Enriched whole genome sequencing identified compensatory mutations in the RNA polymerase gene of rifampicin-resistant *Mycobacterium leprae* strains

Mallika Lavania 1
Itu Singh 1
Ravindra P Turankar 1
Anuj Kumar Gupta 2
Madhvi Ahuja 1
Vinay Pathak 1
Utpal Sengupta 1

1Stanley Browne Laboratory, The Leprosy Mission Trust India, TLM Community Hospital Nand Nagari, New Delhi, India
2Agilent Technologies India Pvt Ltd, Jasola District Centre, New Delhi, India

**Abstract:** Despite more than three decades of multidrug therapy (MDT), leprosy remains a major public health issue in several endemic countries, including India. The emergence of drug resistance in *Mycobacterium leprae* (M. leprae) is a cause of concern and poses a threat to the leprosy-control program, which might ultimately dampen the achievement of the elimination program of the country. Rifampicin resistance in clinical strains of M. leprae are supposed to arise from harboring bacterial strains with mutations in the 81-bp rifampicin resistance determining region (RRDR) of the *rpoB* gene. However, complete dynamics of rifampicin resistance are not explained only by this mutation in leprosy strains. To understand the role of other compensatory mutations and transmission dynamics of drug-resistant leprosy, a genome-wide sequencing of 11 M. leprae strains—comprising five rifampicin-resistant strains, five sensitive strains, and one reference strain—was done in this study. We observed the presence of compensatory mutations in two rifampicin-resistant strains in *rpoC* and *mmpL7* genes, along with *rpoB*, that may additionally be responsible for conferring resistance in those strains. Our findings support the role for compensatory mutation(s) in RNA polymerase gene(s), resulting in rifampicin resistance in relapsed leprosy patients.

**Keywords:** leprosy, rifampicin resistance, compensatory mutations, next generation sequencing, relapsed, MDT, India

**Introduction**

Since the introduction of effective multidrug therapy (MDT) in 1985, India is still inhabited by more than 60% of the leprosy patients of the world—a situation that has remained unchanged despite massive decline in the global prevalence of the disease. However, the attainment of the prevalence rate (PR) of 0.66 per 10,000 population indicated the progress of elimination.1 At this crucial stage of elimination, the emergence of drug resistance in *Mycobacterium leprae* (M. leprae) to rifampicin is being reported from different endemic populations of India.2,3 If this emergence of rifampicin resistance is not checked immediately with the administration of an alternative drug regimen, the elimination program, which relies on rifampicin-based MDT, will delay the achievement of the leprosy control program of the country. Although rifampicin-based MDT, recommended by the World Health Organization (WHO; 1982), has been successful in reducing the number of leprosy cases around the world, the global registered prevalence of leprosy at the end of 2015 was still 176,176, and the new cases reported globally in 2015 was 211,973. The number of registered cases and new cases almost remained steady for many years.4 MDT was originally designed to prevent the emergence and spread of drug-resistant strains and contained rifampicin,
dapsone, and clofazimine. Rifampicin, the backbone of MDT, is a bactericidal drug, which binds to the beta subunit of the RNA polymerase encoded by rpoB and inhibits transcription. Recent reports have, however, indicated the occurrence of rifampicin resistance in several endemic areas such as Brazil, India, and People’s Republic of China. At this present stage, it is important to monitor the emergence of rifampicin-resistant mutants in endemic populations where transmission of M. leprae infection is continuing. On the other hand, resistance to dapsone has been reported since the late 1960s. Moreover, a strain showing resistance to both dapsone and rifampicin was reported in 1993. At present, further reports indicating the emergence of M. leprae strains resistant to multiple drugs are also emerging.

As M. leprae cannot be cultured in vitro, molecular drug susceptibility testing, including PCR-sequencing, for determining resistant strains provides a practical alternative. At present, the rapid detection and control of such drug-resistant strains is essential in countries approaching leprosy elimination levels, such as India. As new cases are continuing, which is considered a major indicator for disease transmission, there is still a need for better measures to control leprosy. Furthermore, relapsed cases with drug-resistant strains could be a new source of disease transmission. Therefore, at this juncture of elimination, it will be essential to conduct a thorough investigation of relapsed multibacillary (MB) cases and all new cases to identify the pattern of existence of primary and secondary resistance to anti-leprosy drugs such as rifampicin and dapsone. Whole genome sequencing (WGS) of M. leprae directly from clinical samples is a good tool for identifying mutations that cause drug resistance, especially those that are not known or well characterized. In biopsy samples, M. leprae genome copies can vary in number, but the relative proportion of bacterial DNA is minute in comparison to host DNA. Direct sequencing of mixed human and bacterial DNA does not yield significant proportions of sequence reads that map to the bacterial genome. For this reason, we used a capture probe-based bacterial DNA enrichment method which selectively isolates bacterial DNA from host DNA prior to sequencing. With the help of this approach, the present study focuses on identifying compensatory mutations in genes other than, but related to, rpoB in rifampicin-resistant strains of M. leprae.

Materials and methods

Ethical approval

Ethical clearance for this study was approved by the The Leprosy Mission Trust India Ethical Committee, held on August 29, 2014, under the regulations of the Indian Council of Medical Research. Written informed consent was obtained from all subjects before collection of biological samples.

Source of biopsies for DNA extraction

A total of ten relapsed leprosy patients from hospitals of The Leprosy Mission Trust India (TLMTI) and one reference control strain (Br4923) from BEI Resources, USA, were included in this study. Among these ten strains, five were resistant to rifampicin, characterized by molecular methods and PCR-sequencing. The remaining five strains were drug sensitive. M. leprae DNA was extracted from these ten clinical biopsy specimens using the Blood and Tissue DNA extraction kit (QIAGEN, USA). Single-nucleotide polymorphism (SNP) subtyping data was generated by PCR sequencing from DNA before carrying out WGS for the remaining DNA.

Library preparation and sequencing

Target capture sequencing libraries were prepared with Illumina-compatible SureSelect QXT Library Prep Reagent Kit (Agilent Technologies, Santa Clara, CA, USA) at Genotypic Technology Pvt. Ltd., Bangalore, India. Briefly, 50 ng Qubit-quantified genomic DNA was fragmented and adapter-tagged using SureSelect QXT Enzyme (Table S1). The fragmented and adapter-tagged DNA was purified with HighPrep magnetic beads and then amplified by eight cycles of PCR. The PCR-enriched products were purified with HighPrep beads, followed by library quality control check using the Agilent 2100 Bioanalyzer.

Target enrichment was undertaken according to the manufacturer’s instructions using SureSelect RNA capture baits. In-solution hybridization was done as per the SureSelect QXT hybridization method. After hybridization, the captured biotinylated probe-target hybrids were pulled down by using streptavidin-coated magnetic beads (Dynabeads MyOne Streptavidin T1, ThermoFisher Scientific, Waltham, MA, USA). The magnetic beads were washed according to the manufacturer’s instructions and resuspended in 15 μl nuclease-free water. The captured DNA libraries were amplified by 12 cycles of PCR with the inclusion of the appropriate indexing primer for each sample.

The final PCR product (sequencing library) was purified with HighPrep beads, followed by quantification by a Qubit fluorometer (ThermoFisher Scientific, MA, USA), and fragment size distribution was analyzed on an Agilent 2100 Bioanalyzer.

Sequencing was carried out on an Illumina Hi-Seq 2500 instrument and the Illumina paired-end raw reads were quality checked using FastQC. Illumina raw reads were processed for adapters and trimming of low-quality bases. The processed...
reads were aligned with *M. leprae* TN reference using a Burrows–Wheeler Aligner (BWA) 0.7.5 algorithm, and the variants were identified with Samtools 1.2 and Bcftools 1.2 filtered for a read depth threshold of greater than 20, and a quality threshold of more than Q30. Filtered variants were further annotated with SnpEff to get the information on the genes, protein change, and the impact of the variation.

**Results**

**Demographics and diagnosis**

Among the five rifampicin-resistant patients, one patient was a defaulter who took MDT for 2 months only and remaining four patients were those who relapsed after completion of a full course of MDT. One of the patients (1027) had a dual infection of leprosy and tuberculosis. The demographic and clinical details of these five resistant patients together with the five sensitive ones are mentioned in Table 1A and B.

All resistant strains were characterized for drug-susceptibility testing by PCR sequencing. We found the already published *rpoB* mutations in three patients (507, 42, and 80) at codon positions Leu436Meth, Asp441Tyr, and Ser456Leu, except in two (1027 and 439) who had mutations at codon positions Gln442His and Ser437Pro (Table 2). These two patients did not respond to the treatment. Gln442His mutation in patient 1027 and Ser437Pro in *rpoB* gene did not confer resistance to rifampicin, which was confirmed by mouse foot pad growth earlier. The patient 1027 continuously visited the hospital with recurrent Type 2 reaction and other complications despite being administered thalidomide. He was later diagnosed with a pulmonary tuberculosis infection along with leprosy. He was started on anti-tuberculosis treatment (ATT) as well, but the infection was unalleviated. We tested the same specimen with *Mycobacterium tuberculosis*-specific *rpoB* primers, undertook DNA sequencing, and

**Table 1** Demographic and clinical details of rifampicin-resistant cases (A). Demographic and clinical details of rifampicin-sensitive cases (B).

<table>
<thead>
<tr>
<th>Patients’ identification</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis (years)</td>
<td>23</td>
<td>38</td>
</tr>
<tr>
<td>Sex</td>
<td>Male</td>
<td>Male</td>
</tr>
<tr>
<td>Completion of MDT (month/year)</td>
<td>10/2011</td>
<td>10/2009</td>
</tr>
<tr>
<td>Relapse date</td>
<td>08/2016</td>
<td>12/2015</td>
</tr>
<tr>
<td>RJ Classification</td>
<td>At diagnosis</td>
<td>BL</td>
</tr>
<tr>
<td></td>
<td>At relapse</td>
<td>BB</td>
</tr>
<tr>
<td>Bacillary index</td>
<td>First diagnosis</td>
<td>0.66+</td>
</tr>
<tr>
<td></td>
<td>Relapse</td>
<td>2.0+</td>
</tr>
<tr>
<td>Clinical complication at the time of relapse</td>
<td>None</td>
<td>Type 1 reaction</td>
</tr>
</tbody>
</table>

Abbreviations: BB, borderline leprosy; BL, borderline lepromatous; BT, borderline tuberculoid; LL, lepromatous leprosy; MDT, multidrug therapy; NA, not available; RJ, Ridley–Jopling.
ferring mutations during inadequate therapy (secondary resis-
strains (primary resistance) or by selection of resistance-con-
infection can be caused by transmission of already resistant
setback to the public health program. Drug-resistant leprosy
during the phase of elimination may cause a major
occurrence of dapsone resistance also took nearly the same
after
The emergence of rifampicin resistance in
Discussion
mutation in the
rpoB
rifampicin-resistant strains, we found that two strains showed
its associated genes. After analyzing the data of the five
SNPs and 958 (8.0%) Indels. As there is no previous data-
11,791 variants were identified, including 10,833 (92 %)
isolates, and the reference strain for
M. leprae
from the 10
Whole genome read coverage was adequately generated
genetic polymorphisms
play a role in conferring resistance in such cases.

found him susceptible for the TB
rpoB
gene. This led us to
consider the possibility that other mutations at other genes
play a role in conferring resistance in such cases.

Genetic polymorphisms
Whole genome read coverage was adequately generated from
the 10
M. leprae
isolates, and the reference strain for
comparative genomic analysis. Among these ten samples, 11,791 variations were identified, including 10,833 (92 %)
SNPs and 958 (8.0%) Indels. As there is no previous database to filter out common variants, we focused on
rpoB
and its associated genes. After analyzing the data of the five rifampicin-resistant strains, we found that two strains showed
mutation in the
rpoC
gene (Patient ID 1027 and 439) along with the
rpoB
gene (Table 2) as well as a non-synonymous mutation in the resistance nodulation division (RND) family transporter efflux pump gene
mmpL
7
. These two genes are mainly responsible for causing resistance among the rest of the SNPs present in these strains.

Analysis of rifampicin-resistant strains revealed 1,308 variations, of which there were 44 synonymous, 101 non-
synonymous, 112 intergenic, and 12 frame-shift mutations. On genome comparison of all ten strains, we observed seven unique SNPs. While comparing resistant strains with other sensitive strains, it revealed 16 unique SNPs that were present only in these strains (Table 3). An insertion of eight nucleo-
tides (TTTCTTAT at position 508802) was found in ML0411, encoding the Pro-Pro-Glu (PPE) serine-rich antigen.

Table 2 Mutations in
rpoB
and
rpoC
genes by WGS in molecularly characterized rifampicin-resistant strains

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Resistance</th>
<th>Geographical area</th>
<th>rpoB mutations (WGS)*</th>
<th>rpoC mutations (WGS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1027</td>
<td>RIF</td>
<td>Delhi</td>
<td>Arg1168Leu; Arg505Trp</td>
<td>Val517Leu</td>
</tr>
<tr>
<td>507</td>
<td>RIF</td>
<td>Chhattisgarh</td>
<td>Iso433Ser; Leu436Met</td>
<td>NM</td>
</tr>
<tr>
<td>42</td>
<td>RIF</td>
<td>Delhi</td>
<td>Leu458Phe</td>
<td>NM</td>
</tr>
<tr>
<td>80</td>
<td>RIF</td>
<td>Delhi</td>
<td>Tyr-Cys</td>
<td>NM</td>
</tr>
<tr>
<td>439</td>
<td>RIF</td>
<td>Chhattisgarh</td>
<td>Tyr-Cys</td>
<td>Leu807Met; Leu244Phe</td>
</tr>
</tbody>
</table>

Note: *Amino acid.
Abbreviations: RIF, rifampicin; NM, no mutation; WGS, whole genome sequencing.

Discussion
The emergence of rifampicin resistance in
M. leprae
after more than 30 years of MDT is not an unexpected outcome. The occurrence of dapsone resistance also took nearly the same
duration. However, the emergence of rifampicin-resistant
M. leprae
during the phase of elimination may cause a major
setback to the public health program. Drug-resistant leprosy
infection can be caused by transmission of already resistant
strains (primary resistance) or by selection of resistance-con-
ferring mutations during inadequate therapy (secondary resis-
tance). Our own work has shown the presence of rifampicin,
dapsone, and ofloxacin resistance cases in highly endemic
areas in several states of India. 
M. leprae
isolates resistant to
single and multiple drugs have already been encountered.
Although the genome of
M. leprae
has undergone massive
gene decay, the rifampicin resistance determining region
(RRDR) region of the organism and the associated genes are
still functional. In order to monitor the transmission dynamics of
drug-resistant leprosy, a comparison of
M. leprae
strains
was carried out by WGS of strains derived from rifampicin-
resistant leprosy patients in this study.

rifampicin inhibits bacterial RNA synthesis by binding strongly to RNA polymerase, preventing attachment of the enzyme to DNA and thus blocking the initiation of transcription. Bacterial resistance to rifampicin is caused by mutations leading to a change in the structure of the beta subunit of RNA polymerase. Resistance to rifampicin is one of the main causes of treatment failure. Various studies have shown that resistance to rifampicin in
M. tuberculosis
as well as in
M. leprae
is mainly due to a single mutation in an
81-bp region in the
rpoB
gene. Furthermore, it has been reported earlier that there are some alternative mechanisms of rifampicin resistance in various bacterial species such as
rox
gene-mediated mono-oxygenation of rifampicin and duplication of the
rpoB
gene as
rpoB2
in
Nocardia
. In our study, all strains belonged to the SNP subtype 1D, which is the predominant genotype prevalent in many countries such as India, Bangladesh, Nepal, Madagascar, Malawi, and the French West Indies.

The coexistence of the two mycobacteria has been demonstrated in archaeological samples by Donoghue et al., who identified DNA from both tuberculosis and leprosy in the archeological samples from the Roman period to the 13th century from several sites around the world. On review of data from three leprosy referral centers in Hyderabad, India, from 2000 to 2013, Rawson et al. identified three cases of concomitant disease. Similarly, we observed in one of our samples a patient (Pt1027) who is suffering from both TB and leprosy.
Compensatory mutations identified in rifampicin-resistant M. leprae

This study provides a better explanation for the presence of the rifampicin-resistance phenotype in leprosy cases in India. It has identified unique SNPs, including non-synonymous SNPs in gltB, gltD, and esxB and mutation in ctpC and mmpL7 transporter genes. The mmpL7 gene in M. tuberculosis was studied and characterized by Pasca et al. It encodes a protein which belongs to RND drug transporter family. The MmpL7 protein is of 920 amino acids with a predicted molecular mass of 95.1 kDa, contains 12 transmembrane domains (TMDs) having two large hydrophilic extracytoplasmic domains between TMD-1 and TMD-2 and between TMD-7 and TMD-8. All of these characteristics have been described as typical of RND efflux pumps. The MmpL7 protein confers a high level of resistance to isoniazid. However, a functional assay is required to determine whether either of these variant genes/transporters confers any degree of rifampicin resistance in M. leprae. In our study, the rpoC gene has a non-synonymous SNPs sequence in two patients – 1027 and 439. Mouse foot pad assay was done for the patient 1027 to check whether the mutation in 442 codon position was conferring resistance. In our previous study, we observed that this 442 codon mutation was not conferring resistance in mice. Therefore, it may be suggested that these compensatory mutations in rpoC and/or transporter genes play a role in conferring resistance in this strain. Thus, our study suggested the role of compensatory mutation in rpoC and possibly of drug transporter genes in contributing to acquired resistance in rifampicin-resistant strains in M. leprae.

Acknowledgments

The authors thank the Indian Council of Medical Research (Grant No. ECD/Ad-hoc/leprosy/2014-113/Fy. 14-15/19/ Delhi/NGO-ECD-I) and England and Wales Foundation (205T03) for the financial support. The authors are likewise grateful to Atul Roy and Manish Gardia for assisting us in the sample collection, and thank the superintendent and staff of TLM Hospitals for their help and assistance during the work.

Table 3 List of unique 16 SNPs present in resistant strains

<table>
<thead>
<tr>
<th>Position</th>
<th>Name of the gene</th>
<th>TN base (reference)</th>
<th>Variant</th>
<th>SNP effect</th>
<th>Amino acid change</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>80206</td>
<td>gltB</td>
<td>C</td>
<td>T</td>
<td>Non-synonymous coding</td>
<td>T/M</td>
<td>Probable ferredoxin-dependent glutamate synthase</td>
</tr>
<tr>
<td>83120</td>
<td>gltD</td>
<td>G</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>G/E</td>
<td>Probable NADH-dependent glutamate synthase</td>
</tr>
<tr>
<td>61866</td>
<td>esxB</td>
<td>C</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>G/C</td>
<td>Possible 10 KDa culture filtrate antigen homolog EsxB (lp) (cfp10)</td>
</tr>
<tr>
<td>257506</td>
<td>MLO192</td>
<td>C</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>Q/H</td>
<td>Conserved membrane protein</td>
</tr>
<tr>
<td>626352</td>
<td>MLO514</td>
<td>C</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>P/Q</td>
<td>Probable conserved membrane protein</td>
</tr>
<tr>
<td>666171</td>
<td>priA</td>
<td>C</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>H/N</td>
<td>Putative primosomal protein N’ PriA (Replication factor Y)</td>
</tr>
<tr>
<td>672897</td>
<td>ribG</td>
<td>C</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>L/M</td>
<td>Probable bifunctional riboflavin-specific deaminase/reductase</td>
</tr>
<tr>
<td>140830</td>
<td>embA</td>
<td>G</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>P/L</td>
<td>Probable integral membrane indolylacetylinositol arabinosyltransferase EmbA (arabinosylindolylacetylinositol synthase)</td>
</tr>
<tr>
<td>145310</td>
<td>embC</td>
<td>G</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>T/I</td>
<td>Probable integral membrane indolylacetylinositol arabinosyltransferase EmbC (arabinosylindolylacetylinositol synthase)</td>
</tr>
<tr>
<td>2271147</td>
<td>rpoC</td>
<td>C</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>V/L</td>
<td>DNA-directed RNA polymerase (beta’ chain)</td>
</tr>
<tr>
<td>190301</td>
<td>mmpL7</td>
<td>G</td>
<td>A</td>
<td>Non-synonymous coding</td>
<td>S/L</td>
<td>Conserved transmembrane transport protein MmpL7</td>
</tr>
<tr>
<td>2600317</td>
<td>M2186</td>
<td>+CAG</td>
<td>Codon change plus codon insertion</td>
<td>R/RR</td>
<td>Possible transcriptional regulatory protein</td>
<td></td>
</tr>
<tr>
<td>241842</td>
<td>MLO175</td>
<td>G</td>
<td>T</td>
<td>Non-synonymous coding</td>
<td>R/L</td>
<td>Probable two component sensor kinase MprB</td>
</tr>
<tr>
<td>148945</td>
<td>MLO107</td>
<td>C</td>
<td>Synonymous coding</td>
<td>P/P</td>
<td>Probable conserved transmembrane protein</td>
<td></td>
</tr>
<tr>
<td>3237399</td>
<td>panA</td>
<td>-A</td>
<td>Frame shift</td>
<td>-T</td>
<td>Probable alanyl-tRNA synthetase AlaS (Alanine–tRNA ligase) and AlaRS (Alanine translatase)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: SNP, single-nucleotide polymorphism.
Disclosures
The authors report no conflicts of interest in this work.

References
### Supplementary material

**Table S1 Concentration of DNA from biopsies and after library preparation**

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Sample ID</th>
<th>DNA concentration by Qubit (ng/μL)</th>
<th>Library concentration (ng/μL)</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>570</td>
<td>56.7</td>
<td>26.7</td>
<td>13</td>
</tr>
<tr>
<td>2.</td>
<td>507</td>
<td>93.2</td>
<td>0.62</td>
<td>13</td>
</tr>
<tr>
<td>3.</td>
<td>42</td>
<td>85.6</td>
<td>0.74</td>
<td>13</td>
</tr>
<tr>
<td>4.</td>
<td>80</td>
<td>35.6</td>
<td>0.38</td>
<td>13</td>
</tr>
<tr>
<td>5.</td>
<td>1027</td>
<td>204</td>
<td>0.54</td>
<td>13</td>
</tr>
<tr>
<td>6.</td>
<td>281</td>
<td>33.1</td>
<td>0.36</td>
<td>13</td>
</tr>
<tr>
<td>7.</td>
<td>439</td>
<td>25.6</td>
<td>0.41</td>
<td>13</td>
</tr>
<tr>
<td>8.</td>
<td>4923</td>
<td>15.3</td>
<td>78.5</td>
<td>13</td>
</tr>
<tr>
<td>9.</td>
<td>395</td>
<td>20.5</td>
<td>0.32</td>
<td>13</td>
</tr>
<tr>
<td>10.</td>
<td>834</td>
<td>432</td>
<td>0.37</td>
<td>13</td>
</tr>
<tr>
<td>11.</td>
<td>588</td>
<td>235</td>
<td>0.41</td>
<td>13</td>
</tr>
<tr>
<td>12.</td>
<td>SM</td>
<td>45.4</td>
<td>0.33</td>
<td>13</td>
</tr>
<tr>
<td>13.</td>
<td>728</td>
<td>24.9</td>
<td>0.57</td>
<td>13</td>
</tr>
<tr>
<td>14.</td>
<td>1335</td>
<td>25.4</td>
<td>2.85</td>
<td>13</td>
</tr>
<tr>
<td>15.</td>
<td>730</td>
<td>33.2</td>
<td>0.34</td>
<td>13</td>
</tr>
<tr>
<td>16.</td>
<td>BH</td>
<td>64.2</td>
<td>1.33</td>
<td>13</td>
</tr>
</tbody>
</table>