Transmission Dynamics of Tuberculosis in a High-Incidence Country: Prospective Analysis by PCR DNA Fingerprinting

W. H. HAAS,1* G. ENGELMANN,3 B. AMTHOR,1 S. SHYAMBA,2 F. MUGALA,2 M. FELTEN,1 M. RABBOW,3 M. LEICHSERING,1 O. J. OOSTHUIZEN,2 and H. J. BREMER1

Molecular Genetic Laboratory, Department of General Pediatrics, Children’s Hospital, University of Heidelberg, Heidelberg, Germany,1 and Department of Medicine and Tuberculosis Reference Laboratory, Central Hospital,2 and National Tuberculosis Control Programme, Ministry of Health and Social Services,3 Windhoek, Namibia

Received 10 May 1999/Returned for modification 10 July 1999/Accepted 21 August 1999

We have prospectively analyzed the DNA fingerprints of Mycobacterium tuberculosis strains from a random sample of patients with newly diagnosed smear-positive tuberculosis in Windhoek, Namibia. Strains from 263 smear-positive patients in whom tuberculosis was diagnosed during 1 year were evaluated, and the results were correlated with selected epidemiological and clinical data. A total of 163 different IS6110 fingerprint patterns were observed among the 263 isolates. Isolates from a high percentage of patients (47%) were found in 29 separate clusters, with a cluster defined as isolates with 100% matching patterns. The largest cluster included isolates from 39 patients. One predominant strain of M. tuberculosis caused 15% of cases of smear-positive pulmonary tuberculosis in Windhoek. That strain was also prevalent in the north of the country, suggesting that in contrast to other African countries with isolates with high levels of diversity in their DNA fingerprint patterns, only a restricted number of different strains significantly contribute to the tuberculosis problem in Namibia.

Control of tuberculosis relies on well-defined tools such as case finding, contact tracing, completion of successful treatment, and vaccination with Mycobacterium bovis BCG (6, 20). The impact of control measures on the course of the tuberculosis epidemic in high-incidence countries is often difficult to assess.

An alternative approach to describe the transmission dynamics of tuberculosis has been based on identification of chains of infection by molecular typing of the organism. This method has been very successful in outbreak investigations (12, 14, 17, 22, 28), and recently, DNA fingerprinting of Mycobacterium tuberculosis has been applied to analyze the spread of tuberculosis within circumscribed populations (3, 7, 21, 26).

We applied this approach using DNA fingerprinting by mixed-linker PCR (9) during a prospective study of patients with smear-positive tuberculosis in Windhoek, Namibia. The incidence of tuberculosis in Namibia is about 400/100,000 population, with 38% smear-positive cases, corresponding to an estimated annual risk of infection of 4% (13). The aim of this study was to determine the number of new cases due to recent infection and to identify hot spots of transmission, thus analyzing the transmission dynamics of the current tuberculosis epidemic in Namibia.

MATERIALS AND METHODS

Study population. The study was performed in Windhoek, the capital of Namibia. Namibia has an estimated population of 1.41 million. Of these, 82% live in rural areas and 12% live in Windhoek. All cases of tuberculosis in the Windhoek region are referred for treatment and registration to the Tuberculosis Unit of the Windhoek Central Hospital.

All patients with newly diagnosed smear-positive tuberculosis that were registered between 1 April 1995 and 31 March 1996 were included in a prospective study. To exclude chronic carriers and treatment failures, only patients with no prior history of tuberculosis treatment within the preceding 6 months were admitted to the study.

Patient data concerning sex, age, smear result, previous history of tuberculosis, human immunodeficiency virus (HIV) infection status, and hometown were collected at admission by using standard questionnaires.

Mycobacterium strains. Culture isolation was carried out at the Windhoek Tuberculosis Reference Laboratory. The sputum was confirmed to be smear positive by direct microscopy, decontaminated by the NaOH method, and cultured for 4 to 8 weeks on Löwenstein-Jensen slants (with and without glycerol) by standard protocols (11). Samples of the colonies were placed into tubes containing 500 μl of DNA-free, sterile water, heat killed for 45 min at 80°C in a water bath, and shipped to the Molecular Genetic Laboratory at Heidelberg University. Cross-contamination between samples was controlled for by inclusion of mock samples at the steps of decontamination and cell sampling.

Cell membranes were obtained by freezing-thawing the specimen three times in an 80°C water bath and a −60°C methanol-dry ice bath. Subsequently, 4 μl of the lysate was subjected to mixed-linker PCR fingerprinting as described previously (9). In brief, mycobacterial DNA was digested with the restriction enzyme HhaI (Gibco Life Technologies, Gibco-BRL, Egggenstein, Germany), and the resulting fragments were ligated (T4 ligase; Gibco Life Technologies, Gibco-BRL) to the mixed linker, which consists of a double-stranded 20-bp oligonucleotide, with one strain containing uracil instead of thymidine. For correct ligation the mixed linker has a GC overhang at the 3′ end complementary to that at the HhaI-cutting site. In the next step, uracil is removed by adding uracil-N-glycosylase (Gibco Life Technologies, Gibco BRL). Thus, in a first PCR only fragments containing the 3′ end of IS6110 are amplified with a primer compatible with a 20-bp fragment at the 3′ end of IS6110 and a primer complementary to the uracil-containing oligonucleotide strand. Finally, a seminested PCR with a second primer specific for IS6110 is performed to increase the specificity of the amplification. The mixed-linker PCR products were separated by gel electrophoresis on a 6% polyacrylamide gel together with three external size markers (100-bp DNA ladder; Gibco-BRL) to the mixed linker. Cross-contamination between samples was controlled for by inclusion of mock samples at the steps of decontamination and cell sampling.

Spoligotyping. Spoligotyping was performed as previously described by Kamerbeek et al. (10). Briefly, 10 ng of mycobacterial DNA was amplified with 5′ biotinylated oligonucleotides designated DRA and DBR. PCR was performed with PCR buffer containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), each deoxynucleoside triphosphate at a concentration of 200 mM, 20 pmol of each primer, and 0.5 μl of AmpliTaq DNA polymerase (Perkin-Elmer, Weiterstadt, Germany). A total of 20 μl of the amplification product was hybridized with 43 oligonucleotides covalently bound to a membrane (Isogen Bioscience B. V., Maarssen, The Netherlands). Each of these oligonucleotides represented a

* Corresponding author. Mailing address: Molecular Genetic Laboratory, Department of General Pediatrics, Children’s Hospital, University of Heidelberg, Im Neuenheimer Feld 150, 69120 Heidelberg, Germany. Phone: 49 (6221) 562311. Fax: 49 (6221) 564624. E-mail: Walter_Haas@med.uni-heidelberg.de.
confirmed that 14 isolates were sequenced. Among these 16 isolates, sequencing of the 16S rRNA gene was done by eye.

**Computer-assisted analysis of DNA fingerprints.** Polaroid pictures of the gels were digitized with an HP LaserJet scanner. Conversion of the digitized gels, normalization, and cluster analysis were carried out with GelCompar software (Applied Maths BVBA, Kortrijk, Belgium). The similarities of the strains were calculated by the unweighted pair group method with arithmetic averages with the Dice coefficient and 0.7% position tolerance. Cluster frequencies were calculated as \( \frac{\text{number of isolates in clusters}}{\text{number of patients}} \).

**RESULTS**

Between April 1995 and March 1996, 378 of 577 patients diagnosed with smear-positive pulmonary tuberculosis fulfilled the inclusion criteria of the study. Of the patients included in the study, 99 were culture negative, and no DNA fingerprint pattern could be produced for isolates from 16 patients. Among these 16 isolates, sequencing of the 16S rRNA gene confirmed that 14 isolates were *M. tuberculosis* and that 1 isolate was *Mycobacterium kansasii*. One sample was negative by both PCR methods, probably due to a sampling error or inhibition of the amplification reaction. The presence of IS6110 was shown for the 14 *M. tuberculosis* isolates by using a diagnostic PCR assay based on the detection of this insertion sequence, excluding the possibility of an IS6110-negative cluster. Thus, the DNA fingerprinting results and data could be analyzed for isolates from 263 (70%) patients. Comparison of the available data for culture-negative and culture-positive patients in the study group showed no significant difference (data not shown). The distribution of female-to-male patients was 1:2.6, with a mean age at the time of diagnosis of 36.0 years. Of the analyzed patients, 82% lived in Windhoek; however, only 27% were born there. HIV infection status data were available for 212 (81%) patients, and the rate of 34% HIV-*M. tuberculosis* coinfection was in good agreement with the results of other studies (15a). A prior history of treatment for tuberculosis was reported for 103 (40%) patients.

The DNA fingerprints of the 263 patient isolates were initially separated according to the number of IS6110 insertion elements. Numbers over the columns according to number of IS6110 copies and those with 11 to 15 IS6110 copies. The number of isolates per-pattern ratio was highest for isolates with 1 to 5 IS6110 copies and those with 11 to 15 IS6110 copies (1.94 and 1.90, respectively), corresponding to a low variability of patterns, especially among low- and high-copy-number strains (Fig. 1).

Pattern analysis revealed 123 isolates (47%) in 29 clusters with identical DNA fingerprints, corresponding to a cluster frequency of 36%. The sizes of clusters varied between 2 and 39 isolates, with nine clusters containing more than 2 isolates (Table 1). Comparison of isolates from patients with known HIV status demonstrated a significant predominance of isolates from HIV-positive patients in clusters with small (fewer than three) numbers of isolates (15 of 29 versus 18 of 67 isolates, respectively, \( P = 0.02 \)). HIV infection status data were available for 78% of patients whose isolates were in clusters, and the number of isolates from patients whose HIV infection status was known did not differ significantly between clusters with 2 isolates and clusters with larger numbers of isolates. No difference between small and large clusters was found for sex, age, and prior treatment for tuberculosis.

Of the patients living in Windhoek, 45% were infected with a strain in a cluster and 55% were infected with individual isolates (isolates not in a cluster). The proportion of clustered strains among isolates from patients from other areas did not differ significantly (Fig. 2).

Isolates in 16 of the 29 clusters were most likely from local outbreaks, as the affected patients lived in the Windhoek area. These clusters included isolates from 33 patients, corresponding to 27% of patients whose isolates were in clusters. The remaining clustered strains were not restricted to local outbreaks. All clusters with isolates from more than three patients contained at least one isolate from a patient who lived outside of Windhoek (Fig. 3).

No difference in age and sex distribution, history of prior treatment for tuberculosis, and HIV-*M. tuberculosis* coinfection was observed between clustered patients and patients infected with individual strains. These data and demographic parameters are summarized in Table 2.

The strain that formed the largest cluster comprised isolates from 39 patients with 100% identical DNA fingerprint patterns by mixed-linker PCR. Inclusion of 26 strains with similar patterns that differed by one IS6110 fragment increased the size of this main cluster so that it comprised 24% of all isolates from study patients.

Secondary typing revealed a clonal spoligotype for 58 of these 65 isolates. Two isolates in the main cluster could not be analyzed by spoligotyping due to a lack of DNA. All but 3 of the remaining 37 isolates in the main cluster were clonal by both typing methods. Remarkably, these three isolates exhibited identical spoligotypes. A total of 92% (24 of 26) of the isolates with one band difference had the same spoligotype as that for the isolates in the main cluster.

Comparison of patients whose isolates were in the main cluster with patients who harbored individual strains did not reveal any significant differences with respect to age, sex, prior tuberculosis treatment, HIV infection status, or demographic data. Of the patients whose isolates were in the main cluster, 82% of patients received a diagnosis of tuberculosis within the

**FIG. 1.** Distribution of isolates (black columns) and patterns (white columns) according to number of IS6110 insertion elements. Numbers over the columns indicate numbers of isolates. Number of isolates per pattern ratio = number of isolates/number of patterns.

**FIG. 2.** Percentage of patients infected with strains in clusters (white columns) and patients infected with strains showing individual patterns (black columns) in three different regions of Namibia.
first 6 months of the study period, whereas the intake of all patients in that time period corresponded to 53% of the study patients \((P < 0.001)\) (Fig. 4). The appearance of this large cluster was not due to cross-contamination, as all but nine isolates were cultured on 28 different days. Together with five of these nine isolates, other samples with isolates with different DNA fingerprints were decontaminated on the same date. In addition, negative controls included with the samples showed no growth after 8 weeks. The equal distribution of positive cultures for all patients during the study period further excludes a bias due to changing sensitivity in the culture isolation methods applied in the study.

**DISCUSSION**

In a prospective study of patients with smear-positive tuberculosis in Windhoek, demographic data were available and isolates could be obtained by culture for 279 (74%) of 378 patients. The finding of smear-positive samples without growth on culture media was probably due to the early start of antituberculosis treatment before obtaining samples for microbiologic analysis. More than 94% of the culture-positive isolates produced a clear DNA fingerprint pattern by mixed-linker PCR. This loss of typeable isolates is comparable to the results obtained by standard restriction fragment length polymorphism analysis with IS6110 (3, 18). The demographic data obtained for our study population including age and sex distribution and HIV-M. tuberculosis coinfection were compatible with data available from the Ministry of Health and Social Services and were in accordance with data from published studies conducted in other countries in the South African region (13, 25).

DNA fingerprint analysis of the cultured isolates showed a cluster frequency of 36%. In comparison, cluster frequencies obtained for isolates from other developing countries ranged between 20 and 30% (24, 27). In contrast to other studies, the cluster analysis in this study was based on a 100% identity of patterns. Thus, the specificity of the analysis was increased, while the sensitivity of our results was sacrificed. As the number of isolates analyzed corresponded to about 46% of all isolates from patients with smear-positive tuberculosis in the study period (including patients with reinfections and chronic infections) the cluster frequency determined in the present study reflects the minimal degree of ongoing spread of tuberculosis.

Typing of the isolates showed less than six IS6110 fragments in 10% of strains or 13% of patient isolates, respectively. In a study by Warren et al. (24) from South Africa, 5% of strains were low-copy-number strains. Studies from other geographic regions revealed that up to 40% of strains had one to five copies of IS6110 (16, 27). The number of isolates-per-pattern ratio showed that the variability of patterns was lowest in this group. However, the small number of isolates in this group did not significantly influence the overall cluster frequency.

A total of 29 clusters with 100% identical DNA fingerprint patterns were identified, with cluster sizes ranging between 2 and 39 isolates. Among these clusters, 9 (31%) included more than 2 isolates (Table 1). Remarkably, the percentage of isolates from HIV-positive patients in small clusters with two isolates was significantly higher than the percentage in clusters with larger numbers of isolates \((P = 0.02)\). However, for patients, infection with a clustered strain was independent of HIV infection status \((P = 1.0)\). In contrast, in industrialized countries HIV-positive patients have predominantly been found to be involved in large outbreaks (1, 4, 5, 18). This

![FIG. 3. Hometowns of 123 patients infected with clustered isolates. Hatched columns represent patients from Windhoek, black columns represent patients from Ovamboland, and white columns represent patients from other regions of Namibia.](http://jcm.asm.org/)

<table>
<thead>
<tr>
<th>No. of clusters</th>
<th>No. of isolates per cluster</th>
<th>No. of IS6110 bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2</td>
<td>3–14</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2–7</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4–7</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>39</td>
<td>11</td>
</tr>
</tbody>
</table>

**FIG. 3.** Hometowns of 123 patients infected with clustered isolates. Hatched columns represent patients from Windhoek, black columns represent patients from Ovamboland, and white columns represent patients from other regions of Namibia.
difference might be explained by incomplete sampling or the poor prognosis for patients with HIV-M. tuberculosis coinfection in developing countries. In addition, patients were not differentiated according to the stage of HIV disease in this study. All clusters with more than three isolates included isolates from patients living in different areas of Namibia, even though 82% of all patients reported Windhoek as their hometown \((P \leq 0.001)\). This observation together with small regional variations in cluster frequencies suggests a lack of hot spots of tuberculosis transmission. This might be due to the high prevalence of tuberculosis. Also, the important function of Windhoek as an economic and administrative center leads to the periodic presence of large numbers of Namibian citizens and might be contributing to the spread of tuberculosis.

A striking result was the detection of an unsuspected outbreak involving 15% of the 263 patients in the study. The isolates in this cluster exhibited 100% identical patterns consisting of 11 IS6110 fragments, thus providing a characteristic and specific DNA fingerprint \((23)\). Patients affected by this outbreak did not differ significantly in their demographic characteristics from patients whose isolates were in other clusters or patients who harbored individual strains. The finding of 25 strains with highly related DNA fingerprint patterns raises the possibility that this outbreak resulted from a strain of tuberculosis endemic to Namibia, as has been described in the context of a low-prevalence country \((2)\). This is further supported by the finding of isolates with identical DNA fingerprints in a rural health district in the north of the country (data not shown). The results of secondary typing by spoligotyping analysis further prove that we have identified the ongoing transmission of a strain endemic to Namibia. This hypothesis of the spread of an endemic strain could be tested by the development of a rapid test for the specific detection of this strain, similar to the approach taken for the W strain in New York City \((15)\). Rapid recognition of patients infected with this highly transmissible strain might represent an important step in breaking the chain of transmission of smear-positive tuberculosis.

The number of patients involved in the outbreak peaked within the first 6 months without increasing the number of smear-positive patients diagnosed with tuberculosis. Thus, the outbreak was missed by classical epidemiological tools such as case detection rate. Similar results have been reported for studies of tuberculosis transmission in low-prevalence countries, even including thorough epidemiological follow-up and contact tracing \((18, 21)\). However, the dynamic of the outbreak with diagnosis of tuberculosis in more than 80% of patients within the first half of the study period is remarkable.

In conclusion, this molecular study of smear-positive tuberculosis in Windhoek, Namibia, confirmed the significant contribution of new infections to the burden of tuberculosis in a high-prevalence country, despite an estimated BCG vaccine coverage rate of more than 80%. In addition to the detection of large, unsuspected outbreaks, the results disproved the hot-spot hypothesis of transmission for Namibia that had been derived from the situation in low-prevalence countries. We suggest that determination of the cluster frequency might serve...
as a useful addition to other more sophisticated ways of estimating the risk of infection, thus providing an independent tool for the identification of changes in the efficacy of tuberculosis control. The equal distribution of cluster frequencies for different areas analyzed in this study will allow the design of follow-up studies at the community level, including classical epidemiological investigations, to test this hypothesis.

ACKNOWLEDGMENTS

This study was supported by research grant 01 KA 9301 from the Bundesministerium für Bildung, Wissenschaft und Technologie and grant HA 1921/3-2,3 from the Deutsche Forschungsgemeinschaft.

REFERENCES


16. Raabow, M. Personal communication.