

Population-Based Molecular Epidemiology of Leprosy in Cebu, Philippines[∇]

Rama Murthy Sakamuri,¹ Miyako Kimura,¹ Wei Li,¹ Hyun-Chul Kim,² Hyeoung Lee,²
Madanahally D. Kiran,¹ William C. Black IV,¹ Marivic Balagon,³ Robert Gelber,³
Sang-Nae Cho,² Patrick J. Brennan,¹ and Varalakshmi Vissa^{1*}

Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, Colorado¹; Department of Microbiology, College of Medicine, Yonsei University, Seoul, Republic of Korea²; and Leonard Wood Memorial, Cebu Skin Clinic, Cebu, Philippines³

Received 18 October 2008/Returned for modification 22 May 2009/Accepted 8 June 2009

To address the persisting problem of leprosy in Cebu, Philippines, we compiled a database of more than 200 patients who attend an established referral skin clinic. We described the patient characteristics in conventional demographic parameters and also applied multiple-locus variable-number tandem-repeat (VNTR) analysis (MLVA) and single nucleotide polymorphism (SNP) typing for *Mycobacterium leprae* in biopsied skin lesion samples. These combined approaches revealed that transmission is ongoing, with the affected including the young Cebuano population under 40 years of age in both crowded cities and rural areas of the island. The emergence of multicase families (MCF) is indicative of infection unconstrained by standard care measures. For the SNPs, we designed a low-cost PCR-restriction fragment length polymorphism typing method. MLVA in *M. leprae* was highly discriminatory in this population yet could retain broad groups, as defined by the more stable SNPs, implying temporal marker stability suitable for interpreting population structures and evolution. The majority of isolates belong to an Asian lineage (SNP type 1), and the rest belong to a putative postcolonial lineage (SNP type 3). Specific alleles at two VNTR loci, (GGT)5 and 21-3, were highly associated with SNP type 3 in this population. MLVA identified *M. leprae* genotype associations for patients with known epidemiological links such as in MCFs and in some villages. These methods provide a molecular database and a rational framework for targeted approaches to search and confirm leprosy transmission in various scenarios.

During the last 4 to 5 years, genetic variation in *Mycobacterium leprae* has been investigated for the purpose of strain typing. Although the *M. leprae* genome (4) has undergone reductive evolution and is highly mutated, limited genome variability has been found between global isolates, and except for loci prone to mutation, such as variable-number tandem repeats (VNTR) (7, 8, 12, 13, 16, 19, 22, 24, 25, 26), *M. leprae* strains are highly clonal species. Three single nucleotide polymorphisms (SNPs) were subsequently discovered by further genome sequencing efforts that allowed the separation of global isolates into four subtypes (14).

Formal, systematic study of *M. leprae* diversity by the application of the known polymorphic markers in defined endemic settings for studying extant population structures and leprosy transmission is limited. Previously, we presented the outcome of a focused study in Qiubei County in Yunnan Province, South West China (22). Using VNTR loci, we discovered subgroups within a major lineage. A differential geographical distribution of these subgroups was seen across the county. Furthermore, we noted the conservation of *M. leprae* genotypes carried by patients of multicase families (MCFs), which is indicative of localized transmission from shared sources.

We now extend such approaches to Cebu, Philippines, an island in the Central Visayas where leprosy is still in existence. From a case detection rate of 5.1 in 2001 to the current rate of

4.19, with an actual rate of about 300 registered cases for the entire island of Cebu, multibacillary cases have consistently comprised 85% to 90% of the total number of registered cases. We describe observations from conventional epidemiological and novel molecular studies. Currently, we can routinely map up to 15 VNTR loci using multiplex PCR and fragment length analysis (FLA) methods (8). In this study, we demonstrate the development, feasibility, and applicability of a low-cost technology, a PCR-restriction fragment length polymorphism (RFLP) scheme, for rapid SNP subtyping of *M. leprae* DNA (14).

MATERIALS AND METHODS

ADML strains of human origin. Four armadillo-derived *M. leprae* (ADML) genomic DNAs (Thai-53, 3039/321, NHDP-63, and BR4923) were used as reference strains and taken from a panel of archived samples at Colorado State University (CSU). These four isolates were previously characterized as belonging to *M. leprae* SNP types 1, 2, 3, and 4 (14). *M. leprae* DNA from the ADML isolates was isolated by use of a procedure described previously (7).

Description of clinical study site, patients, and biological samples. The patients in this study were drawn from those who consulted at the Cebu Skin Clinic (CSC), Leonard Wood Memorial Leprosy Research Centre, Cebu, Philippines. For *M. leprae* strain typing purposes, we accessed skin biopsy specimens and/or DNA extracts from three separate studies performed in collaboration with CSC staff. A total of 228 samples were thus compiled. These samples included DNA extracted from skin biopsy specimens of patients from the 1980s (30 biopsy samples) (8), a collection of 100 frozen sectioned skin biopsy specimens (from newly diagnosed untreated patients or those in the early stages of leprosy treatment) during the 2003–2004 period, and 98 ethanol-fixed skin biopsy specimens from an ongoing study, which started in 2006. These three sets of samples from the 1980s, 2003 to 2004, and 2006 to 2007 are henceforth referred to as groups “R,” “M,” and “L,” respectively.

Besides obtaining basic details of the medical histories of the patients including the leprosy condition and treatment regimen, information about the patients' geographical area of residence (village, city, and province); type/place(s) of

* Corresponding author. Mailing address: Department of Microbiology, Immunology, and Pathology, Colorado State University, 1682 Campus Delivery, Fort Collins, CO 80523. Phone: (970) 491-0752. Fax: (970) 491-1815. E-mail: vvisa@colostate.edu.

[∇] Published ahead of print on 1 July 2009.

employment and education; nature of dwelling; and source of drinking, cooking, and bathing water spanning a prior 20-year period was attained with a standardized questionnaire at the time of recruitment since the "L" study began. All studies involving biological sample collections were performed following approval from the governing human research ethical committee and informed consent procedures as necessary.

DNA extraction, multiplex PCR, and FLA for 15 VNTRs. DNA was extracted from biopsy specimens using a Qiagen (Valencia, CA) DNeasy tissue kit as described previously (7). A multiplex PCR protocol described previously by Kimura et al. comprising four reactions for the amplification of 15 VNTR loci was achieved using a multiplex PCR enzyme kit (Qiagen) and fluorescent 5'-labeled forward primers and reverse unlabeled primers (8). These 15 VNTRs include nine microsatellites [(AC)8a, (AC)8b, (AC)9, (AT)15, (AT)17, (TA)18, (GGT)5, (GTA)9, and (TTC)21] and six minisatellites (6-7, 12-5, 18-8, 21-3, 23-3, and 27-5). Multiplex PCR products were separated by capillary electrophoresis (Applied Biosystems Genetic Analyzer 3130 at the Macromolecular Resource Facility, CSU) and FLA to determine the major allele for each VNTR locus using GeneMapper, version 3.7, software. When DNA sequencing was performed, a multiplex PCR sample was diluted 10-fold; 1 to 2 μ l was combined with 10 pmol of the unlabeled forward primer for the ABI Big Dye cycle sequencing reaction at the Macromolecular Resource Facility, CSU (8).

Development of a PCR-RFLP method for SNP subtyping of *M. leprae*. Known *M. leprae* SNP loci 1, 2, and 3 (nucleotide positions 14676, 1642875, and 2935685, respectively, on the sequenced TN strain) were amplified using previously reported primer sequences (14). Single or multiplex PCRs were set up using 1 to 2 μ l of the DNA extract. An additional forward primer was designed for locus 2 for a shorter product of 114 bp instead of 190 bp. High Fidelity super mix (Invitrogen, Carlsbad, CA) was used for PCRs (final volume of 20 μ l). The DNA products were resolved by agarose gel electrophoresis and detected by ethidium bromide staining.

The three SNP-containing sequences were examined for restriction sites overlapping the SNP using NEBcutterV2.0 (<http://tools.neb.com/NEBcutter2/index.php>). Restriction digest assays were developed using SmlI, CviKI-1, and BstUI for SNP loci 1, 2, and 3, respectively (New England Biolabs, MA). The PCR products (5 μ l) were digested with 1 unit of the enzymes. The SmlI, BstUI, and CviKI-1 reactions were performed at 55°C, 60°C, and 37°C, respectively, for 1 h. The SmlI- and BstUI-uncut and -cut DNAs were subjected to electrophoresis on 3% agarose gels. The uncut and CviKI-1-digested SNP locus 2 products were resolved on a 12% acrylamide gel. The gels were stained with ethidium bromide and visualized by UV transillumination.

Diversity and phylogeny. The Microsatellite Tool Kit (<http://animalgenomics.ucd.ie/sdepar/ms-toolkit/>) was used to calculate allele frequency and average allelic diversity. PAUP4.0b software was used to apply the maximum parsimony (MP) algorithm with equal or unequal weighting for all the VNTR loci (17) for a maximum of 1,000 trees. An allele difference of 1 repeat unit (increase or decrease) per locus was calculated as 1 mutation step, and a difference of 3 was calculated as 3 steps, etc., as described previously (22). The 50% consensus trees are reported. The figure for the 50% majority-rule tree was generated with the PHYLIP program by using the "consensus" application (6).

RESULTS

Leprosy in Philippines and in the island province of Cebu.

In 1998, about a decade after multidrug therapy (MDT) was implemented nationwide, the Philippines reached its elimination goal, with a prevalence of 0.9 cases/10,000 persons, with a total of 6,872 registered cases. After another decade of leprosy elimination campaigns, this has further gone down to a prevalence of 0.36 cases/10,000 persons, with a total of 3,156 registered cases. For Cebu, data from the 2001–2006 period show that the prevalence has been sustained: 0.48 cases/10,000 persons in 2001 and 0.46 cases/10,000 persons in 2006.

The CSC diagnoses around 200 patients a year, more than half of the new cases from the island. The numbers of patients according to the Ridley-Jopling clinical classification are shown in Table 1. These patients come from all parts of the island province, voluntarily or as referrals. A large proportion of data from CSC cases since 2003 are included in the current study.

TABLE 1. New leprosy cases diagnosed at CSC, LWM, 2000 to 2007

Type of case ^a	No. of cases diagnosed for yr							
	2000	2001	2002	2003	2004	2005	2006	2007
IND	7	8	6	10	8	4	4	6
TT	6	4	0	5	2	4	1	0
BT	38	33	29	27	22	11	25	23
BB	15	10	8	4	11	13	2	0
BL	70	70	65	68	84	61	49	50
LL	69	62	72	64	86	80	83	94
Total	205	187	180	178	213	173	164	173

^a IND, indeterminate; TT, tuberculoid; BT, borderline tuberculoid; BB, borderline; BL, borderline lepromatous; LL, lepromatous leprosy.

Cebu is a narrow island, 225 km long and 25 km across at the widest point (Fig. 1A). Metro Cebu comprises seven cities (Carcar City, Cebu City, Danao City, Lapu-Lapu City, Mandaue City, Naga City, and Talisay City) and six other municipalities on the east. Each city is administratively divided into barangays (akin to a county or village). Cebu City per se is located centrally within the Metro Cebu area, while Lapu-Lapu City is in Mactan Island and is connected to the main island by a bridge. The CSC is in Cebu City, while the leprosarium is located in Mandaue City. It is slightly over 10 km north of the CSC. Of the 228 patients in this study, one-third reside in Cebu City (Fig. 1B).

The majority of the patients were bacteriological index (BI) positive, as judged by microscopic examination of slit skin smear (SSS) samples or biopsy specimens and were prescribed the standard multibacillary WHO MDT treatment (1-year duration in Philippines). Considering the "L" and "M" groups ($n = 178$), the ages of the patients ranged from 5 to 71 years, with a mean of 30 years. The proportions of patients under 30 and 40 years of age were 66% and 80%, respectively. The ratio of male to female patients was 3:1.

MLVA-based genomic variation of *M. leprae* in patients in Cebu, Philippines. Skin biopsy specimens taken from leprosy patients at the time of diagnostic procedures or during treatment (as in the case of 38 of the 100 M group samples) were the source of *M. leprae* DNA in these studies. Multiple-locus VNTR analysis (MLVA) was performed for a total of 228 specimens using multiplex PCR and FLA procedures described previously (8) to determine the number of repeats (hereafter known as alleles). The samples from patients ($n = 21$) with a low BI (BI < 1.5) did not yield data for all loci and were excluded from the subsequent analysis.

The MLVA data for 15 loci from the 207 Cebu samples, NHDP63 reference DNA, and sequenced *M. leprae* strain TN (4) were analyzed for allelic diversity and population structure. The allelic diversity across the 15 loci ranged from 0.1 to 0.94, and the heterozygosity value for the entire data set is 0.5334 (21).

Comparison of *M. leprae* VNTR patterns in SSS samples and skin biopsy specimens. To compare the stabilities of VNTRs in different body compartments, VNTR data for biopsy specimens and SSS samples from 10 high-BI patients were also analyzed (Table 2) and were found to be mostly similar. We noted extra product peaks for the SSS for some of the

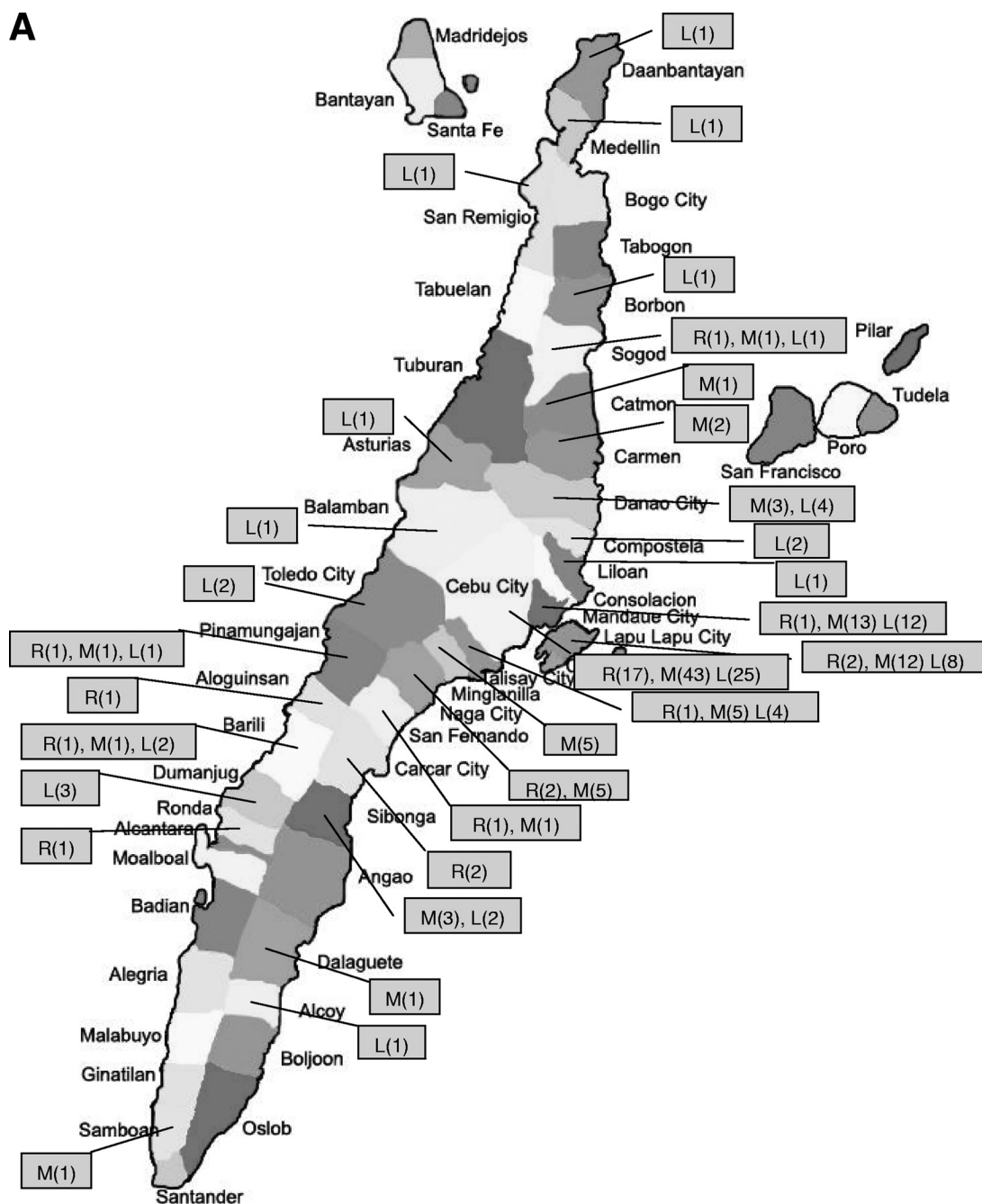


FIG. 1. (A) Map of Cebu Province, Philippines, with patients' cities of residence. One-third of the patients in this study population are from Cebu City. R, M, and L represent each of the sample collections, followed in parentheses by the number of patients in this study. (B) Distribution of patients from Cebu City according to the barangay (village) of residence. R, M, and L represent each of the sample collections, followed in parentheses by the number of patients in this study. (Maps adapted from <http://www.cebuph.gov.ph/> with permission.)

more variable stutter-prone microsatellites. Such patterns may represent variants of a dominant isolate. SSS is usually a mix of smears from the lesion site (biopsy specimen) and as well as from other body sites.

Rapid PCR-RFLP SNP subtyping of *M. leprae* in patients in Cebu, Philippines. By using the NEBcutter2 program, we were able to identify commercially available enzymes that are suitable for the detection and separation of the four SNP subtypes

according to the scheme shown in Fig. 2 and 3. The recognition sequences for SmlI, BstUI, and CviKI-1 are CTYRAG, RGCY, and CGCG, respectively.

At first, we verified the feasibility of using one, two, or three enzymes for subtyping of SNPs 1 to 4 for reference *M. leprae* strains as shown in Fig. 3. The restriction enzyme products could be resolved on a 3% agarose gel for SmlI and BstUI. A novel primer set was designed for the SNP 2 locus to eliminate

B

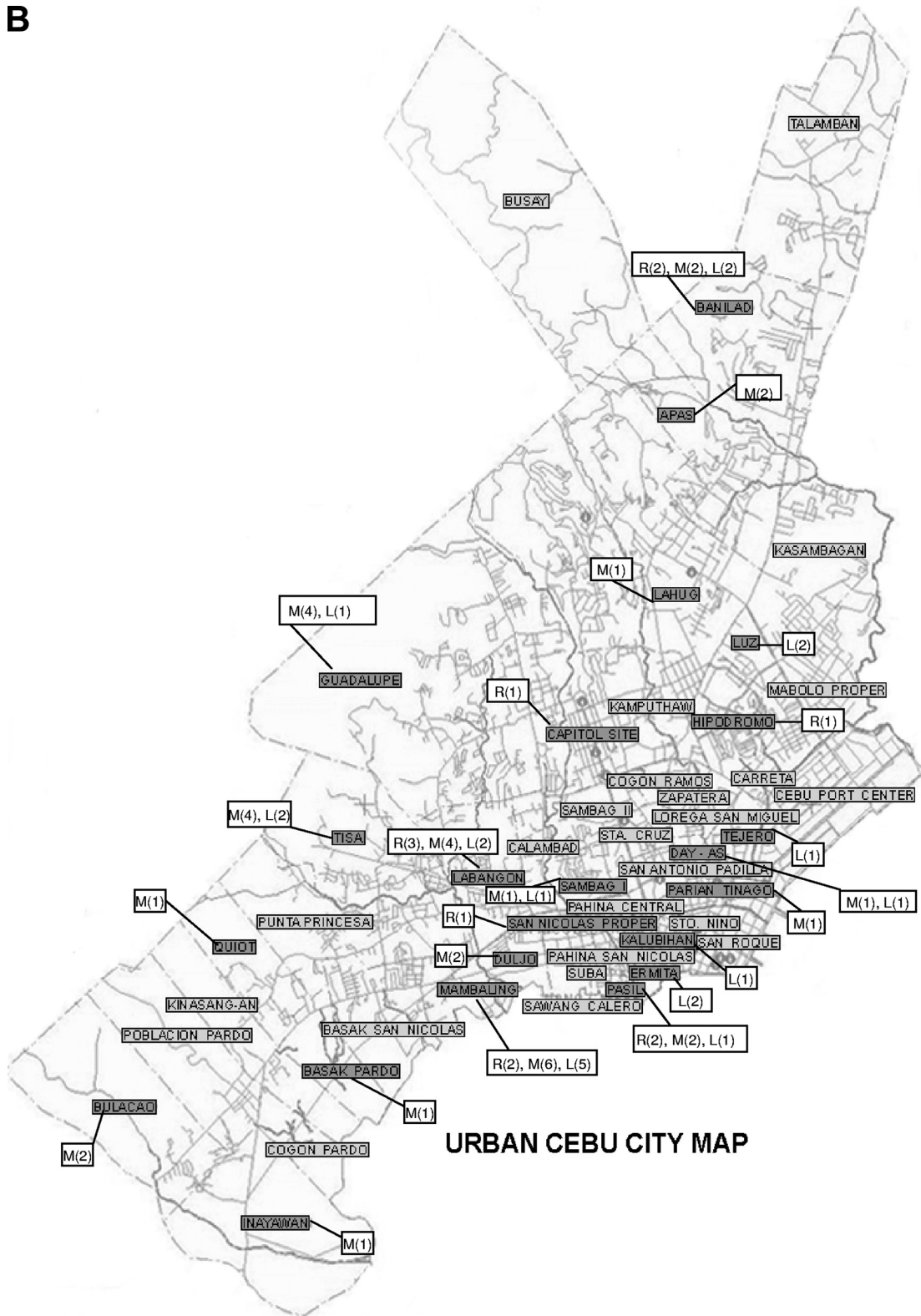


FIG. 1—Continued.

non-SNP CviKI-1 cutting sites. However, there are still two recognition sites: one from the SNP “T” allele would yield fragments of 61, 11, and 42 bp, while amplicons lacking the SNP “T” allele would produce 72- and 42-bp fragments. The

3% agarose gel system was not adequate to reliably differentiate the 72- and 61-bp CviKI-1 fragments; therefore, a 12% acrylamide gel system was used.

After establishing the PCR-RFLP method for separating the

TABLE 2. Comparison of *M. leprae* VNTR patterns between skin biopsy specimens and pooled SSS samples for 10 patients

Sample ^a	No. of repeats for locus ^b														
	(AC)8b	(GTA)9	(GGT)5	(AT)17	21-3	(AC)9	(AT)15	(AC)8a	27-5	6-7	(TA)18	(TTC)21	18-8	12-5	23-3
B01	7	9	4	12	2	8	18	8	5	7	13	18	ND	4	2
SS01	7	9	4	12	2	8	18	8	5	7	13	18	ND	4	2
B03	6	9	5	10	2	8	17	9	5	7	16	17	7	5	2
SS03	6	9	5	10	2	8	17	9	5	7	16	17	7	5	2
B07	8	9	5	16	3	9	14	8	5	7	18	30	8	5	2
SS07	8	9	5	16	3	9	14	8	5	7	18	30/26	8	5	2
B08	8	8	ND	16	3	9	15	9	5	9	17	20	8	4	2
SS08	8	8	ND	16	3	9	15	9	5	9	17	20	8	4	2
B09	8	9	5	15	3	9	16	9	5	7	19	26	ND	5	2
SS09	8	9	5	15	3	9	ND	9	5	7	19	26/27	ND	5	2
B10	6	9	5	10	2	8	20	10	5	7	17	13	8	5	2
SS10	6	9	5	10	2	8	ND	10	5	7	16	14	8	5	2
B11	8	9	5	15	3	9	14	9	5	7	18	28	4	5	2
SS11	8	9	5	15/16	3	9	14	9	5	7	18-20	27/24	4	5	2
B14	7	11	5	13	3	8	16	9	5	8	18	22	8	4	2
SS14	7	11	ND	13	3	8	16	9	5	8	18	22	8	4	2
B19	8	12	5	14	3	8	15	9	5	7	18	36	ND	4	2
SS19	8	12	5	14	3	8	15	9	5	7	18	36	ND	4	2
B20	8	11	5	13	3	9	16	9	5	7	17	25	ND	4	2
SS20	8	11	5	13	3	9	16	9	5	7	17	25	ND	4	2

^a B indicates biopsy samples, and SS indicates SSS samples.

^b ND, not detected.

isolates of SNP types 1 to 4, we applied it to all 100 samples from the “M” study. The first step of PCR amplification of SNP locus 3 revealed that 99% of the DNAs yielded products, and 18 samples could be digested with BstUI. One of the samples did not yield PCR products for all the three SNP loci.

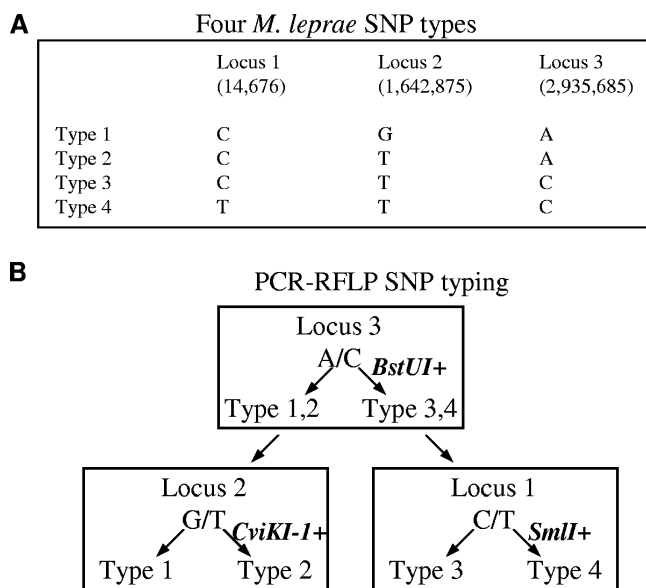


FIG. 2. Schematic of *M. leprae* SNP subtyping of clinical leprosy samples based on PCR-RFLP. (A) The four major SNP types. The numbers below the SNP loci refer to the nucleotide positions in the sequenced TN strain. (B) Scheme of *M. leprae* SNP subtyping.

DNA sequencing of the PCR amplicons of the remaining samples confirmed that all BstUI-sensitive amplicons carried SNP allele “C,” while a random selection of the BstUI-resistant samples carried the “A” allele. In the next step, we examined the susceptibility of SNP locus 1 to SmlI. None of the amplicons were digested, as was corroborated by the sequence of the products, and thus, SNP type 4 *M. leprae* strains are absent in this clinical sample set. Finally, to classify the strains carrying the SNP locus 3 BstUI-resistant phenotype into type 1 or type 2, CviKI-1 enzyme digestion of the locus 2 PCR products was performed. One sample of type 2 was confirmed by sequencing at SNP loci 2 and 3. Another isolate that was restriction positive was sequenced was confirmed to carry the “T” allele; however, a PCR product for locus 3 was not obtained, so we were unable to classify it as being type 2, 3, or 4. One of the samples gave mixed cut and uncut CviKI-1 RFLP patterns. Furthermore, both alleles G and T were observed in the DNA sequence. Interestingly, short tandem repeat (STR) typing of this DNA (M55) also revealed mixed signals at multiple loci even when independent aliquots of biopsy specimen sections of this sample were tested, indicating a mixed infection or contamination of the specimen. In summary, we were able to SNP subtype 97 of the 100 specimens.

VNTR alleles predictive of SNP type 3 for Philippine isolates. Correlating the SNP subtypes with the VNTR profiles of 15 loci revealed that the alleles at the (GGT)5 and 21-3 loci were 4 and 1 or 2, respectively, for the SNP type 3 isolates (Fig. 4A to C), while other combinations of alleles such as 4:3, 5:3, 5:1, and 5:2 were seen in isolates of SNP types 1 and 2.

To determine if this allelic signature is predictive of SNP

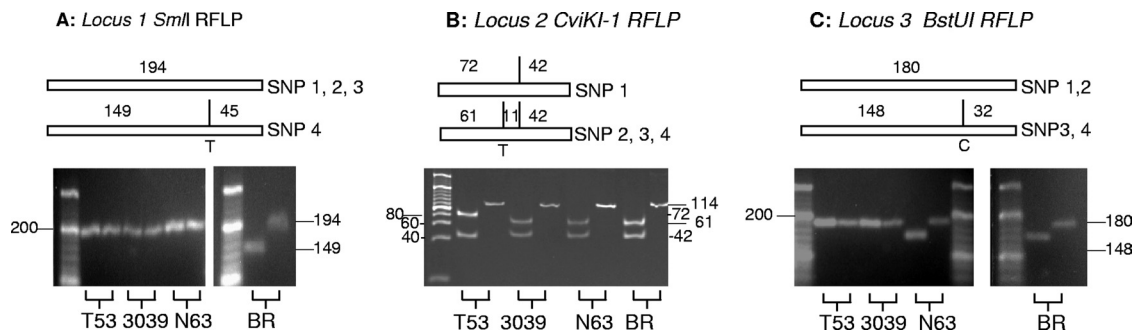


FIG. 3. PCR-RFLP patterns of four *M. leprae* reference strains (Thai-53 [T53], 3039/21 [3039], NHDP63 [N63], and BR4923 [BR]). The PCR products of SNP loci 1, 2, and 3 were subjected to enzyme digestion. The digested and undigested PCR products were resolved on agarose (A and C) or acrylamide (B) gels and run in pairs; the digested products for each of the strains are shown first. DNA sizing ladders are shown in the left lane in each gel. The numbers on the right side of each gel refer to the lengths (bp) of the PCR and digestion products.

type 3 in this population, we reviewed the 207 VNTR patterns and identified 14 other SNP 3 candidates (10 L and 4 R samples). Of these, we first tested eight L and four R samples. As controls, we randomly tested other “L” ($n = 6$) and “R” ($n = 1$) samples that did not have the VNTR pattern of interest. All eight L samples were of SNP type 3 (Fig. 5), while the “controls” were not. The “R” samples had degraded, or very little was remaining at the time when these new assays were being tested, and we could recover PCR amplicons for one sample (SNP type 3). Subsequently, we completed the PCR-RFLP SNP typing of the remaining 64 L samples and demonstrated that the SNP type 3 prediction rule was valid.

Population structure of *M. leprae* isolates from Cebu, Philippines. There are several computation models and software to infer clustering and/or phylogenetic relationships with the molecular data set (1, 9, 10, 11, 15, 17). We applied MP principles, assumed a stepwise mutation model, and executed the analysis using the PAUP software program (17). By arbitrarily defining the sequenced strain TN as the outgroup taxon, 50% consensus trees were generated. Since the allelic diversity of each locus was found to be different in this study population, we weighted

the loci according to the inverse of the individual allelic diversity index for MP. Neighbor joining and unweighted-pair group method using average linkages were also executed from the same data sets (data not shown) in PAUP.

Five major groups emerged in phylogenetic trees using MP (Fig. 6), neighbor joining, and the unweighted-pair group method using average linkage algorithms, which we have named groups A, B, C, D, and T. The T group, comprising 10 isolates, is seen proximal to the sequenced TN strain. This grouping may be attributed to the three copies of the 23-3 VNTR locus that are rare in the entire data set and other alleles shared with strain TN.

The majority of the remaining isolates fall within groups A, B, C, and D. Groups A and B are separated from groups C and D due to the (AC)8b locus. Groups A and B have 8 copies of the VNTR, whereas in groups C and D, the allele is 7 copies. Group A is then branched into groups A1 and A2; the (AC)9 allele is 8 or 9 copies, respectively. The 18-8 locus, with alleles of 8 or 4 copies, separates B1 from B2.

Group C, comprising 45 clinical isolates, is characterized by low allele numbers for (AC)8b, (TA)18, (GGT)5, (TTC)21,

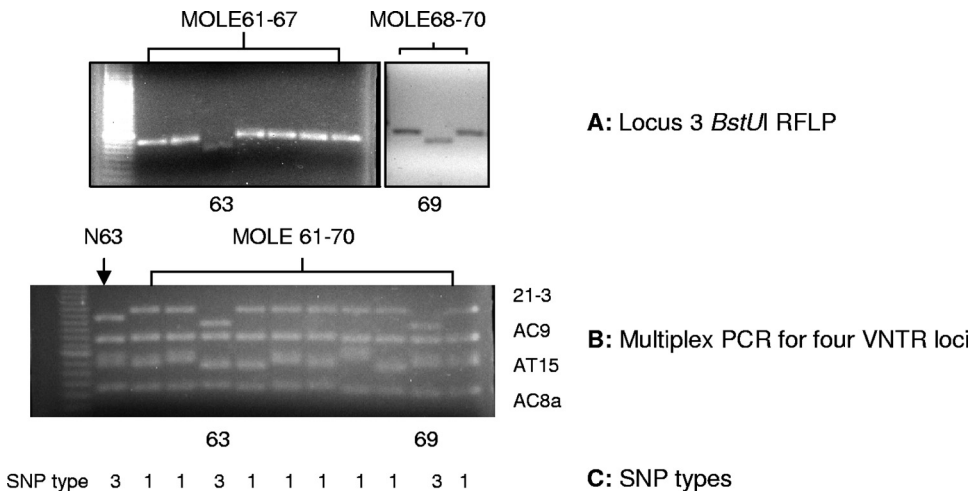


FIG. 4. The relationship between the 21-3 VNTR allele and the BstUI cutting pattern for 10 Philippine *M. leprae* samples is shown in A and B, respectively. (A) BstUI-RFLP gel. (B) Agarose gel showing products of multiplex PCR for four VNTR loci. The 21-3 product is the largest. VNTRs at all loci can be seen within this sample set. (C) SNP types. Samples 63 and 69 are indicated below A and B. N63 is NHDP63 (SNP type 3).

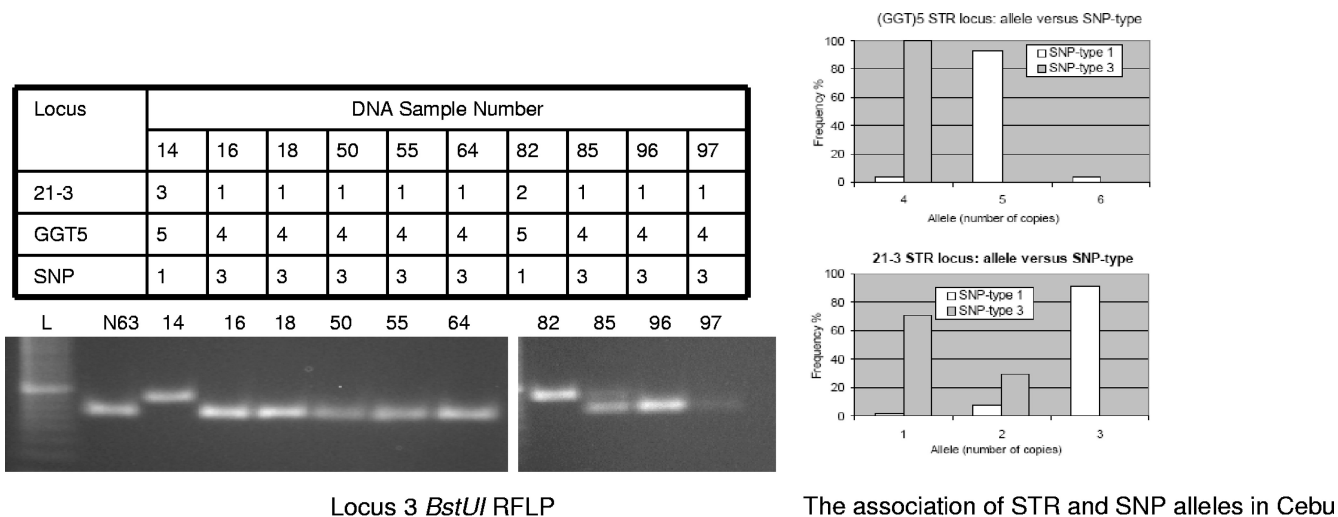


FIG. 5. The 21-3 and (GGT)5 alleles are indicative of SNP types 1 and 3 in Philippine *M. leprae* strains. The table indicates the VNTR alleles for 21-3 and (GGT)5 for 10 *M. leprae* specimens. The gel shows the corresponding SNP locus 3 *Bst*UI cutting patterns for these 10 specimens. N63 is NHDP63 (SNP type 3). The graphs show the allele frequency of these loci for SNP types 1 and 3 for 100 *M. leprae* samples.

21-3, and 18-8 VNTR loci. The dominant alleles for (AC)8b, (GGT)5, and 18-8 VNTR loci are 7, 4, and 7 copies, respectively. The range of alleles for (TA)18, (TTC)21, and 21-3 are 12 to 17, 14 to 20, and 1 to 2 copies, respectively, while in groups A, B, and D, they are 21 to 26, 21 to 32, and 3 copies, respectively (Fig. 6 and 7).

Group D is characterized by a distinctive six-copy allele of the (GGT)5 locus, and all isolates have 10 repeats at the (AC)8a locus.

The majority of the Cebu *M. leprae* isolates are of SNP type 1. Of the 207 samples for which VNTR patterns were mapped, 100 M, 78 L, and 5 R samples were tested for SNP types. From these samples, 10 L, 18 M, and 1 R samples were confirmed as being of SNP type 3. In the rest of the “R” samples for which DNA is no longer available for testing, three more carry the 4:1 alleles for (GGT)5:21-3. Therefore, we estimate that 16% (32/207) of the study population are of SNP type 3. Interestingly, in trees that were derived based solely on VNTR alleles, the SNP type 3 isolates converged into group C described above, separating from the remaining isolates at a deep or intermediate branch level. The allele profiles for the SNP type 3 versus SNP type 1 isolates are shown in Fig. 7. A few non-SNP type 3 (of SNP type 1; *n* = 13) isolates also group with the SNP type 3 isolates in group C, probably due to their low allele numbers, which is typically seen in SNP type 3 isolates (Fig. 7) but not in the majority of the SNP type 1 isolates.

Identification of plausible transmission clusters in Cebu by genotyping. Exposure to the *M. leprae* source can occur within and outside the household, such as in the community, place of employment, and place education, etc. The barangay of residence at the time of sampling was known for nearly 200 patients. However, daily, weekly, or monthly commute for work or other reasons is not uncommon for residents in the Metro Cebu region.

The patient population studied included several MCFs. MCFs presenting to the clinic within a period of 4 years of each other serve as a model for recent infections from one to the other or from another common source. The 15-locus VNTR

profiles of *M. leprae* from these patients are shown in Table 3. In general, the mismatched alleles in the MCFs occurred in the more polymorphic and stutter-prone VNTR loci. In the two pairs family 3 (F3) and F4, there were differences in many loci. Accordingly, four of the six MCF pairs appeared to be close in the consensus tree despite allelic differences at one or more loci (Fig. 6). Based on these trends, we tentatively define transmission clusters in Cebu as patients/isolates whose *M. leprae* VNTR alleles are shared at 12/15 loci (includes all minisatellites) and colocalize in the weighted consensus tree.

Besides examining the genotypes of *M. leprae* strains in MCFs, we sorted the VNTR profiles according to barangays. There is not much evidence of geographic segregation of strains according to VNTR or SNP profiles. However, several matched pairs emerged (Table 4), which are suitable for formal secondary epidemiological investigations of these cases and an active search for other new cases in these barangays, activities not within the scope of the current study.

DISCUSSION

Global standardized guidelines for the diagnosis and treatment of leprosy with a goal of elimination have been available and in practice for more than 2 decades (23). However, new patients are being detected at unchanged rates in some regions of endemicity, leaving fundamental questions about sources and routes of transmission and efficacy of leprosy control measures unanswered. Classical epidemiology and leprosy control programs have frequently identified household contacts of patients as being at a higher risk of infection than the general population (2, 3, 20); explanations or tools are not routinely in place for tracing the origins and relationships of these and remaining new cases (5, 18).

In susceptible individuals, leprosy manifests as single or multiple skin lesions often remaining unnoticed or not reported and undiagnosed for extended durations. Laboratory testing including SSS testing and biopsy specimen examinations is not

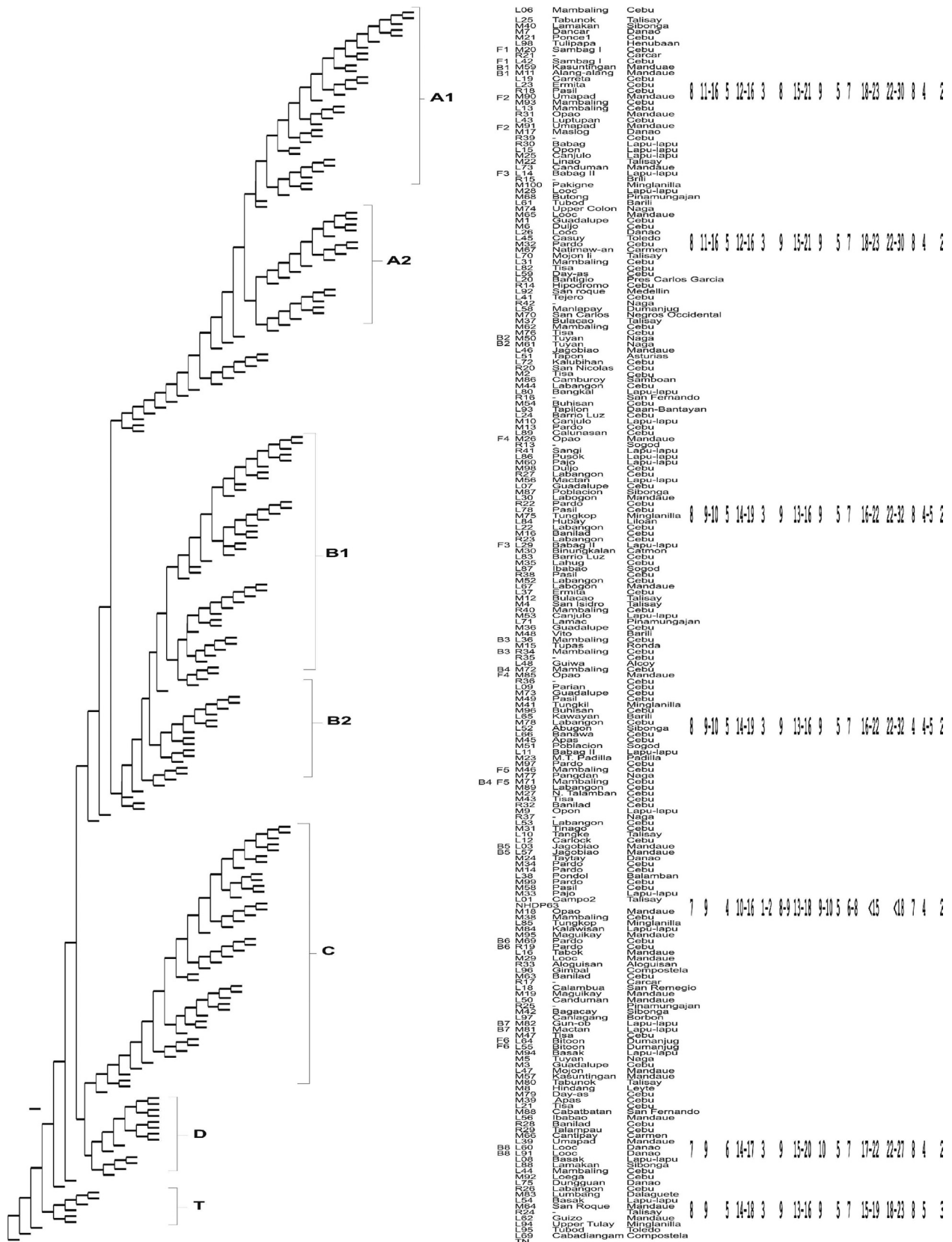


FIG. 6. *M. leprae* population structure based on VNTR markers and the dominant genotypes identified in Cebu, Philippines. On the left, a 50% consensus phylogenetic tree generated using the MP algorithm is shown, and the major branches are marked A, B, C, D, and T. The source of each *M. leprae* isolate is indicated in the three columns to the right: study (R, M, and L) sample, barangay, and city, respectively. The conserved allelic patterns within each branch are indicated at the extreme right. The locus order is (AC)8b, (GTA)9, (GGT)5, (AT)17, 21-3, (AC)9, (AT)15, (AC)8a, 27-5, 6-7, (TA)18, (TTC)21, 18-8, 12-5, and 23-3. Clustering of genotypes within MCFs and/or barangays is seen, represented by codes F1 to F6 (VNTR genotypes are shown in Table 3) and B1 to B8 (as listed in Table 4).

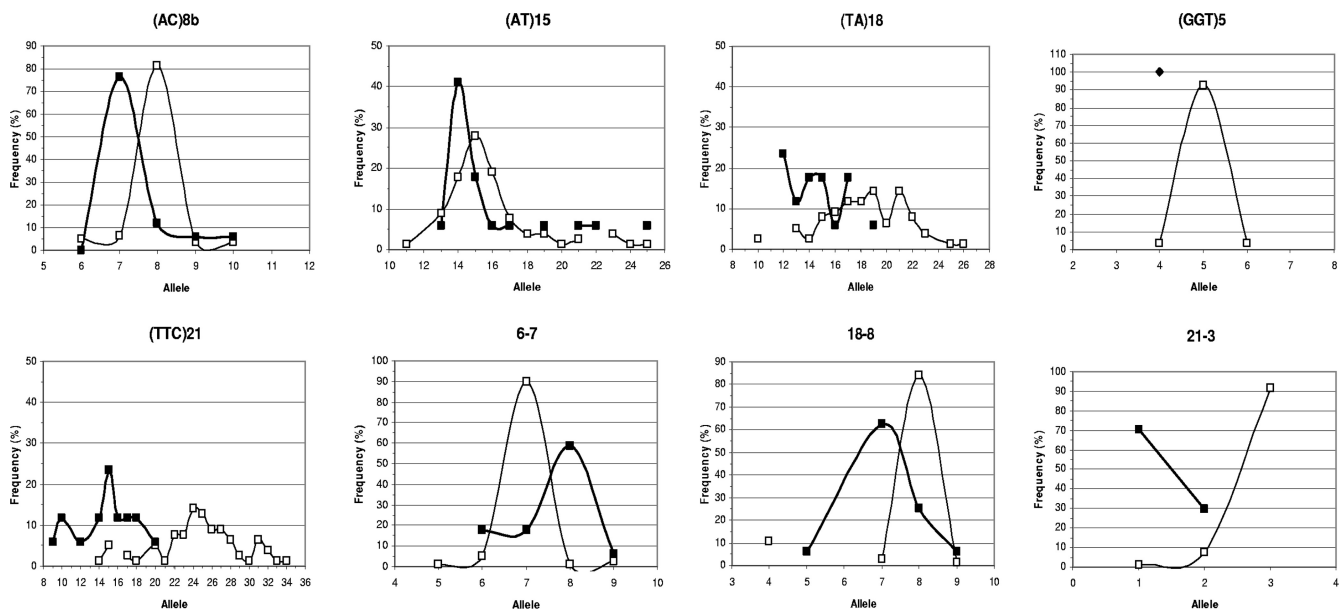


FIG. 7. Comparison of VNTR patterns of SNP type 3 compared to those of SNP type 1 isolates from Cebu, Philippines. The allelic frequencies (y axis) were plotted against the VNTR alleles (x axis) for the locus indicated on the top of each panel. Solid squares and hollow squares represent SNP type 3 and SNP type 1 isolates, respectively.

universally required for diagnosis. *M. leprae* is noncultivable in the laboratory even after detection. These and social factors impede tracing leprosy infections in communities. Nevertheless, molecular technologies of strain typing of the infectious agent directly from skin lesions and SSS, when available, are emerging and being explored to provide novel information and overcome these challenges in traditional leprosy control.

We present findings and interpretations for molecular strain typing of *M. leprae* based on 15 VNTR genomic markers by studying nearly 170 patients residing in Cebu, Philippines, from a recent 4-year period. Another 30 patients from 20 years ago were also included. Subtyping on the basis of three SNP markers was applied to approximately half the study population. In this regard, the Cebu study represents a comprehensive and descriptive molecular analysis encompassing a large cohort of patients.

Skin biopsy specimens from a single lesion per patient were collected and used as the source of the *M. leprae* DNA

for genotyping. For the Cebu patient group, the panel of 15 VNTR markers exhibited slow, medium, and rapid rates of evolution. Therefore, all loci were informative, revealing a heterogeneous, diverse *M. leprae* population. In contrast, we observed that in Qiubei, China, the *M. leprae* isolates appeared to have descendants of a dominant genotype, with only a few of the 15 VNTR loci being polymorphic. Thus, not all *M. leprae* VNTR loci are highly variable; markers suitable for genotyping can be exploited, but these markers need to be determined empirically for each population or region of endemicity.

Finding conserved VNTR profiles in MCF pairs is consistent with the potential for MLVA in tracing sources of infection or identifying a transmission network. While MLVA was able to demonstrate that genotypes were shared for patients with known epidemiological links, the estimated phylogenetic trees contain clusters with no known links together with those with disparate residential histories. Multiple factors including the

TABLE 3. VNTR profiles of multicase families in the 207 Cebu patient database

Family-sample ID	No. of repeats for STR locus													Barangay	City		
	(AC)8b	(GTA)9	(GGT)5	(AT)17	21-3	(AC)9	(AT)15	(AC)8a	27-5	6-7	(TA)18	(TTC)21	18-8			12-5	23-3
F1-L42	8	11	5	13	3	8	19	9	5	7	22	29	8	4	2	Sambag 1	Cebu
F1-M20	8	11	5	13	3	7	20	9	5	7	21	29	8	4	2	Sambag 1	Cebu
F2-M90	8	12	5	13	3	8	16	9	5	7	21	34	8	3	2	Umapad	Mandaue
F2-M91	8	12	5	14	3	8	16	9	5	7	22	27	8	3	2	Umapad	Mandaue
F3-L14	7	11	5	13	3	8	16	9	5	8	16	22	8	4	2	Babag II	Lapu-Lapu
F3-L29	9	9	5	16	3	9	14	9	5	7	16	29	8	5	2	Babag II	Lapu-Lapu
F4-M26	8	9	5	14	3	9	14	9	5	7	25	32	8	4	2	Opao	Mandaue
F4-M85	8	9	5	14	3	9	17	9	5	7	21	25	8	5	2	Lanao	Daanbantayan
F5-M46	8	10	5	18	3	9	15	9	5	7	20	26	8	5	2	Alaska	Mambaling
F5-M71	8	9	5	18	3	9	15	9	5	7	19	25	8	5	2	Alaska	Mambaling
F6-L55	6	9	4	18	1	9	15	9	5	8	15	16	7	4	2	Bitoon	Dumanjug
F6-L64	6	9	4	19	1	9	15	9	5	8	15	16	7	4	2	Bitoon	Dumanjug

TABLE 4. Identification of pairs of *M. leprae* isolates with matching VNTR patterns in various barangays in Cebu

Residence barangay	Barangay-sample ID	No. of repeats of STR locus															
		AC8b	GTA9	GGT5	AT17	21-3	AC9	AT15	AC8a	27-5	6-7	TA18	TTC21	18-8	12-5	23-3	
Alang-alang	B1-M11	8	13	5	13	3	8	17	9	5	7	19	29	8	4	2	
	B1-M59	8	13	5	13	3	8	17	9	5	7	21	31	8	4	2	
Tuyan	B2-M50	8	11	5	13	3	9	16	9	5	7	21	22	8	4	2	
	B2-M61	8	11	5	13	3	9	16	9	5	7	22	22	8	4	2	
Mambaling	B3-R34	8	9	5	15	3	9	15	9	5	7	21	27	8	5	2	
	B3-L36	8	9	5	14	3	9	15	9	5	7	22	27	8	5	2	
Mambaling	B4-M71	8	9	5	18	3	9	15	9	5	7	19	25	8	5	2	
	B4-M72	8	9	5	16	3	9	16	9	5	7	22	25	8	5	2	
Jagobiao	B5-L03	6	9	5	10	2	8	17	9	5	7	16	17	7	5	2	
	B5-L57	6	9	5	10	2	8	18	9	5	7	17	17	7	5	2	
Pardo	B6-R19	7	8	4	14	1	9	17	9	5	8	14	16	7	4	2	
	B6-M69	7	8	4	14	1	9	19	9	5	8	14	16	7	4	2	
Mactan	B7-M81	7	9	4	14	1	9	14	9	5	8	17	18	7	4	2	
	B7-M82	7	9	4	14	1	9	14	9	5	8	19	16	7	4	2	
Looc	B8-L60	7	9	6	17	3	9	17	10	5	7	19	22	8	4	2	
	B8-L91	7	9	6	16	3	9	17	10	5	7	18	20	8	4	2	

long incubation period of leprosy, missing data due to partial patient coverage (capture) in this study (Table 1), the lack of sensitivity of MLVA in paucibacillary biopsy specimens, and patient mobility are probable causes. The search for new genomic markers should continue and be evaluated as they are discovered.

Monot et al. previously described a classification of global isolates into four types based on the discovery of three SNPs (14). SNP type 1 *M. leprae* was shown to be prevalent in Asia (Nepal, India, South Korea, Thailand, and Philippines), the Pacific islands of New Caledonia, and East Africa, while type 3 was found in the European, North African, and American continents and also in New Caledonia. In our study, we also found that the major circulating isolate in Cebu is of SNP type 1; 16% belong to SNP type 3. Cebu was colonized by the Spaniards in the 16th century, and therefore, the finding of SNP type 3 isolates could be attributed to this history. However, SNP type 3 isolates are also prevalent in Japan, South Korea, and China (14, 22). Therefore, the exact origin of the Cebu strains is not definite. Interestingly, VNTR markers could separate the SNP type 3 isolates from the majority of the SNP type 1 isolates. Distinctive allelic patterns were found for SNP type 3 isolates, typically contractions rather than expansions at several microsatellite and minisatellite loci compared to SNP type 1 isolates. In addition, a two-locus signature that is highly predictive of SNP type 3 within the Cebu population was detected. These data indicate that there is sufficient stability in the VNTR markers to retain phylogenetic relationships during transmission. The detection of the polymorphism for the 21-3 minisatellite requires simply a 3% agarose gel system, with the detection of one- or two-copy alleles being a good indicator of the SNP type 3 in the studied population. The two SNP subtypes were not found to be localized to any

part of the island, suggesting that there has been considerable circulation and dispersion of leprosy in this community.

Inpatient variability of VNTRs has been raised as a concern in two previously reported studies (13, 24). Those studies examined VNTR profiles of more than one clinical sample per patient taken at the same and/or different time points. In a study of 42 patients in Mali, Africa, 36 patients showed no allelic variation in all five loci, four of which were microsatellites (13). For six patients, one- or two-copy variations were noted at one or all loci. It was suggested that a larger panel of markers exhibiting a broader range of allelic diversities would be necessary to assess if there were any genotypic relationships within this patient population. SNP typing was unable to further discriminate the isolates, as they were all of type 4.

There are several indications from our findings in Cebu that VNTRs can be used as markers for the study of both strain evolution and leprosy transmission: the distinctive clustering of SNP type 3 isolates by VNTRs, the contrasting allelic properties between the Cebu and Qiubei leprosy patients, and the closely matching profiles in MCFs despite the overall diversity of the population and the analysis of one-time-point single biological samples.

Regarding the possibility of the detection of mixed infections, we found one biopsy sample (M55) from which two distinct alleles were detected at several VNTR loci (11/15) in electropherograms by the FLA technique, which would have been missed with conventional DNA sequencing. Separate aliquots of frozen sections gave the same results.

The absolute definition of strain identity and the a priori choice of markers are not yet clear because these will be dependent on strain diversity in the target population and have yet to be experimentally described. The mechanisms by which *M. leprae* isolates are established during infection (and rein-

fection) and evolve are not known and could be influenced by the endemicity (incidence rate), population size, host susceptibility, leprosy control measures, and other unknowns such as environmental reservoirs. Furthermore, the methodology and analysis can vary on whether the application is for discerning localized transmission, for delving into broader evolutionary relationships, or for both as we have discerned in this study using 15 markers for Cebu cases. The baseline molecular epidemiology database and approaches from this work will aid in designing targeted studies to monitor selected genotypes/clusters, perform active case finding, and trace and thus uncover transmission arising in settings beyond the household such as school, employment, and socialization or from environmental sources. The collection and banking of samples such as skin biopsy specimens and SSS and a tiered approach to strain typing such as that described here could be adopted in other reference clinics.

ACKNOWLEDGMENTS

These studies were funded by NIH-NIAID grants AI-59644 and AI-063457 and contract NO1-AI-25469.

We thank staff at CSC, LWM, for clinical work; patients who volunteered for the research studies; and Macromolecular Resources, CSU, for DNA sequencing and typing services. We thank Michelle Price and Niccole Carner for the graphics.

REFERENCES

- Bohannon, B. J., and J. Hughes. 2003. New approaches to analyzing microbial biodiversity data. *Curr. Opin. Microbiol.* **6**:282–287.
- Cardona-Castro, N., J. Beltrán-Alzate, and R. Manrique-Hernández. 2008. Survey to identify *Mycobacterium leprae*-infected household contacts of patients from prevalent regions of leprosy in Colombia. *Mem. Inst. Oswaldo Cruz* **103**:332–336.
- Cellona, R. V., G. P. Walsh, T. T. Fajardo, Jr., R. M. Abalos, E. C. dela Cruz, L. Guido-Villahermosa, M. V. Felicio-Balagon, G. J. Steenbergen, and J. T. Douglas. 1993. Cross-sectional assessment of ELISA reactivity in leprosy patients, contacts, and normal population using the semisynthetic antigen natural disaccharide octyl bovine serum albumin (ND-O-BSA) in Cebu, the Philippines. *Int. J. Lepr. Other Mycobact. Dis.* **61**:192–198.
- Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honoré, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell. 2001. Massive gene decay in the leprosy bacillus. *Nature* **409**:1007–1011.
- Deps, P. D., B. V. Guedes, J. B. Filho, M. K. Andreatta, R. S. Marcari, and L. C. Rodrigues. 2006. Delay in the diagnosis of leprosy in the metropolitan region of Vitória, Brazil. *Lepr. Rev.* **77**:41–47.
- Felsenstein, J. 1989. PHYLIP—phylogeny inference package (version 3.2). *Cladistics* **5**:164–166.
- Grothouse, N. A., B. Rivoire, H. Kim, H. Lee, S.-N. Cho, P. J. Brennan, and V. D. Vissa. 2004. Multiple polymorphic loci for molecular typing of strains of *Mycobacterium leprae*. *J. Clin. Microbiol.* **42**:1666–1672.
- Kimura, M., R. M. Sakamuri, N. A. Grothouse, B. L. Rivoire, D. Gingrich, S. Krueger-Koplin, S.-N. Cho, P. J. Brennan, and V. Vissa. 2009. Rapid variable-number tandem-repeat genotyping for *Mycobacterium leprae* clinical specimens. *J. Clin. Microbiol.* **47**:1757–1766.
- Kolaczowski, B., and J. W. Thornton. 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature* **431**:980–984.
- Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* **5**:150–163.
- Martin, A. P. 2002. Phylogenetic approaches for describing and comparing the diversity of microbial communities. *Appl. Environ. Microbiol.* **68**:3673–3682.
- Matsuoka, M., S. Maeda, M. Kai, N. Nakata, G.-T. Chae, T. P. Gillis, K. Kobayashi, S. Izumi, and Y. Kashiwabara. 2000. *Mycobacterium leprae* typing by genetic diversity and global distribution of genotypes. *Int. J. Lepr.* **68**:121–128.
- Monot, M., N. Honoré, N. Balière, B. Ji, S. Sow, P. J. Brennan, and S. T. Cole. 2008. Are variable-number tandem repeats appropriate for genotyping *Mycobacterium leprae*? *J. Clin. Microbiol.* **46**:2291–2297.
- Monot, M., N. Honoré, N. Garnier, R. Araoz, J. Y. Coppée, C. Lacroix, S. Sow, J. S. Spencer, R. W. Truman, D. L. Williams, R. Gelber, M. Virmond, B. Flageul, S. N. Cho, B. Ji, A. Paniz-Mondolfi, J. Convit, S. Young, P. E. Fine, V. Rasolofy, P. J. Brennan, and S. T. Cole. 2005. On the origin of leprosy. *Science* **308**:1040–1042.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
- Shin, Y.-C., H. Lee, H. Lee, G. P. Walsh, J.-D. Kim, and S.-N. Cho. 2000. Variable numbers of TTC repeats in *Mycobacterium leprae* DNA from leprosy patients and use in strain differentiation. *J. Clin. Microbiol.* **38**:4535–4538.
- Swofford, D. L. 2003. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, MA.
- Tadesse Argaw, A., E. J. Shannon, A. Assefa, F. S. Mikru, B. K. Mariam, and J. B. Malone. 2006. A geospatial risk assessment model for leprosy in Ethiopia based on environmental thermal-hydrological regime analysis. *Geospat. Health* **1**:105–113.
- Truman, R., A. B. Fontes, A. B. De Miranda, P. Suffys, and T. Gillis. 2004. Genotypic variation and stability of four variable-number tandem repeats and their suitability for discriminating strains of *Mycobacterium leprae*. *J. Clin. Microbiol.* **42**:2558–2565.
- van Beers, S. M., M. Hatta, and P. R. Klatser. 1999. Patient contact is the major determinant in incident leprosy: implications for future control. *Int. J. Lepr. Other Mycobact. Dis.* **67**:119–128.
- Weir, B. S. 1990. Genetic data analysis: methods for discrete population genetic data analysis. Sinauer Associates, Inc., Sunderland, MA.
- Weng, X., Z. Wang, J. Liu, M. Kimura, W. C. Black IV, P. J. Brennan, H. Li, and V. D. Vissa. 2007. Identification and distribution of *Mycobacterium leprae* genotypes in a region of high leprosy prevalence in China: a 3-year molecular epidemiological study. *J. Clin. Microbiol.* **45**:1728–1734.
- World Health Organization. 2008. Global leprosy situation, beginning of 2008. *Wkly. Epidemiol. Rec.* **83**:293–300.
- Young, S. K., J. M. Ponnighaus, S. Jain, S. Lucas, S. Suneetha, D. N. Lockwood, D. B. Young, and P. E. Fine. 2008. Use of short tandem repeat sequences to study *Mycobacterium leprae* in leprosy patients in Malawi and India. *PLoS Negl. Trop. Dis.* **2**:e214.
- Young, S. K., G. M. Taylor, S. Jain, L. M. Suneetha, S. Suneetha, D. N. Lockwood, and D. B. Young. 2004. Microsatellite mapping of *Mycobacterium leprae* populations in infected humans. *J. Clin. Microbiol.* **42**:4931–4936.
- Zhang, L., T. Budiawan, and M. Matsuoka. 2005. Diversity of potential short tandem repeats in *Mycobacterium leprae* and application for molecular typing. *J. Clin. Microbiol.* **43**:5221–5229.