used to identify species
124 mostly cultured on LJ [9, 10]; our data suggest that some *M. gordonae* strains could have
125 been misidentified by SBT as *M. scrofulaceum*, some being associated with features of
126 disease. The importance of this observation lies in the perceived pathogenicity of these two
127 organisms and in our understanding of NTM species distribution in this population.

128

129 **CONCLUSIONS**

130 Some *M. gordonae* strains can be misclassified by SBT as *M. scrofulaceum*.

131 Misidentification of NTM may lead to suboptimal clinical management, particularly in

132 settings with HIV prevalence. Sequencing should be used where available to accurately
133 identify NTM and where SBT is used, the possibility of misidentification should be
134 considered.

135

DRAFT

136 **ACKNOWLEDGEMENTS**

137 Thibela TB is funded by CREATE, with grants from the Bill and Melinda Gates Foundation,
138 and the Safety in Mines Research Advisory Committee (South Africa).

139 Violet Chihota was supported by NIH Fogarty ICORTA TB/AIDS (Grant 5U2RTW007370
140 and 5U2RTW007373).

141 Clare van Halsema was funded by a grant from the Colt Foundation, UK

142 Alison Grant was supported by a Public Health Career Scientist Award from the Department
143 of Health, UK.

144 Katherine Fielding is part-funded by the Biostatistics core of the Consortium to Respond
145 Effectively to the AIDS and TB Epidemics (CREATE) with a grant from the Bill and
146 Melinda Gates Foundation.

147 We thank Minty van der Meulen for her laboratory work

148

149 **REFERENCES**

- 150 1. **Chihota V N, Grant A D, Fielding K, et al.** Liquid vs. Solid culture for tuberculosis:
151 performance and cost in a resource-constrained setting. *Int J Tuberc Lung Dis* 2010;
152 **14(8):1024-31.**
- 153 2. **Rogall T, Flohr T, Böttger E C.** Differentiation of *Mycobacterium* species by direct
154 sequencing of amplified DNA. *J Gen Microbiol* 1990; **136**: 1915-1920.
- 155 3. **Benson D A, Karsch-Mizrachi I, Lipman D J, Ostell J and Sayers E W.** 2011.
156 GenBank. *Nucleic acid research* 2011; 39: D32-D37.
- 157 4. **Harmsen, D, Dostal S, Roth A et al.** RIDOM: comprehensive and public sequence
158 database for identification of *Mycobacterium* species. *BMC Infect Dis* 2003; **3:26.**
- 159 5. **Griffith, D E, Aksamit T, Brown-Elliott B A et al.** An official ATS/IDSA statement:
160 Diagnosis, treatment and prevention of non-tuberculous mycobacterial diseases. *Am J*
161 *Resp Crit Care Med* 2007; **175:367-416.**
- 162 6. **Itoh S, Kazumi Y, Abe C, Takahashi M.** Heterogeneity of RNA polymerase gene
163 (*rpoB*) sequences of *Mycobacterium gordonae* Clinical isolates identified with a DNA
164 probe kit and by conventional methods. *J Clin Micro* 2003; **41(4):1656-63.**
- 165 7. **Barber T W, Craven D E, Farber H W.** *Mycobacterium gordonae*: a possible
166 opportunistic respiratory tract pathogen in patients with advanced human
167 immunodeficiency virus, type 1 infection. *Chest* 1991; **100:716-20.**
- 168 8. **Lessnau K D, Milanese S Talavera W.** *Mycobacterium gordonae*: a treatable disease in
169 HIV-positive patients. *Chest* 1993; **104:1779-85.**
- 170 9. **Corbett E L, Hay M, Churchyard G J, et al.** *Mycobacterium kansasii* and *M.*
171 *scrofulaceum* isolates from HIV-negative South African gold miners: incidence, clinical
172 significance and radiology. *Int J Tuberc Lung Dis* 1999; **3:501-507.**

173 10. **Corbett E L, Churchyard G J, Clayton T, et al.** Risk factors for pulmonary
174 mycobacterial disease in South African gold miners. A case-control study. *Am J Resp*
175 *Crit Care Med* 1999; **159**:94-99.

176

177

178

179

180

DRAFT

181 **Figure Legend:**

182 Figure 1: 16S rRNA sequence alignment of *M. gordonae* and *M. scrofulaceum*. Sequences of
183 *Mycobacterium* type strains are shown in row 1-12. *M. scrofulaceum* type strain sequences
184 (ATCC 19981 and DSM 43992) are shown in row 1-2, while row 3-12 show *M. gordonae*
185 type strain sequences (ATCC 14470, DSM 44160, agha3, Tropicalis, NIPHL050404TB,
186 M138, M120, M223, Tropicalis-2 and Tropicalis-3). Rows 13-15 are examples of clinical
187 isolates from our study that were identified as *M. scrofulaceum* on initial standard
188 biochemical testing and *M. gordonae* on sequencing. Two of these clinical isolates (rows 14-
189 15) show TC instead of a CC at position 412.

190

191

DRAFT

192 **Tables**

193 Table 1: Identification of non-tuberculous mycobacteria using standard biochemical testing
 194 and 16S rDNA sequencing: discordant results on initial testing

Standard biochemical testing	n	16S rDNA sequencing	N
<i>M. scrofulaceum</i>	17	<i>M. gordonae</i>	14
		<i>M. szulgai</i>	2
		<i>M. fortuitum</i>	1
<i>M. avium</i> complex	16	<i>M. parascrofulaceum</i>	9
		<i>M. paraffinicum</i>	3
		<i>M. fortuitum</i>	1
		<i>M. kyorinense</i>	1
		<i>M. palustre</i>	1
		New mycobacterial species	1
<i>M. kansasii</i>	3	<i>M. gordonae</i>	1
		<i>M. parascrofulaceum</i>	1
		<i>M. szulgai</i>	1
<i>M. gordonae</i>	1	<i>M. asiaticum</i>	1
<i>M. flavescens</i>	1	<i>M. gordonae</i>	1
Total	38	Total	38

195

196

