Original articles

Genetic diversity of *Mycobacterium tuberculosis* in Mbarara, South Western Uganda

*Bazira J*, Matte M, Asiimwe BB, Joloba LM

1. Department of Microbiology, Faculty of Medicine, Mbarara University of Science and Technology
2. Department of Microbiology & Immunology, Faculty of Medicine, Muhimbili University of Health and Allied Sciences
3. Department of Medical Microbiology, College of Health Sciences, Makerere University

Abstract

**Background:** We determined the genetic diversity of mycobacteria isolated from tuberculosis patients in Mbarara Uganda, using region of difference (RD) analysis and spacer oligonucleotide typing (spoligotyping).

**Methods:** Sputum samples were cultured on Lowenstein Jensen media. The isolates were characterized using RD analysis and spoligotyping.

**Results:** The majority (92.8%) of the patients were new cases, 60% were males and 44% were HIV positive with a mean age of 33.7 years. All the 125 isolates were identified as *M.tuberculosis* sensu stricto. Most (92.8%) of the isolates were modern strains. Spoligotyping revealed 79 spoligotype patterns, with an overall diversity of 63.2%. Sixty (48%) isolates formed 16 clusters each consisting of 2-15 isolates. Most (59.2%) of the isolates were Uganda genotype strains. The major shared spoligotypes in our sample were SIT 135 (T2-Uganda) with 12 isolates and SIT 128 (T2) with 5 isolates. Sixty nine (87%) patterns had not yet been defined in the SpolDB4.0 database.

**Conclusion:** The TB epidemic in Mbarara is caused mainly by modern *M.tuberculosis* strains of the Uganda genotype. The wide diversity of strains may indicate that the majority of the TB cases are reactivation rather than re-infection. However this needs to be ascertained with more discriminative finger printing techniques.

**Key words:** Mycobacteria, spoligotypes, Region of difference

Introduction

In spite of the wide availability of cost effective interventions for its control, tuberculosis (TB) is still a major global health problem and the leading cause of death from a curable infectious disease. Uganda ranks 16th among the world’s 22 countries with the highest tuberculosis burden in the world. The country had more than 40,000 TB cases in 2007, with an estimated incidence rate of 330 per 100,000 people and a directly observed therapy short course (DOTS) treatment success rate of 70 percent in 2007, with greater Mbarara contributing about 26% of all TB cases.

An increasing amount of evidence indicates that *M. tuberculosis*’ ability to spread varies from strain to strain and that different strains have different geographical and/or host specificities. Since the discovery of DNA polymorphisms in *M. tuberculosis*, molecular typing of strains has become an invaluable tool for the study of epidemiology of TB. Some of the applications include; predicting transmission rates and identifying dominant strains associated with outbreaks, severe disease and drug resistance.

Genotyping of *M. tuberculosis* using IS6110 restriction fragment length polymorphism has for a long time been the gold standard. However this technique is labor-intensive and requires culturing of the slow-growing mycobacteria and the results are sometimes difficult to compare among laboratories. More recently, spoligotyping based on the variability in the direct repeat locus of *M. tuberculosis*, is useful for tracking TB epidemics, detecting new outbreaks, and better defining high-risk populations to focus prevention strategies. Important advantages of spoligotyping are that it is cheap, easy to perform and fast. In addition, it has been demonstrated that the results are highly reproducible.

*Corresponding author:*
Joel Bazira
Department of Microbiology
Faculty of Medicine
Mbarara University of Science and Technology
Mbarara, Uganda
Email: jap2k2001@yahoo.co.uk
advantages of spoligotyping are that it is cheap, easy to perform and fast. In addition, it has been demonstrated that the results are highly reproducible\(^9\). Comparative-genomics approaches greatly enhanced our understanding of the mechanisms of insertion and deletion of DNA and the resulting distribution of variable regions around the genomes of tubercle bacilli\(^{10,12}\). There are 20 variable regions, of which 14 regions of difference (RD1 to RD14) were found to be absent from Bacillus Calmette-Guérin (BCG) Pasteur relative to \(M.\) \(tuberculosis\) H37Rv\(^{12-14}\). Six regions, H37Rv-related deletions (RvD1 to RvD5 and \(M.\) \(tuberculosis\) specific deletion 1 (TbD1), are absent from the \(M.\) \(tuberculosis\) H37Rv genome relative to other members of the \(M.\) \(tuberculosis\) complex. Based on the presence or absence of the TbD1 region, \(M.\) \(tuberculosis\) strains can be divided into “ancestral” and “modern” types. The Beijing, Haarlem, and African strains responsible for major epidemics are modern types\(^9,10\).

The Beijing and Beijing-like strains are most prevalent in the Far East-Asia. In Europe, the Harlem lineage represents about 25% of the isolates. In South America, about 50% of the strains belong to the LAM family. Three major genotypic families (Haarlem, LAM, and T) are the most frequent in Africa, Central America, Europe and South America. Outside Europe, The Haarlem strains were mainly found in Central America and Caribbean. The “ill-defined genetic family, was found in all continents. The East-African-Indian (EAI) family is also highly prevalent in the Far-East-Asia, in the Middle East and Central Asia. The East-African-Indian (EAI) family is more prevalent in South-East Asia, particularly in the Philippines\(^{16}\), in Myanmar and Malaysia\(^{17}\), in Vietnam and Thailand\(^{18}\) The CAS1-Delhi family is essentially localized in the Middle-East and Central Asia, more specifically in South-Asia, and preferentially in India\(^{19,20}\), Iran, and Pakistan\(^{21,22}\). Lastly, the X family is highly prevalent in North America and Central American regions\(^{23}\).

The aim of this study was to evaluate the genotypic diversity of the \(Mycobacterium\) \(tuberculosis\) complex (MTC) strains using Region of Difference analysis and spoligotyping in this high HIV prevalence community.

**Methods**

**Study setting**

The study was conducted between May 2007 and April 2008 in the four districts of greater Mbarara namely Ibanda, Isingiro, Kiruhura and Mbarara. This region has been heavily affected by the TB/HIV epidemic. In 2005, the case notification for Mbarara was 175/100,000 people compared to the 147/100,000 people for Uganda\(^{24}\). In 2008 the TB/HIV co-infection rate for Mbarara was 65%\(^{25}\).

**Study design**

This was a cross sectional study which was conducted between May 2007 and April 2008 and enrolled consecutively all smear-positive newly diagnosed and retreatment TB patients aged ≥ 18 years. Three consecutive sputum samples (spot, early morning and spot) were taken from each patient according to the Uganda National TB and Leprosy guidelines. Only the sample with the highest ZN smear grade was further analyzed for each patient. Samples were stored at 4°C at the recruitment clinics, in any case for not more than 48 hours, until transported in a cold box to the NTRL in Kampala for processing and culture.

**Sputum sample processing**

Specimens (2.5–10 ml) were processed by the standard N-acetyl L-cystein (NALC)-NaOH method\(^{26}\) and concentrated at 4000 × g for 15 minutes. The sediment was reconstituted to 2.5 ml with phosphate buffer pH 6.8, to make the inoculum for the smears and cultures.

**Culture and identification**

Two Lowenstein-Jensen slants, one containing 0.75% glycerol and the other containing 0.6% pyruvate, were inoculated with the sediment and incubated at 37°C and examined weekly for growth. Cultures were considered negative when no colonies were seen after 8 weeks incubation.

**DNA Extraction**

Isolates were harvested and DNA extracted using a standardized protocol\(^{27}\).

**RD analysis**

RD analysis for speciation of the isolates was done at the department of Medical Microbiology, Makerere University College of Health Sciences as previously described\(^{28}\).

**Spoligotyping**

Standard spoligotyping was performed as previously described\(^{29}\) using a commercially available kit (Isogen Bioscience BV, Maarssen, The Netherlands).
Data analysis
Socio-demographic data was entered into the computer using Microsoft Excel 2000 software and was then exported to SPSS version 10 for analysis. All genotyping data were digitized and analyzed with the BioNumerics software, version 5.0 (Applied Maths, Kortrijk, Belgium). Genotype identification was carried out according to SpolDB4 and by the freely assessable MIRU-VNTRplus web database. The latter comprises genotyping data of a reference strain collection of validly described mycobacterium tuberculosis complex (MTBC) genotypes. Labels for major phylogenetic clades were assigned according to signatures provided in SpolDB4. Univariate analysis of categorical variables utilized percentages while means and standard deviations were calculated for continuous variables. The median and range was presented whenever continuous data was skewed.

Ethical considerations
This study received ethical clearance from the faculty research and ethics committee of the faculty of medicine of Mbarara University of Science and Technology, the institutional review board of Mbarara University of Science and Technology and the National Council for Science and Technology. Patients were identified and managed according to Uganda NTLP guidelines. Informed consent to participate in the study as well as permission to use isolates from samples provided were obtained from all enrolled participants.

Results
We enrolled a total of 167 sputum positive tuberculosis patients who were presenting for TB diagnosis at the various TB clinics in the greater Mbarara during the study. Most (92.8%) of the patients were newly diagnosed and nine (7.2%) had a history of previous treatment. Their mean age was 33.7 years and seventy five (60%) of them were males. Fifty five of the 125 patients were HIV sero-positive, twenty six were HIV sero-negative and forty four had unknown sero-status because they did not consent to testing.

Of the 167 samples that were cultured 140 (84%) grew, 14 (8%) had no growth and 13 (8%) were contaminated. Of the 140 cultures 15 (10.7%) were no available for DNA extraction. All the 125 isolates available for RD analysis were identified as M. tuberculosis sensu stricto. Seven isolates were found to have all the RD loci conserved including TbD1 and were therefore ancestral strains (Table 1).

Table 1: Species identification and PCR results of the M. tuberculosis complex isolates

<table>
<thead>
<tr>
<th>Targeted PCR locus</th>
<th>M. tuberculosis complex</th>
<th>Ancestral M. tuberculosis*</th>
<th>Modern M. tuberculosis‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>RD4</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>RD9</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>RD12</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>TbD1</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

*Defined as possessing the TbD1 locus intact
‡Defined as possessing the TbD1 deletion
+ = Locus intact
- = Locus deleted

The isolates gave 79 different spoligopatterns, with an overall diversity of 63.2%. A total of 62 (49.6%) isolates, were grouped into 16 clusters consisting of 2-15 isolates each, while the remaining 63 (51.4%) of the strains did not cluster.

The family assignment showed that 59.2% of the isolates belonged to the Uganda family, 7.2% to the CAS_DELHI Family, 6.4% to the Latin American Mediterranean (LAM) family, and 5.6% to the East-African Indian (EAI) family, 4% belonged to the Cameroon family, 2.4% to the Ghana family and 13.6% were not assigned to any family. Among the 16 clusters, three clusters included five or more isolates each and were defined as major spoligotypes, while minor spoligotypes, on the other hand, were defined as spoligotype international types (SITs) that contained less than five isolates per cluster (Figure 1).
Figure 1: Spoligotype pattern of clustered *M. tuberculosis* strains in the study.

*As identified in SpolDB4.0; SIT, spoligotype international type; N, number of isolates (as a percentage of total *M. tuberculosis* strains in the study); filled boxes represent positive hybridization while empty boxes represent absence of spacers; label defining the lineage/sub lineage; ND, not yet determined in SpolDB4.0.*

<table>
<thead>
<tr>
<th>N</th>
<th>Spoligopatterns</th>
<th>SIT</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparison of the spoligotypes in our study with SpolDB4.0 allowed differentiation between ubiquitous types (SIT 26-CAS_DELHI, SIT 52-T2, SIT 53-T1, SIT 54-MANU2, SIT 59-LAM11_ZWE, SIT356-CAS, SIT 358-T1, SIT 420-T2, SIT 815-LAM11_ZWE, SIT 1090 – EA15 and PubMed IT 1572-T2) and those believed to be endemic in Uganda (SIT 128-T2 and SIT 135 – T2).

Spoligotypes that did not match any existing pattern in the SpolDB4 database were defined as orphans. Of the 79 patterns observed in this study, 69 were true orphans having no counterpart in the database.

Majority (59.2%) of the isolates lacked hybridization to either spacer 40 alone or to both 40 and 43. These Uganda genotype strains formed 8 clusters (ranging 2 to 15 isolates). Nine strains were Uganda genotype I while seventy seven were Uganda genotype II (Data not shown). Forty eight Uganda genotype strains did not cluster.

Discussion

To have a better knowledge of moving and expanding clones of *M. tuberculosis* within the Mbarara, we have used RD analysis and spoligotyping to characterize mycobacteria isolated from newly diagnosed and previously treated sputum smear positive TB patients in Mbarara, South Western Uganda, a country in which both HIV infection and TB are endemic.

Deletion analysis using 16s, RD4, RD 9, RD12 and TbD1 revealed that all the strains investigated were *M. tuberculosis*. Seven of 125 *M. tuberculosis* strains tested had the TbD1 region intact, indicating these were ancestral strains. Most of the ancestral strains had the following spoligotype signatures: absence of spacers 29 to 32, presence of spacer 33, and absence of spacer 44.

The diversity of the *M. tuberculosis* found in the present study (63.2%) is high compared to 17.8% described from a study in central Uganda33, which may indicate geographical differences in the epidemiology of TB (reactivation versus reinfections) in Uganda. This is not easily determined using spoligotyping alone and would therefore require more discriminative techniques to confirm this. Similar studies in Tanzania have shown genetic diversity that ranges from 36%- 52%34,35. The effect of this diversity on transmission is not clear and need to be studied further.

In this study 69 (87%) of the spoligopatterns could not be typed on the basis of the existing SpolBD4 database. These strains warrant further characterization to better understand them and to add our knowledge to the molecular epidemiology of TB in this area of Uganda. Majority (59.2%) of the strains in our sample were of the Uganda genotype, a finding which is in keeping those of earlier studies in Uganda36-39 as opposed to the surrounding East African countries40.

For example a study in Kenya40 found only eight (11%) of 73 isolates to be of the Uganda family while in northern Tanzania34 only four (3%) of 130 strains were T2-Uganda. Collectively these findings are in keeping with those of other studies showing a tendency for local genotypes to form a greater proportion of the circulating strains. For example in Guinea Bissau, 51% of the isolates belonged to the Guinea Bissau family41, Cameroon where 46.7% of the isolates belonged to the Cameroon family, LAM10-CAM42, Harare, Zimbabwe where 31.8%-47.2% of the isolates were LAM-ZWE variants43,44.
and in Zambia where 65% of the isolates were also of the LAM-ZWE family. It is thus tempting to speculate that local strains are more likely to transmit in a given local setting compared to others.

Conclusions
We have shown that the majority of the Mycobacteria in Mbarara are modern *M. tuberculosis* with a wide diversity of spoligotypes and a predominance of the Uganda family. Fingerprinting of these isolates with more discriminative techniques will give a better insight of the molecular epidemiology of TB in Mbarara.

Acknowledgements
This study received funding from DAAD and WHO/Tropical disease research. We thank Mr. Charles Ndaula of the TB laboratory at Mbarara University and the staff at National reference tuberculosis laboratory at Wandegeya and the molecular biology laboratory staff at Makerere University Medical School for technical assistance.

References


