Polymorphism in the RD1 locus and its effect on downstream genes among South Indian clinical isolates of *Mycobacterium tuberculosis*

Ahmed Kabir Refaya, Shanmugam Sivakumar, Balaji Sundararaman and Sujatha Narayanan

Department of Immunology, National Institute for Research in Tuberculosis, Chetput, Chennai 600 031, India

INTRODUCTION

*Mycobacterium tuberculosis*, one of the most successful pathogens in the world for thousands of years, remains a major public health burden around the globe in the human immunodeficiency virus/AIDS era. Whole-genome hybridization between *Mycobacterium bovis* BCG and virulent clinical strains of *M. tuberculosis* has identified a novel deletion in the genome of the attenuated strain named the region of difference (RD1) (Mahairas et al., 1996; Talbot et al., 1997; Behr et al., 1999). Since the identification of this region, RD1 has become the most investigated and attractive locus for vaccine research, and also for understanding the pathogenesis of mycobacteria. There are nine predicted ORFs (Rv3871–Rv3879c) spanning the 9.5 kb RD1 region (Cole et al., 1998; Behr et al., 1999). Partial deletion of the RD1 region or insertion mutation in the Rv3874 (CFP10) gene, and mutations in Rv3870, Rv3871, Rv3876 and Rv3877 attenuate *M. tuberculosis* and reduce its cytolytic effects on macrophages and its spread within them. Gene disruption and deletion studies have demonstrated that Rv3871 is necessary for the secretion of ESAT6/CFP10 (Hsu et al., 2003), whilst Stanley et al. (2003) illustrated that Rv3871 interacts directly with Rv3874 and this interaction is necessary for successful secretion of the latter protein.

We have been exploring the genetic polymorphisms of clinical isolates of *M. tuberculosis* using various genetic markers such as IS6110 RFLP, spoligotyping and deletion microarrays (Narayanan et al., 2008). South Indian *M. tuberculosis* strains have a lower number of IS6110 copies and have been reported previously to be lower in virulence compared with *M. tuberculosis* isolates from London in a guinea pig model (Mitchison et al., 1960). Rao et al. (2005) used a single set of primers to amplify the 9.2 kb RD1 region and showed that all clinical strains of Indian origin tested in their study lacked this region. This is technically demanding, and any mutations in the primer-binding site might give false-negative results. Therefore, Soman et al. (2007) checked whether the RD1 genes were really absent in clinical strains of Indian origin by using three sets of primers to amplify the coding regions alone. The results of the above two conflicting reports from India tempted us to explore polymorphisms of the RD1 genes in our isolates, which were reported previously to be of low virulence (Mitchison et al., 1960).
During screening of the RD1 region, we found a polymorphic intergenic region comprising a 53 bp repeat element between the two genes Rv3870 and Rv3871, which was identified previously as the mycobacterial interspersed repetitive unit (MIRU) 39 locus and was excluded from analysis of 15 loci by mycobacterial interspersed repetitive unit-variable number of tandem repeat (MIRU-VNTR) studies due to its low Hunter–Gaston discriminatory index (HGDI) and low allelic diversity index. The functional role of this repeat element in transcription and translation has been studied in a few pathogenic bacteria. Such repeat elements residing in coding regions may result in both frame-shift and non-frameshift mutations, allowing the bacteria to adapt optimally to different environments. Repeat elements have been implicated in the pathogenesis of *M. tuberculosis* and other pathogenic bacteria (Whatmore, 2001; Erwin et al., 2006). For example, polymorphisms in the VNTR 3690 locus of *M. tuberculosis* were detected in the intergenic region between Rv3304 and Rv3303c resulted in 12.5-fold upregulation of *lpdA* expression (Akhtar et al., 2009). Therefore, we investigated the presence or absence of the RD1 genes in south Indian strains and also studied whether this polymorphism had any effect on the transcription of downstream genes in the RD1 region.

**METHODS**

**Reference strains and clinical isolates.** *M. tuberculosis* H37Rv and *M. bovis* BCG were used as reference strains. A total of 407 mycobacterial isolates received from pulmonary tuberculosis patients as part of the model DOTS project from the Tiruvallur district, south India, between 1999 and 2003, were included in this study. Sputum samples were processed using a modified Petroff’s method, cultured on Lowenstein–Jensen medium and examined weekly for up to 8 weeks for growth (Canetti et al., 1969). Positive cultures were subjected to identification tests – niacin production, growth in 7H10 albumin/dextrose saline and 0.05 % Tween 80. Trizol reagent (Invitrogen) was added to a cell pellet obtained from 20–30 ml culture and disrupted with 0.1 mm zirconia beads in a mini-bead beater. Total RNA was purified using an RNeasy purification kit (Qiagen) following the manufacturer’s instructions. Contaminated DNA in the RNA sample was digested with RNase-free DNase I. The purity of the RNA was determined by measuring the absorbance at 260 and 280 nm. The first-strand cDNA was synthesized from 1 μg total RNA using an Improm Reverse Transcriptase II kit (Promega).

Real-time quantitative RT-PCR (qRT-PCR) for Rv3871 and Rv3874 was performed in an ABI 7500 system (Applied Biosystems) using *Taq*Man assays. 16S rRNA was used for normalization. Each reaction was repeated three times with independent RNA samples. Negative controls consisting of no reverse transcriptase and no template mixtures were run with all reactions. After baseline correction and determination of threshold settings, calculations were carried out using the 2^−ΔΔCt method of Livak & Schmittgen (2001), as PCR efficiencies were found to be similar. Results were expressed as fold induction, which denoted the fold change in expression of the gene.

**Statistical analysis.** The HGDI described by Hunter & Gaston (1988) was used as a numerical index for RD1 locus discriminatory power. The HGDI was calculated using the following formula:

$$HGDI = 1 - \frac{1}{N(N-1)} \sum_{i=1}^{s} n_i(n_i - 1)$$

where *N* was the total number of strains in the typing scheme, *s* was the total number of different patterns in locus RD1 and *n* was the number of strains belonging to the *j*th pattern. A one-way analysis of variance table with Turkey’s multiple comparison tests was used to calculate the change in fold induction between each group using GraphPad Prism software.

**RESULTS**

PCR conditions for the five sets of primers were standardized using *M. tuberculosis* H37Rv and *M. bovis* BCG each dNTP, 1 × PCR buffer (Invitrogen), 1.5 U *Taq* polymerase (Invitrogen) and 1.5 mM MgCl₂, made up to 25 μl with nuclease- and protease-free water (Gibco). Genomic DNA (1–10 ng) from the clinical isolates served as template. PCR amplicons were resolved by 1.5% agarose gel electrophoresis, except for the RD-INS amplicons, which were resolved in a 2% agarose gel, and visualized on a UV trans-illuminator after staining with ethidium bromide.

**Automated DNA sequencing.** PCR products were gel purified using Amersham GFX (Wipro GE Healthcare) columns according to the manufacturer’s protocol, and the DNA concentration was determined by comparison with known standards following agarose gel electrophoresis. Approximately 200 ng DNA was cycle sequenced using 3 pmol primer with a BigDye Terminator Ready Reaction Mix (Applied Biosystems), and excess unincorporated dye was removed using a DyeEx kit (Qiagen) following the manufacturer’s instructions. The final product was vacuum dried, suspended in 13 μl HiDi formamide and sequenced using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Sequenced data were analysed using DNA Sequencing Analysis Software (Applied Biosystems). DNA sequences were analysed using Genetools software (http://www.syngene.com/genetools/).

**RNA isolation and real-time PCR.** All strains were grown to exponential growth phase in Middlebrook 7H9 broth containing albumin/dextrose saline and 0.05% Tween 80. Trizol reagent (Invitrogen) was added to a cell pellet obtained from 20–30 ml culture and disrupted with 0.1 mm zirconia beads in a mini-bead beater. Total RNA was purified using an RNeasy purification kit (Qiagen). The purity of the RNA was determined by measuring the absorbance at 260 and 280 nm. The first-strand cDNA was synthesized from 1 μg total RNA using an Improm Reverse Transcribease II kit (Promega).

A one-way analysis of variance table with Turkey’s multiple comparison tests was used to calculate the change in fold induction between each group using GraphPad Prism software.

**PCR amplification for genes in RD1.** Initially, five sets of primers were designed to amplify the intergenic regions between genes in the RD1 region along with a few hundred bp of the flanking genes. Another set of internal primers for the RD1A region was designed to amplify exactly the intergenic region between Rv3870 and Rv3871 genes, which was named RD-INS. Primer sequences, annealing efficiencies were found to be similar. Results were expressed as fold induction, which denoted the fold change in expression of the gene.

**Genotyping by spoligotyping.** Genomic DNA from the isolates was extracted using a cetyl trimethylammonium bromide/NaCl method described elsewhere (Baess, 1974). Spoligotyping was performed with a commercially available kit (Ocimum Biosolutions) according to the manufacturer’s instructions. The analysis defined ‘related strains’ as two or more isolates with identical spoligotyping patterns, and these patterns were analysed using Spotclust software (http://thinsight.cs.rpi.edu/about_spotclust.html). The most common strain in our study region is the EAI strain, which can be defined as isolates lacking spots 29, 30, 31 and 32 by spoligotyping. EAI 3 strains lack spots 2, 3, 37, 38 and 39 in addition to the family characteristic spoligotype, whilst EAI 5 strains lack spot 34 in addition to the family characteristic spoligotype.

**RESULTS**

PCR conditions for the five sets of primers were standardized using *M. tuberculosis* H37Rv and *M. bovis* BCG...
Table 1. Primer sequences used in the study and regions they amplify

<table>
<thead>
<tr>
<th>Fragment name</th>
<th>Primer or probe name</th>
<th>Sequence (5’→3’)</th>
<th>Annealing conditions (°C/min)</th>
<th>Expected product size (bp)</th>
<th>Genomic coordinates*</th>
<th>Genes amplified</th>
</tr>
</thead>
<tbody>
<tr>
<td>RD1A</td>
<td>RD1AF</td>
<td>CCGGGCTGCAAACGCTTTT</td>
<td>60/1</td>
<td>810</td>
<td>4348245–4349055</td>
<td>Rv3870–Rv3871</td>
</tr>
<tr>
<td></td>
<td>RD1AR</td>
<td>GGGGGCGAGAAGCTGCTTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD1B</td>
<td>RD1BF</td>
<td>CCGGGCAGCGCGGTTAAGA</td>
<td>60/1</td>
<td>736</td>
<td>4350585–4351321</td>
<td>Rv3872–Rv3873</td>
</tr>
<tr>
<td></td>
<td>RD1BR</td>
<td>GGGCTGCTGTTGACGCGGTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD1C</td>
<td>RD1CF</td>
<td>GAGCTGGGACGAGGAGGACG</td>
<td>58/1</td>
<td>714</td>
<td>4352154–4352868</td>
<td>Rv3874–Rv3875</td>
</tr>
<tr>
<td></td>
<td>RD1CR</td>
<td>GGGCTGACTGGAGCGAAGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD1D</td>
<td>RD1DF</td>
<td>CAGGAGTCGGGGGAAGGGTCG</td>
<td>60/1</td>
<td>718</td>
<td>4353645–4352928</td>
<td>Rv3875–Rv3876</td>
</tr>
<tr>
<td></td>
<td>RD1DR</td>
<td>GGGGGAGGAGGAGGAGGACG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD1E</td>
<td>RD1EF</td>
<td>CGGGATGGTGGTGGCTGGAAGT</td>
<td>60/1</td>
<td>775</td>
<td>4354671–4355446</td>
<td>Rv3876–Rv3877</td>
</tr>
<tr>
<td></td>
<td>RD1ER</td>
<td>GCCCCAACAAGGGATTCAATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDIN</td>
<td>RDINF</td>
<td>CGCGGCATTTCACAGGTTCACC</td>
<td>60/1</td>
<td>180–420</td>
<td>4348667–4348897</td>
<td>Rv3870–Rv3871</td>
</tr>
<tr>
<td></td>
<td>RDINR</td>
<td>CGCGATTCAGTGCCGAGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qRT-PCR primers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RV3871</td>
<td>Forward</td>
<td>AGTTGGACCAGGTGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>ACCCCCGGGTGATC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fluorogenic probe</td>
<td>FAM-CATAGCCGCGCAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RV3871</td>
<td>Forward</td>
<td>TAGACCACAAAATTCACGCGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>TGCTTTTGGCAGTTCTTGGAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fluorogenic probe</td>
<td>FAM-GTCACGCCAAGTCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Coordinates as given by the *M. tuberculosis* H37Rv genome sequence in the TubercuList database (http://genolist.pasteur.fr/TubercuList/).
genomic DNA as templates. Seventy clinical isolates were selected randomly and analysed for deletions and polymorphisms in the RD1 region. There was no major insertion or deletion polymorphism observed in RD1B, -C, -D or -E, as all the clinical isolates tested generated the expected product sizes of 736, 714, 718 and 775 bp, respectively. The standardized PCR conditions for the primers (RD1-A, -B, -C, -D and -E) are shown in Table 1.

We observed significant polymorphisms in the RD1A region, which comprises part of Rv3870 and Rv3871 and the intergenic region between them in the clinical isolates. The PCR product size varied from approximately 50 to 200 bp, as determined by 1.5% agarose gel electrophoresis. These products were further sequenced, and multiple alignments were performed using GeneTools software. The intergenic region between Rv3870 and Rv3871 is illustrated as a multiple sequence alignment in Fig. 2. Multiple sequence analysis revealed that there was a 53 bp direct repeat element that occurred with a frequency of one to six repeats. At the 5’ end, 6 bp of the first repeat (CCGTGA) was different from that of the second to the sixth repeats (CAACGG). Single-nucleotide polymorphisms were observed in the repeat elements, and were confirmed by sequencing with forward and reverse primers.

In the case of the RD1A primers, the discriminatory power using 1.5% agarose gel electrophoresis for the PCR product of ~750–950 bp was comparatively poor and there was a
bias in determining the correct size, as two independent observers read the size of the PCR product differently. As the sequencing results showed, the polymorphism observed was only in the intergenic region of Rv3870 and Rv3871. Hence, a set of primers was designed in the flanking region that would amplify the last 57 bp of Rv3870, the 103 bp intergenic region and the first 70 bp of Rv3871. The standard amplification product for M. tuberculosis H37Rv using these primers was 230 bp and was named RD-INS. We repeated the PCR using this set of primers for the 70 test samples, which were already differentiated into four clusters by the RD1A primers. The RD-INS primers differentiated these samples into five groups, namely H1 (~180 bp), H2 (~230 bp), H3 (~280 bp), H4 (~330 bp) and H5 (~420 bp) when resolved in a 2% agarose gel (Fig. 3). This indicated that there was an ~50 bp deletion (H1) or an ~50 bp insertion (H3), ~100 bp insertion (H4) or ~200 bp insertion (H5) in the intergenic region compared with H37Rv (H2). The insertion of ~200 bp is a novel finding in this study.

We analysed a total of 407 clinical isolates from four major families namely, EAI3 (102), EAI5 (93), Beijing (77) and CAS (84), and 51 samples from the 33, T and LAM families (Table 2) with the RD1-INS primers to look for the frequency of polymorphisms. Almost two-thirds (266/407) of the total tested clinical isolates had a 53 bp insertion ($P<0.0001$), whilst 80% of the CAS and Beijing isolates had this 53 bp insertion. The clinical isolates belonging to the EA15 family had a higher percentage of the 53 bp deletion ($P<0.001$, compared with other samples), whilst only the clinical isolates with the EA13 signature had the 200 bp insertion. There was no significant difference in terms of insertions or deletions between the CAS, Beijing, 33, LAM and T1 families.

The HGDI was calculated for each spoligotype cluster and was found to be 0.48, 0.61, 0.29 and 0.26 for the EA13, EA15, CAS and Beijing families, respectively. The HGDI for the EAI family was 0.58.

### DISCUSSION

We assessed the presence of RD1 genes in south Indian isolates of M. tuberculosis and observed that there were no large sequence alterations or polymorphisms in the coding genes of the RD1 region except for single-nucleotide polymorphisms. In the present study, we observed the presence of a 53 bp tandem repeat in the intergenic region between Rv3870 and Rv3871. The number of repeats of MIRU 39 differed among M. tuberculosis isolates. The internal primer used in the study, which amplified the

### Expression of Rv3871 and Rv3874 among the different groups

The expression of Rv3871 and Rv3874 was monitored in six isolates each from the H1, H2, H3 and H4 groups and three isolates from the H5 group by qRT-PCR using gene-specific primers. RNA was extracted from these isolates, and qRT-PCR was performed using 16S rRNA as an internal control for normalization of transcriptional analysis. Expression of Rv3871 was found to be 8- and 12-fold higher in samples containing two and four copies of the repeat element, respectively (Fig. 4a), whilst Rv3874 showed very high expression with a ~40-fold increase in samples containing six copies of the repeat element compared with samples having only one copy (Fig. 4b). The results observed were statistically significant ($P<0.05$), suggesting an association between the number of repeat elements and transcription of the downstream genes.

### Table 2. Comprehensive results of strain differentiation by the RD-INS primer

<table>
<thead>
<tr>
<th>RD-INS</th>
<th>EAI3 (n=102)</th>
<th>EAI5 (n=93)</th>
<th>CAS (n=84)</th>
<th>Beijing (n=77)</th>
<th>Others (n=51)</th>
<th>Total (n=407)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H5</td>
<td>3 (2.9%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>3 (0.7%)</td>
</tr>
<tr>
<td>H4</td>
<td>11 (10.8%)</td>
<td>0 (0%)</td>
<td>3 (3.6%)</td>
<td>6 (7.8%)</td>
<td>1 (1.9%)</td>
<td>21 (5.2%)</td>
</tr>
<tr>
<td>H3</td>
<td>75 (73.5%)</td>
<td>43 (46.2%)</td>
<td>69 (82.1%)</td>
<td>66 (85.7%)</td>
<td>13 (25.6%)</td>
<td>266 (65.4%)</td>
</tr>
<tr>
<td>H2</td>
<td>13 (12.8%)</td>
<td>36 (38.7%)</td>
<td>11 (13.1%)</td>
<td>5 (6.5%)</td>
<td>34 (66.6%)</td>
<td>99 (24.3%)</td>
</tr>
<tr>
<td>H1</td>
<td>0 (0%)</td>
<td>14 (15.1%)</td>
<td>1 (1.2%)</td>
<td>0 (0%)</td>
<td>3 (5.9%)</td>
<td>18 (4.4%)</td>
</tr>
</tbody>
</table>
intergenic region along with short regions of the flanking genes, has an advantage over the other sets of primers that amplify several hundred bases of the flanking genes.

In order to characterize this repeat element, we used BLAST to analyse the sequence in TubercuList and found that the repeat corresponded to the MIRU 39 locus. The MIRU 39 coordinate in the *M. tuberculosis* H37Rv chromosome is between 4348718 and 4348823 (Supply et al., 2000). These authors reported that MIRU 39 has a 53 bp repeat element occurring at a frequency of one to four copies, but we observed this frequency to be one to six copies, which is a novel finding. The standard MIRU-VNTR typing protocol used by these authors included 12 and 15 MIRU loci, which did not include MIRU 39 (Supply et al., 2006; Phyu et al., 2009). The HGDI of the EAI isolates for MIRU 39 was 0.58, which is higher than the previous report of 0.41 (Supply et al., 2006). Thus, MIRU 39 should be included in the MIRU-VNTR typing protocol for the south Indian strains where EAI isolates correspond to 84 % of the total circulating strains (Shanmugam et al., 2011). Our findings also suggest that the MIRU polymorphism may vary among geographically diverse clinical isolates, so region-specific MIRU locus standardization will be necessary to achieve a high discriminatory index as reported previously (Comas et al., 2009). Here, we demonstrated that MIRU 39 occupying the intergenic region between Rv3870 and Rv3871 is present in variable copy numbers ranging from one to six.

It has already been reported that genome-wide insertion is prominent in the Beijing strain (McEvoy et al., 2007). We also observed that 93.5 % of the tested Beijing strains prevalent in south India retained insertions of either 53 or 100 bp compared with the standard strain. Insertion of the 53 bp repeat in the EAI5 family was similar to that of the Beijing and CAS family, whilst the EAI3 strains had fewer insertions and more deletions common among ancient strains, as reported by Gutierrez et al. (2006). Two-thirds of the tested strains retained the 53 bp insertion, stressing the functional importance of this repeat element and its effect on the flanking genes.

There are several well-documented examples of the role of VNTRs and point mutations in gene regulation of bacteria, most notably in *Haemophilus influenzae* and *Neisseria* and *Streptococcus* species, in which changes in the copy number of VNTRs results in phase variation and changes in pathogenic properties (van Belkum et al., 1998; Moxon et al., 2006). As many VNTRs occupy the entire region between two genes, it is also possible that they may contain the entire regulatory element, which will consequently be repeated in multiple but variable copies among clinical isolates (Tantivitayakul et al., 2010). Pinto Júnior et al. (2007) used PCR amplification of the intergenic region between the *plcB* and *plcC* genes to differentiate *M. tuberculosis* from *M. bovis*. They also tested these primers to detect and differentiate clinical isolates of the *M. tuberculosis* complex. Point-mutation polymorphisms in intergenic regions have been shown to affect drug susceptibility and resistance in *M. tuberculosis*. Sreevatsan et al. (1997) demonstrated a polymorphism in the 105 bp *oxyR*-ahpC intergenic region resulting in altered expression of KatG and correlated with resistance to isoniazid (Sreevatsan et al., 1997). Ramaswamy et al. (2000) illustrated a polymorphic intergenic region between the *embC* and *embA* genes involved in ethambutol resistance (Ramaswamy et al., 2000).

There is a good correlation between VNTR copy number and transcription of the genes, as cited by Akhtar et al. (2009), where there was a 12.5-fold upregulation in the expression of *IpdA* with a flanking region containing four VNTR 3690 repeats compared with a region containing just one repeat (Akhtar et al., 2009). It has also been reported that VNTR 0960c, which occupies the intergenic region between Rv0862c and *errc3*, provided promotion for expression of a downstream *gfp* reporter gene (Tantivitayakul et al., 2010). Similarly, we also observed a significant difference in the gene expression of Rv3871 and Rv3874 among the different groups containing different copy numbers of MIRU 39. As the study by Tantivitayakul et al. (2010) revealed, multiple copies might have an effect on the level of expression, albeit in an indirectly proportionate manner. Our study also found differential expression patterns of the two genes with respect to copy number. Thus, the expression level of Rv3871 was...
higher in strains containing two and four copies when compared with one copy of the repeat element, whereas Rv3874 expression levels were proportionate to the increase in the number of copies present in the intergenic region. Thus, mutations and polymorphisms in the intergenic region have been found to influence the expression of flanking genes and consequently their associated functions. From our results, we conclude that the presence of repeat elements has an effect on the transcription levels of Rv3871 and Rv3874, thereby influencing the regulation of gene expression in *M. tuberculosis*.

**ACKNOWLEDGEMENTS**

S. B. and A. K. R. are grateful to Department of Biotechnology (DBT) and International Center for Excellence in Research (ICER) and Indian Council of Medical Research (ICMR) for the Research Fellowship. We thank Ms Aarty for her help during this project. We acknowledge the support from the Department of Bacteriology of National Institute for Research in Tuberculosis (NIRT) for providing samples.

**REFERENCES**


