HAL open science

Mycobacterium tuberculosis polyclonal infections and microevolution identified by MIRU-VNTRs in an epidemiological study

Elisabeth Streit, Julie Millet, Nalin Rastogi

To cite this version:

Elisabeth Streit, Julie Millet, Nalin Rastogi. Mycobacterium tuberculosis polyclonal infections and microevolution identified by MIRU-VNTRs in an epidemiological study. International Journal of Mycobacteriology, 2015, 4 (3), pp.222-227. 10.1016/j.ijmyco.2015.05.005 . pasteur-02949744

HAL Id: pasteur-02949744

https://riip.hal.science/pasteur-02949744

Submitted on 28 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

International Journal of Mycobacteriology 4 (2015) 222–227

HOSTED BY

ELSEVIER

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/IJMYCO

Mycobacterium tuberculosis polyclonal infections and microevolution identified by MIRU-VNTRs in an epidemiological study

Elisabeth Streit, Julie Millet, Nalin Rastogi *

WHO Supranational TB Reference Laboratory, TB & Mycobacteria Unit, Institut Pasteur de la Guadeloupe, F97183 Abymes, Guadeloupe

ARTICLE INFO

Article history:
Received 24 April 2015
Received in revised form
5 May 2015
Accepted 6 May 2015
Available online 3 June 2015

Keywords:
Mycobacterium tuberculosis
Epidemiology
Spoligotyping
MIRU-VNTRs
Polyclonal infections
Microevolution

ABSTRACT

Introduction: The advent of molecular typing using MIRU-VNTR mini-satellites has largely facilitated tuberculosis (TB) molecular epidemiological studies. Apart from detecting the chains of transmission and risk factors, these markers have also allowed to study the phenomena of mixed strain infections versus microevolutionary events.

Methods: An initial set of *Mycobacterium tuberculosis* strains (n = 161) genotyped using spoligotyping and MIRU-VNTRs in Guyana and Suriname was evaluated for indications mixed strain infections (characterized by the detection of double alleles in 2 or more MIRU loci) versus "in-patient" microevolutionary events (characterized by the detection of double alleles in a single locus).

Results: The present study hereby reports evidence of microevolution in 3.7% ($n = 6/161$) of the studied population, vs. 0.6% (n = 1/161) for mixed infection. The strains belonged to three different spoligotyping-based lineages, namely the T (SITs 44, 53, and 1081), Haarlem (SIT47), and EAI (SITs 72 and 349) lineages, while 1 isolate (SIT237) could not be assigned to any lineage.

Discussion: By comparing these results on microevolutionary cases ($n = 6$) to 112,000 strains present in the SITVIT2 database, evidence is presented that in 2/6 cases (each case corresponding to 2 patterns due to MIRU double bands), one of the patterns corresponded to a shared type found exclusively in Suriname or Guyana. Phylogenetic analysis showed that no spoligotyping lineage in particular was more prone to microevolutionary events in this study's sample. Overall, the observations fortify the awareness regarding the existence of microevolution and polyclonal TB infections which has important implications for patient care.

© 2015 Asian African Society for Mycobacteriology. Production and hosting by Elsevier Ltd. All rights reserved.

Introduction

Tuberculosis (TB) remains a major global health problem, having caused 9 million new cases and 1.5 million deaths in 2013 [1]. It has long been assumed that a patient is infected by a single strain of *Mycobacterium tuberculosis* (MTB) at a time. Nonetheless, the advent of molecular epidemiological techniques has led to increasing reports of mixed strain infections

* Corresponding author at: WHO Supranational TB Reference Laboratory, TB & Mycobacteria Unit, Institut Pasteur de la Guadeloupe, Morne Jolivière, BP 484, F97183 Abymes, Guadeloupe. Tel.: +590 590 897661.
E-mail address: nrastogi@pasteur-guadeloupe.fr (N. Rastogi).

Peer review under responsibility of Asian African Society for Mycobacteriology.
http://dx.doi.org/10.1016/j.ijmyco.2015.05.005
2212-5531/© 2015 Asian African Society for Mycobacteriology. Production and hosting by Elsevier Ltd. All rights reserved.

(patients harboring more than one MTB clone at a time) over the past 15 years [2–6; reviewed in 7]. MIRU-VNTR typing [8,9], widely applied in molecular epidemiologic studies, facilitates the detection of polyclonal infections as the presence of more than one strain in a sample is likely to lead to the detection of multiple alleles in a number of loci. Depending on the number of implicated loci, two distinct mechanisms are considered responsible for the presence of such multiple alleles; "in-patient" microevolution of the infecting clone in case of the detection of double alleles in a single locus as opposed to simultaneous infection with two distinct MTB strains when two or more loci are concerned [2–4,10]. Microevolution can of course affect any genotypic marker, for example, a strain may evolve by losing a spacer of the initial spoligotype profile. As the molecular clock of spoligotyping profiles is lower than that of MIRU-VNTRs [11], evolutionary changes are much less frequent in the former. Moreover, "in-patient" microevolution within the DR-locus (i.e., the locus targeted by spoligotyping technique) is not detectable unless single colonies from serial isolates (containing the initial isolate vs. recently evolved isolates) are obtained and analyzed separately. Since the present report presents essentially MIRU-VNTR data, hereafter the focus will be on microevolution concerning this particular marker. Most often, studies on polyclonal and multiple strain infections were conducted in settings with high TB burden [2–5]. Nevertheless, they have also been observed in areas with moderate TB incidence [6]. The present study reports evidence for clonal heterogeneity and mixed-infection observed in a recent epidemiologic study on isolates from Guyana and Suriname [12] based on MIRU-VNTRs.

Materials and methods

Clinical isolates

The MTB bacterial isolates ($n = 7$) described in this study were genotyped in the course of a recent epidemiological study on clinical isolates ($n = 161$) from Guyana and Suriname [12]. They were part of a convenience sample of clinical isolates sent to the Caribbean Epidemiology Centre (CAREC), Trinidad, for identification and drug susceptibility testing (DST) [13].

Genotyping and database comparison

The initial set of strains ($n = 161$) was subjected to standard spoligotyping [14] and 15-loci MIRU-VNTR typing [9] in the following order: MIRU-4, MIRU-10, MIRU-16, MIRU-26, MIRU-31, MIRU-40, ETR-A, ETR-C, QUB-11b, QUB-26, QUB-4156, Mtub04, Mtub21, Mtub30, and Mtub39. The results from each of the 15 loci were combined to create a 15-digit allelic profile, and a cluster was defined as two or more strains sharing identical spoligotypes and 15-loci MIRU patterns. The obtained profiles were compared with SITVIT2, a proprietary database of the Pasteur Institute of Guadeloupe which is an updated version of SpolDB4 [15] and SITVITWEB [16], and assigned to a SIT (spoligo-international-type) and/or 15-MIT (15-loci MIRU-international-type) respectively, if they matched at least one other profile in the database or classified as an orphan (no match found).

From the initially published study [12], strains for which multiple bands were repeatedly obtained for $\geqslant 1$ locus were excluded from epidemiological analysis ($n = 7$), and were further investigated in the present study to elucidate the phenomena of polyclonal infections versus microevolutionary events identified by MIRU-VNTRs. For this purpose, all polymerase chain reactions (PCRs) were repeated twice to confirm the presence of multiple alleles in a given MIRU locus.

Phylogenetical analysis

Major phylogenetic clades were assigned according to the signatures provided in the database defining 62 genetic lineages/sublineages. These include various MTB complex members, as well as rules defining major lineages/sublineages for MTB *sensu-stricto*, namely: Beijing clade, the Central Asian (CAS) clade and 2 sublineages, the East African-Indian (EAI) clade and 9 sublineages, the Haarlem (H) clade and 3 sublineages (including H3/Ural-1 and H4/Ural-2 sublineages), the Latin American-Mediterranean (LAM) clade, its 12 sublineages (note that sublineages LAM7-TUR and LAM10-CAM are now referred to as Turkey and Cameroon lineages), the ancestral "Manu" family and 3 sublineages, the S clade, the IS6110-low-banding X clade and 3 sublineages, and an ill-defined T clade with 5 sublineages.

BioNumerics v6.6 (Applied Maths NV, Sint-Martens-Latem, Belgium) was used to construct a minimum spanning tree (MST) based on spoligotypes and 15-loci MIRU VNTR profiles and nodes differing in a maximum of two genotypic markers (i.e., VNTR locus or spoligo-spacer) were grouped in order to highlight single locus variants (SLVs) and double locus variants (DLVs).

Results

The results obtained along with schematic representation of microevolution and polyclonal infection, the mechanisms involved, and genotyping results are summarized in Table 1 and Figs. 1 and 2. As listed in Table 1, 7 patient isolates ($n = 5$ from Suriname and $n = 2$ from Guyana) repeatedly displayed multiple alleles in one or more MIRU locus. As schematically illustrated in Fig. 1, they were considered suggestive of microevolution if a single locus ($n = 6$ strains) was implicated, and mixed infection if several loci were concerned ($n = 1$ strain). Given that the original epidemiological study [12] covered $n = 154$ isolates with unambiguous profiles, the rate of patients harboring clonal subpopulations was 3.7% ($n = 6/161$) in the studied population, and 0.6% ($n = 1/161$) showing evidence of mixed strain infection. Their genotypic profiles and basic demographic data are shown in Table 1; the strains belonged to three different spoligotyping-based lineages, namely the T (SITs 44, 53, and 1081), Haarlem (SIT47), and EAI (SITs 72 and 349) lineages. One isolate (SIT237) could not be assigned to any lineage. The patients' mean age was 48 years, ranging from 14 to 83 years. With the exception of the youngest patient (a female), all other patients were male. All strains were pan-susceptible.

In $n = 6/7$ samples double alleles were detected in a single locus. In accordance with current conventions [2–5,10], these

Table 1 – Isolates showing multiple alleles for one or more VNTR loci.

Origin	Age/ Sex	DST*	HIV**	SIT***	Clade	VNTR profile	MIT15	Conclusion
SUR	38/M	S	+	44	T5	1'/2 3 1/2 52 3/4 2/3 4/5/6 352222 2/3	NA	Mixed infection/ contamination
SUR	83/M	S	NA	47	H1	242533 2/3 34732343	Orphan	Clonal variant
SUR	48/M	S	—	53	T1	23153 2/3 342522222	Orphan	Clonal variant
GUY	54/M	S	NA	72	EAI1-SOM	543253725412 6/9 24	Orphan/ MIT872	Clonal variant
SUR	43/M	S	—	237	Unknown	24323334 3/4 732343	Orphan	Clonal variant
GUY	14/M	S	NA	349	EAI1-SOM	543253927412 8/C 26	Orphan/ MIT876	Clonal variant
SUR	56/M	S	—	1081	T	243 6/8 31336833442	Orphan	Clonal variant

* DST (drug-susceptibility testing): S, pansusceptible.
** HIV serology: (+) positive, (–) negative, NA, not available.
*** Note that strains with SIT44 and SIT47 further showed a third, unspecific band for the multiple allele loci, observed right above the larger of the two specific amplification products.

samples were therefore classified as clonal variants of the same strain. The respective patients correspond to what is depicted as an index case for the establishment of a new SLV in a given population in Fig. 1A (i.e., patient harboring stain A and A1). On the other hand, a single isolate (SIT44/T5) showed multiple alleles in six loci indicating a simultaneous infection with two different MTB strains (strain A + strain B), classified as polyclonal infection (Fig. 1B).

In parallel, it was also checked to see if any of the two MIRU profiles detected in each of the patients harboring microevolved strains (Table 1, except SIT44/T5) corresponded to profiles that had been observed before by comparing to 112,000 strains present in SITVIT2 database. In case one of the detected profiles matched a 15-MIT profile detected earlier in the region, one could tentatively assume that this profile corresponded to the original infecting strain, while the other profile represents the newly evolved variant. Interestingly, one of the VNTR profiles detected in a patient harboring the microevolved strain (SIT72/EAI1-SOM) matched only 2 strains (Fig. 2, shown by red arrow), being exclusively reported in Guyana (a previous study [12]). Interestingly, a third strain in the same study represented a SLV of this same 15-MIT pattern, indicating that strains belonging to the clonal group SIT72/EAI1-SOM/15-MIT872 have been implicated in microevolutionary events in the past. Likewise, the profile of another microevolved strain (SIT349/EAI1-SOM) matched exclusively 2 strains reported in Suriname (a previous study [12]), belonging to the clonal group SIT349/EAI1-SOM/15-MIT876 (Fig. 2, shown by blue arrow). The fact that these profiles (i.e., 15-MIT872 and 15-MIT876, Table 1) correspond to patterns actively circulating in the region (Fig. 2) suggests that the patient initially contracted the strains with a shared-type pattern which later evolved to give an orphan pattern. The profiles detected in the 4 other patients (Table 1) were exclusively orphan patterns. Note that Fig. 2 illustrates the population structure of the global sample studied ($n = 154$ strains from Guyana and Suriname), excluding the 7 strains showing multiple copies in one or more MIRU loci (Table 1). In the MST shown, various lineages are shown by different colors, and profiles differing by a change in 1 marker (SLV), vs. 2 markers (DLV) are shown by bold vs. thin connecting lines, globally highlighted by a gray background.

Discussion

In this study, 5 loci have been implicated in in-patient microevolution events, namely Mtub 21 (VNTR 1955, changing twice) and MIRU 26 (VNTR 2996), MIRU 40 (VNTR 802), ETR A (VNTR 2165) and QUB 11b (VNTR 2163b) all changing once. These loci were found to be among the 7 most variable loci of the MIRU-typing scheme in a large-scale evaluation conducted by Supply et al. [9]. These authors furthermore reported the observation of changes in MIRU 40 and QUB 11b, in a series of epidemiologically linked isolates.

In order to verify, if there is a particular group of spoligotyping profiles more prone to microevolution in the 15 VNTR loci than others, a spoligoforest (http://www.emi.unsw.edu.au/spolTools/) was constructed using the spoligotypes of all $n = 161$ clinical isolates from Guyana and Suriname, and spoligotypes observed in microevolved isolates were marked. The visual assessment showed that spoligotypes of microevolved strains were scattered over different branches of the spoligoforest without any obvious branch preferences. It was, therefore, concluded that no spoligotyping lineage is particularly prone to be involved in microevolution in this sample (data not shown). Regarding other potential predisposing factors for in-patient microevolution, diagnostic delay has been reported in connection with patients harboring clonal variants in a study conducted in Spain [6]. However, seeing as similar mutation rates have been observed during latency, active disease and in logarithmically growing cultures in a recent study [17], the time passed since the primo infection might be more influential than the lapse of time between disease outbreak and diagnosis.

The transmission of microevolved strains may entrain the appearance SLVs in a series of epidemiologically linked cases (Fig. 1, transmission of A1 strains; [9]). However, these events

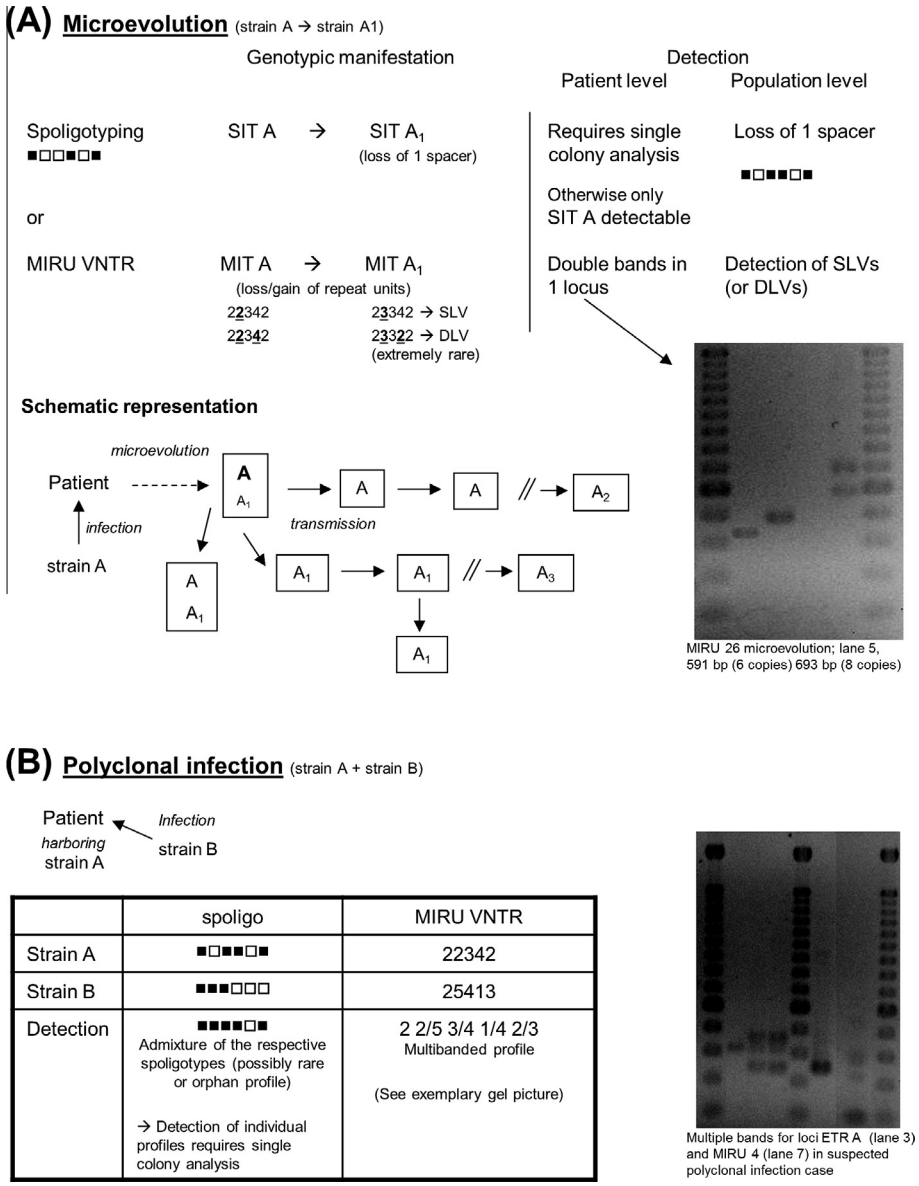

Fig. 1 – Microevolution (A) and polyclonal infection (B): schematic representation of involved mechanisms and genotyping results.

are rare, and the detection of SLVs usually indicates epidemiologic unrelatedness [18,19]. Nevertheless, the presence of SLVs in a given strain population is an echo of past evolutionary events, and the abundance of closely related patterns suggests that the clones in question and their ancestors have been circulating in the region for a long time and/or subject to heavy active transmission, which permitted the accumulation of numerous changes to the original genotype. To check for such traces in this study population, an MST was constructed based on spoligotypes and 15-loci MIRU profiles (Fig. 2), and nodes that differed in a maximum of two genotypic markers were grouped. While most groups contain only two nodes, the two most frequent spoligotypes of the T family (SIT53 and SIT131) stand out as they group a number of SLVs (marked by bold connecting lines) around a central MIRU-pattern. In order to evaluate the genetic diversity of both SITs on a global level, available data were extracted from SITVIT2, and the number of distinct MIRU15 profiles (shared and orphan) was determined. The same assessment was carried out for a ubiquitous pattern SIT50/H3 in SITVIT2, acting as a control. Interestingly, the diversity of VNTR profiles observed among SIT53 strains is considerably higher than that observed for the two other patterns (SIT50, SIT131). This observation is probably due to a higher ongoing evolution of SIT53; indeed, in the SITVIT2 database, as much as 47.6% ($n = 325$) of SIT53 strains with 15-loci MIRU data ($n = 683$) are recorded as orphan vs. only 29% ($n = 171$) of the SIT50 strains with 15-loci MIRU data ($n = 589$). Future studies should elucidate if this is simply due to the widespread occurrence and transmission of SIT53/T1, or if they are more prone to genetic diversification. It is also possible that the two go hand-in-hand, since subtle differences in the infectivity of microevolved MTB variants co-infecting the same patient have been observed [20], and microevolution has been linked

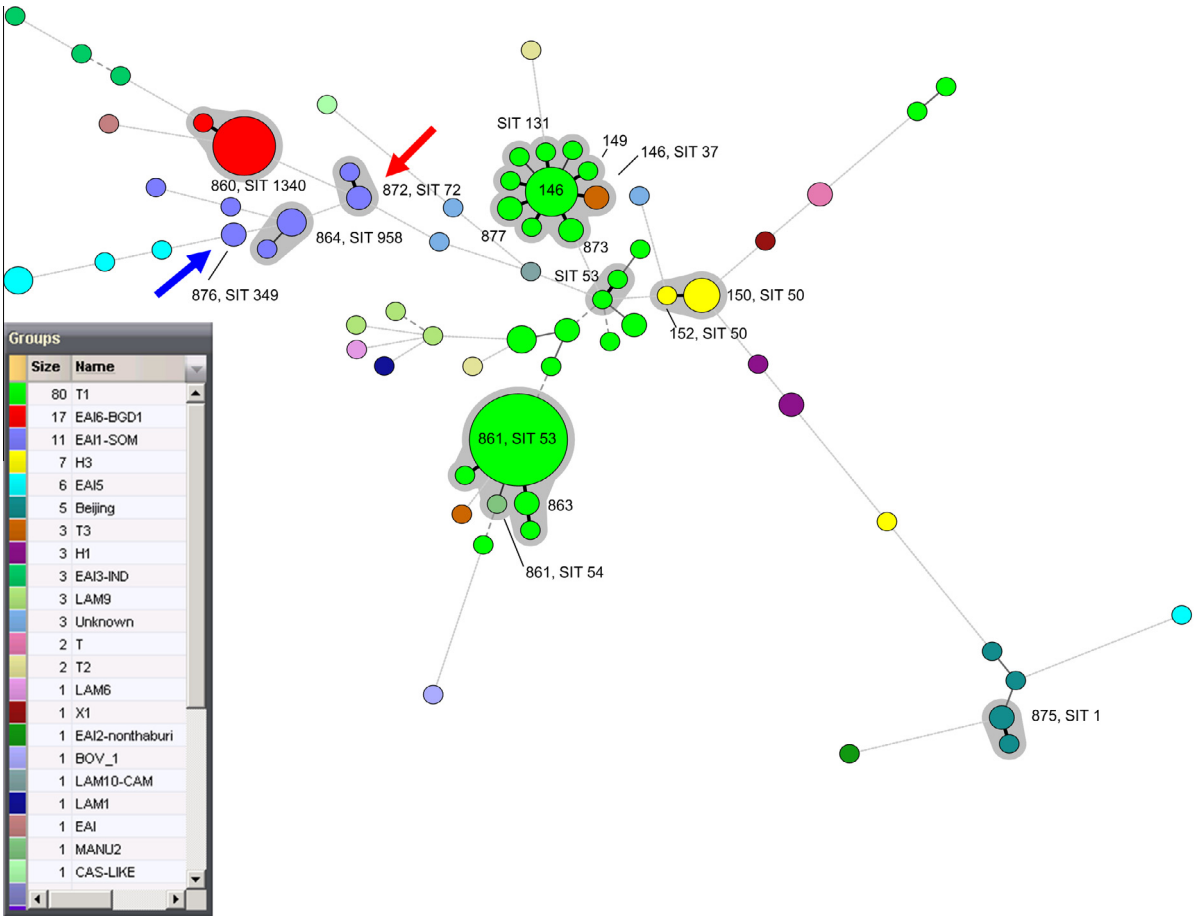

Fig. 2 – MST of isolates from patients from Guyana and Suriname harboring monoclonal *M. tuberculosis* infections: profiles differing by a change in one marker (bold connecting line) or two markers are depicted on gray background (when applicable, SIT and MIT15 are specified).

to differences in gene expression between clonal variants [21]. Thus, predisposition for evolution of SIT53/T1 strains may have favored their ubiquitous spread in all human populations and macro-regions [16]. Finally, the last isolate showed multiple alleles in 6 loci indicating a simultaneous infection with two different MTB strains. This isolate displayed SIT44/T5, a profile that can emerge as a result of almost 11,000 possible combinations of spoligotypes [22].

Little is known about predisposing factors or patient characteristics associated with mixed strain infections as the number of patients found to harbor such infections is usually too low to allow for meaningful statistical analysis. In this sense, the present study is no exception and did not allow for drawing any statistically meaningful conclusions. On the same lines, it could not conclude if HIV infection increased the vulnerability for mixed strain infections as suggested in a previous study from Kampala, Uganda [2], even though the only patient harboring a mixed-strain infection in this study was HIV+ (Table 1).

It has been previously argued that mixed infections, particularly in cases where co-infecting strains display different drug susceptibility patterns (i.e., heteroresistance), have a negative impact on treatment outcome [7]. However, the single case of mixed infection represented by SIT44/T5 in the present study was pansusceptible, not allowing for a conclusion on this particular aspect.

Conclusions

Overall, the observations in the present study fortify the awareness regarding the existence of clonally complex TB infections that has been growing in recent years. The realization that polyclonal infections constitute a reality has important implications for patient care considering the potential occurrence of heteroresistance [7]. Further research is needed to better document the frequency of mixed infections and their relevance for disease progression and treatment outcome in affected patients, as well as their implication for vaccine development.

Nonetheless, it is important to keep in mind that MIRU-VNTRs analyze only a very small part of the mycobacterial genome and therefore might underestimate the extent of heterogeneity within a patient or population. As whole genome sequencing (WGS) is becoming more accessible, studies using this technique are expected to provide a more accurate assessment (reviewed in [23]).

Conflict of interest

No conflict of interest is perceived and none declared.

Acknowledgments

We are highly grateful to Dr. Shirematee Baboolal since this work benefited from the MTB DNA collection sent to Institut Pasteur de Guadeloupe for epidemiological studies included in her Ph.D. thesis (awarded by the University of the West Indies, St. Augustine, Trinidad and Tobago; see References [12] and [13]). Permission to retrospectively genotype these DNAs was kindly granted by Dr. Persaud (Chief Medical Officer, Guyana), and Dr. Eersel (Director, Ministry of Health, Suriname). E.S. would like to thank the International Division of the Institut Pasteur, Paris, France for her Ph.D. fellowship.

REFERENCES

[1] World Health Organization. Global Tuberculosis Report 2014. WHO Library Cataloguing-in-Publication Data, ISBN 978 92 4 156480 9 (WHO/HTM/TB/2014.08), Geneva. <http://apps.who.int/iris/bitstream/10665/137094/1/9789241564809_eng.pdf>.

[2] K.R. Dickman, L. Nabyonga, D.P. Kateete, F.A. Katabazi, B.B. Asiimwe, H.K. Mayanja, et al, Detection of multiple strains of *Mycobacterium tuberculosis* using MIRU-VNTR in patients with pulmonary tuberculosis in Kampala, Uganda, BMC Infect. Dis. 10 (2010) 349, http://dx.doi.org/10.1186/1471-2334-10-349.

[3] C. Mulenga, I.C. Shamputa, D. Mwakazanga, N. Kapata, F. Portaels, L. Rigouts, Diversity of *Mycobacterium tuberculosis* genotypes circulating in Ndola, Zambia, BMC Infect. Dis. 10 (2010) 177, http://dx.doi.org/10.1186/1471-2334-10-177.

[4] T. Cohen, D. Wilson, K. Wallengren, E.Y. Samuel, M. Murray, Mixed-strain *Mycobacterium tuberculosis* infections among patients dying in a hospital in KwaZulu-Natal, South Africa, J. Clin. Microbiol. 49 (2011) 385–388, http://dx.doi.org/10.1128/JCM.01378-10.

[5] I.C. Shamputa, L. Jugheli, N. Sadradze, E. Willery, F. Portaels, P. Supply, et al, Mixed infection and clonal representativeness of a single sputum sample in tuberculosis patients from a penitentiary hospital in Georgia, Respir. Res. 7 (2006) 99, http://dx.doi.org/10.1186/1465-9921-7-99.

[6] Y. Navarro, M. Herranz, L. Pérez-Lago, M. Martínez Lirola, M.J. Ruiz-Serrano, E. Bouza, et al, Systematic survey of clonal complexity in tuberculosis at a populational level and detailed characterization of the isolates involved, J. Clin. Microbiol. 49 (2011) 4131–4137, http://dx.doi.org/10.1128/JCM.05203-11.

[7] T. Cohen, P.D. van Helden, D. Wilson, C. Colijn, M.M. McLaughlin, I. Abubakar, et al, Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control, Clin. Microbiol. Rev. 25 (2012) 708–719, http://dx.doi.org/10.1128/CMR.00021-12.

[8] P. Supply, E. Mazars, S. Lesjean, V. Vincent, B. Gicquel, C. Locht, Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome, Mol. Microbiol. 36 (2000) 762–771.

[9] P. Supply, C. Allix, S. Lesjean, M. Cardoso-Oelemann, S. Rüsch-Gerdes, E. Willery, et al, Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*, J. Clin. Microbiol. 44 (2006) 4498–4510, http://dx.doi.org/10.1128/JCM.01392-06.

[10] A. Martín, M. Herránz, M.J.R. Serrano, E. Bouza, D. García de Viedma, Rapid clonal analysis of recurrent tuberculosis by direct MIRU-VNTR typing on stored isolates, BMC Microbiol. 7 (2007) 73, http://dx.doi.org/10.1186/1471-2180-7-73.

[11] J.F. Reyes, M.M. Tanaka, Mutation rates of spoligotypes and variable numbers of tandem repeat loci in *Mycobacterium tuberculosis*, Infect. Genet. Evol. 10 (2010) 1046–1051, http://dx.doi.org/10.1016/j.meegid.2010.06.016.

[12] E. Streit, S. Baboolal, P.E. Akpaka, J. Millet, N. Rastogi, Finer characterization of *Mycobacterium tuberculosis* using spoligotyping and 15-loci MIRU-VNTRs reveals phylogeographical specificities of isolates circulating in Guyana and Suriname, Infect. Genet. Evol. 30 (2015) 114–119, http://dx.doi.org/10.1016/j.meegid.2014.12.015.

[13] J. Millet, S. Baboolal, E. Streit, P.E. Akpaka, N. Rastogi, A First Assessment of *Mycobacterium tuberculosis* genetic diversity and drug-resistance patterns in twelve Caribbean territories, Biomed. Res. Int. 2014 (2014) 1–11, http://dx.doi.org/10.1155/2014/718496.

[14] J. Kamerbeek, L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, S. Kuijper, et al, Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology 35 (1997) 907–914.

[15] K. Brudey, J.R. Driscoll, L. Rigouts, W.M. Prodinger, A. Gori, S.A. Al-Hajoj, et al, *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology, BMC Microbiol. 6 (2006) 23, http://dx.doi.org/10.1186/1471-2180-6-23.

[16] C. Demay, B. Liens, T. Burguière, V. Hill, D. Couvin, J. Millet, et al, SITVITWEB – a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology, Infect. Genet. Evol. 12 (2012) 755–766, http://dx.doi.org/10.1016/j.meegid.2012.02.004.

[17] C.B. Ford, P.L. Lin, M.R. Chase, R.R. Shah, O. Iartchouk, J. Galagan, et al, Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection, Nat. Genet. 43 (2011) 482–486, http://dx.doi.org/10.1038/ng.811.

[18] M.C. Oelemann, R. Diel, V. Vatin, W. Haas, S. Rüsch-Gerdes, C. Locht, et al, Assessment of an optimized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing system combined with spoligotyping for population-based molecular epidemiology studies of tuberculosis, J. Clin. Microbiol. 45 (2007) 691–697, http://dx.doi.org/10.1128/JCM.01393-06.

[19] C. Allix-Béguec, M. Fauville-Dufaux, P. Supply, Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of *Mycobacterium tuberculosis*, J. Clin. Microbiol. 46 (2008) 1398–1406, http://dx.doi.org/10.1128/JCM.02089-07.

[20] Y. Navarro, L. Pérez-Lago, F. Sislema, M. Herranz, L. de Juan, E. Bouza, et al, Unmasking subtle differences in the infectivity of microevolved *Mycobacterium tuberculosis* variants coinfecting the same patient, Int. J. Med. Microbiol. 303 (2013) 693–696, http://dx.doi.org/10.1016/j.ijmm.2013.10.002.

[21] L. Pérez-Lago, Y. Navarro, M. Herranz, E. Bouza, D. García-de-Viedma, Differences in gene expression between clonal variants of *Mycobacterium tuberculosis* emerging as a result of microevolution, Int. J. Med. Microbiol. 303 (2013) 674–677, http://dx.doi.org/10.1016/j.ijmm.2013.09.010.

[22] L.C. Lazzarini, J. Rosenfeld, R.C. Huard, V. Hill, J.R. Lapa e Silva, et al, *Mycobacterium tuberculosis* spoligotypes that may derive from mixed strain infections are revealed by a novel computational approach, Infect. Genet. Evol. 12 (2012) 798–806, http://dx.doi.org/10.1016/j.meegid.2011.08.028.

[23] C. Ford, K. Yusim, T. Ioerger, S. Feng, M. Chase, M. Greene, et al, *Mycobacterium tuberculosis* – heterogeneity revealed through whole genome sequencing, Tuberculosis (Edinb.) 92 (2012) 194–201, http://dx.doi.org/10.1016/j.tube.2011.11.003.