Genetic Diversity of *Mycobacterium tuberculosis* Strains in Salvador, Bahia State, Brazil: needs for further research and implications for regional policy

A. C. do Nascimento Silva*1,4, L. Ferrazoli2, J. Neves Reis3, S. Martins Pereira4, E. L. Andrade Mota4 and M. Galvão Reis1

1The Gonçalo Muniz Research Foundation/The Oswaldo Cruz Foundation,(CPGM/Fiocruz), Avda. Waldemar Falcão, 121, Brotas, 40296-710, Salvador, Bahia, Brazil.
2The Adolf Lutz Institute/the Secretary of Health of the State of São Paulo (IAL/SES-SP), Avda. Dr. Arnaldo, 355 - Cerqueira César, 01246-000, São Paulo, Brazil.
3Departament of Pharmacy, University Federal of Bahia, Avda. Barão do Jeremoabo, 147, Ondina, 40.170-115, Salvador, Bahia, Brazil.
4Collective Health Institute, University Federal of Bahia, Avda. Basílio da Gama, s/n, Campus Universitário Canela, 40110-040, Salvador, Bahia, Brazil
*Corresponding author

Little is known about transmission dynamics and genetic diversity of tuberculosis (TB) in Salvador, Bahia State. TB is not fought only with medication, is strongly related to social factors, and is an indicator of poverty. This study constitutes a first attempt to describe the genetic population structure of *Mycobacterium tuberculosis* circulating in Salvador, Brazil. A total of 56 confirmed cases of pulmonary tuberculosis, identified between March and June 2008, were analyzed using restriction fragment length polymorphism (IS6110-RFLP). The study population was characterized by a predominance of males (71.43%) over 30 years of age (68.75%). Forty-one isolates were found to belong to a single pattern (73.2%), while 15 (26.7%) were found in group patterns, forming six clusters. The higher level of diversity observed is much more suggestive of endogenous reactivation than recent transmission.

Keywords: Tuberculosis; Transmission; Molecular Epidemiology; Public Health

**1. Mycobacterium tuberculosis: a problem of health public and old challenges of the disease**

TUBERCULOSIS (TB) is a public health issue worldwide, with 9.27 million new cases and 1.3 million deaths in 2007 (WHO, 2011). World Health Organization recommends, among the strategies for reducing the burden of TB, the expansion of knowledge about the identification of resistance of *M. tuberculosis* to drugs, poverty and the relative populations vulnerable to the disease, and the incentive to conduct research. Brazil is one of the 22 countries with the highest rates of tuberculosis worldwide, 100,000 new cases of TB being registered annually. In 2006, 6,500 cases of tuberculosis were registered in the state of Bahia, which is currently the third state in absolute numbers of cases in Brazil. Of these, 2,300 cases occurred in the city of Salvador (Brazil, 2006).

It is known that socio-economic factors (social class, education, family income, age, race) and etiological (agent drug resistance) are related to the development of TB. And notably, studies report a strong negative association between socioeconomic characteristics and disease (Silva *et al.*, 2011), and still others show that the increase in poverty, the number of homeless, urbanization, malnutrition and loss of quality of programs TB control, have contributed to the high incidence of the disease (Dye & Williams, 2010).

Recently, molecular techniques have been used to explain epidemiological aspects of TB, and have contributed towards improving understanding the transmission dynamics of the disease (Small *et al.*, 1999).

No many studies have yet about on the epidemiologic and genetic diversity of circulating *Mycobacterium tuberculosis* in Salvador, a city in northeastern Brazil, in which the incidence of this disease is high. The aim of the present study was to describe the genetic population structure of *M. tuberculosis* circulating in Salvador.

**2. TB and evaluation of the resistance: some hypothesis**

Some authors argue that the changes in the patterns of transmission of *M. tuberculosis* attributed to the prevalence of drug-resistant TB associated with low socioeconomic conditions of individuals and deficiency of health public services, and it is assumed that the majority of TB cases is the result of reactivation (Small *et al.*, 1999; Dye & William, 2010). Therefore, it is noteworthy that TB disease can be the manifestation of both the primary form of the disease (primary infection) and the reactivation of a latent infection acquired in the past (Silva *et al.*, 2011).
Considering that, TB remains a major cause of death in adults by an infectious agent, the concern in some countries has led to the reformulation of the antituberculosis programs with application of the measures, supervised treatment, investment in discovery new techniques for accelerating the diagnosis, and the use of molecular epidemiology studies transmission (WHO, 2011). Regarding the definition of the objectives of molecular epidemiology, some authors (van Embden et al, 1993; Borges et al, 2004) emphasize, from the identification of microorganisms causing infectious disease, its transmission route until the use of molecular markers to describe the distribution of strains in populations and verify the risk factors of hosts and parasites, the spread of disease.

The social and epidemiological implications of the lack of this disease are of great relevance to our environment and non-adherence to treatment, with variations peculiar to each region, but with unacceptable levels in many parts of Brazil, and is an important cause of treatment failure and development resistant strains (Brazil, 2006).

Note also that, as in many other countries, there is a discrepancy between the number of reported cases and estimated incidence (WHO, 2011). This fact, associated with high mortality rate, indicates the condition of delay in diagnosis and initiation of treatment, also limiting factors to proper TB control.

### 3. Evolution of the Genetic Diversity of Mycobacterium tuberculosis in Salvador, State Bahia: analysis of the RFLP, needs for further research and epidemiological dates

Molecular typing of *M. tuberculosis* isolates is an efficient tool for studying the dynamics of TB transmission. The current standardized strain typing method is based on Southern blot analysis of the restriction fragment length polymorphisms (RFLP) associated with the insertion sequence IS6110, in accordance with the standardized protocol of van Embden et al (1993).

In recent research in the Brasil (Silva et al, 2011), and as the first study conducted in Bahia, was used the RFLP analysis of clinical isolates of *M. tuberculosis* in combination with clinical and epidemiological analysis to describe the genetic population structure of this agent, as regards recent infection or reactivation, among TB patients in referral hospital of the Salvador City.

The genetic profile of each strain was analyzed and visualized with GelCompar II version 4.0 (Applied Maths, Kortrijk, Belgium) using *M. tuberculosis* 14323 as the reference strain. All the strains for which the similarity coefficient was >90% were considered part of a “family”. Within the molecular typing system, cluster strains were defined as those with the same RFLP pattern (number of bands and the position of the fragment) in two or more isolates of *M. tuberculosis*. Isolates whose RFLP patterns were distinctly different from any of the other patterns identified were considered as non-clusters. An isolate forming part of a cluster pattern was defined as representing a recently acquired infection and a standard non-cluster isolate as representing reactivation of an infection acquired in the past (Ferrazoli, 2000).

This study included a total of 56 confirmed cases of pulmonary tuberculosis, from patients over 18 years of age diagnosed between March and June, 2008 and receiving care at the Otavio Mangabeira Specialist Hospital (HEOM/SESAB), a referral unit for pneumopathies situated in the city of Salvador, Brazil. The research project was approved by the Health Surveillance National Agency/ Ministry of Health in Brasilia, and was approved by the ethics committee of the Collective Health Institute, Federal University of Bahia.

The *M. tuberculosis* isolates were collected and processed from sputum in accordance with the protocol established at the Mycobacteriology Department of the Professor Gonçalo Muniz Central Public Health Laboratory (LACEN-BA). Exploratory data management was performed using the STATA version 7.0 (STATA, College Station, USA). Prevalence ratios (PR) and their respective 95% confidence intervals (95%CI) were calculated.

Of the 56 patients, the majority (71.43%) were male and over 30 years of age (68.75%) (Table 1). Regarding ethnicity, most of the patients (80%) classified themselves as non-white. With respect to education, 63.89% of the patients were illiterate or had failed to complete primary school. A quarter of the patients did not reside in Salvador and 23.40% had been submitted to a Category III treatment regimen (indicated in Brazil for the treatment of cases in which treatment with previous regimens I, IR [recurrence after cure] and II, has failed).

A positive association was found between being male (PR = 1.87; 95%CI: 0.62-5.63), age > 30 years (PR = 1.67; 95%CI: 0.54-5.11), self-classification as non-white (PR = 3.00; 95%CI: 0.46 – 19.35), being illiterate or having failed to complete primary school (PR = 1.70; 95%CI: 0.55-5.17); not living in Salvador (PR = 1.67; 95%CI: 0.69-4.00); having been submitted to a Category III treatment regimen for tuberculosis (PR = 1.31; 95%CI: 0.51-3.36) and the formation of cluster patterns. The associations found were not statistically significant.

In the analysis of the RFLP, of the 56 patients, Forty-one of these isolates (73.21%) had single patterns, while 15 (26.78%) had group patterns. Overall, all the isolates had 2-16 copies of IS6110, and 10.7% of all the samples analyzed had 6 or fewer copies of this repetitive element. 15 isolates formed 6 distinct clusters denominated CL
I, CL II, CL III, CL IV, CL V and CL VI. Each cluster was composed of 2 isolates except for CL V, which was composed of 5 isolates (Figure 1).

Thus, Silva et al (2011) emphasize that the principal finding refers to the fact that 26.78% of isolates formed cluster patterns, suggesting recent transmission of TB in this group of patients (Table 1). This proportion is slightly higher than the findings of studies conducted in Rio de Janeiro using similar methodology (19%) and similar to the results of studies carried out in isolates in Porto Alegre and São Paulo (Ferrazoli, 2000; Fandinho et al, 2000) of 29.1% and 33%, respectively. In the present study, the percentage of cluster strains found was slightly lower than those reported in studies conducted in France (35%) and San Francisco (40%) (Silva et al, 2011).

The proportion of non-cluster strains was 73.21%, suggesting the need to improve TB control, directing preventive actions towards dealing with its reactivation. In Brazil, recent studies show similar findings, the majority of cases, around 70%, resulting from old infections (non-cluster strains), showing that the occurrence of new cases in the country originates predominantly from endogenous reactivation (Ferrazoli, 2000; Borges et al, 2004).

In study of Silva et al (2011), the patients were found to be concentrated heterogeneously in the capital city, within the greater metropolitan region of Salvador and also in other towns within the state, in addition to two samples from patients living in another state (São Paulo). This finding shows mobility of genetic population with tuberculosis in different regions and contribute to implement adequate measures either in the case of recent infection or reactivation of an old infection.

One of the limitations of the present study refers to the small sample size, which did not permit statistically significant associations to be obtained. Another limitation refers to the fact that the findings on the proportion of cluster patterns may be affected by the incidence of TB, by the heterogeneity of the population and by the fraction of the sample included in the study, aspects that would contribute towards obtaining a more accurate calculation of recent infection in a given community (Glynn, 1999).

Despite these limitations, a wide diversity was found in the strains of M. tuberculosis analyzed in TB patients in this referral hospital in the city of Salvador. Nevertheless, the approach can support public health policies in the definition of preventive and control TB program in health services, the dynamics of disease transmission in high risk subgroups, even when performed at the institutional level. Further studies should be carried out to acquire more data on this subject in Salvador, Brazil, enabling associations to be sought among the various co-variables and the proportion of cluster strains in this city.
Table 1  Bivariate analysis of the association between the co-variables and RFLP pattern (cluster or non-cluster) of *M. tuberculosis* isolates from patients infected with TB in Salvador, Brazil.

<table>
<thead>
<tr>
<th>Predictive variables</th>
<th>Cluster (%) (N = 15)</th>
<th>Non-cluster (%) (N = 41)</th>
<th>Total (%) N = 56</th>
<th>PR b</th>
<th>95%CI b</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>14 (35.00)</td>
<td>26 (65.00)</td>
<td>40 (71.43)</td>
<td>1.87</td>
<td>0.62-5.63</td>
</tr>
<tr>
<td>Female</td>
<td>3 (18.75)</td>
<td>13 (81.25)</td>
<td>16 (28.57)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;30</td>
<td>11 (33.33)</td>
<td>22 (66.67)</td>
<td>33 (68.75)</td>
<td>1.67</td>
<td>0.54-5.11</td>
</tr>
<tr>
<td>18-29</td>
<td>3 (20.00)</td>
<td>12 (80.00)</td>
<td>15 (31.25)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Self-reported ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-white</td>
<td>12 (42.86)</td>
<td>16 (57.14)</td>
<td>28 (80.00)</td>
<td>3.00</td>
<td>0.46-19.35</td>
</tr>
<tr>
<td>White</td>
<td>1 (14.29)</td>
<td>6 (85.71)</td>
<td>7 (20.00)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Education level d</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiterate/did not finish primary school</td>
<td>9 (39.13)</td>
<td>14 (60.87)</td>
<td>23 (63.89)</td>
<td>1.70</td>
<td>0.56-5.17</td>
</tr>
<tr>
<td>Completed primary school/ attended at least some high school</td>
<td>3 (23.08)</td>
<td>10 (76.92)</td>
<td>13 (36.11)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Residence in Salvador</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>5 (41.67)</td>
<td>7 (58.33)</td>
<td>12 (25.00)</td>
<td>1.67</td>
<td>0.69-4.00</td>
</tr>
<tr>
<td>Yes</td>
<td>9 (25.00)</td>
<td>27 (75.00)</td>
<td>36 (75.00)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment regimen</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category III (therapeutic failures)</td>
<td>4 (36.36)</td>
<td>7 (63.64)</td>
<td>11 (23.40)</td>
<td>1.31</td>
<td>0.51-3.36</td>
</tr>
<tr>
<td>Category I (new cases of TB)</td>
<td>10 (27.78)</td>
<td>26 (72.22)</td>
<td>36 (76.60)</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

*a* Total varies in accordance with the availability of data.

PR: prevalence ratio; CI: confidence interval.

*b* PR and CI were calculated after excluding missing data.

*c* Missing data: Age (missing in 8 cases); Self-reported ethnicity (missing in 21 cases); Education level (missing in 20 cases); Residence in Salvador (missing in 8 cases); Treatment regimen (missing in 9 cases).

*d* Classification adopted in the registers of the Notification and Investigation of Cases of Tuberculosis (SINAN/MS).
Figure 1 Dendrogram representing the different RFLP patterns found in isolates of *M. tuberculosis* from patients receiving care at a referral hospital, according to the identification number of the sample, date of isolation, sex, age, place of residence and treatment regimen. Salvador, Bahia State, Brazil.
REFERENCES


