**DISPATCHES**

**Predominant Tuberculosis Spoligotypes, Delhi, India**

Urvashi Balbir Singh,* Naga Suresh,* N.Vijaya Bhanu,* Jyoti Arora,* Hema Pant,* Sanjeev Sinha,* Ram Charan Aggarwal,* Sushma Singh,* Jitendra Nath Pande,* Christophe Sola,† Nalin Rastogi,† and Pradeep Seth*

One hundred five *Mycobacterium tuberculosis* clinical isolates from the Delhi area were typed by spoligotyping; 45 patterns were identified. Comparison with an international spoligotype database showed type 26, Delhi type (22%), type 54 (12%), and type 1, Beijing type (8%), as the most common. Eighteen spoligotypes did not match any existing database pattern.

India accounts for 30% of tuberculosis (TB) cases worldwide. Each year, this disease develops in approximately 2 million people in India. Approximately 0.5 million people die, a figure likely to increase with the emergence of multidrug-resistant tuberculosis (MDR-TB) and the HIV epidemic (1). Delhi alone has an annual risk for infection of 2.4% (state TB officer, Delhi, pers. comm.). Traditional methods for tracing transmission of TB are imprecise and ineffective in controlling the disease. The impact of control programs is often difficult to assess for infection of 2.4% (state TB officer, Delhi, pers. comm.). Traditional methods for tracing transmission of TB are imprecise and ineffective in controlling the disease. The impact of control programs is often difficult to assess.

The aim of our study was to identify predominant spoligotypes with an international designation responsible for transmission and prevalence of TB in Delhi. The spoligotypes obtained were compared with an international spoligodatabase, spolDB3.0 (5,6).

The Study

The study included patients with culture-confirmed TB whose cases were reported to a district TB center, a primary health center, and an outpatient department, All India Institute of Medical Sciences (AIIMS), Delhi, the coordinating center. AIIMS is the top tertiary-care hospital with referrals from the entire city of Delhi, as well as from other states, with a daily outpatient attendance of 4,000 to 5,000. The district TB and primary health centers serve a well-defined population, including all TB patients in their areas.

Patients with a diagnosis of new, smear-positive pulmonary TB or those with a high suspicion for TB on clinical or radiologic grounds were included in the study. In all, 1,500 patients who met inclusion and exclusion criteria were recruited from the Delhi area over 2 years. Demographic data were collected on the patient’s sex, age, present address, employment, economic status, literacy, living conditions, household contacts, chest radiologic findings, and HIV infection. Three sputum specimens were collected in the early morning on consecutive days and transported to the AIIMS TB laboratory. Sputum specimens were processed by Petroff’s method and plated on duplicate Lowenstein-Jensen slants. A smear was examined after Ziehl-Neelsen staining. Species confirmation was followed by tests for drug susceptibility (proportion method) for rifampicin (RIF, 40.0 µg/mL), ethambutol (ETH, 2.0 µg/mL), streptomycin (STR, 4.0 µg/mL), and isoniazid (INH, 0.2 µg/mL) (7). The isolates were processed on the respective days of collection with the date and other details recorded. During the initial 3 months, January 2001–March 2001, 105 (55%) of 190 patients enrolled were culture-positive. The 105 individual *M. tuberculosis* isolates from different patients (34 from the district center, 12 from the primary health center, and the rest from AIIMS), both smear-negative and smear-positive samples, were spoligotyped. Work on the isolates collected later in the study is ongoing.

Spoligotyping to detect 43 known spacers in the direct repeat locus was performed with a commercially available kit, according to the instructions supplied by the manufacturer (Isogen Bioscience B.V., Maarsen, the Netherlands). To avoid any possibility of artificial hybridization spots on the commercial membranes, appropriate controls included DNA from *M. bovis* and *M. tuberculosis* H37Rv and autoclaved purified water for adequate number of negative controls in each experiment. Reproducibility of spoligotyping was confirmed by repeating the test with the DNA extracted again from a few isolates (data not shown). None of the negative controls demonstrated carryover DNA.

Results were doublechecked visually by an experienced operator to eliminate any systematic artifact caused by using commercial membranes. The results obtained were entered in a binary format as Excel spreadsheets (Microsoft, Redmond, WA) and compared to the

*All India Institute of Medical Sciences, New Delhi, India; and †Institut Pasteur de Guadeloupe, Pointe-à-Pitre, Guadeloupe
spolDB3.0 of the Pasteur Institute of Guadeloupe. At the
time of the matching analysis, spolDB3.0 contained
13,008 patterns distributed into 813 shared types (patterns
reported at least twice that grouped 11,708 clinical
isolates) and 1,300 orphan patterns from >90 countries (6).
The results were also computed into Recognizer files of
the Taxotron package (P.A.D. Grimont, Taxolab, Institut
Pasteur, Paris) to calculate the 1-Jaccard Index (8) and to
allow the construction of dendrograms by using the
unweighted pair-group method with arithmetic averages
(UPGMA [9]). Odds ratios for clustering with 95%
confidence intervals were calculated to compare
characteristics of clustered and nonclustered patients.
Differences were considered significant if values were
<0.05 (Table 1).

Table 1 shows the detailed demographic data of the
study population. None of the patients’ families had
symptoms suggestive of TB during the study period. In 2%
of families, death attributable to pulmonary TB was
reported.

A total of 45 distinct spoligopatterns were obtained
from the study population of 105 isolates (Figure). Twenty-nine (28%)
clinical isolates were represented by a
unique pattern, whereas 76 (72%) isolates were clustered
in 16 clusters, i.e., 2 predominant clusters of 23 (22%) and
13 (12%) isolates (ST26 and ST54), followed by 1 cluster
of 9 isolates (ST1), 2 clusters of 4 isolates each (ST11
and ST119), 1 cluster of 3 isolates (ST1088), and 10 clusters
of 2 isolates (ST100, ST276, ST1089-1092 and ST1094-
1097). The isolates in all the clusters had different drug-
resistance profiles and had been collected, processed, and
amplified into separate batches on different days; hence,
carryover contamination was ruled out.

Spoligotypes were compared with those in spolDB3.0.
The three most prevalent spoligotypes from this study
were type 26, 54, and 1. ST26 (22%) was initially
described in 1997 in a study performed in the United
Kingdom (4). It was later shown to belong to the major
genetic group I of \( M. tuberculosis \) complex organisms
(10,11), i.e., an ancestral group of human TB, as evidenced
by the presence of the TbD1 region in these strains (12).
Until now, this family of strains has been reported in 11
countries of the Middle East (Iran, Pakistan, and India),
Oceania (Australia), the United States, and Europe (United
Kingdom, the Netherlands, France, Sweden, Italy, and
Austria; [5]). In Europe and Australia, these strains were
regularly found to be linked with immigrants from the
Middle East and Central Asia, hence the name of Central
Asian 1 or CAS1 family, which was recently given to all
isolates characterized by the absence of spacers 4–7 and
23–34 (6). Indeed, a total of 20 shared-type variants linked
to this family of strains have been found to date (5). In a
recent article, this predominant clade of bacilli and some
of their variants were called the Delhi type (13). For
spoligotype patterns, see Table 2.

Type 54 (12%) is also likely to belong to group I
organisms, as suggested by the existence of a closely
linked profile, found within the Houston study (clinical
isolate S179; [10]). This shared-type is less widespread,
and its distribution is different. It has been reported to be

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No. (%) of patients in</th>
<th>OR for clustering (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clustered group</td>
<td>Nonclustered group</td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td>15–45</td>
<td>63 (81)</td>
<td>4.52 (1.6 to 12.96)</td>
</tr>
<tr>
<td></td>
<td>≥46</td>
<td>13 (48)</td>
<td>0.88 (0.30 to 2.49)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>53 (72)</td>
<td>21 (28)</td>
<td>UD</td>
</tr>
<tr>
<td>Female</td>
<td>23 (74)</td>
<td>8 (26)</td>
<td></td>
</tr>
<tr>
<td>HIV status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seropositive</td>
<td>2 (100)</td>
<td>0 (0)</td>
<td>UD</td>
</tr>
<tr>
<td>Seronegative</td>
<td>74 (72)</td>
<td>29 (28)</td>
<td></td>
</tr>
<tr>
<td>Previous history of TB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No previous therapy</td>
<td>52 (80)</td>
<td>13 (20)</td>
<td>2.53 (0.89 to 7.22)</td>
</tr>
<tr>
<td>Previously treated</td>
<td>19 (61)</td>
<td>12 (39)</td>
<td></td>
</tr>
<tr>
<td>Drug resistance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug resistance*</td>
<td>22 (71)</td>
<td>6 (29)</td>
<td>1.56 (0.51 to 4.97)</td>
</tr>
<tr>
<td>Susceptible to all drugs</td>
<td>54 (70)</td>
<td>23 (30)</td>
<td></td>
</tr>
<tr>
<td>Radiologic findings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extensive cavitary</td>
<td>12 (76)</td>
<td>5 (24)</td>
<td>1.6 (0.33 to 11.22)</td>
</tr>
<tr>
<td>Limited cavitary</td>
<td>6 (60)</td>
<td>4 (40)</td>
<td></td>
</tr>
<tr>
<td>Sputum smear positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–10 AFB/10–100 fields</td>
<td>30 (73)</td>
<td>11 (27)</td>
<td>0.96 (0.34 to 2.70)</td>
</tr>
<tr>
<td>&gt;1 AFB per field</td>
<td>37 (74)</td>
<td>13 (26)</td>
<td></td>
</tr>
</tbody>
</table>

*OR, odds ratio; CI, confidence interval; TB, tuberculosis; NS, not statistically significant; UD, undefined; AFB, acid-fast bacilli.

*Resistance to one or more drugs.
present in Africa (Guinea-Bissau and Senegal) and in Europe (France, United Kingdom, and the Netherlands). This type may be an ancestor of both the CAS and the Beijing family. It is characterized by the absence of spacers 33 and 34, two spacers likely to be of high phylogenetic importance in group I organisms since they permit distinguishing between 1) East African Indian (EAI) superfamily (presence of spacer 33, absence of 34), 2) CAS family (absence of 33 and 34), and 3) M. bovis (presence of 33 and 34). Combined prevalence of these two spoligotypes (34%) indicates that these two families are highly prevalent in our high-incidence area and may play an important role in disease transmission in Delhi (Table 2).

The third most prevalent type is the Beijing type shown in nine isolates (8%). This type was originally described by van Soolingen et al. in China and is highly prevalent throughout Asia and Eurasia (15), with a reported prevalence of approximately 3% in India (13). Among these isolates, seven were resistant to ETH, STR, and INH, and one isolate (strain 45) was resistant to all four drugs tested. An epidemiologic link could be established for eight patients who resided in one area (Faridabad) and were referred to the district TB center. Two shared types (ST11 and ST119) belong to the EAI and X families, respectively (14). The presence of the X family in India could be linked to the past British history in this region.

Spoligotypes that did not match any existing pattern in the database were defined as orphans. Of 45 patterns observed in this study, 18 (observed for isolates 4665, 90, 92, 1020, 470, 100, 3257, 3570, 337, 274, 2290, 368, 1063, 1024, 2127, 565, 314, 581) were true orphans (no counterpart in the database). This percentage referred to patterns (40%) is high and reflects both the current absence of knowledge on the genetic diversity of Indian M. tuberculosis strains and the microevolutionary genetic driving forces active in TB-epidemic dynamics in India. Of 105 isolates, 18 (17%) had orphan spoligopatterns. New shared types were also created for newly identified types (ST1088–1092 and 1094–1097), which may either reflect homoplasia (creation of common genetic structures without common ancestor, also called convergence; [16]) or true synapomorphy (common ancestors). Except for one clinical isolate belonging to ST1092, isolated in 2002 in New York (S863, J. Driscoll, unpub. data), all the shared types mentioned above have not been reported elsewhere in the world. A total of 10 isolates unique to this study (Table 2) did match with strains already reported elsewhere (ST50, 52, 53, 138, 141, 357, 381, 427, 458, and 1093). The origin of their counterparts from different parts of the world is described in spolDB3.0. However, many of these patterns (particularly ST 138, 381, 1093) were originally reported in the United Kingdom by Goyal et al. in 1997 (4), who described many strains harboring Indian genetic characteristics.

Little information is available from India or the neighboring countries on the molecular epidemiology of TB. Our study has demonstrated that the epidemiology of TB in India is much different than TB epidemiology elsewhere (5). Our results showed significant clustering in the 15- to 45-year age group (p = 0.001) and significant
lack of clustering in the older age groups. Patients with no history of previous treatment also showed significant clustering (p = 0.05). No significant differences were observed for other parameters such as age <60 years, gender, HIV status, drug resistance, radiologic findings, and sputum smear positivity.

Conclusions

Spoligotyping is based on the variability in the direct repeat locus of *M. tuberculosis*, which most likely occurs by one of three mechanisms—homologous recombination between neighboring or distant direct variable repeats, IS-mediated transposition, and DNA replication slippage (17). Spoligotyping is useful for tracking TB epidemics, detecting new outbreaks, and better defining high-risk populations to focus prevention strategies (6,18). Spoligotyping may also constitute a potential tool for global TB epidemiology, population genetics, and phylogeny, although it should be used with another independent genotyping method in many settings to prove clonality (5,6,10). Another poorly investigated theoretical limitation issue is the level of convergence that may jeopardize phylogenetic reconstruction when spoligo
typing is used (16).

To have a better knowledge of moving and expanding clones of *M. tuberculosis* within the Indian subcontinent, we attempted to identify the predominant spoligotypes prevalent in Delhi and determine their specific signature. Comparison with the spolDB3.0 database enabled us to compare spoligotypes generated in a New Delhi laboratory with different laboratories around the world. The present study included few isolates from patients enrolled at one center, in the beginning of a multicentric study, with no bias for patient selection; hence, it was truly representative of the community. As a tertiary-care referral center, AIIMS gets patients from throughout Delhi. In addition, isolates from the district TB and the primary health centers (which cover all the patients in their respective areas) were also included. It seemed imperative to study the population structure of Indian *M. tuberculosis* strains by using additional genetic markers to better comprehend the origin and evolutionary genetics of TB in this geographic region. The study strains have been further subjected to IS6110-RFLP and double repetitive element PCR. Also the spoligo clusters are being subjected to mycobacterial interspersed repetitive units—variable number of DNA tandem repeats. Investigation of strains from Delhi and those collected from the other centers from all over the country is ongoing.

Clustering was much higher within the age group of 15 to 45 years but decreased with increasing age and was lowest in patients >60 years (odds ratio for clustering 4.52, p = 0.001). This finding may suggest active transmission of TB among younger persons as opposed to possible reactivation of disease in the older age group. A high rate of clustering in a population with no history of TB treatment further corroborated the notion of active transmission of certain prevailing genotypes, despite the existence of genetically diverse strains in our high-incidence community. These prevailing genotypes may play an important role in the propagation of the TB epidemic in Delhi and surrounding regions in coming years. Extensive cavitary disease and degree of

Table 2. Predominant spoligotype prevalent in Delhi and shared type designations for 10 unique isolates, clade designation, and initial description (reference)

<table>
<thead>
<tr>
<th>Spoligotype designation</th>
<th>Spoligotype pattern</th>
<th>Clade†</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 26</td>
<td></td>
<td>CAS1 (Delhi)</td>
<td>13</td>
</tr>
<tr>
<td>Type 54</td>
<td></td>
<td>Manu</td>
<td>5</td>
</tr>
<tr>
<td>Type 1</td>
<td></td>
<td>Beijing</td>
<td>10</td>
</tr>
<tr>
<td>Shared spoligotype</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 50</td>
<td></td>
<td>H3</td>
<td>4</td>
</tr>
<tr>
<td>Type 52</td>
<td></td>
<td>T2</td>
<td>4</td>
</tr>
<tr>
<td>Type 53</td>
<td></td>
<td>T1</td>
<td>4</td>
</tr>
<tr>
<td>Type 138</td>
<td></td>
<td>EAI</td>
<td>4</td>
</tr>
<tr>
<td>Type 141</td>
<td></td>
<td>CAS1</td>
<td>5</td>
</tr>
<tr>
<td>Type 357</td>
<td></td>
<td>CAS1</td>
<td>5</td>
</tr>
<tr>
<td>Type 381</td>
<td></td>
<td>CAS1</td>
<td>4</td>
</tr>
<tr>
<td>Type 427</td>
<td></td>
<td>CAS1</td>
<td>5</td>
</tr>
<tr>
<td>Type 458</td>
<td></td>
<td>EAI</td>
<td>5,14</td>
</tr>
<tr>
<td>Type 1093</td>
<td></td>
<td>CAS</td>
<td>4</td>
</tr>
</tbody>
</table>

†Clades were defined according to the definitions in spolDB3.0 (6); = no hybridization, = positive hybridization. CAS1 (Central Asian 1) family (also termed as the “Delhi type”) is characterized by D4–7 and D23–34; Manu (derived from the name of a Hindu mythological figure supposed to be the world’s first king and father of the human race) is characterized by D33–34 (5), and presumed to be the probable ancestor of both the CAS and EAI (East African Indian). Beijing family is characterized by D1–34. Both EAI and CAS1 belong to the Major Genetic Group 1; H3 (Haarlem 3) belongs to the Major Genetic Group 2; and T1 and T2 are poorly defined families that belong to the Major Genetic Group 2 or 3 (11). T1 and T2 need other markers for a better characterization (T1, D33–36 and T2, D33–36 and D40 [6]).

Emerging Infectious Diseases • www.cdc.gov/eid • Vol. 10, No. 6, June 2004
infectiousness (smear positivity) indicative of large bacterial populations did not affect clustering.

Studies focusing on the polymorphism of *M. tuberculosis* isolates from developing countries, where TB is highly prevalent, would provide new insights into epidemiology, transmission dynamics, phylogenetic analysis, and virulence. Similar studies with detailed epidemiologic data that would reflect on the TB control programs are needed to understand the current epidemic in India.

**Acknowledgments**

We acknowledge Ingrid Filliol for her assistance in the consultation of the spolDB3.0 database, and Christel Delfino for the implementation of the spolDB3.0 automatization system at Institut Pasteur de Guadeloupe.

The strains of mycobacteria typed in this study were obtained during a study sponsored by India Clinical Epidemiology Network. The international spoligotype database construction projects are supported through grants by the Délégation Générale au Réseau International des Instituts Pasteur et Instituts Associés, Institut Pasteur, Paris, and Fondation Française Raoul Follereau, Paris, France. SpolDB3.0 may be consulted on line at http://www.pasteur-guadeloupe.fr/tb/spoldb3.

Dr. Singh is assistant professor and chief of the Tuberculosis Laboratory, Department of Microbiology, All India Institute of Medical Sciences, New Delhi, where she established molecular epidemiology and molecular resistance detection techniques for *Mycobacterium tuberculosis*. Her research interests include molecular diagnostics and molecular epidemiology of tuberculosis for clinical applications.

**References**


Address for correspondence: Urvashi B. Singh, Tuberculosis Laboratory, Department of Microbiology, All India Institute of Medical Sciences, New Delhi-110 029, India; fax: 91-11-26588663; email: urvashi00@hotmail.com

Search past issues of EID at www.cdc.gov/eid