Invest Clin 55(4): 332 - 351, 2014

Mycobacterium tuberculosis population structure and molecular epidemiological analysis in Sucre municipality, Miranda state, Venezuela.

Margareth A. Patiño[1], Edgar Abadía[1]. Solalba Gómez[2], Mailis Maes[3], Mariana Muñoz[1], Daniela Gómez[1], Patricia Guzmán[1], María Victoria Méndez[1], Carmen Ramirez[4], Mercedes España[4], Jacobus de Waard[3] and Howard Takiff [1]

[1]Laboratorio de Genética Molecular, Instituto Venezolano de Investigaciones Científicas. Altos de Pipe. Venezuela.
[2]Laboratorio "Luis Quiñones". TB Sucre, Distrito #7 de Salud, Edo. Miranda, Venezuela. Laboratorio de Tuberculosis del Instituto de Biomedicina. Caracas, Venezuela.
[3]Laboratorio de Tuberculosis del Instituto de Biomedicina. Caracas, Venezuela.
[4]Instituto de Tuberculosis. Algodonal, Caracas, Venezuela.

Abstract. Sucre municipality is a large, densely populated marginal area in the eastern part of Caracas, Venezuela that consistently has more cases of tuberculosis than other municipalities in the country. To identify the neighborhoods in the municipality with the highest prevalence of tuberculosis, and determine whether the *Mycobacterium tuberculosis* strain distribution in this municipality is different from that previously found in the western part of Caracas and the rest of Venezuela, we collected data on all tuberculosis cases in the municipality diagnosed in 2005-6. We performed two separate molecular epidemiological studies, spoligotyping 44 strains in a first study, and spoligotyping 131 strains, followed by MIRU-VNTR 15 on 21 clustered isolates in the second. With spoligotyping, the most common patterns were Shared International Type SIT17 (21%); SIT42 (15%); SIT93 (11%); SIT20 (7%); SIT53 (6%), a distribution similar to other parts of Venezuela, except that SIT42 and SIT20 were more common. MIRU-VNTR 15 showed that six of seven SIT17 strains examined belonged to a large cluster previously found circulating in Venezuela, but all of the SIT42 strains were related to a cluster centered in the neighborhoods of Unión and Maca, with a MIRU-VNTR pattern not previously seen in Venezuela. It appears that a large percentage of the tuberculosis in the Sucre municipality is caused by the active transmission of two strain families centered within distinct neighborhoods, one reflecting communication with the rest of the country, and the other suggesting the insular, isolated nature of some sectors.

Corresponding author: Howard Takiff. Laboratorio de Genética Molecular, CMBC, IVIC. Km 11. Carr. Panamericana, Caracas 1020 A, Venezuela. E-mail: htakiff@gmail.com

Estructura poblacional y análisis epidemiológico molecular de Mycobacterium tuberculosis en el municipio Sucre, estado Miranda, Venezuela.

Invest Clin 2014: 55(4): 332 - 351

Palabras clave: epidemiología molecular, genotipo, *spoligotyping*, *Mycobacterium tuberculosis,* MIRU-VNTR, epidemiología, tuberculosis.

Resumen. El municipio Sucre es un área densamente poblada del este de Caracas, Venezuela, con más casos de tuberculosis que otros municipios del país. Para establecer las áreas en el municipio Sucre con la mas alta prevalencia de tuberculosis y determinar sí la distribución de cepas de *Mycobacterium tuberculosis* es diferente de las encontradas previamente en el Oeste de Caracas y el resto de Venezuela, se recolectaron los datos de todos los casos diagnosticados de tuberculosis en el municipio en el 2005-6. Además, se aplicaron dos estudios de epidemiología molecular, el primero con 44 aislados en 2006 y el segundo con 131 aislados del 2006 al 2011, todos caracterizados por *spoligotyping*. Fue aplicada la técnica MIRU VNTR15 sobre 21 aislados agrupados. Con *spoligotyping*, los patrones encontrados fueron SIT17 (21%); SIT42 (15%); SIT93 (11%); SIT20 (7%); SIT53 (6%), presentando una distribución similar en otras partes de Venezuela, con la diferencia de que el SIT42 y el SIT20 fueron comunes en el municipio. MIRU VNTR15 mostró que seis de las siete cepas SIT17 pertenecían a un gran grupo encontrado previamente en Venezuela, mientras las cepas SIT42, estaban relacionados a un grupo concentrado en los *Barrio*s Unión y Maca, con un patrón MIRU VNTR no visto previamente en Venezuela. Los resultados indicarían que un gran porcentaje de tuberculosis en el municipio Sucre es causada por transmisión activa de dos familias, una reflejando comunicación con el resto del país, y otra sugiriendo que es un aislado propio de algunos *Barrio*s del municipio.

Recibido: 8-1-2014 Aceptado: 19-6-2014

INTRODUCTION

Despite the availability of effective anti-tuberculosis chemotherapy for over 50 years, Tuberculosis (TB) remains a major global health problem. Each year there are nearly 9 million cases and an estimated 1.7 million deaths worldwide, and the number of new cases is higher than at any other time in history (1). Fortunately, the incidence of TB in Venezuela is only moderate, with about 33 cases per 100.000 of inhabitants per year (2) but marginal and indigenous populations have TB rates as much as two to three times higher than the national TB rate (data from the Venezuelan National TB program).

To combat TB there is a need for new therapeutics, diagnostics and vaccines, in conjunction with improved operational guidelines to enhance current TB control strategies (3). The only markers available to study TB epidemiology in the 1980's were drug susceptibility profiles and phage types (4-6), but since the 1990's new markers and new methods have opened the doors to the molecular epidemiology of *M. tuberculosis* and made it possible to understand the population structure of this successful pathogen (7-15). Particular strains or strain families have been found to be disproportionate contributors to both local and global TB, and the careful study of dominant compared to rare strains should provide clues to the determinants of their success and possibly yield strategies for control (16, 17).

Molecular epidemiological analysis of TB epidemics in different populations usually encounters a number of phylogenetically unrelated lineages (18). For pathogenic microorganisms, fitness generally refers to heritable traits among members of a given species or phylogenetic lineage that determine its relative growth or pathogenic capacity compared to all other members (3). Different clinical *M. tuberculosis* strains can induce different host immune responses, as well as differences in relative pathogenesis and virulence that can be demonstrated in animal models (19-22). Knowledge of the local *M. tuberculosis* epidemiology and molecular epidemiology should help to define outbreaks and identify dominant clones that could be more virulent, more easily transmissible or drug resistant, and which may be amenable to tailored strategic interventions to control their spread.

We previously characterized the molecular epidemiology of *M. tuberculosis* strains isolated in diverse regions of Venezuela, including many strains isolated from the marginal neighborhoods in the western part of the capital Caracas (23). We found dominant Shared International Types (SITs) by spoligotyping and identified a few highly prevalent genotypes by MIRU-VNTR analysis. Caracas, the capital of Venezuela, is a city of about 6 million inhabitants with several geographically distinct marginal regions. In the present work we sought to extend our study to the eastern part of Caracas located in the Sucre municipality of the Miranda state, a very large, densely populated marginal area with about a 1,133,000 inhabitants, many with low socioeconomic status, and more tuberculosis cases per year than any other municipality in the country. The aim of this work was to determine which neighborhoods have the greatest burden of tuberculosis cases and to delineate the population structure of *M. tuberculosis* strains circulating in the Sucre municipality. Furthermore, we sought to determine whether the distribution of strain families is different from that previously found in the western part of Caracas, and whether particular genotypes predominated in different neighborhoods, commonly termed "*Barrios,*" of the Sucre municipality.

MATERIALS AND METHODS

Cases 2005-2006

To determine the sectors of Sucre municipality with the most cases of tuberculosis, a retrospective study was undertaken to identify all of the residents of this municipality who had been diagnosed with tuberculosis in 2005-2006. This was done by searching the records of: the tuberculosis control program of Health District #7, which includes Sucre municpality; the Venezuelan National TB Control Program; the TB diagnostic laboratory of the Venezuelan National Institute of Hygiene; and the Tuberculosis Laboratory of the Institute of Biomedicine in Caracas. A total of 981 cases were identified, 492 diagnosed in 2005, and 489 in 2006. Of these, it was possible to determine residence within a particular sector or *Barrio* of Sucre municipality for 829 cases.

Clinical isolates

A preliminary molecular epidemiology study performed in 2006 included a convenience sample of 44 available *M. tuberculosis* strains from patients with pulmonary tuberculosis isolated from January-July 2006, either in the bacteriology lab of the Petare Maternal-Infant Hospital "Joel Valencia Parpacén," or in the National Mycobacterial Reference lab. The strains were spoligotyped and categorized by the *Barrios* listed as the patients' residence (Table I).

In 2009-2010, a more extensive molecular epidemiology study was performed. The strains analyzed were a convenience sample of 131 available *M. tuberculosis* strains isolated from 2006 to 2010. All patients were from municipality Sucre and diagnosed either in the Laboratory Luis Quiñones, located in the annex of the Petare Maternal-Infant Hospital, or the TB Laboratory of the Instituto de Biomedicina. Sputum samples were treated with the Petroff method (24) and inoculated onto Lowenstein Jensen slants, or with the Kudoh Method and inoculated onto Ogawa slants (25), and incubated at 37°C for 3 to 6 weeks until colonies were visible. Patient location and strain distribution were as follows: Petare 29, Unión 19, San Blas 13, Maca 11, Carpintero 7, José Félix Rivas 6, Filas de Mariche 5, Mirador del Este 5, Agricultura 3, 5 de Julio 2, Antonio José de Sucre 2, El Campito 2, Isaías Medina 2, La Línea 2, San Isidro 2, Guaicoco 2, La Bombilla 2, La California 2, La Dolorita 2, 19 de Abril 1, Bolívar 1, Cuatricentenario 1, El Dorado 1, El Nazareno 1, San Miguel 1, Vista Alegre 1, Boleíta 1, Buena Vista 1, Campo Rico 1, El Llanito 1, La Urbina 1 and Turumo 1 (Fig. 1 and Table I).

DNA extraction

Briefly, 2 to 3 loopfulls of colonies from original slants were resuspended in 250 μL of Tris- EDTA 1X buffer and heated at 95°C for 15 minutes in 1.5 mL plastic tubes. The tubes were then centrifuged at 13.000 rpm and 5-10 μL of the supernatant was used for each PCR reaction.

Spoligotyping

The CRISPR region (26, 27) was used as a target for membrane-based Spoligotyping, as previously described (9, 23, 28). The hybridization patterns for the 43 spoligotype spacers were transcribed from the films into a previously formatted and labeled Excel spreadsheet and double-

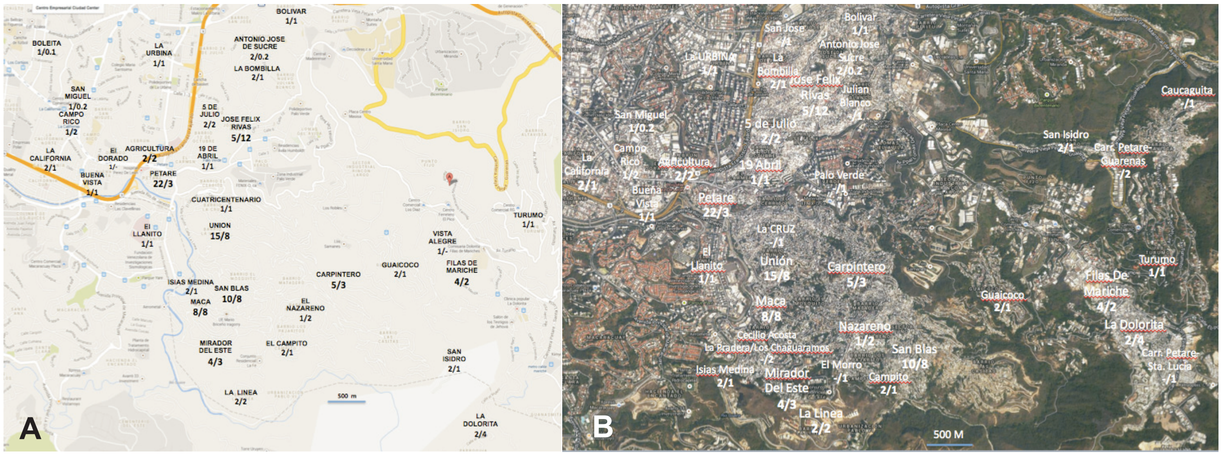


Fig. 1. A. Map; and B. Satellite photo of municipio Sucre in Miranda state, showing the sectors or "Barrios" with the highest number of cases of tuberculosis. The numbers below names of the sectors indicate the percentage of strains from that sector included in the molecular epidemiology study of 131 isolates/the percentage of TB cases diagnosed in patients from that sector in 2005-2006, out of the total 839.

TABLE I
COMPARISON OF THE DISTRIBUTION OF ISOLATES BY SECTOR OF PATIENT RESIDENCE IN THE THREE STUDIES

	2005-6 (% of 829)	2006 (% of 44)	2006-2010 (% of 131)		2005-6 (% of 829)	2006 (% of 44)	2006-2010 (% of 131)
Jose Felix Rivas	11.5%	15.9%	4.6%	La Bombilla	0.7%	-	1.5%
San Blas	7.7%	2.3%	9.9%	Lebrun	0.7%	-	-
Barrio Union	7.5%	9.1%	14.5%	Mezuca	0.7%	-	-
Maca	7.5%	2.3%	8.4%	Av. Romulo Gallegos	0.6%	-	-
La Dolorita	3.5%	4.5%	1.5%	Barrio Cuatricentenario	0.6%	-	0.8%
Carpintero	3.3%	2.3%	5.3%	Barrio La Lucha	0.6%	-	-
Mirador del Este	2.7%	9.1%	3.8%	El Marquez	0.6%	-	-
Petare	2.7%	20.5%	22.1%	Guaicoco	0.6%	-	1.5%
5 Julio	2.4%	2.3%	1.5%	Mirador del Cafetal	0.6%	-	-
Carretera Petare-Guarenas	2.4%	-	-	12 Febrero (Nazareno)	0.6%	-	-
Barrio La Linea	1.9%	2.3%	1.5%	12 Octubre	0.6%	2.3%	-
Mariche	1.8%	-	3.8%	24 Julio	0.6%	-	-
Barrio Agricultura	1.7%	-	2.3%	Barrio Nuevo	0.6%	-	-
Campo Rico	1.7%	4.5%	0.8%	Los Ruices	0.6%	-	-
Barrio Nazareno	1.6%	-	0.8%	Paulo VI	0.6%	-	-
Carretera Petare-Santa Lucia	1.4%	4.5%	-	Zona Colonial Petare	0.6%	2.3%	-
Caucaguita	1.4%	-	-	24 Marzo	0.4%	-	-
Palo Verde	1.4%	2.3%	-	Baloa	0.4%	-	-
Turumo	1.4%	-	0.8%	Sebucan	0.4%	-	-
Barrio El Morro	1.3%	-	-	Barrio Antonio Jose Sucre	0.2%	-	1.5%
Barrio La Cruz	1.3%	-	-	Barrio El Cerrito	0.2%	2.3%	-

TABLE I (*Continuación*)

	2005-6 (% of 829)	2006 (% of 44)	2006-2010 (% of 131)		2005-6 (% of 829)	2006 (% of 44)	2006-2010 (% of 131)
La Urbina	1.3%	-	0.8%	Barrio El Esfuerzo	0.2%	-	-
Barrio San Jose	1.2%	2.3%		Barrio Piritu	0.2%	-	-
San Isidro	1.2%	2.3%	1.5%	Barrio Valle Alto	0.2%	-	-
El Campito	1.1%	-	1.5%	Los Cortijos	0.2%	-	-
La California	1.1%	2.3%	1.5%	San Miguel	0.2%	-	0.8%
Barrio Isaias Medina	1..0%	-	1.5%	Urb Miranda	0.2%	-	-
Buena Vista	1.0%	-	0.8%	Vegas de Petare	0.2%	-	-
Cecilio Acosta (Sector)	1.0%	-	-	Vista Hermosa	0.2%	-	-
Santa Cruz del Este	1.0%	-	-	Los Chaguaramos	0.2%	-	-
19 Abril	0.8%	-	0.8%	Barrio El Carmen	0.1%	-	-
Barrio Bolivar	0.8%	2.3%	0.8%	Barrio El Milagro	0.1%	-	-
Barrio La Alcabala	0.8%	-			0.1%	-	-
Barrio Las Praderas	0.8%	-			0.1%	-	0.8%
Bello Campo	0.8%	-			0.1%	-	-
El Llanito	0.8%	2.3%			0.1%	-	-
Julian Blanco	0.8%	-			0.1%	-	-
Urb Manuel Gonzalez Carvajal	0.8%	-			0.1%	-	-
Barrio El Encantado	0.7%	-			-	-	0.8%
Barrio San Pascual	0.7%	-			-	-	0.8%

checked to insure the accuracy of the resultant table (23). Spoligotypes were compared with the most recent release of the International Spoligo Database, SITVITWEB (29). Our national *Mycobacterium tuberculosis* database (MTBdbase) was used for strain comparisons in order to find genotypes circulating in Venezuela, and SPOTCLUST (http://tbinsight.cs.rpi.edu/run_spotclust.html) (30), was used to tentatively assign phylogeographic lineages to spoligotype patterns not reported in SITVITWEB.

Homemade membranes for hybridization

Byodine-C Nylon membranes were prepared for spoligotyping as previously described (9). The oligonucleotide probes were generally the same as described (9), with some corrected sequences (31). Homemade membranes were subject to quality control measures to insure that adequate hybridization signals could be obtained from all spacer oligonucleotides fixed to the membranes.

MIRU-VNTR genotyping

Mycobacterial interspersed repetitive-unit-variable-number tandem-repeat analysis was performed on strains belonging to the principal clusters defined by spoligotyping to accurately establish strain identity. Fifteen loci were amplified using multiplex PCR and run with Rox labeled Map Marker 1000 size standards (Bioventures) on an ABI3130xl sequencer as described previously (26). Sizing of the PCR fragments and assignment of MIRU- VNTR alleles were done with the Gene Mapper software package (PE Applied Biosystems).

Dendrogram

The online MIRU-VNTR*plus* tool (32, 33) was used with the UPGMA algorithm (34) for cluster analysis. The UPGMA Dendrogram treated the VNTR values as categorical data (35) and the spoligotyping data was treated as binary with the Jaccard's index (36).

RESULTS

This report describes three separate studies of the traditional and molecular epidemiology of tuberculosis in the Sucre municipality of the Miranda state of Venezuela. The first study retrospectively searched the records of the local and national TB control programs, as well as the diagnostic mycobacteriology labs in Caracas and found 829 patients who were diagnosed with tuberculosis in 2005-2006 and had an address within the municipality. A pilot molecular epidemiology study performed spoligotyping on 44 strains isolated in 2006 from patients living in the municipality, and then a larger study on molecular epidemiology analyzed 131 *M. tuberculosis* clinical strains isolated from Sucre residents between 2006-2010. The distribution of the cases by the sector, or "*Barrio*" of residence in the municipality is shown in Table I and Fig. 1. Combining the three studies, the patients came from 80 different geographic sectors of the municipality, indicating that tuberculosis appears to be an endemic problem in the entire municipality. In all three studies, however, more than 45% of the cases resided in the same sectors: Petare, Unión, San Blas, Maca, Carpintero, José Félix Rivas, Filas de Mariche and Mirador del Este. Fig. 1A is a conventional map of the region, and Fig. 1B is a satellite map of the municipality (Google maps), showing the percentages of cases from the two larger studies that resided in the different sectors. From the satellite image it is clear that the most affected *Barrios* are also the largest and most densely populated. Petare was the sector most commonly listed as place of residence in the two molecular epidemiology studies, with 20.5% and 22.1% of cases respectively (Table I), but this may not be accurate, as Petare is not a large *Barrio* but is a name commonly used to non-specifically indicate the marginal, low resource sectors of Sucre municipality, and may have been written in the patient records when the precise sector of residence in the municipality was unknown.

In the pilot study of 44 strains, there were 22 spoligotypes, of which 3 were not present in SITVITWEB. There were 4 clusters: SIT17, 4 strains; SIT42, 7 strains; SIT20, 3 strains; and SIT1692, 2 strains (Table IV).

In the study of 131 strains, there were 41 different spoligotypes (Table II), of which 27 were in the international spoligotype database SITVITWEB. The remaining 14 spoligotype patterns, representing 18 patients (14%), were not in the database, but SPOTCLUST determined that most belonged to the LAM (Latin American Mediterranean) lineage. The isolates grouped into 15 spoligotype clusters with 2 to 27 strains each, representing a total of 105 (80%) strains (Table II). Thirteen of the clusters had spoligotypes present in SITVITWEB, while the spoligotypes of the remaining two groups, composed of 4 and 2 strains respectively, were not in SITVITWEB. SPOTCLUST assigned them to the T1 and LAM9 lineages respectively. Of the 14 spoligotypes not assigned SITs, five were previously submitted to SITVITWEB, one from Senegal and four that were isolated in a prior survey of Venezuelan strains by Aristimuño et al (37).

Table III shows the percentages of the different SITpatterns in the two molecular epidemiology studies of municipality Sucre strains, compared to the previously published percentages of the SITpatterns in strains isolated in the Instituto de Biomedicina located in western Caracas and the states of Carabobo, Delta Amacuro, Amazonas and Sucre. In both molecular epidemiology studies, SIT's 42 and 20 are more common in municipality Sucre than in any other region, including the western side of Caracas.

Table IVA shows the distribution by sector of the 26 clustered isolates in the smaller study, and Table IVB shows the distribution by sector of the 105 clustered spoligotypes from the study of 131 isolates. The four most commonly isolated spoligotypes, SITs 17, 42, 93 and 20, all belonging to the Latin American Mediterranean (LAM) lineage, were found in patients from at least 7 different *Barrios*. The most frequent spoligotype pattern in the municipality, SIT17, was found in 12 different *Barrios*. In the larger study the Unión sector has 7 of the 19 SIT42 clustered strains, while Petare had 4 and Maca 3 isolates. In the smaller study Unión had 2 of the 7 SIT42 strains, the only sector with more than one SIT42 isolate. Of the SIT20 cluster in the larger study, 4 of 9 were from Unión and 2 from San Blas. There was also a cluster of 3 SIT20 strains in the smaller study, one from each of Petare, Palo Verde and the Colonial region of Petare.

Although several of the clustered spoligotypes were isolated from more than one patient in particular *Barrios*, spoligotyping is not sufficiently discriminative to determine whether this was evidence of local spread of a specific strain within the *Barrio*. Therefore, in the larger molecular epidemiology study we used 15 MIRU-VNTR loci to analyze a sample of 21 spoligotyped strains from the largest spoligotype clusters, SIT17 and 42, and also 4 clustered strains with spoligotype patterns not present in the SITVITWEB database. A dendrogram built with the results (Fig. 2) showed that 4 of the SIT42 isolates from *Barrio* Unión had an identical MIRU-VNTR pattern that was not seen in previous studies of other regions of Venezuela. This SIT42 MIRU-VNTR pattern was also not

TABLE II

SPOLIGOTYPE PATTERNS PRESENT IN MUNICIPIO SUCRE IN A COHORT OF 131 STRAINS OF *M. tuberculosis* ISOLATED FROM 2006-2010

#	Spoligotypes	Octal	SIT	Lineage	FREQ	Prob*
1	■■□■■■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	677737607760771	17	LAM2	27	
2	■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	777777607760771	42	LAM9	19	
3	■■■■■■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	777737607760771	93	LAM5	14	
4	■■□■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	677777607760771	20	LAM1	9	
5	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	777777777760771	53	T1	8	
6	■■■■■■■■□□■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	776377777760771	34	S	5	
7	□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■□□□□■■■■■■■	000000007760771	4	U	5	
8	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■□□□□	777777777760700	51	T1	2	
9	■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■□■■■	777777774020731	62	Haarlem1	2	
10	■■■■■■■■■■■■□■■■■■■■□□□□■■■■■■□□□□□□□□□□■■■	777737607700031	605	LAM5-LAM4	2	
11	■■■■■■■■□□□□□□■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	776017607760771	209	LAM12_MAD1	2	
12	■■□■■■■■■■■■□■■■■■■■□□□□■■■■■■□□□□□□■■■■■■■	677737607700771	905	U	2	
13	□■■■■■■■□□□■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	376177607760771	376	LAM3	2	
14	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■	777777777720771	50	Haarlem3	1	
15	■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■	777777774020771	47	Haarlem1	1	
16	■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■□□	777777607760760	738	LAM9	1	
17	■■■■■■■■■■■■■■■■■■■■□□□□□□□■□■■■□□□□■■■■■■■	777777600560771	1147	T1	1	
18	■■■■■■■■■■■■■■■□■■■■□□□□■■■■■■■■□□□□■■■■■■■	777773607760771	1832	LAM9	1	
19	■■■■■■■■■■■■■■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■	777767607760771	150	LAM9	1	
20	■■■■■■■■■■■■□■■■■■■■□□□□■■□■■■■■□□□□■■■■■■■	777737606760771	1709	LAM8	1	
21	■■■■■■■■□□□■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	776177607760771	33	LAM3	1	
22	■■□■■■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■□■■■■	677737607760671	1661	LAM2	1	
23	■□■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	577777607760771	866	LAM9	1	

TABLE II (*Continuación*)

#	Spoligotypes	Octal	SIT	Lineage	FREQ	Prob*
24	■□■■■■■■■■■■■■□■■■■■■■■■■■■■■■■■□□□□■■■■■■■	577767777760771	1318	T1	1	
25	■□■■■■■■□□■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	576377777760771	1211	S	1	
26	■■■■■■■■■■■■■■□■■■■■■■■■■■■■■■■■□□□□■■■■■■■	777767777760771	118	T1	1	
27	■■■■■■■■■■■□■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	777677607760771	770	LAM9	1	
28	■■■■■■■■■■□□□□□□□□□■□□□□■■■■■■■■□□□□■■■■■■■	777400207760771	ND	LAM9	1	1.00
29	■■■■□■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■□□■■	757777777760711	ND	T1	4	0.99
30	■■■■□■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■□■■■	757737607760731	ND	LAM9	2	1.00
31	■■■□□□□□□□□□□□□□■■■■□□□□■■■■■■■■□□□□■■■■■■■	700003607760771	ND	T3	1	0.78
32	■■■□□□□□□□□□□□□□□■■■□□□□■■■■■■■■□□□□■■■■■■■	700001607760771	ND	T3	1	1.00
33	■■□■■■■■■■■■□■■■■■■■□□□□■■■■■□□■□□□□■■■■■□□	677737607620760	ND	LAM2	1	0.82
34	■□■■■■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	577737607760771	ND	LAM9	1	1.00
35	■□■■■■■■■■■■□■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■	577727607760771	ND	LAM9	1	1.00
36	■□■■■■■■■■■□■■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■	577667607760771	ND	LAM9	1	1.00
37	■□□■■■■■■■■■□■□■■■■■□□□□■■■■■■■□□□□□■■■■■■■	477727607740771	ND	LAM2	1	0.81
38	■□□□□■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	417737607760771	ND	LAM2	1	0.81
39	□□□■■■■□□□■■■■■■■■■■□□□□■■■■■■■■□□□□■■■□□□□	074377607760700	ND	LAM1	1	0.69
40	□□□■■■■□□□■■■■□■■■■■□□□□■■■■■■■■□□□□■■■□□□□	074367607760700	ND	LAM1	1	0.69
41	■□■■□■■■■■■■□■□■■■■■□□□□■■■■■■■□□□□□■■■■■■■	557727607740771	ND	LAM9	1	1.00
				Total	131	

"#" numbering is just for visual localization purpose in this Table, it does not identify the profile.
"ND" stands for Non determined SITs according to the SITVITWEB database.
*Prob indicates for the probability given by SPOTCLUST to belong to a determined lineage.

TABLE III
COMPARISON OF THE FREQUENCY OF DIFFERENT SHARED INTERNATIONAL SPOLIGOTYPES IN DIFFERENT REGIONS OF VENEZUELA

	Caracas West	Municip Sucre 2006-11	Municip Sucre 2006	Carabobo State	Delta Amacuro State	Amazonas State	Sucre State
SIT	572	131	44	467	84	100	54
17	20.1%	20.6%	31.8%	13.9%	27.4%	27.0%	13.0%
93	8.7%	10.7%	2.3%	11.1%	8.3%	7.0%	16.7%
605	3.0%	1.5%	-	15.0%	-	1.0%	7.4%
42	7.3%	**14.5%**	**15.9%**	6.6%	4.8%	2.0%	9.3%
53	5.1%	6.1%	2.3%	2.6%	10.7%	1.0%	1.9%
20	3.3%	**6.9%**	**6.8%**	1.5%	6.0%	3.0%	
50	2.6%	0.8%	-	0.6%	1.2%	2.0%	1.9%
1696	0.5%	-	-	3.9%	-	-	-
34	2.8%	3.8%	-	0.6%	1.2%	-	-
51	2.4%	1,5%	-	0.6%	1.2%	-	-
376	1.4%	1.5%	-	1.3%	-	-	-
33	1.0%	0.8%	-	1.5%	-	-	-
47	0.5%	0.8%	-	2.1%	-	-	-
4	1.9%	3.8%	2.3%	0.2%	-	-	-
64	0.9%	-	-	0.2%	-	1.0%	5.6%
106	0.3%	-	-	-	9.5%	-	-
1702	0.5%	-	-	1.5%	-	-	-
NA	-	-	-	0.4%	-	-	14.8%
62	0.3%	-	-	0.4%	-	5.0%	-
960	1.6%	-	-		-	-	-
1691	0.7%	-	-	1.1%	-	-	-

TABLE III (*Continuación*)

	Caracas West	Municip Sucre 2006-11	Municip Sucre 2006	Carabobo State	Delta Amacuro State	Amazonas State	Sucre State
162	-	-	-	0.4%	-	6.0%	-
194	0.3%	-	2.3%	0.4%	-	4.0%	-
291	1.2%	-	-		-	-	-
375	0.9%	-	-	0.4%	-	-	-
1694	0.9%	-	-		1.2%	1.0%	-
866	0.9%	0.8%	-	0.2%	-	-	-
1	0.9%	-	-		-	-	-
86	-	-	-		6.0%	-	-
150	0.2%	0.8%	-	0.9%	-	-	-
167	0.5%	-	2.3%		-	-	3.7%
334	0.9%	-	-		-	-	-
905	0.2%	-	-		3.6%	-	1.9%
1692	0.3%	-	4.5%	0.6%	-	-	-
1905	-	-			-	5.0%	-

The data for all regions except municipality Sucre in the east of Caracas were previously reported (23). The two columns from municipality Sucre represent two separate samplings. The numbers at the top of each column indicate the total strains studied in each region, and below these are the percentages of these totals that had to the respective sits. Bold and underlining indicate that sits 42 and 20 were present in a higher percentage of strains from municipality Sucre than from the other regions.

TABLE IV

A. DISTRIBUTION BY SECTOR OF THE 26 SPOLIGOTYPE CLUSTERED ISOLATES IN THE STUDY OF 44 *M. tuberculosis* STRAINS.
B. DISTRIBUTION BY SECTOR OF 105 SPOLIGOTYPE CLUSTERED ISOLATES IN THE STUDY OF 131 STRAINS

A

		Sector of Residence															
		Petare	JFRivas	La Unión	M. del Este	Carr. Petare-Santa Lucia	Campo Rico	La Dolorita	San Blas	Maca	Bolivar	12 de octubre	5 de Julio	Zona Colonial Petare	La Línea	Palo Verde	Total
SIT	**FREQ**	9	7	4	4	2	2	2	1	1	1	1	1	1	1	1	44
17	14	3	4	1		1	1	1	1		1	1					31.8%
42	7	1		2	1					1			1		1		15.9%
20	3	1												1		1	6.8%
1692	2	2															4.5%
Total in clusters	26	7	4	3	1	1	1	1	1	1	1	1	1	1	1	1	59.1%

B

			Sector of Residence																											
			Petare	La Unión	San Blas	Maca	Carpintero	Jose F. Rivas	F. de Mariche	M. del Este	Agricultura	5 de Julio	A. Jose Sucre	Campito	I. Medina	La Línea	San Isidro	Guaicoco	La Bombilla	La Dolorita	19 Abril	Bolivar	Dorado	Nazareno	San Miguel	Vista Alegre	Boleíta	Llanito	Urbina	Turumo
Spoligotypes	**SIT**	**FREQ (%)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**19**	**20**	**21**	**23**	**24**	**25**	**26**	**27**	**30**	**31**	**32**
■■□■■■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	17	**27 (20.6)**	9	1		2	3	3	3	1						1				1		1		1					1	
■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	42	**19 (14.5)**	4	7	1	3									1			1			1							1		
■■■■■■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	93	**14 (10.7)**	4		1		1				1	1	1	1	1		1		1						1					
■■□■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	20	**9 (6.9)**		4	2												1		1								1			
■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	53	**8 (6.1)**	2	1	2	2	1																							
■■■■■■■■□□■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	34	**5 (3.8)**	1	1		1	1				1																			
□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■□□□□■■■■■■■	4	**5 (3.8)**	1					1		1						1				1										
■■■■□■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■□□■■	ND	**4 (3.1)**				2	1		1																					
■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■□□□□	51	2 (1.5)										1	1																	
■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■□■■■	62	2 (1.5)			1																		1							
■■■■■■■■■■■■□■■■■■■■□□□□■■■■■■□□□□□□□□□□■■■	605	2 (1.5)			1																					1				
■■■■■■■■□□□□□□■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	209	2 (1.5)							1																					1
■■□■■■■■■■■■□■■■■■■■□□□□■■■■■■□□□□□□■■■■■■■	905	2 (1.5)				1								1																
□■■■■■■■□□□■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	376	2 (1.5)			1					1																				
■■■■□■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■□■■■	ND	2 (1.5)						2																						
		105 (80.2)	21	14	9	11	7	6	5	3	2	2	2	2	2	2	2	1	2	2	1	1	1	1	1	1	1	1	1	1

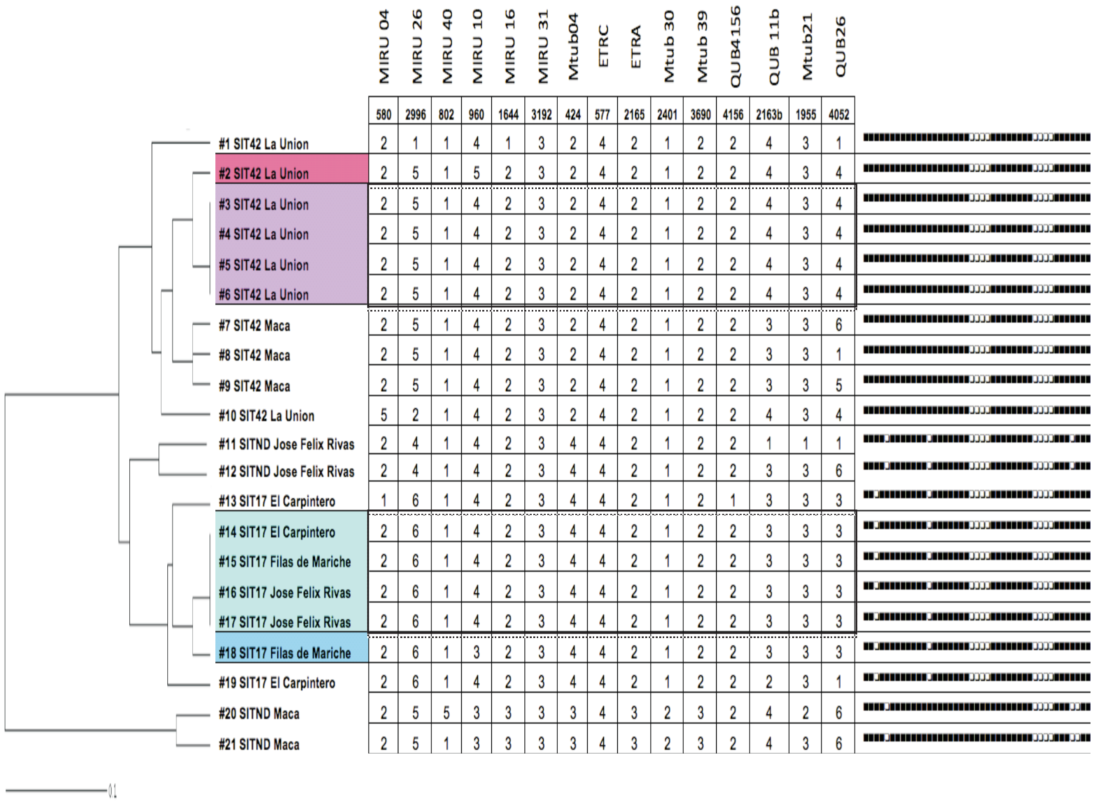
present in the MIRU-VNTRplus database, where the most similar SIT42 strain differed at 5 of 15 MIRU-VNTR loci. Three other SIT42 strains from the same *Barrio* were closely related, differing by only one or two loci (strains 10, 11 and 17). An additional three very closely related strains from Maca (14, 15 and 19), differ from those in Unión only at locus 4052.

Four of the SIT17 strains—two from *Barrio* Jose Felix Rivas (6 and 7), one from Filas de Mariche (5) and one from *Barrio* Carpintero (3)—had MIRU-VNTR patterns that were identical to a large cluster previously found to be widely distributed in Venezuela. Another SIT17 strain from Filas de Mariche (5) differed only at locus 960, and one from *Barrio* Carpintero (1) was different only at two loci - 580 and 4156.

Two strains from Maca with an identical spoligo pattern, not found in SITVITWEB (Fig. 2 strains 18 and 21), had related MIRU-VNTR patterns that varied in just two loci-802 and 1955.

DISCUSSION

This study found that the distribution of SITs in the strains of *M. tuberculosis* circulating in the large marginal region in the east of Caracas-Sucre municipality was similar to the distribution previously reported for the west of Caracas and other parts of Venezuela, showing a wide diversity of *M. tuberculosis* genotypes, but differed in having greater proportions of SIT42 and SIT20 strains. The MIRU-VNTR analysis found that the SIT42 strains appeared to be related,



	MIRU 04	MIRU 26	MIRU 40	MIRU 10	MIRU 16	MIRU 31	Mtub04	ETRC	ETRA	Mtub 30	Mtub 39	QUB4156	QUB 11b	Mtub21	QUB26
	580	2996	802	960	1644	3192	424	577	2165	2401	3690	4156	2163b	1955	4052
#1 SIT42 La Union	2	1	1	4	1	3	2	4	2	1	2	2	4	3	1
#2 SIT42 La Union	2	5	1	5	2	3	2	4	2	1	2	2	4	3	4
#3 SIT42 La Union	2	5	1	4	2	3	2	4	2	1	2	2	4	3	4
#4 SIT42 La Union	2	5	1	4	2	3	2	4	2	1	2	2	4	3	4
#5 SIT42 La Union	2	5	1	4	2	3	2	4	2	1	2	2	4	3	4
#6 SIT42 La Union	2	5	1	4	2	3	2	4	2	1	2	2	4	3	4
#7 SIT42 Maca	2	5	1	4	2	3	2	4	2	1	2	2	3	3	6
#8 SIT42 Maca	2	5	1	4	2	3	2	4	2	1	2	2	3	3	1
#9 SIT42 Maca	2	5	1	4	2	3	2	4	2	1	2	2	3	3	5
#10 SIT42 La Union	5	2	1	4	2	3	2	4	2	1	2	2	4	3	4
#11 SITND Jose Felix Rivas	2	4	1	4	2	3	4	4	2	1	2	2	1	1	1
#12 SITND Jose Felix Rivas	2	4	1	4	2	3	4	4	2	1	2	2	3	3	6
#13 SIT17 El Carpintero	1	6	1	4	2	3	4	4	2	1	2	1	3	3	3
#14 SIT17 El Carpintero	2	6	1	4	2	3	4	4	2	1	2	2	3	3	3
#15 SIT17 Filas de Mariche	2	6	1	4	2	3	4	4	2	1	2	2	3	3	3
#16 SIT17 Jose Felix Rivas	2	6	1	4	2	3	4	4	2	1	2	2	3	3	3
#17 SIT17 Jose Felix Rivas	2	6	1	4	2	3	4	4	2	1	2	2	3	3	3
#18 SIT17 Filas de Mariche	2	6	1	3	2	3	4	4	2	1	2	2	3	3	3
#19 SIT17 El Carpintero	2	6	1	4	2	3	4	4	2	1	2	2	2	3	1
#20 SITND Maca	2	5	5	3	3	3	3	4	3	2	3	2	4	2	6
#21 SITND Maca	2	5	1	3	3	3	3	4	3	2	3	2	4	3	6

Fig. 2. UPGMA dendrogram built with the MLVA and Spoligotype information using the MIRU- VNTR *plus* site. Each code is made of a composition of the arbitrary position number that was taken from a file and then used for the tree, the strain code itself, the SIT code, the Lineage identification and the patient location.

and were found principally in the Unión and Maca *Barrios*. In addition, there were strains that belong a previously identified nationwide SIT17 cluster, which were more prevalent in the *Barrios* Carpintero, Jose Felix Rivas and Filas de Mariche.

In our previous study of 1298 strains from the western region of Caracas and 6 geographically distributed states (23), the most common spoligotype patterns (Table III) were: SIT17 (18.6%); SIT93 (9.9%); SIT605 (7.2%); SIT42 (6.5%); SIT53 (4.1%); SIT20 (2.7%); and SIT50 (1.7%) (23). The overall distribution was skewed, however, by a very large number of SIT605 strains in Carabobo, where it was the most common SIT found, while it was much less common in all other regions, including Sucre municipality. The two molecular epidemiology studies of strains from Sucre municipality showed some differences in the prevalence of different SITs but basic strain distribution was generally similar to the other regions of Venezuela studied, with SIT17 the most common spoligotype (Table III). Surprisingly though, both studies showed that SIT42 and SIT20 were more common in Sucre municipality than in any other region of Venezuela studied, and the percentages in the two studies were remarkably similar: 14.5% and 15.9% for SIT42; 6.9% and 6.8% for SIT20. This was the only region we have examined where SIT42 is the second most common pattern (Table III), and the percentages of strains with SIT42 and SIT20 in this municipality were twice that found in strains isolated in western Caracas.

When the distribution of strain spoligotypes was examined by *Barrio*, it appeared that some clustered strains were more common in particular *Barrios*, but spoligotyping is not sufficiently discriminative to identify transmission or outbreaks because the identical spoligotypes can be shared by strains with different genotypes. Therefore, a limited number of the clustered strains were examined by the more discriminative technique of MIRU-VNTR 15 loci. We previously (23) found a large MIRU-VNTR cluster of a SIT17 strain (261423442122333) that was distributed throughout the country, and 5 strains from Sucre municipality belonged to this cluster, two from Filas de Mariche, two from Jose Felix Rivas, and one from Carpintero, and a second strain from Carpintero was different in only two MIRU-VNTR loci.

The most striking finding was that the eleven SIT42 strains studied by MIRU-VNTR 15 were all closely related. Four strains isolated from patients living in *Barrio* Unión had an identical MIRU-VNTR pattern, 251423242122434, which had not been previously found in Venezuela. One additional strain differed from these four identical strains by one locus, another by two loci, and a third by 3 loci. Three other strains differing by only two loci came from *Barrio* Maca, and one of the differences, at MIRU-VNTR locus 2163b, was the same in all of these three strains, suggesting that they may be an offshoot strain circulating in Maca. A strain with three differences was from Jose Felix Rivas. Except for Petare, the only *Barrios* with more than one SIT42 isolate were Unión and the adjacent Maca (Fig. 1), so the spread of this and closely related strains, especially in these *Barrios*, could account for the higher prevalence of SIT42 in the Sucre municipality compared to the rest of the country. There were 4 SIT42 strains isolated in Petare, but, as mentioned above, "Petare" is a general term for the marginal neighborhoods in Sucre municipality, and is often listed as residence when the correct sector or *Barrio* is unknown; the actual geographic region "Petare" does not contain a large marginal population. It is thus possible that some of these 4 "Petare" SIT42 strains might also be from patients living in Unión or Maca. They were not analyzed by MIRU-VNTR 15, so it is not known if they share the genotype of the SIT42 strains from Unión.

From these small samples of strains, it appears that perhaps 20-25 % of the tuberculosis in Sucre municipality is caused by dominant strains of SIT17 and SIT42 and their offshoots, and that the SIT42 cluster is a strain particular to this municipality that is being actively transmitted, especially in the densely populated region of Unión and the adjacent *Barrio* Maca. The presence of strains with differences at 2-3 MIRU-VNTR loci suggests that the SIT42 strain may have been present and circulating in the municipality for many years. Unión and Maca are home to many immigrants from Colombia and other South American countries, who often live clustered in country-specific sectors of these *Barrios*. The SIT42 strain circulating in these *Barrios* was not found in other parts of Venezuela and could have been introduced by an immigrant and spread within the immigrant's community. This MIRU-VNTR pattern is not similar to any strain in the MIRU-VNTRplus database, so without more detailed epidemiologic investigation, it is not possible to speculate on a possible country of origin. A recent report described the predominance of SIT42 in the neighboring country of Colombia (38), but none of the Colombian isolates had VNTR patterns similar to that of the SIT42 cluster in Sucre municipality (Dr. Jaime Robledo, personal communication).

The presence of strains belonging to the large nationwide SIT17 cluster is perhaps further evidence that this strain may have some advantage in virulence or transmissibility over others circulating in Venezuela (23). The high percentage of SIT20 strains in this municipality is also interesting, but we were not able to perform MIRU-VNTR analysis on them to see if they belong to a single dominant genotype. It appears that the extensive marginal areas of Sucre municipality, such as Unión, are regions with that have had a persistent, long-term tuberculosis presence resulting in the microevolution of their own family of strains, especially SIT42, that are not found in other regions, not even in the large marginal regions of western Caracas. As can be seen in Fig. 1, *Barrio* Unión is at the center of a very large, densely populated region, and may be relatively isolated.

This study found a large diversity of spoligotype patterns, many of which were not found in the SITWEB international data base or previously seen in Venezuela, and were identified in only one isolate from the Sucre municipality. Although it is possible that technical errors were responsible for this diversity, this seems unlikely, as all isolates were tested at least twice on different membranes and the results were consistent on more than one reading. Strains that differ by just one or two MIRU-VNTR differences were regarded as the same genotype, but recent studies using whole genome sequencing found that these strains can be only distantly related (39). However, as the strains were from patients living in the same geographic region and had patterns that differ only slightly from a predominant pattern, a process of microevolution seems likely.

From our retrospective search for cases diagnosed in 2005-6, as well as data from more recent records of the TB control program, it appears that each year there are between 500 and 600 new cases of tuberculosis in Sucre municipality, yielding an incidence of about 45-50 cases per 100,000 inhabitants per year. This figure is likely to be an underestimate because not all cases are diagnosed, although it is difficult to calculate an accurate incidence because of the fluid and growing nature of the population. Nevertheless, the true incidence is probably significantly higher than the overall estimated TB incidence in Venezuela of ~33/100,000.

This study had several shortcomings and limitations. The sample of 131 strains presented here represents only about a quarter of one year's cases, and the earlier cohort of 44 strains represents less than a tenth. These were convenience samples of the available strains isolated from patients presenting to the municipality's only TB clinic, and we have no way of assuring they are accurately or proportionately representative of TB cases in the municipality. In fact, there is evidence that our molecular epidemiology study of 131 cases was not proportionately representative, because Jose Felix Rivas is a very large, densely populated *Barrio* that was represented with only 5 cases in this cohort. While the *Barrios* with the most TB cases are generally the most densely populated in the municipality, we don't have accurate population data for these *Barrios* to determine whether the TB incidence is higher in these than in other sectors of the municipality. However, in the larger study of all cases from 2005-6, Jose Felix Rivas had many more cases than any other sector, presumably reflecting its very large population and crowded housing. It is unlikely that changes in the incidence of tuberculosis in this *barrio* over a few years could account for this difference, so the molecular epidemiology study with 131 isolates may have oversampled *Barrio* Unión, which had the highest percentage of TB cases (14.5%) in that study. It is also possible that this reflects an ongoing outbreak of the SIT42 strain in this sector, or a tendency for patients from Jose Felix Rivas to be diagnosed elsewhere.

The molecular epidemiology studies would have been more complete if the cohorts were larger and all of the strains clustered by spoligotyping could have been analyzed by MIRU, but this was not possible due to limitations of time, resources and personnel. It would also be worthwhile looking at the residency data for those patients with similar or identical MIRU-VNTR patterns, to see if they live in close proximity within the *Barrios*, or belong to the same families or social groups. Further epidemiologic investigation in the high incidence *Barrios* might reveal that patients infected with the same strains have links amenable to strategic interventions to interrupt the on-going transmission, but this was not possible within the scope and resources for this study. These sectors are difficult to reach and challenging for outsiders. A more feasible approach might to reducing the incidence might be to work with community groups within the highest prevalence sectors to promote the early diagnosis of cases, the search for epidemiologic associations, investigation of their close contacts, and outreach to patients who abandon therapy before its completion.

ACKNOWLEDGMENTS

This work would not have been possible without the efforts of CorpoMir to build the TB diagnostic laboratory "Luis Quiñones-TB Sucre". The authors acknowledge the important contributions of the staff of the laboratory "Luis Quiñones–TB Sucre" in diagnosing tuberculosis in the municipality and providing cultures of the patient isolates; the direction of Health District #7; and the director and maintenance staff of the Hospital Materno-Infantil de Petare "Joel Valencia Parpacén". The authors also thank Sandra Fernández and the Venezuelan National Institute of Hygiene, and Carmen Ramirez and the National Tuberculosis Reference lab for contributing strains analyzed in this study. This work was supported by FONACIT projects 2001001851 and G02005000393, as well as LOCTI projects "Tuberculosis en Venezuela," and "Programa Integrado Comunitario para el Control de la Tubercu-

losis en el Municipio Sucre del Estado Miranda.

REFERENCES

1. **Lawn SD, Zumla AI.** Tuberculosis. Lancet. 2011; 378(9785):57-72. doi: 10.1016/S0140-6736(10)62173-3.
2. **World Health Organization.** Global tuberculosis control: WHO report 2011. Geneva: World Health Organization; 2011.
3. **Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN.** Molecular epidemiology of tuberculosis: current insights. Clin Microbiol Rev 2006; 19(4):658-685.
4. **Rado TA, Bates JH, Engel HW, Mankiewicz E, Murohashi T, Mizuguchi Y, Sula L.** World Health Organization studies on bacteriophage typing of mycobacteria. Subdivision of the species *Mycobacterium tuberculosis*. Am Rev Respir Dis 1975; 111(4):459-468.
5. **Snider DE, Jr., Jones WD, Good RC.** The usefulness of phage typing *Mycobacterium tuberculosis* isolates. Am Rev Respir Dis 1984; 130(6):1095-1099.
6. **Gruft H, Johnson R, Claflin R, Loder A.** Phage-typing and drug-resistance patterns as tools in mycobacterial epidemiology. Am Rev Respir Dis. 1984; 130(1):96-97.
7. **van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, Hermans P, Martin C, McAdam R, Shinnick TM.** Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol 1993; 31(2):406-409.
8. **Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, Schecter GF, Daley CL, Schoolnik GK.** The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. N Engl J Med 1994; 330(24):1703-1709.
9. **Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J.** Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol 1997; 35(4):907-914.
10. **Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Locht C.** Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol Microbiol 2000; 36(3):762-771.
11. **Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Locht C, van Soolingen D.** Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J Clin Microbiol 2006; 44(12):4498-4510.
12. **Cowan LS, Diem L, Brake MC, Crawford JT.** Transfer of a *Mycobacterium tuberculosis* genotyping method, Spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system. J Clin Microbiol 2004; 42(1):474-477.
13. **Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, Hilty M, Hopewell PC, Small PM.** Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA. 2006; 103(8):2869-2873.
14. **Zhang J, Abadia E, Refregier G, Tafaj S, Boschiroli ML, Guillard B, Andremont A, Ruimy R, Sola C.** *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay. J Med Microbiol 2010; 59(Pt 3):285-294. doi: 10.1099/jmm.0.016949-0.
15. **Comas I, Homolka S, Niemann S, Gagneux S.** Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. PLoS ONE 2009; 4(11):e7815. doi: 10.1371/journal.pone.0007815.
16. **Victor TC, de Haas PE, Jordaan AM, van der Spuy GD, Richardson M, van**

Soolingen D, van Helden PD, Warren R. Molecular characteristics and global spread of *Mycobacterium tuberculosis* with a western cape F11 genotype. J Clin Microbiol 2004; 42(2):769-772.

17. **Drobniewski F, Balabanova Y, Ruddy M, Weldon L, Jeltkova K, Brown T, Malomanova N, Elizarova E, Melentyey A, Mutovkin E, Zhakharova S, Fedorin I.** Rifampin- and multidrug-resistant tuberculosis in Russian civilians and prison inmates: dominance of the beijing strain family. Emerg Infect Dis 2002; 8(11): 1320-1326.
18. **Warren R, Richardson M, van der Spuy G, Victor T, Sampson S, Beyers N, van Helden P.** DNA fingerprinting and molecular epidemiology of tuberculosis: use and interpretation in an epidemic setting. Electrophoresis 1999; 20(8):1807-1812.
19. **Dormans J, Burger M, Aguilar D, Hernandez-Pando R, Kremer K, Roholl P, Arend SM, van Soolingen D.** Correlation of virulence, lung pathology, bacterial load and delayed type hypersensitivity responses after infection with different *Mycobacterium tuberculosis* genotypes in a BALB/c mouse model. Clin Exp Immunol 2004; 137(3): 460-468.
20. **Lopez B, Aguilar D, Orozco H, Burger M, Espitia C, Ritacco V, Barrera L, Kremer K, Hernandez-Pando R, Huygen K, van Soolingen D.** A marked difference in pathogenesis and immune response induced by different *Mycobacterium tuberculosis* genotypes. Clin Exp Immunol 2003; 133(1):30-37.
21. **Manabe YC, Dannenberg AM, Jr., Tyagi SK, Hatem CL, Yoder M, Woolwine SC, Zook BC, Pitt ML, Bishai WR.** Different strains of *Mycobacterium tuberculosis* cause various spectrums of disease in the rabbit model of tuberculosis. Infect Immun 2003; 71(10):6004-6011.
22. **Krishnan N, Malaga W, Constant P, Caws M, Tran TH, Salmons J, Nguyen TN, Nguyen DB, Daffe M, Young DB, Robertson BD, Guilhot C, Thwaites GE.** *Mycobacterium tuberculosis* lineage influences innate immune response and virulence and is associated with distinct cell envelope lipid profiles. PLoS ONE 2011; 6(9): e23870. doi: 10.1371/journal.pone.0023870.
23. **Abadia E, Sequera M, Ortega D, Mendez MV, Escalona A, Da Mata O, Izarra E, Rojas Y, Jaspe R, Motiwala AS, Alland D, de Waard J, Takiff HE.** *Mycobacterium tuberculosis* ecology in Venezuela: epidemiologic correlates of common spoligotypes and a large clonal cluster defined by MIRU-VNTR-24. BMC Infect Dis 2009; 9:122. doi: 10.1186/1471-2334-9-122.
24. **Petroff SA.** A new and rapid method for the isolation and cultivation of tubercle bacilli directly from the sputum and feces. J Exp Med. 1915; 21(1):38-42.
25. **Kudoh S, Kudoh T.** A simple technique for culturing tubercle bacilli. Bull World Health Organ 1974; 51(1):71-82.
26. **Jansen R, Embden JD, Gaastra W, Schouls LM.** Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol 2002; 43(6): 1565-1575.
27. **Sorek R, Kunin V, Hugenholtz P.** CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat Rev Microbiol 2008; 6(3):181-186. doi: 10.1038/nrmicro1793.
28. **Sequera M, Delgado V, Araque W, Torrealba M, Nuñez R, Da Mata O, Abadía E, Takiff H, de Waard J.** *Mycobacterium tuberculosis:* Espoligotipos en el estado Carabobo, Venezuela. Rev Chil Infectol 2008; 25(5):362-367.
29. **Demay C, Liens B, Burguiere T, Hill V, Couvin D, Millet J, Mokrousov I, Sola C, Zozio T, Rastogi N.** SITVITWEB—a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. Infect Genet Evol 2012; 12(4):755-766. doi: 10.1016/j.meegid.2012.02.004.
30. **Vitol I, Driscoll J, Kreiswirth B, Kurepina N, Bennett KP.** Identifying *Mycobacterium tuberculosis* complex strain families using spoligotypes. Infect Genet Evol. 2006; 6(6):491-504.

31. **Van Embden J, Van Soolingen D.** Molecular epidemiology of tuberculosis: coming of age. Int J Tuberc Lung Dis 2000; 4(4): 285-286.
32. **Allix-Beguec C, Harmsen D, Weniger T, Supply P, Niemann S.** Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. J Clin Microbiol 2008; 46(8):2692-2699.
33. **Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D.** MIRU-VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. Nucleic acids Res 2010; 38(Web Server issue):W326-331. doi: 10.1093/nar/gkq351.
34. **Sneath A, Sokal R.** Numerical Taxonomy. The principles and practice of Numerical Classification. San Francisco: W. H. Freeman and Co.; 1973.
35. **Sola C, Filliol I, Legrand E, Lesjean S, Locht C, Supply P, Rastogi N.** Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. Infect Genet Evol 2003; 3(2):125-133.
36. **Jaccard P.** Nouvelles reserches sur la distribution florale. Bulletin de la Société Vaudoise des Sciences Naturelles 1908; 44:223-270.
37. **Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, Eyre DW, Wilson DJ, Hawkey PM, Crook DW, Parkhill J, Harris D, Walker AS, Bowden R, Monk P, Smith EG, Peto TE.** Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. The Lancet Infect Dis 2013; 13(2):137-146. doi: 10.1016/S1473-3099(12)70277-3.