Probable Zoonotic Leprosy in the Southern United States

Richard W. Truman, Ph.D., Pushpendra Singh, Ph.D., Rahul Sharma, Ph.D., Philippe Busso, Jacques Rougemont, Ph.D., Alberto Paniz-Mondolfi, M.D., Adamandia Kapopoulou, M.S., Sylvain Brisse, Ph.D., David M. Scollard, M.D., Ph.D., Thomas P. Gillis, Ph.D., and Stewart T. Cole, Ph.D.

ABSTRACT

BACKGROUND
In the southern region of the United States, such as in Louisiana and Texas, there are autochthonous cases of leprosy among native-born Americans with no history of foreign exposure. In the same region, as well as in Mexico, wild armadillos are infected with Mycobacterium leprae.

METHODS
Whole-genome resequencing of M. leprae from one wild armadillo and three U.S. patients with leprosy revealed that the infective strains were essentially identical. Comparative genomic analysis of these strains and M. leprae strains from Asia and Brazil identified 51 single-nucleotide polymorphisms and an 11-bp insertion–deletion. We genotyped these polymorphic sites, in combination with 10 variable-number tandem repeats, in M. leprae strains obtained from 33 wild armadillos from five southern states, 50 U.S. outpatients seen at a clinic in Louisiana, and 64 Venezuelan patients, as well as in four foreign reference strains.

RESULTS
The M. leprae genotype of patients with foreign exposure generally reflected their country of origin or travel history. However, a unique M. leprae genotype (3I-2-v1) was found in 28 of the 33 wild armadillos and 25 of the 39 U.S. patients who resided in areas where exposure to armadillo-borne M. leprae was possible. This genotype has not been reported elsewhere in the world.

CONCLUSIONS
Wild armadillos and many patients with leprosy in the southern United States are infected with the same strain of M. leprae. Armadillos are a large natural reservoir for M. leprae, and leprosy may be a zoonosis in the region. (Funded by the National Institute of Allergy and Infectious Diseases and others.)
Leprosy (Hansen’s Disease) is a chronic infectious disease caused by Mycobacterium leprae. Though often considered a disease of antiquity, it is found most commonly today in tropical and semitropical regions, and a total of 249,007 new cases were reported globally in 2008. Genomic polymorphisms have allowed us to trace the historical spread of leprosy around the world, as human populations migrated. The disease was not present in the New World before Columbus discovered the American continents but rather appears to have been introduced here from Europe and Africa during colonization. Early case reports suggest that leprosy was already well established among settlers in the vicinity of New Orleans by the 1750s, and autochthonous transmission of the infection continues in the region today.

Leprosy is rare in the United States, with only about 150 new cases reported each year. The majority of these affected people lived or worked abroad in leprosy-endemic areas and may have acquired their disease there. However, about a third of all patients in the United States report no foreign residence and appear to have acquired their disease from local sources — although most are unable to recall any known contact with a person who had leprosy. These cases arise most frequently in Texas and Louisiana, but the range of endemic involvement appears to be expanding to other states.

M. leprae is an obligate intracellular pathogen that cannot be cultivated on artificial laboratory mediums. The only animal in which leprosy is reliably recapitulated is the nine-banded armadillo (Dasypus novemcinctus). Its unique susceptibility to experimental infection with M. leprae was first demonstrated in the 1970s, and armadillos have been the primary animal model for leprosy since. However, M. leprae infection also occurs naturally among some free-ranging armadillos.

The origins of M. leprae infection among armadillos, the geographic range of the infected animals, and the potential risks infected armadillos present to people have been topics of concern. The infection originated among armadillos decades before they were ever used in leprosy research, and numerous surveys have confirmed that armadillos in the southern United States are a large natural reservoir for M. leprae; its prevalence exceeds 20% in some locales. Infected armadillos have been reported in Alabama, Arkansas, Louisiana, Mississippi, Texas (Fig. 1), and Mexico. Several case reports have suggested that armadillos may be a source of M. leprae for some U.S. patients and contact with armadillos has been shown to be a significant risk factor for leprosy in three U.S. case–control studies.

Here, we describe our use of whole-genome sequencing, single-nucleotide polymorphism (SNP) typing, and variable-number tandem-repeat (VNTR) analysis to compare M. leprae obtained from wild armadillos and patients in the United States with leprosy and to understand better the role of armadillos in perpetuating leprosy in this country.

**METHODS**

**STUDY DESIGN**

We performed an ecologic cohort study to determine the type and frequency of M. leprae strains in U.S. patients, assess the degree to which they are genetically similar to strains from wild armadillos, and compare the geographic locations of the patients and armadillos. Statistical analysis was performed with the use of SAS software (version 9.2). The studies of the patients did not require board approval (on the basis of exemption category 4) and the requirement of written informed consent was waived by the institutional review board of Louisiana State University (Baton Rouge).

**M. leprae Strains and DNA**

Details of the M. leprae strains obtained from 50 patients with leprosy and 33 wild armadillos in the United States are given in Table S1 in the Supplementary Appendix (available with the full text of this article at NEJM.org). DNA from four genome-sequenced reference strains described previously and two foreign reference standards (LWM26 Philippines and 43926 Brazil) were also included.

**Armadillo DNA**

Armadillos were captured from the wild in five southern U.S. states (Fig. 1, and Table S1 in the Supplementary Appendix). DNA was isolated from 1-g specimens of liver, spleen, or lymph-node tissue by means of homogenization and was extracted with the use of a DNeasy kit (Qiagen).
Skin biopsy in Patients

Skin-biopsy specimens are regularly obtained from patients who attend the National Hansen’s Disease Program outpatient clinic (Baton Rouge, LA) or are referred there for diagnosis. The specimens are stored in optimum-cutting-temperature compound (VWR) and frozen. We used frozen specimens from unrelated patients with multibacillary leprosy who presented between 1993 and 2007. Each sample was thawed, washed with phosphate-buffered saline, and processed by means of a DNeasy kit. Skin-biopsy specimens from 64 Venezuelan patients with leprosy living in 12 provinces were collected between 2007 and 2009, preserved in 70% ethanol, and processed for use in this study as described previously.4

Genome Resequencing and Analysis

Genome sequences of the human reference strains of *M. leprae* NHDP-98 (from a Mexican-born patient residing in Texas) and NHDP-55 (from another patient residing in Texas) (see Table S1 in the Supplementary Appendix) and the wild-armadillo-derived strain I-30 were obtained from DNA-fragment libraries, sequenced (36-Cycle Sequencing Kit v.1), and analyzed ( Genome Analyzer II [Illumina]), as described previously.22 Sequence reads were mapped onto the consensus genome sequence of the *M. leprae* reference strain TN with the use of mapping quality scores.23 All sequence variations were confirmed by another means (BigDye Terminator v.3.1 cycle sequencing on an ABI3130XL DNA sequencer [Applied Biosystems]).

Sequence and Phylogenetic Analyses

Details of the primers used for genotyping are given in Table S2 in the Supplementary Appendix, and the genotyping scheme is shown in Figure 2. Polymerase-chain-reaction (PCR) assays were performed as described previously,22 and amplicons
were treated with exonuclease I and shrimp alkaline phosphatase (USB) before sequencing (BigDye Terminator). Sequence data were compared as described previously,\(^24,25\) and minimum-spanning-tree analysis based on SNPs and VNTRs was performed (BioNumerics software, version 6.1 [Applied Maths]).

**VNTR Typing**

The primers used for VNTR typing at 10 loci\(^26\) are given in Table S3 in the Supplementary Appendix. Nei’s index of diversity\(^27\) (\(D = 1 − \sum [\text{allele frequency}]^2\)) based on 475 *M. leprae* strains\(^28\) was used to eliminate hypervariable loci (i.e., those with \(D > 0.85\)) from consideration. PCR amplicons were sequenced (BigDye Terminator) to determine copy number.

**RESULTS**

**GENOME SEQUENCING AND SNP IDENTIFICATION**

Using Illumina technology, we obtained deep coverage of the genome sequences from *M. leprae* reference strains NHDP-55, NHDP-98, and the wild-armadillo–derived strain I-30 (Table 1, and Table S4 in the Supplementary Appendix). The resultant sequence reads were mapped onto the genome sequence of the TN reference strain (from India, SNP type 1A)\(^29\) and compared with the other sequenced *M. leprae* reference strains: Br4923 (from Brazil, SNP type 4P), NHDP-63 (from the United States, SNP type 3I), and Tha153 (from Thailand, SNP type 1A).\(^22\) This analysis confirmed the exceptionally high level of sequence conservation (99.995% identity), even among *M. leprae* strains.
of widely different geographic origins, and identified all four U.S.-derived genomes as SNP type 3I.

On detailed comparison of these seven genome sequences, 52 markers were found only in the SNP type 3I strains. These 3I strains differed among themselves at 21 positions (Table 1, and Table S4 in the Supplementary Appendix). One 11-bp indel (indel_17915) was particularly important, since the 3I strains have only one copy of the sequence (TTGGTGGTGTA, in pseudogene ML0014), whereas all other M. leprae strains have two copies.

**Table 1. Single-Nucleotide Polymorphisms (SNPs) and Indels in 3I-Type Mycobacterium leprae Genomes Found in the United States.**

<table>
<thead>
<tr>
<th>M. leprae Strain</th>
<th>Average Coverage†</th>
<th>Markers Specific to 3I Type</th>
<th>Markers Differing within 3I Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SNP bp</td>
<td>11 bp</td>
<td>SNP bp</td>
</tr>
<tr>
<td>NHDP-63</td>
<td>46</td>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>NHDP-55</td>
<td>57</td>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>NHDP-98</td>
<td>78</td>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>I-30</td>
<td>22</td>
<td>49</td>
<td>2</td>
</tr>
</tbody>
</table>

* The SNPs were identified by means of comparative genomic methods based on M. leprae reference standards (strains TN, Br4923, and Thai53). NHDP denotes National Hansen’s Disease Program.
† Average coverage is defined as the average number of consensus sequence reads obtained from the strain; there were no gaps except for the dispersed repeats that could not be distinguished owing to the short read length from Illumina sequencing.

Among patients with possible exposure by means of foreign residence only, the SNP type was typical for strains previously reported from the foreign location. SNP type 3I, generally associated with European–American populations, was most abundant in our samples, found in those from all 33 armadillos and 26 of the 29 patients with no history of foreign residence.

To improve the resolution of our data for M. leprae 3I strains, we also surveyed for 30 of the 52 newly discovered markers. In addition to the 11-bp indel, 24 of 30 markers were restricted to SNPs of type 3I, irrespective of the source of the strain. However, in four 3I strains identified in patients, five SNPs contained ancestral bases and may represent intermediate sequences arising during the evolutionary divergence of 3I strains from their common ancestor. The strains with ancestral bases were classified as having SNPs of the subtype 3I-1 to differentiate them from the more divergent strains classified as 3I-2 strains found in all armadillos and most indigenous U.S. patients (Fig. 3, and Table S5 in the Supplementary Appendix).

To gain more insight into the distribution of 3I-1 strains in the Americas, we examined biopsy specimens from 64 Venezuelan patients infected with M. leprae strains containing SNP type 3I. Of the 64 specimens, 48 (75%) belong to the strain subtype 3I-1 and 16 (25%) belong to strain subtype 3I-2. In addition, we found 3I-1 subtypes in patients from Brazil, Puerto Rico, and the Dominican Republic. Therefore, the prevalence of 3I-1 and 3I-2 strains in North America is significantly different (P<0.001) from their prevalence in the Caribbean and South America (see Table S6 in the Supplementary Appendix).

**VNTR and Minimum-Spanning-Tree Analyses**

Owing to the remarkable conservation of the M. leprae genome, SNP analysis is of limited power. Accordingly, we used 10 polymorphic VNTR markers to enhance discrimination (see Table S4 in the Supplementary Appendix) of strain subtypes. Minimum-spanning-tree analysis of the combined SNP and VNTR profiles was performed to examine relationships among strains (Fig. 3). The resulting SNP–VNTR genotypes (see Table S1 in the Supplementary Appendix) confirm a high degree of homogeneity between M. leprae from ar-
madillos and most indigenous U.S. cases of leprosy (Fig. 3). Among wild armadillos, 28 of the 33 strains showed complete genetic identity with respect to the SNP–VNTR genotype (3I-2-v1); the remaining 5 strains comprised two genotypes, 3I-2-v14 (2 specimens) and 3I-2-v13 (3 specimens) that varied at one VNTR locus (see Table S5 in the Supplementary Appendix). Similarly, 25 of the 39 patients with a history of residence in areas in which exposure to M. leprae from armadillos was possible — including 22 of the 29 patients with no history of foreign residence — also carried the 3I-2-v1 strain. The 3I-2-v1 genotype appears to be unique and highly distinctive. This combination of VNTR alleles was not found within a database of VNTR genotypes identified around the world, and allele frequencies in that database suggest a probability of random reassortment of the VNTR v1 genotype of only 1 in 10,000.

In our study, 3I-2-v1 was the only genotype found in more than two patients. The combination of SNP and VNTR genotyping is highly discriminatory and confirms a significant association between the M. leprae strain infecting armadillos and many U.S. patients. The 3I-2-v1 strain was significantly associated with a history of residence in areas where M. leprae–infected armadillos have been found (P<0.001). People with leprosy who live in areas with infected armadillos and who have no history of foreign residence have a significantly increased risk of presenting with infection with the 3I-2-v1 strain, as compared with any other strain (odds ratio, 16.5; 95% confidence interval [CI], 4.2 to 64.7; P<0.001) (see Table S6 in the Supplementary Appendix).

A history regarding contact with armadillos was available for 15 patients. Although 7 recalled no contact, 8 recalled having contact, including 1 who reported frequently hunting, cooking, and eating armadillos. Nine of the 15 patients were infected with the 3I-2-v1 strain. These data confirm interaction with armadillos by some of our patients, and suggest an increased likelihood of infection with the 3I-2-v1 strain as a result (odds ratio vs. having no contact, 4.0; 95% CI, 0.5 to 35.8; P=0.314).

**Discussion**

Combining the high discriminatory power of VNTR analysis with the robust common-roots approach of SNP typing has proved effective in molecular epidemiologic studies of tuberculosis, and was used in our study of leprosy to similar benefit. We show that a high percentage of unrelated leprosy cases in the southern United States involve infection with the same unique strain of M. leprae that occurs naturally among

---

**Figure 3. Minimum-Spanning Phylogenetic Tree of Mycobacterium leprae Genotypes Based on Analysis of Single-Nucleotide Polymorphisms (SNPs) and Variable-Number Tandem Repeats (VNTRs).**

Minimum-spanning-tree analysis was performed with the use of combined VNTR and SNP data from human and armadillo M. leprae strains. Each circle represents a genotype (human unless marked as armadillo) based on the combined data, with the circle size directly proportional to the number of strains with the corresponding genotype. Numbers along the links between circles indicate the number of loci that differ between the genotypes on either side of the link. Three fully sequenced reference M. leprae strains (TN, Thai53, and Br4923) are labeled, as are two other reference strains (LWM26 and 43926) of foreign origin. Samples from patients with a history of foreign residence are indicated with an asterisk (with three asterisks indicating three patients). The 114 polymorphisms investigated include 84 SNPs described previously and 30 identified during our study; 10 VNTRs were also analyzed. The large circle illustrates the predominance of the 3I-2-v1 M. leprae genotype in our study, with 25 patients and 28 armadillos having this identical genotype.
Armadillos are found in the western hemisphere only. Their potential effect on leprosy control in the Americas depends on the rate at which the infection can spread to contiguous armadillo populations and its likelihood to emerge among other high-density armadillo populations. Biomarkers of *M. leprae* have been reported in armadillos in Argentina, Brazil, and Colombia,\(^18,19,21\) and contact with armadillos has been reported to be a risk factor for leprosy in Brazil.\(^33\) Application of advanced genotyping techniques in such areas may help elucidate leprosy transmission in other human populations and increase our understanding of the specific risk factors perpetuating this infection globally.

Susceptibility to leprosy is modulated by a number of complex genetic traits,\(^34\) and the majority of people appear to be naturally immune to *M. leprae* infection. There are currently no recognized pathological variant *M. leprae* strains, and the outcome of infection is influenced primarily by the individual host response. Early diagnosis and prompt drug therapy remain the most effective means to avoid the undesirable complications of leprosy.\(^1\) Physicians caring for patients with potential exposure to *M. leprae* by means of armadillos should consider leprosy in their differential diagnosis of chronic cutaneous lesions, especially those not responsive to common treatments.

Leprosy remains one of the most socially stigmatizing diseases; diagnosis can evoke profound anxiety in the patient, especially when there is no identifiable index case and the source of infection is unknown. Armadillos are the only known non-human reservoir of *M. leprae*. Recognition of this zoonotic link may provide relief to some patients by identifying a biologically plausible source of their disease.

Supported by Fondation Raoul Follereau, the American Leprosy Missions, the Society of Saint Lazarus, the IDEAL (Initiative for Diagnostic and Epidemiological Assays for Leprosy) Consortium, the Health Resources and Services Administration, and the National Institutes of Health, National Institute of Allergy and Infectious Diseases (grants ROI-AI47197-01A1 and NIAID IAA-2646).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank all the participants in this study, particularly J. Adams, K. Andrews, A. Carter, P. Brennan, B. Hall, K. Harshman, M. Kearney, S. Keas, J. Krahenbuhl, G. McCormick, N. Robbins, J. Spencer, R. Stevenson, B. Strzyewiska, S. Uplekar, and H. Zhang; F. Knight, J. Kumaresan, J. Loughry, and C. McDonough for their assistance in armadillo surveys; and the Vital-IT Center of the Swiss Institute of Bioinformatics for computational support.
REFERENCES


Copyright © 2011 Massachusetts Medical Society.