Large-Scale DNA Fingerprinting of *Mycobacterium tuberculosis* Strains as a Tool for Epidemiological Studies of Tuberculosis

DENISE CHEVREL-DELLAGLI,1,* AMEL ABDERRAHMAN,1 RAJA HALTIT,2 HICHEM KOUBAJI,2 BRIGITTE GICQUEL,3 AND KOUSSAY DELLAGI1

Laboratoire des Mycobactéries, Institut Pasteur de Tunis 13, place Pasteur, B.P. 74, 1002 Tunis-Belvedere,1 and Chest Department, District Hospital of Menzel Bourguiba, Menzel Bourguiba,2 Tunisia, and Unité de Génie Microbiologique, Centre National de la Recherche Scientifique, Unite de Recherche Associée 1300, Institut Pasteur, 75724 Paris Cedex 15, France3

Received 10 February 1993/Returned for modification 29 March 1993/Accepted 22 May 1993

We conducted a large-scale DNA fingerprinting analysis of *Mycobacterium tuberculosis* strains in a country in which tuberculosis is endemic (Tunisia) in order to evaluate the importance of microepidemics in the maintenance of the disease within the population. The genetic polymorphisms of 201 strains of *M. tuberculosis* isolated from 196 unrelated patients living in four districts of northern Tunisia during a 3-year period were studied by restriction fragment length polymorphism (RFLP) analysis by using the insertion sequence IS6110 as a probe. Seventy-three strains isolated from 68 patients living in the districts of Tunis, Nabeul, and Jendouba generated 67 different RFLPs, indicating a high degree of polymorphism of the *M. tuberculosis* strains within these areas. In contrast, the 128 strains isolated from individuals in the district of Menzel Bourguiba appeared much less heterogeneous since they often generated identical or very similar fingerprints. Seventeen of 29 cases (58%) of active tuberculosis in the city of Menzel Bourguiba could be traced to as few as four *M. tuberculosis* strains. These results indicate the persistence of underestimated microepidemics in this region. The RFLP typing of a large number of randomly collected strains provides a general picture of the strains involved in tuberculosis. The systematic study of limited areas where tuberculosis is endemic can provide evidence for the existence of persisting epidemics. This stresses the different problems which remain to be solved in order to improve the control of tuberculosis.

Tuberculosis remains a major public health problem worldwide. This disease affects about 40 million individuals and results in approximately 3 million deaths per year (8). Identification and subsequent treatment of sputum-positive patients, who constitute the main source of contamination, as well as finding active cases of infection among their close contacts are the keystones of programs for the elimination of tuberculosis. However, despite the implementation of an efficient multidrug therapy and the *Mycobacterium bovis* BCG vaccine, the incidence of this disease has changed little in developing countries. This indicates that the parameters of disease transmission have yet to be sufficiently understood to improve global tuberculosis control.

Epidemiological studies have been hampered by the lack of markers that are able to differentiate the various *Mycobacterium tuberculosis* strains. The recent development of restriction fragment length polymorphism (RFLP) analysis for the identification of mycobacteria has made it feasible to evaluate RFLP analysis as an epidemiological tool (7, 9, 11). Using the IS6110 insertion sequence fragment, which is specific to the *M. tuberculosis* complex, as a probe, several investigators have detected genotypic variations among *M. tuberculosis* strains (7, 14, 15). It appears that a variable number of IS6110 sequences are scattered in the *M. tuberculosis* genome, where they are arranged differently from strain to strain.

Determination of these IS6110 RFLP patterns has, in turn, the potential to be an extremely useful tool for epidemiological studies of *M. tuberculosis*. For example, it has been demonstrated that strains isolated from patients not involved in the same tuberculosis outbreak exhibit different RFLP patterns, whereas strains isolated from patients involved in the same tuberculosis outbreak have identical RFLP patterns (1, 2, 4, 6, 7, 9, 11, 15, 16). In this report, we describe a systematic analysis of the IS6110 RFLP patterns of a large panel of *M. tuberculosis* complex strains isolated in Tunisia. Comparison of the RFLP patterns of 201 strains isolated from unrelated patients living in four different geographical areas provided a panorama of the *M. tuberculosis* strains circulating within Tunisia and demonstrated the persistence of microepidemics.

**MATERIALS AND METHODS**

*Mycobacterial strains.* Two hundred one strains of *M. tuberculosis* isolated at the Pasteur Institute of Tunis during a 3-year period (1990 to 1992) were selected for DNA fingerprinting. These isolates were grown from clinical specimens from patients admitted to any of the four district hospitals in northern Tunisia: Tunis (n = 25), Jendouba (n = 20), Nabeul (n = 28), and Menzel Bourguiba (n = 128) (Fig. 1). Patients admitted to the last three district hospitals most likely lived within these districts; these districts are situated 70 km north, 70 km south, and 120 km west of Tunis, respectively. Patients admitted to the university hospitals of Tunis were likely to be heterogeneous with regard to their geographical origins. Only strains isolated from apparently unrelated patients were included in the present study, and isolates obtained from close contacts (family contacts) of patients known to have tuberculosis as part of the search for active cases were excluded to avoid bias in the estimation of the number of *M. tuberculosis* strains. The 128 strains from the district of Menzel Bourguiba represent strains from

---

* Corresponding author.
about 60% of all patients with tuberculosis admitted to the Chest Department of the district hospital. Of the 201 strains, 181 were isolated from sputum, 16 were isolated from lymph nodes, 3 were isolated from urine, and 1 was isolated from cerebrospinal fluid. For five patients from Tunis, two different isolates of *M. tuberculosis* were recovered from multiple sites (urine, lymph node, sputum, cerebrospinal fluid) within a few weeks. Two strains were isolated from sputum specimens from the same patient within on two separate occasions 6 months apart.

*M. tuberculosis* strains were identified by standard biochemical tests (niacin, catalase, nitrate reductase) (5). A large panel of tests was needed for complete identification of some *M. tuberculosis* strains with peculiar DNA fingerprints (see below). These tests included biochemical tests (niacin, nitrate reductase, catalase, β-glucosidase, and arylsulfatase), pigment production, growth rates at different temperatures (30, 37, and 42°C), and growth inhibition by various inhibitors (5).

**Culture and genomic DNA isolation.** Culture of mycobacterial strains and DNA extraction were performed as described by Ota et al. (11). Briefly, mycobacterial strains were grown for 3 weeks at 37°C in 20 ml of 7H9 Middlebrook medium supplemented with Albumin Dextrose Catalase (Difco). Cycloserine (1 μg/ml) was added for the last 24 h of culture. Bacteria were pelleted by centrifugation and resuspended for 1 h at 37°C in 500 μl of TE buffer (50 mM Tris [pH 8.1] containing 50 mM EDTA and sucrose [25%; wt/vol]) and 500 μg of lysozyme per liter. Finally, 500 μl of 100 mM Tris (pH 8) containing 400 μg of proteinase K per ml and 1% sodium dodecyl sulfate (SDS) was added and the suspension was incubated at 37°C for 2 h. The DNA was extracted with phenol and chloroform-isoamyl alcohol (24:1; vol/vol) and was precipitated with ethanol by standard methods (12).

Separate colonies were obtained from five *M. tuberculosis* strains by subculturing the strains on solid medium, as follows. Each strain was inoculated in 20 ml of 7H9 Middlebrook liquid medium as indicated above, and then 200-μl aliquots of the bacterial suspension were plated onto petri dishes containing solid 7H10 Middlebrook medium supplemented with ADC (Difco), and the plates were incubated for 3 weeks at 37°C. Well-grown single colonies were picked up, resuspended in 1 ml of liquid 7H9 Middlebrook medium, grown for 2 weeks at 37°C, and then inoculated in 20 ml of 7H9 Middlebrook medium for an additional 3 weeks of incubation. Forty-one colonies were derived from five *M. tuberculosis* strains, and the colonies were further processed for DNA extraction as indicated above.

**Southern blot analysis.** Mycobacterial DNA extracted from isolates or colonies was digested overnight at 37°C with *PstI* (5 U/μg), which has no recognition site within the IS6110 element. Some DNA samples were also digested with *EcoRI*, *PvuII*, or *BamHI* (5 U/μg). DNA fragments were separated by electrophoresis on 0.8% agarose gels containing 500 ng of ethidium bromide per ml. After denaturation, the DNA fragments were transferred to Hybond membranes (Amersham, Amersham, United Kingdom) by the method of Southern (13) and were hybridized with 32P-labelled probes under stringent conditions (65°C). Washing steps (65°C) included 10 min with 2× SSC buffer (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 30 min with 2× SSC-0.1% SDS, and 10 min with 0.1× SSC. Membranes were exposed to X-Omat films (Eastman Kodak Co., Rochester, N.Y.) at −70°C for various lengths of time.

**Labelling of DNA probe.** The IS6110 probe specific for the *M. tuberculosis* complex was isolated from plasmid PMT 03 by *EcoRI* digestion and was purified by electroelution (13). The probe was labelled with [32P]dCTP by using the multiprime DNA labelling kit (Amersham).

**RESULTS**

A total of 201 *M. tuberculosis* strains isolated from clinical specimens were typed by the RFLP technique, which revealed a variable number of hybridizing bands (1 to 14 bands) representing the minimum possible number of insertion sequence (IS) copies (Fig. 2). The vast majority (95%) of the 201 strains had more than six copies. Six strains (3%) contained a single hybridizing band. In five of these six strains the copy was observed in a 3.2-kb *PstI* fragment, and in the other strain the copy was observed in an 8-kb *PstI* fragment identical to that described previously in *M. bovis* (15). Figure 3 illustrates the high degree of genetic polymorphism that could be demonstrated by the IS6110 fingerprinting technique.

**Strains with a high degree of polymorphism.** A high degree of polymorphism was found among the strains isolated from 68 unrelated patients (who were of different geographical origins and who had no known relationship to each other) who were admitted to the district hospitals of Tunis, Nabeul, and Jendouba. Except for two isolates from the district of Jendouba, all strains from these areas had different RFLP profiles. In contrast, paired strains isolated from different sites of infection of five patients always exhibited identical RFLP profiles (data not shown).
Strains with related RFLP patterns. One hundred twenty-eight strains were isolated from unrelated patients who lived in the district of Menzel Bourguiba. In contrast to the strains isolated in the other districts, as described above, many of the strains isolated in Menzel Bourguiba had identical RFLP patterns. In fact, 53 strains shared 12 banding patterns, which are shown in Fig. 3. These identities were confirmed by Southern blotting with the IS6110 probe after digestion of the genomic DNA with restriction enzymes other than PsI (EcoRI, BamHI, and PvuII; data not shown). The 12 RFLP patterns obtained with PsI are described in Table 1.

No correlation was found between RFLP types and the phenotypes of resistance to various antituberculosis drugs. For instance, one strain that was resistant to isoniazid, one strain that was resistant to streptomycin, and one strain that was resistant to both streptomycin and pyrazinamide were shown to belong to types XI, XII, and VIII, respectively, as did strains that were susceptible to these drugs. In addition, two strains isolated from the same patient at a 6-month interval had the same RFLP type, although they exhibited different susceptibility patterns. One of these strains was resistant to isoniazid and streptomycin, and another was resistant to isoniazid, streptomycin, and rifampicin.

TABLE 1. RFLP patterns (types I to XII) shared by
M. tuberculosis isolates from unrelated patients
living in the district of Menzel Bourguiba

<table>
<thead>
<tr>
<th>Pattern type</th>
<th>Lane in Fig. 3</th>
<th>No. of strains</th>
<th>No. of IS copies</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>a</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>II</td>
<td>b</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>III</td>
<td>c</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>IV</td>
<td>d</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>V</td>
<td>e</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>VI</td>
<td>f</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>VII</td>
<td>g</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>VIII</td>
<td>h</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>IX</td>
<td>i</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>X</td>
<td>j</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>XI</td>
<td>k</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>XII</td>
<td>l</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

*RFLPs were generated by PsI restriction enzyme digestion.*

Some RFLP types were very similar in that they differed by the location of a single IS copy. For instance, type VIII differed from type VI by one IS that contained a fragment with a slightly different molecular size (Fig. 3, lanes f and h). This slight difference was more apparent when DNA was digested with EcoRI instead of PsI (data not shown). Except for the presence of an additional IS copy, type XI was similar to the type VIII. The RFLP profiles of strains of types XI and VIII were also similar to the RFLP profiles of seven additional strains from the Menzel Bourguiba district (Fig. 4). All the strains isolated from the districts of Jendouba, Nabeul, and Tunis appeared to be different from those isolated from Menzel Bourguiba. However, the two strains isolated from Jendouba, which shared identical RFLPs, as well as four other strains from this area had profiles similar to those shown in Fig. 4.

Evolution in RFLP patterns. In order to demonstrate that small variations in RFLP patterns could arise during strain cultivation, we isolated colonies from five isolates and analyzed their RFLP patterns. Most of these colonies had RFLP patterns identical to those of the initial isolate. Consistently, the 17 colonies derived from three isolates had RFLP patterns identical to those of the respective parental strains. However, 7 of 10 colonies and 6 of 14 colonies derived from two other isolates, respectively, lost one IS copy. Figure 5 shows the results obtained with one of these isolates and five representative colonies derived from it.

Correlations of RFLP patterns with place of residence. We tried to correlate the distributions of the RFLP types defined above with the place of residence of the patients from whom the M. tuberculosis strains were isolated and who lived within the district of Menzel Bourguiba. Several factors of epidemiological interest were noted. For instance, some RFLP patterns were strictly restricted to one given village or city. Thus, patterns V and X were shared only by some strains isolated from patients living in Sejnane. Patterns III
and VI were characteristic of strains obtained from patients living in Menzel Bourguiba. In contrast, some RFLP patterns were widely distributed. For instance, pattern VIII was detected in strains isolated from patients living in Menzel Bourguiba (five strains), Ras Jebel (four strains), or Mateur (two strains); these cities are within 20 to 30 km of each other. Of the 29 strains of *M. tuberculosis* that were isolated from patients living in Menzel Bourguiba, three strains belonged to pattern XII, five strains belonged to pattern VIII, seven strains belonged to pattern VI, and two strains belonged to pattern III. Therefore, 17 of 29 cases (58%) of active tuberculosis reported in Menzel Bourguiba could be traced to as few as four *M. tuberculosis* strains.

**DISCUSSION**

Tuberculosis is still endemic in Tunisia, despite continuous efforts to control the disease through mass BCG vaccination at birth, active case finding, and treatment with multidrug therapies. These efforts have resulted in some changes in the epidemiological parameters of the disease. The annual risk of infection dropped from 3 to 0.5% between 1960 and 1984, and the incidence of reported cases declined from 48 × 10⁻⁵ in 1975 to 31 × 10⁻⁵ in 1989. However, the persistence of the disease in Tunisia indicates that several components of the control program are still suboptimal. A better knowledge of the dynamics of transmission and dissemination of the tuberculosis bacillus would improve the detection of new cases and help in tracking the sources of contamination, especially if unique methods of transmission linked to local habits are still underestimated. In this context, the recent description of IS6110 as a reliable tool for differentiating strains of *M. tuberculosis* is of great interest. In order to evaluate the number of *M. tuberculosis* strains, we analyzed the RFLP patterns of 201 randomly selected strains of *M. tuberculosis* that were isolated during a 3-year period.

In our study, *M. tuberculosis* strains isolated from the same patient (from different sites of infection or at a 6-month interval) always showed identical RFLP patterns. These findings are in accordance with the findings of Hermans et al. (7), who showed that no RFLP modification was found after growth of *M. tuberculosis* in guinea pigs and our previous findings (11) showing that strains isolated from patients with a relapse after 2 years had RFLP types identical to those of the strains isolated during the first diagnosis of the infection.

Our analysis of the RFLPs of 201 mycobacterial strains showed that all strains from unrelated patients living in the areas of Tunis and Nabeul exhibited different RFLP patterns. Except for two strains, the same results were obtained in analyses of strains from patients of the Jendouba area, resulting in 67 different RFLP types among the 68 strains that were isolated. Since these three areas correspond to regions with mixed populations, we feel that our study may provide a general picture of the strains that exist in Tunisia. To prove this, further epidemiological studies with the aim of following the dissemination of strains that have already been identified in the general population will be undertaken.

Strains isolated from the area of Menzel Bourguiba were much less heterogeneous than those from the other areas described above. Among the 128 strains analyzed, 53 shared only 12 RFLP patterns. Preliminary investigations of the patients whose *M. tuberculosis* strains shared identical fingerprints revealed that although they were apparently unrelated, they frequently lived in the same village or city. This fact was particularly striking for pattern VI, which was characteristic of seven strains isolated only from patients living in Menzel Bourguiba. Seventeen of the 29 strains (58%) isolated in Menzel Bourguiba generated only four different RFLPs. These results suggest that in Menzel Bourguiba, microepidemics of tuberculosis still occur and that contagious patients might have escaped detection by public health authorities who search for active cases. About 70% of the reported cases of tuberculosis in Tunisia are pulmonary tuberculosis; 70% of the patients with pulmonary tuberculosis are sputum positive. Detection of new cases, including active case finding among close contacts of case patients, relies mainly on Ziehl-Neelsen staining. Since this method requires at least 10⁴ to 10⁵ bacilli per ml in the sample for a positive result, many specimens containing lower quantities of bacilli are recorded as negative. Thus, patients with fewer than 10⁴ bacilli per ml are not diagnosed and remain contagious.

Among the 12 RFLP types registered in the Menzel Bourguiba area, several hybridization patterns differed by only one band and, although they were not strictly identical, showed considerable similarities. Similar results were described previously (9) in an analysis of *M. tuberculosis* strains isolated from close contacts or in the context of outbreaks. These differences might be due to the fact that the IS6110 fragment has the ability to transpose and integrate in different sites (14). Its transposition might give rise to the modifications noted in some RFLP patterns. By analyzing *M. tuberculosis* colonies isolated from strains streaked onto solid media, we demonstrated that one IS copy was missing from some of them. Similar changes can occur during the transmission of the bacilli among patients. However, in all cases analyzed so far, this strain evolution is limited to only one IS copy. Since the great majority of *M. tuberculosis* strains harbor more than six IS copies (8), we believe that their RFLP types will be maintained over long periods of time.
Antibiotic resistance is due to the emergence of drug-resistant strains not usually preexisting during the first treatment. However, these resistant strains can be transmitted from close contacts of patients infected with these new mutant strains and also in delineating outbreaks that could result from the dissemination of these resistant strains. The emergence of drug-resistant tuberculosis in the setting of the AIDS epidemic makes this point of utmost importance (3).

Finally, it should be pointed out that the vast majority (95%) of the M. tuberculosis strains analyzed in the present study displayed more than six copies of IS6110 and five strains had only one copy; of these five strains, it was located in one strain at the same position as the IS copy detected in some M. bovis and M. bovis BCG strains (15). Therefore, RFLP analysis cannot be used to unambiguously identify strains of the M. tuberculosis complex, including the BCG vaccine strain.

It must be stressed that M. tuberculosis strains isolated from close contacts of patients known to have tuberculosis were not included in the present study in order to avoid bias in the estimation of the number of different RFLPs. We have found that these strains regularly express identical fingerprints (5a), as reported by others (9, 16). The microepidemics revealed in the present study involved individuals who were apparently unrelated. The route of transmission of M. tuberculosis between these apparently unrelated patients remains to be elucidated. Although it is well known that contamination is higher among close contacts of a sputum-positive patient (i.e., family members or working place contacts), it is clear that sources of contamination depend largely on sociological habits, which may maintain a high rate of circulation of M. tuberculosis strains. For instance, in villages, contamination can take place between apparently unrelated persons in some socially important gathering points, for example, grocer’s and coffee shops, and may be amplified by some local customs (for instance, the common use of the narguile, an Oriental pipe, in coffee shops). Identification of these mechanisms of contamination will play a key role in limiting the dissemination of M. tuberculosis and improving programs to control tuberculosis.

ACKNOWLEDGMENTS

We are grateful to Dahmani Fathallah, Ikram Guizani, and Jean Luc Guesdon for interest and assistance and to Souad Hajjam for typing the manuscript.

This work was supported in part by a grant from La Fondation Nationale à la Recherche Scientifique Tunisienne and the French-Tunisian scientific cooperation.

REFERENCES


