Genotyping of *Mycobacterium leprae* present on Ziehl-Neelsen-stained microscopic slides and in skin biopsy samples from leprosy patients in different geographic regions of Brazil

Amanda Nogueira Brum Fontes, Harrison Magdiner Gomes, Marcelo Ivens de Araujo, Edson Cláudio Araripe de Albuquerque, Ida Maria Fochiani Dias Baptista, Maria Manuela da Fonseca Moura, Denise Silva Rezende, Maria Cristina Vidal Pessolani, Flávio Alves Lara, Maria Araci de Andrade Pontes, Heitor de Sá Gonçalves, Norma Lucena-Silva, Euzenir Nunes Sarno, Varalakshmi D Vissa, Patrick J Brennan, Philip Noel Suffys

1Laboratório de Biologia Molecular Aplicada a Micobactérias 2Laboratório de Hansenase 3Laboratório de Microbiologia Celular, Instituto Oswaldo Cruz-Fiocruz, Rio de Janeiro, RJ, Brasil 4Departamento de Imunologia, Centro de Pesquisas Aggeu Magalhães-Fiocruz, Recife, PE, Brasil 5Instituto Lauro de Souza Lima, Bauru, SP, Brasil 6Departamento de Bioengenharia, Universidade Federal de Rondônia, Porto Velho, RO, Brasil 7Centro de Referência em Dermatologia Dona Libânia, Fortaleza, CE, Brasil 8Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO, USA

We analysed 16 variable number tandem repeats (VNTR) and three single-nucleotide polymorphisms (SNP) in Mycobacterium leprae present on 115 Ziehl-Neelsen (Z-N)-stained slides and in 51 skin biopsy samples derived from leprosy patients from Ceará (n = 23), Pernambuco (n = 41), Rio de Janeiro (n = 22) and Rondônia (RO) (n = 78). All skin biopsies yielded SNP-based genotypes, while 48 of the samples (94.1%) yielded complete VNTR genotypes. We evaluated two procedures for extracting *M. leprae* DNA from Z-N-stained slides: the first including Chelex and the other combining proteinase and sodium dodecyl sulfate. Of the 76 samples processed using the first procedure, 30.2% were positive for 16 or 15 VNTRs, whereas of the 39 samples processed using the second procedure, 28.2% yielded genotypes defined by at least 10 VNTRs. Combined VNTR and SNP analysis revealed large variability in genotypes, but a high prevalence of SNP genotype 4 in the Northeast Region of Brazil. Our observation of two samples from RO with an identical genotype and seven groups with similar genotypes, including four derived from residents of the same state or region, suggest a tendency to form groups according to the origin of the isolates. This study demonstrates the existence of genetically related *M. leprae* genotypes and that Z-N-stained slides are an alternative source for *M. leprae* genotyping.

Key words: leprosy - *Mycobacterium leprae* - Ziehl-Neelsen-stained microscope slides - molecular epidemiology - variable number tandem repeats - single-nucleotide polymorphisms

Leprosy remains a public health problem in Brazil, where 30,298 new cases were detected in 2011 (MS 2011). The disease is caused by infection with *Mycobacterium leprae* and diagnosis is based on an observation of clinical symptoms, eventually supported by bacteriologic and histopathological features. Although leprosy is one of the oldest recorded infectious diseases, little is known regarding the factors that influence disease transmission, the evolution from exposure to infection, the development of active disease and the factors behind clinical outcomes.

Some studies have explored the use of genotyping for differentiation of bacterial strains as a tool during epidemiological investigations and, recently, multiple-locus variable number tandem repeat (VNTR) analysis (MLVA) of a set of micro and mini-satellites of *M. leprae* proved to be useful for differentiation to the strain level (Grothhouse et al. 2004, Truman et al. 2004, Zhang et al. 2005, Kimura et al. 2009). Due to their considerable mutation rate, some of these markers could be helpful in evaluating transmission patterns in communities (Weng et al. 2007, 2011), distinguishing between reactivation and re-infection (da Silva Rocha et al. 2011) and evaluating bacterial population structures at the local and global levels. These applications have been demonstrated in studies that characterised large panels of VNTRs in *M. leprae* from patients in China (Weng et al. 2007, 2011), India (Young et al. 2008), the Philippines (Sakamura et al. 2009a, b), Thailand (Srisungnam et al. 2009), Mexico (Matsuoka et al. 2009) and Colombia (Cardona-Castro et al. 2009). In addition, single-nucleotide polymorphism (SNP) analysis elucidates the introduction, evolution and spread of *M. leprae* in countries on a global scale (Monot et al. 2005, 2009).

Our earlier findings based on an analysis of VNTRs and SNPs in *M. leprae* were mainly based on samples from residents in the states of Rio de Janeiro (RJ) and São Paulo (SP), Southeast Brazil, and exhibited high genetic variability (Fontes et al. 2009). Here, we present...
genotype characteristics on a larger set of samples from RJ, Rondônia (RO), Ceará (CE) and Pernambuco (PE), North and Northeast Brazil; in addition to skin biopsies, we include DNA extracts prepared from Ziehl-Neelsen (Z-N)-stained slides.

**SUBJECTS, MATERIALS AND METHODS**

*Patients and clinical samples* - The samples included in this study were 51 skin biopsies from different leprosy patients who had been subjected to standardised diagnostic procedures based on the determination of the presence of skin lesions, nerve damage and microscopic analysis for the presence of acid-fast bacilli (AFB). These samples were frozen immediately after collection and were stored at -20°C until DNA extraction. In addition, we analysed Z-N-stained slides containing slit skin smears derived from ear lobes, elbows and lesions. Leprosy disease subtype classifications of all patients, the number of samples analysed per state, the years of collection and the bacterial indices of the Z-N-stained slides are shown in Table 1. Both Z-N-stained slides and skin biopsies were available from two patients and were analysed. The first pair was collected in 2007 and the second in 2008.

*Nucleic acid extraction* - Skin biopsies were processed using the DNeasy Blood & Tissue Kit (Qiagen Inc, Valencia, CA), according to the manufacturer’s instructions, whereas two different protocols for the extraction of nucleic acids from material present on Z-N-stained slides were followed. Samples from RO and RJ were treated using the protocol of van der Zanden et al. (1998) with some modifications, including the application of 25 µL of distilled water on each of the three fixed (1998) with some modifications, including the application of 25 µL of distilled water on each of the three fixed and stained AFB-containing tissue sites on the glass slides, removal of fixed material from the glass surface by scraping with a clean microscopic slide and collection in a single microcentrifuge tube. Samples were mixed with an equal volume of 15% Chelex-100 (Sigma-Aldrich) in water, incubated for 30 min at 97°C and centrifuged at 14,462 g for 10 min. The supernatant was then transferred to another microcentrifuge tube and was stored at -20°C. The Z-N-stained slides from PE were processed according to Kamble et al. (2010). Briefly, this method consisted of adding 150 µL of lysis buffer (1 mg/mL proteinase K and 0.05% Tween 20) on slide-bound tissue smears, followed by scraping and collection as described above. After incubation at 60°C overnight, 30 µL of 10% sodium dodecyl sulfate was added, followed by an incubation at 60°C for 1 h and then at 94°C for 15 min. After centrifugation, the supernatant was removed and ethanol was added to a final concentration of 70%, inducing precipitation of nucleic acids by incubation at -20°C. The supernatant was discarded and the pellet was washed again with ethanol, air-dried, re-suspended in 30 μL of water and stored at -20°C.

*Genotyping by MLVA* - A selection of 91 DNA extracts (51 from skin biopsies and 40 from Z-N-stained slides) was submitted for genotyping using 16 VNTRs, as described by Kimura et al. (2009). For each sample, four multiplex polymerase chain reactions (PCR) reactions were performed that generated 16 amplicons. Allele copy number was determined by denaturation of the amplicons followed by capillary gel electrophoresis on an ABI 3130 Genetic Analyzer sequencer using GeneScan-500 LIZ sizing standard (Applied Biosystems, Foster City, CA, EUA). Copy number definition was determined by the Peak Scanner software (Applied Biosystems).

*SNP genotype analysis* - To differentiate between the four genotypes of *M. leprae* based on three SNPs, we used a procedure that combined PCR-restriction fragment length polymorphism (RFLP) and direct sequencing, as described by Sakamuri et al. (2009b). In brief, differentiation of genotypes 1/2 from 3/4 was achieved by subjecting the locus, including the SNP at nucleotide position 2,935,685, to PCR-RFLP analysis, which involved incubation of 5 μL of PCR product with 5 units of BstUI (New England BioLabs, Beverly, MA) at 60°C

**TABLE**

<table>
<thead>
<tr>
<th>State</th>
<th>LL</th>
<th>BL</th>
<th>B</th>
<th>TT</th>
<th>NI</th>
<th>Year of collection</th>
<th>BI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE</td>
<td>SB</td>
<td>9</td>
<td>11</td>
<td>-</td>
<td>1</td>
<td>2008 (n = 22)/2009 (n = 1)</td>
<td>-</td>
<td>23</td>
</tr>
<tr>
<td>PE</td>
<td>SB</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2002 (n = 2)</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Z-N</td>
<td>21</td>
<td>18</td>
<td>-</td>
<td>-</td>
<td>2004 (n = 1)/2007 (n = 6)/2008 (n = 32)</td>
<td>0.5-4.5</td>
<td>39</td>
</tr>
<tr>
<td>RJ</td>
<td>SB</td>
<td>7</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>2007 (n = 4)/2009 (n = 4)</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Z-N</td>
<td>10</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>2008 (n = 10)/2009 (n = 4)</td>
<td>2.0-4.75</td>
<td>14</td>
</tr>
<tr>
<td>RO</td>
<td>SB</td>
<td>7</td>
<td>1</td>
<td>-</td>
<td>9</td>
<td>2007 (n = 8)/2008 (n = 10)</td>
<td>-</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Z-N</td>
<td>27</td>
<td>31</td>
<td>1</td>
<td>3</td>
<td>2001 (n = 1)/2005 (n = 6)/2006 (7)/2007 (n = 30)/2008 (n = 15)/NI (n = 3)</td>
<td>1.75-5.25</td>
<td>62</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>83</td>
<td>16</td>
<td>58</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>-</td>
</tr>
</tbody>
</table>

B: borderline leprosy without specification of the subtype; BL: borderline lepromatous leprosy; TT: borderle tuberculoid leprosy; CE: Ceará; LL: lepromatous leprosy; NI: no information on disease classification or year of collection; PE: Pernambuco; RJ: Rio de Janeiro; RO: Rondônia; TT: tuberculoid leprosy.
for 1 h. Digestion indicates genotype 3 or 4, while no digestion indicates genotype 1 or 2. For further differentiation of genotypes 3 and 4, restriction enzyme analysis of the PCR product, which includes the SNP at nucleotide position 14,676, was achieved by incubating 5 µL of PCR product with 5 U of SmII for 1 h at 50°C; digestion indicates SNP genotype 4, while lack of digestion is typical for genotype 3. Finally, alleles containing either SNP genotype 1 or 2 were identified by sequencing as described by Monot et al. (2005).

Phylogenetic analysis - For phylogenetic analysis, we selected genotypes obtained from 78 of the 91 samples, excluding 13 that lacked allele definition of one or more VNTR or of the SNP loci. Genotype similarity was determined using the program Bionumerics (version 6.5, Applied Maths, Belgium) by construction of a similarity matrix using the unweighted pair group method with arithmetic mean (UPGMA) and the Dice similarity coefficient.

RESULTS

Genotyping results using biopsy samples - The 51 biopsy samples were subjected to genotyping, using both the MLVA procedure and SNP analysis. All samples yielded an SNP genotype, while 48 of them (94.1%) were characterised by 16 VNTRs, two (3.9%) by 15 VNTRs and a single sample (2%) by 11 VNTRs. The samples genotyped for 15 VNTRs were from CE and RO and the analysis failed to yield the VNTRs 18-8 and AT15, respectively. The sample genotyped for 11 VNTRs was from CE and belonged to a paucibacillary patient; in this instance, the VNTRs 21-3, AC9, AT15, TA10 and 12-5 were not defined (data not shown).

Comparison DNA extraction procedures and genotyping on Z-N-stained slides - Among the 76 samples processed using the procedure described by van der Zanden et al. (1998), 23 (30.3%) we could define copy numbers for 16 (n = 20) or 15 (n = 3) VNTRs and two more samples (2.6%) could be genotyped for eight or nine VNTRs (data not shown). The rest did not generate any VNTR copy numbers (data not shown). Although a limited number of samples generated partial results, we did not observe VNTR-associated genotyping failure. An analysis of the possible relationship between genotyping yield and date of Z-N-stained slide preparations revealed that none of the slides from patients diagnosed in 2001, 2005 and 2006 generated PCR products. However, slides that had been prepared from patients diagnosed in 2007, 2008 or 2009 yielded positive PCRs and could be further genotyped for 16 or 15 VNTRs in 10 (33.3%), 11 (44%) and two (50%) of the cases for 2007, 2008 and 2009, respectively, therefore demonstrating a slight increase in the genotyping yield efficiency. Among the 25 samples that yielded results for at least eight VNTRs and 24 (96%) yielded SNP-based genotypes. The bacterial index (BI) of samples genotyped by VNTRs and SNPs ranged from 1.75-5.25, but we did not observe a relationship between genotyping yield and BI.

Among the 39 Z-N-stained slide samples processed according to Kamble et al. (2010), 11 (28.2%) yielded genotypes defined by either 16 or 15 VNTRs, while four samples (10.3%) rendered genotypes with 10, 12 or 13 VNTRs. When considering positivity among samples collected from patients diagnosed more recently, namely in 2007 and 2008, the genotyping yield was also higher, with genotypes defined by 16 or 15 VNTRs for four (80%) and seven (70%) of the samples, respectively. Among the 15 samples with genotypes defined by at least 10 VNTRs (n = 15), 14 (93.3%) yielded SNP genotypes. The BI of samples genotyped ranged from 0.75-4.5 and again, no relationship between BI and genotyping yield was observed.

Frequency of genotypes defined by SNP analysis - We obtained conclusive SNP genotypes in 89 (97.8%) of 91 samples, derived from 89 patients from four Brazilian states. Two Z-N-stained slides (2.2%), one from RJ and another from PE, generated insufficient PCR products. In another two patients with dual samples, the same genotype was obtained, which is suggestive of a single input. Among the 87 patients, SNP genotype 4 was predominant (n = 46, 52.9%), genotype 3 was identified in 37 patients (42.5%) and genotype 1 was detected less frequently in four cases (4.6%). The distribution of each SNP genotype per state is demonstrated in Fig. 1.

Genotypes as defined by analysis of 16 VNTRs and SNPs - All VNTRs used in this analysis were polymorphic. The smallest discriminatory power was observed for locus 21-3 (0.067), consisting of two alleles (1 and 2 copies), whereas the highest allelic diversity was observed at locus AT15 (0.94), consisting of 22 different alleles (between 7-32 copies). The earlier observed correlation between the 27-5 and 12-5 allele combination and SNP genotype was also present here: 40 (88.9%) SNP genotype 4 samples possessed the 5:4 copy number allele combination, while 24 cases (72.7%) of SNP genotype 3 had the 4.5 copy number combination.

The dual samples from the two patients with identical SNP genotypes differed in copy number at three loci each. In both subjects, biopsy samples were collected before Z-N-stained slides. In the first case, where 15 VNTRs were used for comparison (because the allele definition failed), we observed a decrease of five and four copies for TTC21 and TA15, respectively. In the case of AC9, we observed an increase of one copy; the other 12 loci were identical between the samples. This situation exhibited a decrease of one copy for AT17 and an increase of two and three copies for TTC21 and AT15. The other 13 loci were identical.

Genotype clustering and phylogeny - Fig. 2 represents an UPGMA-base dendrogram of 78 samples with complete genotypes in which only one cluster of two patients (3794/L15) from RO was observed. Interestingly, this involved a skin biopsy (SB) and a Z-N-stained slide, confirming the possibility that Z-N-stained slides could be used for cluster definition. After excluding the four highly polymorphic AT/TA repeats for the definition of genotypes, seven more clusters were observed, one consisting of four samples and the remainder of two samples each. The largest cluster was composed of two samples from RO (L21/ L06) and two samples from CE (301005/301001), while
three more clusters were formed by isolates derived from the same state: one from RJ (MFMS/F1903) and two from RO (2818/L24; 3788/715). In addition, one cluster was from the Northeast Region (79-05/301024) and two clusters contained isolates from RO and CE (L51/301030; 902/301004). By relating sample type and clusters, we observed that five groups were composed of Z-N-stained slides and skin biopsies, while two were composed of SB samples alone. In addition, all but one of the groups was composed of isolates with the same SNP genotype; the exception included the two samples from RJ that presented SNP genotypes 3 and 4.

Upon construction of a minimum spanning tree (MST) using the same genotypes (Fig. 3), two major branches were observed, showing that isolates from CE had genotypes that were more closely related to those observed in PE, while genotypes from samples from RJ were more similar to those of RO. However, genotypes from the latter state were similarly distributed in both branches of the tree, perhaps due to the larger sample number, which also suggests a higher genetic diversity than that of isolates from other regions. Interestingly, the only cluster was composed of genotypes from Ron- donian isolates and was the central interactive node between both branches of the phylogenetic tree.

An analysis of the relationship between VNTR and SNP-based genotypes revealed a strong correlation between the positioning of the genotypes within the MST (data not shown).

**DISCUSSION**

The first reliable data on genomic variation between isolates of *M. leprae* were published just over a decade ago (Matsuoka et al. 2000, Shin et al. 2000) and since then, several polymorphic VNTRs and SNPs have been reported and evaluated in material from leprosy patients residing in different regions around the world (Weng et al. 2007, Young et al. 2008, Cardona-Castro et al. 2009, Matsuoka et al. 2009, Sakamuri et al. 2009a, b, Srisungnam et al. 2009). Using this genotyping approach in biopsy samples from unrelated leprosy patients from RJ and SP (Fontes et al. 2009), we observed a high genetic variability of *M. leprae*; however, because both regions are part of Southeast Brazil, a region that covers only a limited area of the country (10.9% of the Brazilian territory), we investigated genotypes from patients who are residents of the North and Northeast Regions of the country. Additionally, as biopsy sampling is an invasive procedure and because preparing Z-N-stained slides is part of the diagnostic procedure for leprosy in many health centres of Brazil, we evaluated the usefulness of Z-N-stained slides as a genotyping source of the causative agent.

Two procedures of DNA extraction were used: the first was previously utilised for genotyping of *Mycobacterium tuberculosis* (*Mt b*) on Z-N slides prepared from sputum samples of tuberculosis patients (van der Zanden et al. 1998) and the second was reported recently for PCR-mediated detection of *M. leprae* DNA on Z-N-stained slides with bacterial counts of 0, but that were PCR-positive in...
FIG. 2: unweighted pair group method with arithmetic mean-based dendrogram based on complete variable number tandem repeat genotypes from 78 Brazilian Mycobacterium leprae isolates. The single-nucleotide polymorphism type was not included for tree building and merely attached to the Figure. CE: Ceará; PE: Pernambuco; RJ: Rio de Janeiro; RO: Rondônia; SB: skin biopsy; Z-N: Ziehl-Neelsen.
32.6% (n = 15) of the samples (Kamble et al. 2010). When comparing typing efficiency, both extraction procedures returned similar results, yielding 28-30% complete genotypes. An inverse correlation between the length of time that the slides had been stored and genotype quality was verified, while no direct correlation between the latter and the BI was observed. Complete M. leprae genotypes were sometimes obtained in material with low BI. Our data demonstrate that Z-N-stained slides are an additional source for DNA fingerprinting of M. leprae, but more studies are needed to better determine the effect of long-term storage of slides on genotype quality.

We observed 90 unique genotypes from 89 patients, including different genotypes in two samples available from two patients and identical patterns in isolates from two residents of RO. When decreasing stringency for genotype similarity was employed, many of the cases that yielded clusters also had similar origins. Taken together, both results strongly suggest region-associated genotype similarity. Unfortunately, we did not have access to epidemiological data of the patients to confirm this association. The need to better define stringency of definition of genotype similarity was also shown by our observation of the differences in some VNTRs between Z-N-stained slides and skin biopsies from the same patients. Genotype differences in clinical specimens of the same individual have been investigated and some studies demonstrated the existence of an identical pattern in minisatellites and a small variation in microsatellites with the highest allelic diversities, such as AT17, TTC21 (Sakamuri et al. 2009a) and TA10 (Xing et al. 2009). These results suggest that some short tandem repeat loci are most likely prone to stuttering during incubation, treatment and transmission.

Our earlier study also demonstrated that SNP genotype 3 is the most frequent among M. leprae isolates in Southeast Brazil (Fontes et al. 2009), also occurring predominantly in Europe, North Africa and the Americas (Monot et al. 2005). Here, we confirm the predominance of SNP genotype 3 in RJ and frequently observed this SNP genotype in RO. Conversely, SNP genotype 4 was more predominant in CE and PE, two bordering states in Northeast Brazil. This SNP genotype is characteristic of Africa, the French West Indies, Brazil (Monot et al. 2005), Mexico (Matsuoka et al. 2009) and Japanese Brazilians, who were likely infected in Brazil (Weng et al. 2007). SNP genotype 4 isolates in Latin American countries are theorised to have originated in Africa, in agreement with the notion that leprosy was carried by the slave trade (Monot et al. 2005). The frequencies of SNP genotypes 1, 3 and 4 and the lack of SNP genotype 2 are consistent with the ethnic groups that were present during Brazilian colonisation, as presented by Monot et al. (2005).

In an earlier study, we observed a 4/5 copy number combination of VNTR loci 27-5 and 12-5 and the SNP genotype 3 in 87.5% of isolates and the 5/4 copy number combination in 85.7% of isolates with SNP genotype 4 (Fontes et al. 2009); this result had also been observed in M. leprae samples from Colombia (Cardona-Castro et al. 2009). Associations between different markers, such as SNPs and VNTRs, have also been detected in Mtb (Filliol et al. 2006). Our data demonstrated that this association was less pronounced for SNP genotype 3 when analysing patients from more regions, now occurring in 72.7% (n = 24) of these genotypes, while the strong association of SNP genotype 4 with the 5/4 combined VNTR allele type was maintained (88.9%, n = 40).

In conclusion, our study strongly suggests the existence of geographically related M. leprae genotypes in Brazil and demonstrates that Z-N-stained slides are an alternative sample source for the genotyping of M. leprae.

ACKNOWLEDGEMENTS

To the Genomic Platform for DNA Sequencing (PDTIS-Fiocruz).

REFERENCES


