Finer Snapshot of Circulating *Mycobacterium tuberculosis* Genotypes in Guadeloupe, Martinique, and French Guiana[†]

Julie Millet, William Laurent, Thierry Zozio, and Nalin Rastogi*

WHO Supranational TB Reference Laboratory, Unité de la Tuberculose et des Mycobactéries,
Institut Pasteur de Guadeloupe, Abymes, Guadeloupe, France

Received 8 April 2011/Returned for modification 12 April 2011/Accepted 2 May 2011

This study shows the benefit of spoligotyping coupled to mycobacterial interspersed repetitive-unit (MIRU) typing to pinpoint circulating *Mycobacterium tuberculosis* genotypes in Guadeloupe, Martinique, and French Guiana. We hereby propose reduced 4-locus and 6-locus subsets for LAM and Haarlem lineage strains that predominate in South America and the Caribbean, retaining 99.35% and 99.64% of the total discriminatory power of the 12-locus scheme, respectively.

The present investigation provides an improved phylogenetic characterization of existing *Mycobacterium tuberculosis* transmission chains in Guadeloupe, Martinique, and French Guiana. It also aimed to estimate the benefit of 12-locus mycobacterial interspersed repetitive-unit (MIRU) typing coupled to spoligotyping and to select a reduced MIRU locus typing scheme for an efficient subtyping of clusters belonging to LAM and Haarlem lineages that predominate in South America and the Caribbean (1, 3–5, 13).

Under a longitudinal universal genotyping program covering a 2-year period (January 2004 to December 2005), all culture-positive tuberculosis (TB) cases ($n = 176$) from Guadeloupe ($n = 54$), Martinique ($n = 32$), and French Guiana ($n = 90$) were studied. Mycobacterial identification, drug susceptibility testing, spoligotyping, and 12-locus MIRU typing were performed as reported earlier (4, 11, 14, 16). Additional typing schemes were used retrospectively to compare a subset of strains by using 5 exact tandem repeat (ETR) loci (8) and additional MIRU loci (15). The genotyping data were compared to the international SITVIT2 database of the Pasteur Institute of Guadeloupe as described previously (12). In this database, SIT (spoligotype international type) and MIT (MIRU international type) designate spoligotyping and MIRU patterns shared by 2 or more patient isolates, respectively, whereas “orphan” designates patterns reported for a single isolate. Major phylogenetic clades were assigned according to signatures provided in SpolDB4, which defines 62 genetic lineages/sublineages (3). Discriminatory power of a typing method (or a combination of methods) was calculated using the Hunter and Gaston discriminatory index (HGDI) (10), and significant differences between percentages were estimated by using a chi-square test. Using the Perl programming language (<http://www.cpan.org>), we also developed and implemented a

dynamic programming algorithm named MIRU-Selector. For a given set of strains, this program calculates the discriminatory ability of each possible combination of MIRUs by subtracting 1 locus at a time (12 loci, then 11 loci, and so on), until a single locus is left.

Spoligotyping allowed a first-line screening of the clinical isolates and lineage assignments (Fig. 1A to C and Table 1; see Tables S1 to S3 in the supplemental material for demographic characteristics of patients, drug resistance of *M. tuberculosis* clinical isolates, and detailed typing results). A total of 76 patterns obtained corresponded to 30 clusters (130 strains [73.9%]; 2 to 25 strains/cluster) and 46 (26.1%) unique patterns with 8 orphans. The bulk of genotypes in our study (129/176 [73.3%]) concerned 3 families: evolutionarily recent LAM (25%), Haarlem (23.9%), and ill-defined T clades (24.4%). MIRU typing was performed on all the 176 strains; however, 22 strains with incomplete profiles were excluded from the analysis (see Table S2 in the supplemental material); 97 patterns obtained corresponded to 24 clusters (81 strains [52.6%]; 2 to 15 strains/cluster) and 73 (47.4%) unique patterns with 58 orphans. Combined spoligotyping and MIRU analysis grouped 58/154 (37.7%) strains in 20 clusters (2 to 11 strains per cluster), while 96 strains were unclustered.

The combined numerical analysis of isolates clustered by spoligotyping for which MIRU data were available ($n = 109$) is summarized in Fig. S1 in the supplemental material. It additionally shows all single-locus (SLVs) and double-locus (DLVs) variants, representing strains differing in copy numbers for 1 or 2 MIRU loci for enhanced cluster analysis, respectively. As can be seen, MIRU loci 26 and 31 were the most important contributors to the locus variations observed. Four clusters were specific for Guadeloupe: 2 with identical patterns (SIT5-MIT15 and SIT99-MIT131) and 2 SLVs (SIT93-MIT25/MIT738 and SIT42-MIT25/MIT738). We also detected introduction of the Beijing genotype (SIT1) in Martinique, in 2 young and epidemiologically linked (epi-linked) patients (a brother and sister <18 years of age), for the first time since 1994 (4). However, the 2 strains differed by 3-copy changes upon 12-locus MIRU typing (MIT721 and MIT1049). Finally, 9 clusters were specific to French Guiana, among which 2 concerned patients born in Brazil but living in French Guiana (SIT176-MIT801 and DLVs SIT95-MIT898/MIT1052). Lastly,

* Corresponding author. Mailing address: WHO Supranational TB Reference Laboratory, Unité de la Tuberculose et des Mycobactéries, Institut Pasteur de Guadeloupe, 97183 Abymes, Cedex, Guadeloupe, France. Phone: 590-590-897661. Fax: 590-590-896941. E-mail: nrastogi@pasteur-guadeloupe.fr.

[†] Supplemental material for this article may be found at <http://jcm.asm.org/>.

[‡] Published ahead of print on 11 May 2011.

FIG. 1. Distribution of the SITs found in this study and corresponding lineage attribution in various settings (A to C) and relative discriminatory powers of individual MIRU loci among major genotypic lineages (D to F). (A) Guadeloupe; (B) Martinique; (C) French Guiana; (D) Haarlem lineage; (E) LAM lineage; (F) total sample. Note that various lineages in panels A to C are shown by different colors. Refer to Table S1 in the supplemental material for detailed results.

none of the clusters was made up of only multiple-drug-resistant (MDR) isolates; however, 2 clusters of 2 isolates each, SIT5-MIT15 and SIT75-MIT42, found in Guadeloupe and French Guiana, respectively, contained single MDR strains as a result of acquired drug resistance.

As summarized in Table 1, the discriminatory power of typing methods used alone or in combination increased in the order 5-locus ETR typing (least discriminatory power), spoligotyping, and then 12-locus MIRU (most discriminatory power), achieving an HGDI of 0.992 for spoligotyping plus 12-locus MIRU typing (subset B, with 154 isolates). In subset C (with 97 isolates), with additional data on 5 ETR loci, the highest HGDI was 0.994 for the combination of all markers pooled together. These results argue in favor of a two-step approach for large-scale, prospective genotyping of *M. tuberculosis* based on spoligotyping and MIRU-variable-number tandem-repeats (VNTRs) (6, 9). We further decided to evaluate the potential of additional MIRU-VNTRs (15) by screening a subset containing 13 SIT-MIT clusters ($n = 36$; 2 to 7 strains per cluster). As summarized in Table S3 in the supplemental material, only 4 out of 13 clusters did not subdivide.

Thus, the use of additional loci reduced the clustering by 52.8% to as few as 6 clusters (17 strains; 2 to 5 strains per cluster).

Last but not least, we attempted to define a minimal subset of MIRU loci allowing satisfactory discrimination of strains of the 2 largest clades in the present study, i.e., LAM and Haarlem. Figure 1D to F summarize the performance of each of the individual MIRU loci compared to the discriminatory power of the 12-locus typing scheme for the Haarlem ($n = 37$; HGDI = 0.836) (Fig. 1D) and LAM ($n = 39$; HGDI = 0.918) (Fig. 1E) lineage strains (as a percentage). Data for the whole sample are shown as a reference in Fig. 1F ($n = 154$; HGDI = 0.983). Using MIRU-Selector, we were able to define reduced 6-locus (MIRU loci 10, 20, 23, 31, 39, and 40), and 4-locus (MIRU loci 10, 26, 31, and 40) typing schemes for studying specifically the Haarlem and LAM lineage strains. In our sample, these schemes allowed us to achieve 99.64% (HGDI = 0.838) and 99.35% (HGDI = 0.912) of the total discriminatory power of the full 12-locus scheme, respectively.

As indicated previously (4), Haarlem (23.9%), LAM (25%), and the T family (24.4%) constituted the 3 major clades in our

TABLE 1. Discriminatory power of various typing methods used alone and in combination on different subsets of *M. tuberculosis* clinical isolates

Subset (no. of isolates)	Typing method(s) ^a	No. of distinct profiles	No. of unique profiles	No. (size) of clusters	No. (%) of clustered isolates	Recent transmission rate (%)	HGDI
A (176)	Spoligotyping	76	46	30 (2-25)	130 (73.9)	56.8	0.964
B (154)	Spoligotyping	71	45	26 (2-22)	109 (70.8)	53.9	0.964
	12-loci MIRU typing	97	73	24 (2-15)	81 (52.6)	37.0	0.983
	Spoligotyping and 12-locus MIRU typing	116	96	20 (2-11)	58 (37.7)	24.7	0.992
C (97)	Spoligotyping	51	34	17 (2-15)	63 (64.9)	47.4	0.960
	5-locus ETR typing	29	17	12 (2-26)	80 (82.5)	70.1	0.895
	12-locus MIRU typing	71	61	10 (2-11)	36 (37.1)	26.8	0.983
	Spoligotyping and 5-locus ETR typing	71	57	14 (2-8)	40 (41.2)	26.8	0.988
	Spoligotyping and 12-locus MIRU typing	77	67	10 (2-8)	30 (30.9)	20.6	0.990
	Spoligotyping, 12-locus MIRU typing, and 5-locus ETR typing	80	70	10 (2-6)	27 (27.8)	17.5	0.994

^a The five ETR loci corresponded to exact tandem repeats A, B, C, D, and E according to Frothingham and Meeker-O'Connell (8).

setting. Strains of the EAI lineage were found mainly in patients living in French Guiana (9/10 patients), with SIT1340 (EAI6-BGD1 sublineage) being predominant 4/9 strains). According to the SITVIT2 database, 41 strains were reported to present SIT1340, and among this number, 43.9% (18/41) originated in Suriname, where it represents as high as 16% of all *M. tuberculosis* isolates (our unpublished observations). Hence, the SIT1340 strains may represent an emerging transborder *M. tuberculosis* clone in French Guiana. The presence of the X strains, of presumed Anglo-Saxon origin (4), in all the 3 French departments of the Americas and the Caribbean territories probably reflects the history of this geographical area. Lastly, despite a decrease in the proportion of drug-resistant strains and MDR compared to the observations in a previous study (4), the detection of the Beijing genotype in Martinique and French Guiana must be taken with utmost care, since it is often associated with development of drug resistance (2, 7).

We thank colleagues at various university hospitals, local clinics, dispensaries, health services, and research institutions for their precious collaboration. We are grateful to colleagues at Institut Pasteur de la Guadeloupe for their precious help regarding identification and drug susceptibility testing of the isolates (M. Berchel, K. S. Goh, and F. Prudenté), for genotyping and database comparison (C. Sola, B. Liens, I. Filliol, S. Ferdinand, and K. Brudey), for retrospective genotyping of clustered isolates using additional MIRU loci (J. Vanhomwegen), and for developing the MIRU-Selector application (C. Demay).

REFERENCES

1. Abadía, E., et al. 2009. *Mycobacterium tuberculosis* ecology in Venezuela: epidemiologic correlates of common spoligotypes and a large clonal cluster defined by MIRU-VNTR-24. *BMC Infect. Dis.* 9:122.
2. Bifani, P. J., B. Mathema, N. E. Kurepina, and B. N. Kreiswirth. 2002. Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. *Trends Microbiol.* 10:45–52.
3. Brudey, K., et al. 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol.* 6:23.
4. Brudey, K., et al. 2006. Long-term population-based genotyping study of *Mycobacterium tuberculosis* complex isolates in the French departments of the Americas. *J. Clin. Microbiol.* 44:183–191.
5. Candia, N., et al. 2007. First insight into *Mycobacterium tuberculosis* genetic diversity in Paraguay. *BMC Microbiol.* 7:75.
6. Cowan, L. S., et al. 2005. Evaluation of a two-step approach for large-scale, prospective genotyping of *Mycobacterium tuberculosis* isolates in the United States. *J. Clin. Microbiol.* 43:688–695.
7. European Concerted Action on New Generation Genetic Markers and Techniques for the Epidemiology and Control of Tuberculosis. 2006. Beijing/W genotype *Mycobacterium tuberculosis* and drug resistance. *Emerg. Infect. Dis.* 12:736–743.
8. Frothingham, R., and W. A. Meeker-O'Connell. 1998. Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. *Microbiology* 144:1189–1196.
9. García de Viedma, D., I. Mokrousov, and N. Rastogi. 2011. Innovations in the molecular epidemiology of tuberculosis. *Enferm. Infecc. Microbiol. Clin.* 29(Suppl. 1):8–13.
10. Hunter, P. R., and M. A. Gaston. 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J. Clin. Microbiol.* 26:2465–2466.
11. Kamerbeek, J., et al. 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J. Clin. Microbiol.* 35:907–914.
12. Millet, J., S. Baboolal, P. E. Akpaka, D. Ramoutar, and N. Rastogi. 2009. Phylogeographical and molecular characterization of an emerging *Mycobacterium tuberculosis* clone in Trinidad and Tobago. *Infect. Genet. Evol.* 9:1336–1344.
13. Molina-Torres, C. A., et al. 2010. *Mycobacterium tuberculosis* spoligotypes in Monterrey, Mexico. *J. Clin. Microbiol.* 48:448–455.
14. Sola, C., et al. 2003. Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. *Infect. Genet. Evol.* 3:125–133.
15. Supply, P., et al. 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable number tandem repeat typing of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 44:4498–4510.
16. Supply, P., et al. 2001. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. *J. Clin. Microbiol.* 39:3563–3571.