Genetic Biodiversity of *Mycobacterium tuberculosis* Complex Strains from Patients with Pulmonary Tuberculosis in Cameroon

Sara Ngo Niobe-Eyangoh, Christopher Kuaban, Philippe Sorlin, Patrick Cunin, Jocelyn Thonnon, Christophe Sola, Nalin Rastogi, Veronique Vincent, and M. Cristina Gutierrez

Laboratoire des Mycobactéries, Centre Pasteur; and Programme National de Lutte contre la Tuberculose, Yaoundé, Cameroon; Laboratoire de Référence des Mycobactéries, Institut Pasteur, 75724 Cedex 15, Paris, France; and Unité de la Tuberculose et des Mycobactéries, Institut Pasteur de Guadeloupe, F-97165 Pointe-à-Pitre Cedex, Guadeloupe

Received 16 December 2002/Returned for modification 5 February 2003/Accepted 7 March 2003

We analyzed DNA polymorphisms in 455 *Mycobacterium tuberculosis* complex isolates from 455 patients to evaluate the biodiversity of tubercle bacilli in Ouest province, Cameroon. The phenotypic and genotypic identification methods gave concordant results for 99.5% of *M. tuberculosis* isolates (413 strains) and for 90% of *M. africanum* isolates (41 strains). *Mycobacterium bovis* was isolated from only one patient. Analysis of regions of difference (RD4, RD9, and RD10) proved to be an accurate and rapid method of distinguishing between unusual members of the *M. tuberculosis* complex. Whereas *M. africanum* strains were the etiologic agent of tuberculosis in 56% of cases 3 decades ago, our results showed that these strains now account for just 9% of cases of tuberculosis. We identified a group of closely genetically related *M. tuberculosis* strains that are currently responsible for >40% of smear-positive pulmonary tuberculosis cases in this region of Cameroon. These strains shared a spoligotype lacking spacers 23, 24, and 25 and had highly related IS6110 ligation-mediated (LM) PCR patterns. They were designated the “Cameroon family.” We did not find any significant association between tuberculosis-causing species or strain families and patient characteristics (sex, age, and human immunodeficiency virus status). A comparison of the spoligotypes of the Cameroon strains with an international spoligotype database (SpolDB3) containing 11,708 patterns from >90 countries, showed that the predominant spoligotype in Cameroon was limited to West African countries (Benin, Senegal, and Ivory Coast) and to the Caribbean area.

In 1993, the World Health Organization declared tuberculosis (TB) a global emergency. One-third of the world’s population is infected by *Mycobacterium tuberculosis* complex strains, the etiologic agents of TB. Although <10% of infected people actually develop active TB during their lifetimes, this represents 8 million new cases of TB each year, including 3.5 million (44%) cases of smear-positive pulmonary disease, leading to 1.9 million deaths per year (5, 6). Ninety-five percent of cases occur in developing countries, where the lack of proper health care systems leads to incomplete case and contact tracing, incomplete treatment, and an increase in drug resistance. Due to the powerful interaction between TB and human immunodeficiency virus (HIV) disease, together with the problems of poverty and malnutrition, the incidence of TB is increasing dramatically in sub-Saharan Africa (22).

In Cameroon, a country with 15 million inhabitants, the incidence of TB in 2000 was estimated at >300 cases per 100,000 inhabitants in the last World Health Organization report (29), with an estimated 21,594 new sputum smear-positive cases. Although there is a paucity of information regarding the distribution of *M. tuberculosis* complex strains in Cameroon, one study performed 30 years ago (14) reported that 56% of cases of TB were due to *Mycobacterium africanum* strains in Ouest and Sud provinces, Cameroon.

Several intervention strategies are expected to reduce the incidence of TB. Molecular epidemiology methods, in particular the genetic typing of *M. tuberculosis* complex strains, can contribute to such strategies. One of these typing methods is spoligotyping, a rapid, simple, and cost-effective system that allows the simultaneous detection and differentiation of *M. tuberculosis* complex strains and provides genotypic information (16). This method analyzes the DNA polymorphism observed in spacer sequences present within the direct-repeat (DR) region of the *M. tuberculosis* complex genome. It has been used as a first-line discriminatory test to type a large number of strains (10). Moreover, when used as a single genotyping method in a new setting, it is a good indicator of strain identity and provides information about epidemiologically important clones (24). Another advantage of spoligotyping is its ability to measure the overall diversity of *M. tuberculosis* complex strain patterns, including differences between regions and countries.

<table>
<thead>
<tr>
<th>Phenotypic identification (n)</th>
<th>Spoligotype (n)</th>
<th>Result for RD4</th>
<th>RD9</th>
<th>RD10</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. tuberculosis</em> (415)</td>
<td><em>M. tuberculosis</em> (410)</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><em>M. africanum</em> (40)</td>
<td><em>M. africanum</em> (36)</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><em>M. bovis</em> (1)</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><em>M. tuberculosis</em> (2)</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ambiguous (1)</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

* +, positive; -, negative.
populations and the prevalence of endemic strains (8, 23). The discriminatory ability of spoligotyping can be improved by applying a second typing method to clustered strains (10). In this context, ligation-mediated (LM) PCR, another rapid genotyping method based on IS6110 polymorphism, is useful to confirm spoligotyping clusters or to discriminate among the isolates that they contain (1).

In the last decade, a large number of molecular epidemiology studies have analyzed the TB situation in developed countries (25). Despite the high prevalence of TB in Africa, little information is available. We applied PCR-based molecular typing methods to a collection of M. tuberculosis complex strains recently isolated in Ouest province, Cameroon. These methods are rapid, simple, and cost-effective and are easy to implement in that country. The aim of this study was to assess the genetic biodiversity of M. tuberculosis complex strains in Cameroon to determine whether the transmission of dominant clones contributes to the high prevalence of TB in the country.

MATERIALS AND METHODS

Study population. The study was performed over a 12-month period (July 1997 to June 1998) in all 15 district hospitals in Ouest province, Cameroon. This region, with an estimated population of 1.8 million, was chosen because it was the first area in which the national TB control program was set up and it was easily accessible. Data concerning each patient’s sex, age, smear result, and previous history of TB were collected upon hospitalization by the use of standard questionnaires for all patients aged 15 years and over with sputum smear-positive pulmonary TB. All of the patients were born in Cameroon. Before the beginning of treatment, the HIV status of each patient was determined, and one new sputum sample was collected from each patient in a sputum container with transport medium (0.65% cetylpromidinium bromide) and kept at room temperature. Every week, specimens were sent to the Provincial Delegation of Health for Ouest province, from where they were transported by road to the central laboratory at the Centre Pasteur du Cameroun at Yaounde for bacteriological analysis and processing as previously described (18). Briefly, each sputum sample was cultured in three Löwenstein-Jensen tubes, one of which was supplemented with a 0.4% solution of sodium pyruvate. The cultures were incubated at 37°C and read weekly for growth for a maximum duration of 10 weeks. This collection was previously used to determine the prevalence of initial and acquired resistance to the main anti-TB drugs (18).

M. tuberculosis complex strains. A total of 455 M. tuberculosis complex strains isolated from 455 patients were analyzed. These strains represent 82% of those isolated during this period in the region. The remaining 18% of strains were not available for DNA extraction because the bacterial cultures were not viable on subculturing or were contaminated. The 455 strains were previously subjected to phenotypic identification by conventional methods (colony morphology and biochemical tests) and were positive for RD4, RD9, and RD10 as described previously (2), using the IS6110 polymorphism as described elsewhere (21). Genomic DNA was digested with SalI, and the digestion products were ligated to an asymmetric double-stranded linker. The template DNA was then amplified, and the PCR products were separated on a 2.5% agarose gel. The gel was photographed and digitized. The LM-PCR patterns were normalized and subjected to cluster analysis using the Biommeres software.

Genotypic confirmation of identification. All M. africanum strains and all other strains for which the phenotypic identification (colony morphology and biochemical tests) was discordant with the genotypic classification by spoligotyping were subjected to additional PCR analysis. This analysis concentrated on the regions of differences RD4, RD9, and RD10 as described previously (2), using internal and external primers for each region. For strain 265, the polymorphisms at position 169 in the pncA gene and at position 285 in the oxyR gene were analyzed by allele-specific amplification as previously described (7).

Analysis of patient data. Patient data were analyzed using EpInfo version 6.04b. Two comparisons were done. First, tuberculous patients infected with M. tuberculosis were compared to those infected with M. africanum. Second, tuberculous patients infected by the “Cameroon family” strains of M. tuberculosis were compared to those infected with non-Cameroon family strains. Proportions were compared by the χ² test with Yates correction or, when appropriate, by Fisher’s exact test. Means were compared by Student’s t test. A difference was considered significant if P was <0.05.

RESULTS

Molecular identification of the strain in the M. tuberculosis complex. Table 1 summarizes the results of the phenotypic and genotypic characterizations of the 455 M. tuberculosis complex strains. According to the phenotypic characteristics, 40 of these strains were previously identified (18) as M. africanum and 415 were identified as M. tuberculosis strains. Thirty-six of the 40 strains with a phenotype typical of M. africanum showed a spoligotype that is typically associated with M. africanum (Fig. 1). As expected, the RD patterns of these 36 strains were typical of M. africanum (negative for RD9 and positive for RD4 and RD10). Two strains showed a spoligotype that is generally associated with M. tuberculosis and were positive for RD9, RD4, and RD10. A review of their phenotypes indicated that both strains were dysgonic and multidrug resistant, which might explain why they were misidentified as M. africanum. One strain showed a spoligotype that is typically associated with Mycobacterium bovis (lack of spacers 8, 9, and 39) (Fig. 1). As expected, the RD patterns of these 36 strains were typical of M. africanum (negative for RD9 and positive for RD4 and RD10). Two strains showed a spoligotype that is generally associated with M. tuberculosis and were positive for RD4, RD9, and RD10. A review of their phenotypes indicated that both strains were dysgonic and multidrug resistant, which might explain why they were misidentified as M. africanum. One strain showed a spoligotype that is typically associated with M. tuberculosis. The degree of similarity of the spoligotypes was calculated by the Jaccard index. The relationships between patterns were assessed by the unweighted pair group method of averages. The spoligotypes listed are the designations in the SpolDB3 database. Spoligotypes 836 to 862 were first described in this study. Orphan patterns did not have matches in the SpolDB3 database.

FIG. 1. Dendrogram and schematic representation of the 95 spoligotype patterns obtained from the 455 M. tuberculosis complex strains. The degree of similarity of the spoligotypes was calculated by the Jaccard index. The relationships between patterns were assessed by the unweighted pair group method of averages. The spoligotypes listed are the designations in the SpolDB3 database. Spoligotypes 836 to 862 were first described in this study. Orphan patterns did not have matches in the SpolDB3 database.
ever, four strains showed both a spoligotype and an RD pattern typical of *M. africanum*. A review of their phenotypes indicated that two of these strains were negative for nitrate reductase. One strain showed an ambiguous spoligotype that did not fit any typical spoligotype of the *M. tuberculosis* complex (lack of spacers 7, 8, and 9 and the last 19 spacers). Its RD pattern confirmed that it was an *M. tuberculosis* strain (positive for RD4, RD9, and RD10).

Thus, 41 of the 455 *M. tuberculosis* complex strains were *M. africanum*, 413 were *M. tuberculosis*, and 1 was *M. bovis*.

**Spoligotyping analysis.** Spoligotyping of the 41 *M. africanum* isolates revealed 23 distinct patterns (Fig. 1). Fourteen unique spoligotype patterns were seen, and the remaining 27 strains were contained within nine clusters: one cluster of 7 strains, four clusters of 3 strains, and four clusters of 2 strains. Twenty-three strains were *M. africanum* subfamily type A2, and 18 strains were *M. africanum* subfamily type A3 according to the criteria of Viana-Niero et al. (27).

Spoligotyping of the 413 *M. tuberculosis* isolates revealed 72 distinct spoligotypes. Thirty-five unique spoligotype patterns were seen, and the remaining 375 isolates grouped into 37 clusters. Most of the clusters contained only two strains (15 clusters). One cluster included 140 isolates. Overall, the clustering rate was much higher among *M. tuberculosis* strains (91%) than among *M. africanum* strains (66%). These results reflect the presence of a predominant genetic group in Cameroon, including 31% of the *M. tuberculosis* complex strains, and show transmission of an endemic strain. The other strains were genetically heterogeneous, showing high local diversity.

**Comparison with an international spoligotype database.** We compared the 95 spoligotypes found in this study with those contained in an international spoligotype database (SpolDB3) (8) containing 11,708 patterns split into 813 spoligotypes shared by more than two patients in any region of the world and 1,300 unique (“orphan”) spoligotypes. Only 51 of our spoligotypes were already described in the SpolDB3. The other 44 spoligotypes were novel. Of these 44 spoligotypes, 14 were already clustered and 30 were unique. The new clusters generated by the incorporation of our results into the SpolDB3 database were designated spoligotypes 836 to 862 (Fig. 1).

The main *M. tuberculosis* cluster (140 isolates) corresponded to spoligotype 61. This spoligotype was previously described in strains from 40 patients, mainly from Benin (19 patients), other West African countries (Ivory Coast and Senegal), and the Caribbean area (Guadeloupe, French Guiana, and Barbados). The second-largest group corresponded to the ubiquitous Haarlem family (spoligotypes 50 and 47), which represents 11% of all *M. tuberculosis* complex strains. The African-type *M. tuberculosis* (spoligotype 52), as defined by Kremer et al. (17), represents 5% of all *M. tuberculosis* complex strains. Finally, 4.5% of the *M. tuberculosis* strains were the ubiquitous spoligotype 53. None of the strains had the spoligotype of the worldwide expanded epidemic Beijing family.

**The prevalent Cameroon genetic group.** The most prevalent spoligotype was 61, which was shown by isolates from 140 patients. This spoligotype represents 34% of all *M. tuberculosis* isolates. The specific characteristic of spoligotype 61 is the lack of spacers 23, 24, and 25 in the DR region. Secondary typing of these 140 strains by IS6110 LM-PCR revealed a clonal population. This group was designated the Cameroon family. All the isolates showed highly related LM-PCR patterns with six common bands and one or two additional bands. The 140 isolates were thus split into seven subgroups. Most of the strains displayed LM-PCR pattern A or B (78 and 50 strains, respectively) (Fig. 2). The remaining 12 strains displayed LM-PCR pattern C, D, E, F, or G. We also included in this group 53 additional isolates with a spoligotype similar to spoligotype 61, lacking spacers 23, 24, and 25 and one or more other spacers (Fig. 3). IS6110 LM-PCR typing confirmed that these isolates were genetically related to isolates displaying spoligotype 61. These 53 strains displayed the LM-PCR pattern common to Cameroon family strains. Most of them presented LM-PCR pattern A or B (20 and 11 strains, respectively). The other 22 strains showed six different LM-PCR patterns (C, G, H, I, J, and K). This increased the size of the main group from 140 to 193 strains, representing 42% of all *M. tuberculosis* complex isolates and 47% of the *M. tuberculosis* isolates included in this study.

**Patient characteristics.** A comparison of tuberculous patients infected by *M. tuberculosis* and those infected by *M. africanum* did not reveal any statistically significant differences with respect to age, sex, HIV status, or drug resistance. No statistically significant difference in age, sex distribution, drug resistance, or HIV-*M. tuberculosis* coinfection was observed between tuberculous patients infected with Cameroon family isolates and the other patients.

**Geographic distribution of strains in Ouest province, Cameroon.** The Cameroon family strains were isolated in all 15 district hospitals. When isolated, *M. africanum* represented a small percentage of *M. tuberculosis* complex strains (Table 2).

**DISCUSSION**

Our study analyzed the biodiversity of *M. tuberculosis* complex isolates collected from the population of Ouest province, Cameroon, a country where TB is hyperendemic. *M. africanum* is generally responsible for TB in patients living in or coming from sub-Saharan countries (3, 9). In the 1970s, most reported
cases of TB in Cameroon were caused by M. africanum (14). At that time, only phenotypic identification methods were available. The phenotypic tests used by Huet et al. (14) were identical to those used in our study. In our study, phenotypic analyses showed that only 9% of strains were M. africanum. A few strains were reclassified after being genotyped, but even after molecular methods were applied, 9% of the isolates were classified as M. africanum. This suggests that the decreasing trend cannot be attributed to identification bias but is in fact due to a genuine regression of M. africanum as the etiologic agent of TB in Cameroon in the last 3 decades (from 56 to 9%). TB caused by M. africanum was unrelated to the age of patients, and >40% of M. africanum-infected patients were younger than 30 years, indicating that although the prevalence of M. africanum is decreasing, it continues to be actively transmitted in the population. Despite the high prevalence of bovine TB in Cameroon (19), pulmonary TB caused by M. bovis in humans remains rare. Nowadays, most cases of TB are caused by M. tuberculosis, the percentage of which has increased from 44 to 91% in 30 years. Among M. tuberculosis strains, we found one highly genetically related group of strains, designated the Cameroon family. These strains were responsible for 42% of TB cases. This group was defined on the basis of the characteristics of spoligotype 61 (lack of spacers 23, 24, and 25). Its clonality was confirmed by the use of an independent genetic marker, IS6110. The success of particular predominant clones related to a high incidence of TB was first reported in studies of the population structure of M. tuberculosis in countries where TB is highly prevalent (11, 12). In these countries, most circulating M. tuberculosis strains belong to a limited number of strain families and are thought to descend from a few clones that expanded in the recent past (12). Our results suggest that the Cameroon family of strains recently expanded in the Cameroon population. The characteristics of the patients affected by strains from this Cameroon family did not differ significantly from those of patients affected by isolates from other clusters or those of patients who harbored strains with unique patterns. Moreover, we did not observe any significant differences in the distribution of the Cameroon family strains in different hospital districts. This suggests that the family expanded homogeneously in the western region of Cameroon. The reasons for its selection and dissemination are unknown. The M. bovis BCG vaccination, which is common practice in Cameroon, may play a role in the selection of Cameroon family strains. The possible role of BCG vaccination in the selection of resistant strains with BCG-induced immunity was previously suggested to explain the expansion of Beijing family strains and the predominance of other families in certain geographic settings (12, 26). This could also be a factor to explain why M. africanum, which is less virulent than M. tuberculosis in experimental models (4), dramatically diminished as a cause of TB after the generalization of BCG vaccination in Cameroon.

The international spoligotype database (SpolDB3) shows that the worldwide distribution of spoligotype 61 is limited. M. tuberculosis isolates with this spoligotype have been found only in countries on the West African coast (Cameroon, Benin, Senegal, and Ivory Coast) and in the Caribbean (Guadeloupe, French Guiana, and Barbados). The finding of this spoligotype in neighboring countries may be due to recent or past cross-border transmission events.
Despite the predominance of the Cameroon family, the other M. tuberculosis strains showed a high degree of genetic diversity. Forty percent of clusters contained only two strains. Such a large number of different circulating strains was unexpected, as it was hypothesized that an epidemic area might have relatively few circulating strains (12). There are two hypotheses to explain this. These strains may represent independent cases of TB, possibly resulting from the reactivation of past TB or individual clones appearing or disappearing over time, as suggested by a study performed in South Africa in a community with a high incidence of TB and high strain diversity (28).

The classic method for identifying members of the M. tuberculosis complex is based on a combination of culture aspects and biochemical characteristics. Whereas classic M. tuberculosis has characteristics opposite to those of M. bovis, M. africanum shares some properties with M. tuberculosis and others with M. bovis. This high degree of variability among M. africanaum strains and the subjectivity of the interpretation of the results makes it difficult to unambiguously identify strains and may lead to misclassification (11, 13). Distinguishing among members of the M. tuberculosis complex is essential for epidemiologic investigations and public health purposes (9, 15). The development of molecular methods resulted in single systems for a more accurate identification of the members of the M. tuberculosis complex (2). We applied molecular methods to confirm the identification of strains previously identified as M. tuberculosis or M. africanum by classic methods. Overall, the classic and molecular methods gave concordant results for the identification of strains in the M. tuberculosis complex. Nearly all M. tuberculosis strains (99.5%) were correctly identified by phenotypic analysis. The only two discordances were due to the dysgenic growth of multidrug-resistant M. tuberculosis, a previously observed phenomenon (M. C. Gutierrez, E. Bouvet, J. Blazquez, and V. Vincent, Letter, Lancet 351:758, 1998). In our experience, molecular methods are especially useful for the identification of M. africanum. Spoligotyping was able to identify 40 of 41 M. africanaum strains unambiguously. In accordance with recent findings, the analysis of RD regions confirmed that all of these strains lacked RD9 (20). Moreover, all of the strains harbored the RD10 region, suggesting that they belonged to one of the two hypothetical evolutionary groups of M. africanum proposed by Brosch et al. (2): one group lacking RD10 and another group harboring RD10. Given the clear results and the simplicity of analyzing RD regions by PCR, this technique proved to be an excellent approach for the rapid identification of M. africanaum strains.

Our investigation of the DNA polymorphism of M. tuberculosis complex strains from humans in Cameroon has shown that M. africanaum strains as the etiologic agents of TB have decreased in this area of West Africa. A group of M. tuberculosis strains designated the Cameroon family is currently responsible for most smear-positive pulmonary cases of TB. Isolates belonging to this group share a spoligotype lacking spacers 23, 24, and 25 and a highly related LM-PCR pattern. The geographic distribution of these isolates seems to be limited to West Africa and the Caribbean. Further studies are needed to understand the expansion in Cameroon of this predominant group of strains.

ACKNOWLEDGMENT

Sara Ngo Niobe-Eyangoh received a doctoral fellowship from the Agence Universitaire de la Francophonie.

REFERENCES


