

# Pulmonary tuberculosis in Harare, Zimbabwe: analysis by spoligotyping

R S Heyderman, M Goyal, P Roberts, S Ushewokunze, S Zizhou, B G Marshall, Robert Makombe, J D A Van Embden, P R Mason, R J Shaw

## Abstract

**Background**—Over the last 10 years there has been a fourfold increase in cases of tuberculosis in Harare, Zimbabwe. The use of molecular epidemiology to understand tuberculosis transmission in this epidemic has been hampered by the availability of suitable culture facilities. A study was therefore undertaken to explore the potential of spoligotyping, a polymerase chain reaction based technique that does not require tuberculosis culture.

**Methods**—Adults attending a chest clinic with clinical or radiological pulmonary tuberculosis and one smear positive sputum were enrolled over one month. Demographic, socioeconomic, and clinical data were gathered using a standardised questionnaire. Molecular fingerprinting of genomic DNA recovered from sputum was performed by spoligotyping.

**Results**—Sixty one subjects (median age 28 years (range 18–73); 61% men) were recruited and 57 provided adequate sputum samples. Recent rural-urban migration or immigration was not common; 40% of subjects lived in crowded living conditions. DNA suitable for spoligotyping was recovered from 28 patients and 20 different genotypes of *Mycobacterium tuberculosis* were identified. Fifteen patients were infected with an *M tuberculosis* strain shared by one or more individuals. Patients infected with a shared spoligotype were not closely linked geographically within Harare, but were more likely to live in overcrowded conditions (69% versus 23%; odds ratio 6.85 (95% CI 1.2 to 47),  $p = 0.026$ ). Analysis of the patients' original rural family homes revealed two geographically related spoligotype clusters.

**Conclusions**—Spoligotyping may yield valuable molecular typing information in populations where tuberculosis culture is not available. This novel technique requires further development and evaluation in larger epidemiological studies.

(Thorax 1998;53:346–350)

**Keywords:** pulmonary tuberculosis; HIV infection; genetic fingerprinting; tuberculosis transmission; Zimbabwe

ment, Zimbabwe, 1995). While co-infection rates with human immunodeficiency virus (HIV) exceeding 40% suggest that HIV-related immunosuppression has played an important part in this dramatic change,<sup>3</sup> a number of other demographic, socioeconomic, and nutritional factors may have influenced the scale and underlying mechanisms of the tuberculosis epidemic.<sup>4–7</sup> It is widely held that the majority of the population in Africa are exposed to and infected with tuberculosis in childhood, but it remains uncertain whether the present epidemic has occurred through reactivation of latent infection in the context of deteriorating immunity or as a result of re-infection of susceptible individuals.<sup>2, 8–11</sup>

Recent epidemiological studies have combined conventional data collection with molecular fingerprinting of *Mycobacterium tuberculosis*. Restriction fragment length polymorphism (RFLP) analysis using fragments of the IS6110 insertion element has proved a reliable and reproducible epidemiological tool with which to explore the mechanisms of transmission of tuberculosis.<sup>8–10–15</sup> However, RFLP is costly, time consuming, and requires culture of the organism prior to analysis. Spoligotyping (from spacer oligotyping) is a newly developed technique based on the polymorphism observed in spacer sequences present within direct repeat regions of the *M tuberculosis* genome.<sup>16</sup> The spacers are amplified using a novel polymerase chain reaction (PCR) technique and then identified using short synthetic oligomeric DNA sequences.<sup>17</sup> Although spoligotyping does not at present have the same discriminatory power as RFLP, the technique does not require culture of the organism and therefore potentially represents a valuable field tool for molecular epidemiology of tuberculosis.

In this prospective cross sectional study we have determined the prevalence of a range of demographic, socioeconomic, nutritional, and clinical factors amongst patients with pulmonary tuberculosis presenting to a Harare chest clinic. We have explored the potential of spoligotyping in a setting where tuberculosis culture is not available, and have used the technique to characterise the diversity of *M tuberculosis* strains underlying tuberculosis in Harare.

## Methods

### PATIENT POPULATION

The study was performed at the larger of the two Harare City Health Department chest clinics which serve a population of approximately 1.25 million people. The clinic is situated in the southern part of the city which

### Department of Medicine

R S Heyderman  
S Ushewokunze  
S Zizhou

### Department of Medical Laboratory Technology

P R Mason

### Medical School, University of Zimbabwe, Zimbabwe

### Department of Respiratory Medicine, National Heart and Lung Institute, Imperial College School of Medicine, London, UK

M Goyal  
B G Marshall  
R J Shaw

### Health Services, City of Harare, Zimbabwe

R Makombe

### Department of Bacteriology, Research Laboratory for Infectious Diseases, National Institute of Public Health and the Environment, Bilthoven, The Netherlands

J D A Van Embden

Correspondence to:  
Dr R S Heyderman,  
Department of Infection and  
Tropical Medicine, Imperial  
College School of Medicine,  
Northwick Park Hospital,  
Harrow, Middlesex HA1  
3UJ, UK.

Received 7 April 1997  
Returned to authors 17 June  
1997

Revised version received 17  
September 1997

Accepted for publication  
5 November 1997

In parallel with the worldwide resurgence of tuberculosis,<sup>1, 2</sup> the incidence of tuberculosis in Harare has increased from 53.5 to 195 per 100 000 population between 1985 and 1995 (Annual Report of Harare City Health Depart-

contains most of the industrial areas and has predominantly high and medium density housing. The city is well served with electricity, piped water, and water-borne sewage. All patients with tuberculosis in this section of the city are referred to the clinic for laboratory confirmation of diagnosis (Ziehl-Neelsen staining of sputum smears), registration and supervision of treatment.

#### ENROLMENT OF PATIENTS, ADMINISTRATION OF QUESTIONNAIRE, AND SPUTUM COLLECTION

In the month of December 1995 consecutive patients above 15 years of age with a clinical presentation and/or a chest radiograph consistent with pulmonary tuberculosis and at least one positive sputum culture for acid alcohol fast bacilli were enrolled into the study following informed consent. A structured questionnaire was designed to assess demography (age, sex, marital status, employment status, income, urban address, location of rural home, country of origin, movement within and between the urban and rural areas); housing (number of people per household, number of people per room, number of children per household, and number of children per room); tobacco and alcohol consumption; and nutrition by assessing the frequency of consumption of four staple proteins (meat, fish, poultry and eggs).<sup>18</sup> The questionnaire was administered in local languages by the clinical investigators and a trained research nurse. Consistency in interview technique was ensured by pre-testing of the questionnaire and regular supervision by one of the principal investigators (RSH). Overcrowding was defined as more than two people per room.<sup>19</sup> In view of the difficulties of quantifying consumption of traditional beers, heavy alcohol consumption was defined as consumption more than three times per week. High protein consumption was defined as consumption of at least one of the recorded protein items on four or more days per week, medium protein consumption was defined as consumption of at least one protein item weekly, and low protein consumption was defined as consumption of at least one protein item less than once weekly.

Following an initial smear positive sputum sample a second sputum specimen was collected from each patient prior to commencement of antituberculous therapy. This sample was decontaminated by boiling in sodium hydroxide. Standard clinical and radiological data were also collected. No additional investigations were performed as part of the study.

This study received ethical approval by the Medical Research Council of Zimbabwe.

#### SPOLOGOTYPE ANALYSIS

Spoligotyping was performed on genomic DNA extracted from decontaminated sputum by standard phenol-chloroform methods.<sup>17 20</sup> The spacers between the direct repeats in the target region were amplified using two 18 nucleotide primers (5'CCAAGAGGGGACG-GAAAC3' and biotinylated 5'GGTTTTGGGTCTGACGAC3'). The PCR products were then hybridised to a Biotyne C membrane.<sup>17</sup> This membrane contains immo-

bilised synthetic oligomeric spacer sequences derived from the direct repeat region of *M. tuberculosis* H37Rv and *M. bovis* BCG. Hybridised DNA was detected using an enhanced chemiluminescence kit (Amersham International plc, Buckingham, UK) which resulted in patterns for each strain reminiscent of a "bar code". A cluster was defined as a group of two or more patients whose molecular fingerprints were identical.<sup>10 11 14</sup> Patients with unrelated spoligotypes were assumed to have different strains of *M. tuberculosis*<sup>17 20</sup> and therefore to have acquired their infection from unrelated sources. Clustered patients were considered likely to have related strains of *M. tuberculosis* which may have arisen from a common or related source.

#### MAPPING OF PATIENTS

At the time of administration of the questionnaire the map coordinates of the subjects' urban and rural homes were recorded using a Harare street atlas (Mapping and Promotional Services, Harare, Zimbabwe) and a map of Zimbabwe (Surveyor General's Office, Harare, Zimbabwe). To analyse the geographical relationship between subjects the coordinates were plotted onto a scattergram using CA-Cricket Graph III software (Computer Associates Inc, Slough, Berks, UK). Individuals within a spoligotype cluster living less than 1 km apart in Harare or having the same rural home were defined as a geographical cluster.

#### STATISTICAL ANALYSIS

The results were compared using the Mann-Whitney U test for continuous variables and the  $\chi^2$  test or Fisher's exact test as appropriate for discrete variables.

## Results

#### OVERALL CHARACTERISTICS OF STUDY

##### POPULATION

In December 1995 141 cases of sputum positive tuberculosis were notified to the City Medical Officer of Health in Harare (City Health Department data). During this period 61 consecutive sputum positive cases (median age 28 years (range 18–73); 37 men) were enrolled into the study from the Harare City Health Department chest clinic. Only one patient refused to participate in the study. Twenty eight of the 61 patients (46%) came from areas of Harare with more than 190 tuberculosis cases per 100 000 population and 10 (16%) of the patients came from an area with 312 tuberculosis cases per 100 000 population (1994 City Health Department Data). Recent rural-urban migration and immigration was uncommon, the median stay in Harare being 11 years (range 0–38) with only four patients (7%) being normally resident outside the city. Two patients were born outside Zimbabwe (length of stay in Harare 17 years and 38 years, respectively). Forty percent of the study population lived in crowded living conditions, exposing 216 children within their households. None of the subjects was homeless. Overall, 57% of the subjects were married, but significantly more men than women had

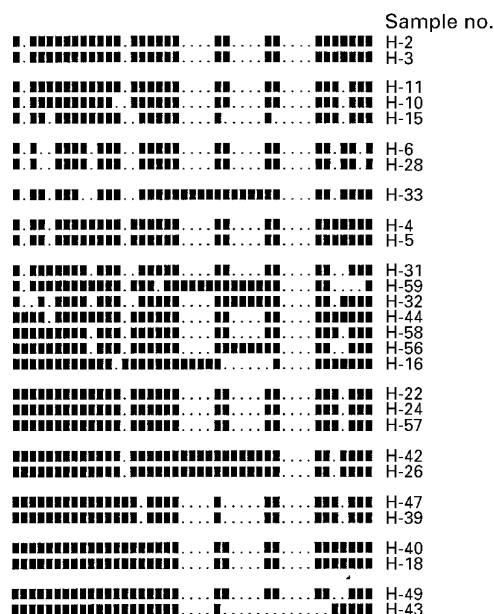


Figure 1 Spoligotype map of the DNA recovered from 28 of 57 sputum samples. The map is a graphical representation of the "bar code" pattern derived from the hybridised DNA detected by chemiluminescence. There were 20 distinct patterns. Unique spoligotype patterns were seen in 13 patients. The remaining 15 patients were contained within seven clusters: six pairs and one group of three individuals sharing identical spoligotypes.

never been married (38% versus 8%;  $p = 0.01$ ). Seventy four percent of the population (81% of men and 63% of women) were in either formal or informal employment, or both. There were no health care workers amongst the subjects recruited. Household incomes were similar to those recorded by the Zimbabwe Central Statistical Office.<sup>21</sup> Only 11% of the subjects admitted to heavy alcohol consumption, and smoking was more common amongst men than women (57% versus 13%;  $p = 0.001$ ). Most of the subjects recruited into the study (83%) had a high protein intake. Clinical and radiological features were similar to those reported in

previous studies from the region.<sup>3 22</sup> Patients presented after a median illness of 68 days (range 14–700), typically with cough (92%), breathlessness (66%), chest pain (74%), and fever (80%). The median body mass index was 18 (range (13–27) and 84% had a BCG scar. Chest radiographs most commonly revealed cavitation (59%) and/or patchy infiltrates (54%).

#### CHARACTERISATION OF POPULATION ACCORDING TO SPOLIGOTYPE

Sputum was obtained from 57 of the 61 patients recruited and genomic DNA suitable for spoligotyping was recovered from 28 of these individuals. Demographic, socio-economic, nutritional, and clinical characteristics did not differ significantly from the overall study population. Spoligotyping of *M tuberculosis* DNA (fig 1) revealed 20 distinct patterns. Unique spoligotype patterns were seen in 13 patients and the remaining 15 patients were contained within seven clusters: six pairs and one group of three individuals sharing identical spoligotypes. Individuals with clustered spoligotypes did not differ significantly from individuals with unique strains with respect to age or sex distribution. Clustered spoligotypes were more commonly identified in patients who lived in overcrowded living conditions (69% versus 23%; Fisher's exact test  $p = 0.026$ ; odds ratio 6.85 (95% CI 1.24 to 47.2)). However, these patients could not be distinguished from patients with unique spoligotypes by other demographic, socioeconomic, nutritional, clinical, or radiological parameters.

#### GEOGRAPHICAL MAPPING OF SPOLIGOTYPES

When cases were analysed according to Harare home address a close geographical relationship between patients with clustered spoligotypes was not identified (data not shown). However, analysis of the location of the patients' original rural family homes revealed one group of three patients and one of two patients with identical spoligotype patterns where the rural but not the urban homes were closely related geographically (fig 2).

#### Discussion

In the context of the rising seroprevalence of HIV in Zimbabwe,<sup>23</sup> tuberculosis has become an increasing health burden. The Harare City Health Department currently spends over 10% of its budget on the treatment and control of tuberculosis. There is therefore a great need to define the epidemiology of the disease in Harare in order to plan strategies effectively and invest scant resources appropriately. For example, a greater understanding of the aetiology of the disease will help to determine whether instituting isoniazid prophylaxis in immunocompromised patients to prevent reactivation or strengthening conventional contact tracing to interrupt transmission has the greatest public health priority.<sup>10 11 24–26</sup>

In this study we have characterised the socioeconomic, clinical, and radiological characteristics of patients with sputum positive

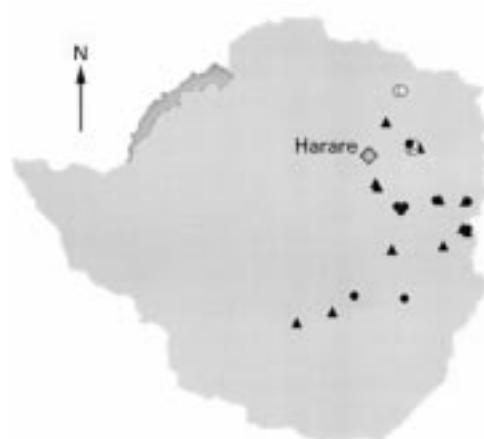


Figure 2 Rural homes: map of Zimbabwe showing the geographical distribution of the clustered (●,  $n = 14$ ) and non-clustered patients (▲,  $n = 12$ ). Two patients had been born outside Harare. On two occasions the rural home of clustered individuals was in the same geographical area (○).

pulmonary tuberculosis presenting to a large Harare chest clinic. Contrary to what might be expected from such an urban population, recent rural-urban migration or immigration were not common.<sup>7</sup> Most of the subjects were not single, homeless, or unemployed. Perhaps surprisingly, while 40% of subjects live in overcrowded conditions, housing and family income compared favourably with Harare census data published two years previously.<sup>21</sup> Heavy alcohol consumption was not implicated in the aetiology of tuberculosis in this population.<sup>5</sup> Previous studies suggest that, although the association between tuberculosis and various socioeconomic factors is complex,<sup>5 7 11 27 28</sup> a clear association remains between low socioeconomic status and a higher incidence of tuberculosis, even when overall poverty levels are improving. Why a relationship between low income and tuberculosis was not observed in Harare is uncertain, but it may be that the high seroprevalence of HIV masks such a relationship.

In recent years molecular fingerprinting techniques have been increasingly employed to define the relationship between mycobacterial strains and to clarify the epidemiology of tuberculosis transmission.<sup>8 10-15</sup> Spoligotyping is a novel rapid technique which can be applied directly to sputum extracts in settings such as Harare where tuberculosis culture is not readily available. In the first study to use this technique under such conditions we have typed *M tuberculosis* strains derived from an epidemiologically defined patient group, demonstrating 20 distinct spoligotype patterns. The identification of 13 patients with unique spoligotype patterns highlights the observation that, in an environment such as Harare where tuberculosis is highly endemic, cases that appear to be geographically linked cannot be assumed to have the same strain of *M tuberculosis*. While unique spoligotypes clearly reflect different strains,<sup>17 20</sup> the spoligotype clustering observed amongst the remaining 15 patients must be interpreted with caution. Two recent RFLP studies concluded that up to 40% of cases of urban adult tuberculosis may be clustered and therefore caused by recent transmission of *M tuberculosis*.<sup>11 14</sup> However, at its current stage of development spoligotyping is not as discriminatory as IS6110 RFLP analysis which has been shown to identify distinct strains within some spoligotype clusters.<sup>20</sup> Further advances in spoligotyping technology are likely to reveal a greater diversity of strains and therefore at present it is only possible to interpret typing results in conjunction with standard epidemiology. Nonetheless, although there was no direct evidence for person to person spread either at home or in the work place (data not shown), we suggest that, for at least some of the potential clusters, recent transmission may account for the association between infection with mycobacteria with a shared spoligotype and overcrowded living conditions. The finding that the patients' original rural family homes, in two potential clusters, were closely related geographically supports the possibility of reactivation of organisms derived in the past from

common sources. Recent transmission following visits to the rural home appears less likely as three of the five subjects denied any recent contact with their rural home (data not shown).

There are a number of limitations to our study. Firstly, we have been unable to determine the interaction between HIV and other aetiological factors because logistic and ethical considerations precluded HIV testing. In the context of tuberculosis it was considered that clinical assessment as an alternative to serology would not have been sufficiently specific.<sup>22</sup> Secondly, the proportion of sputum samples containing sufficient DNA for analysis was low (approximately 50%). Subsequent studies suggest that this relates to variability in the number of mycobacteria in different sputum samples from the same patient, and that the collection technique can be improved (unpublished data). Finally, this study was limited by a small sample size and study period. Since related cases may present months or years apart, the full evaluation of the genetic diversity of *M tuberculosis* in a community such as Harare requires a larger cohort of patients studied over a longer period.

In conclusion, spoligotyping may provide a valuable molecular epidemiological tool for the typing of mycobacterial strains derived from sputum collected in environments where tuberculosis culture is not readily available. This novel methodology is continuously evolving and merits further evaluation in carefully conducted field studies.

We would like to thank Professor Douglas Young for his support and Marc Swan, Buxton Ndemera, and Lovemore Gwanzura for their technical assistance. We thank Dr Lovemore Mbengeranwa and the Health Department of the City of Harare for their support. This project was funded by the University of Zimbabwe Research Board and the British Lung Foundation.

- 1 Haas DW, Des-Prez RM. Tuberculosis and acquired immunodeficiency syndrome: historical perspective on recent developments. *Am J Med* 1994;**96**:439-50.
- 2 De Cock K, Soro B, Coulibaly IM, Lucas SB. Tuberculosis and HIV infection in Sub-Saharan Africa. *JAMA* 1992;**268**:1581-7.
- 3 Houston S, Ray S, Mahari P, *et al*. The association of tuberculosis and HIV infection in Harare, Zimbabwe. *Tubercle Lung Dis* 1994;**75**:220-6.
- 4 Raffalli J, Sepkowitz KA, Armstrong D. Community-based outbreaks of tuberculosis. *Arch Intern Med* 1996;**156**:1053-60.
- 5 Coetzee N, Yach D, Joubert G. Crowding and alcohol abuse as risk factors for tuberculosis in the Mamre population. Results of a case-control study. *S Afr Med J* 1988;**74**:352-4.
- 6 Festenstein F. Spread of tuberculosis within a family. *Lancet* 1981;**i**:603-5.
- 7 Enarson DA, Wang JS, Dirks JM. The incidence of active tuberculosis in a large urban area. *Am J Epidemiol* 1989;**129**:1268-76.
- 8 Godfrey-Faussett P, Githui W, Batchelor B, *et al*. Recurrence of HIV-related tuberculosis in an endemic area may be due to relapse or reinfection. *Tubercle Lung Dis* 1994;**75**:199-202.
- 9 Yang ZH, de Haas P, van Soolingen D, *et al*. Restriction fragment length polymorphism *Mycobacterium tuberculosis* strains isolated from Greenland during 1992: evidence of tuberculosis transmission between Greenland and Denmark. *J Clin Microbiol* 1994;**32**:3018-25.
- 10 Small PM, Hopewell PC, Singh SP, *et al*. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. *N Engl J Med* 1994;**330**:1703-9.
- 11 Alland D, Kalkut GE, Moss AR, *et al*. Transmission of tuberculosis in New York City. An analysis by DNA fingerprinting and conventional epidemiologic methods. *N Engl J Med* 1994;**330**:1710-6.
- 12 Godfrey-Faussett P, Mortimer PR, *et al*. Evidence of transmission of tuberculosis by DNA fingerprinting. *BMJ* 1992;**305**:221-3.
- 13 Daley CL, Small PM, Schecter GF, *et al*. An outbreak of tuberculosis with accelerated progression among persons



- infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms. *N Engl J Med* 1992;**326**:231-5.
- 14 Yang ZH, Mtoni I, Chonde M, et al. DNA fingerprinting and phenotyping of *Mycobacterium tuberculosis* from human immunodeficiency virus (HIV)-seropositive and HIV seronegative patients in Tanzania. *J Clin Microbiol* 1995;**33**: 1064-9.
  - 15 Sepkowitz KA, Friedman CR, Hafner A, et al. Tuberculosis among urban health care workers: a study using restriction fragment length polymorphism typing. *Clin Infect Dis* 1995; **21**:1098-101.
  - 16 Hermans PWM, van Soolingen D, Dale JW, et al. Insertion element IS986 from *M tuberculosis*: a useful tool for the diagnosis and epidemiology of tuberculosis. *J Clin Microbiol* 1990;**28**:2051-8.
  - 17 Kamerbeek J, Schouls L, Kolk A, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 1997;**35**:907-14.
  - 18 Ralph A. Methods for dietary assessment. In: Garrow J, James W, ed. *Human nutrition and dietetics*. 9th ed. Edinburgh: Churchill Livingstone, 1993: 777-81.
  - 19 Schoeman JH, Westaway MS, Neethling A. The relationship between socioeconomic factors and pulmonary tuberculosis. *Int J Epidemiol* 1991;**20**:435-40.
  - 20 Goyal M, Saunders N, van Embden JDA, et al. Comparisons of IS6110-RFLP and spoligotyping for the ability to type strains of *M tuberculosis*. *J Clin Microbiol* 1997;**35**:647-51.
  - 21 Central Statistical Office. Income consumption and expenditure survey report. Central Statistical Office, Zimbabwe, 1994.
  - 22 Elliott AM, Luo N, Tembo G, Halwiindi B, et al. Impact of HIV on tuberculosis in Zambia: a cross sectional study. *BMJ* 1990;**301**:412-5.
  - 23 Mbizvo MT, Machekano R, McFarland W, et al. HIV seroincidence and correlates of seroconversion in a cohort of male factory-workers in Harare, Zimbabwe. *AIDS* 1996; **10**:895-901.
  - 24 Daley CL. Tuberculosis recurrence in Africa: true relapse or re-infection? *Lancet* 1993;**342**:756-7.
  - 25 Pope JW, Jean SS, Ho JL, Johnson WD. Effect of isoniazid prophylaxis on the incidence of active tuberculosis and progression of HIV infection. *Lancet* 1993;**342**:268-72.
  - 26 Porter JDH, McAdam KPWJ. Aspects of tuberculosis in Africa 1: tuberculosis in Africa in the AIDS era - the role of chemoprophylaxis. *Trans R Soc Trop Med Hyg* 1992;**86**: 467-9.
  - 27 Reichman LB, O'Day R. Tuberculous infection in a large urban population. *Am Rev Respir Dis* 1978;**117**:705-12.
  - 28 Tignor MM. Socioeconomic factors in tuberculosis. *N Engl J Med* 1981;**304**:431-2.