MAJOR ARTICLE

Enhanced Molecular Typing of *Treponema pallidum*: Geographical Distribution of Strain Types and Association with Neurosyphilis

Christina M. Marra,[1] Sharon K. Sahi,[1] Lauren C. Tantalo,[1] Charmie Godornes,[2] Tara Reid,[2] Frieda Behets,[5] Anne Rompalo,[6] Jeffrey D. Klausner,[7] Yue-Ping Yin,[8] Fiona Mulcahy,[9] Matthew R. Golden,[2,4] Arturo Centurion-Lara,[2] and Sheila A. Lukehart[2,3]

Departments of [1]Neurology, [2]Medicine (Infectious Diseases), and [3]Global Health, University of Washington School of Medicine, and [4]Public Health–Seattle and King County, Seattle, Washington; [5]Departments of Epidemiology, Gillings School of Global Public Health and Medicine, University of North Carolina Chapel Hill, Chapel Hill, North Carolina; [6]Departments of Medicine (Infectious Diseases) and Obstetrics and Gynecology, Johns Hopkins University, Baltimore, Maryland; [7]San Francisco Department of Public Health, San Francisco, California; [8]Department of Reference Sexually Transmitted Disease Laboratory, National Center for Sexually Transmitted Disease Control, Chinese Academy of Medical Sciences and Peking Union Medical College Institute of Dermatology, Nanjing, China; [9]Department of Genitourinary Medicine and Infectious Diseases, St James Hospital and Trinity College, Dublin, Ireland

Background. Strain typing is a tool for determining the diversity and epidemiology of infections.

Methods. *Treponema pallidum* DNA was isolated from 158 patients with syphilis from the United States, China, Ireland, and Madagascar and from 15 *T. pallidum* isolates. Six typing targets were assessed: (1) the number of 60-bp repeats in the acidic repeat protein gene, (2) restriction fragment length polymorphism (RFLP) analysis of *T. pallidum* repeat (*tpr*) subfamily II genes, (3) RFLP analysis of the *tprC* gene, (4) determination of *tprD* allele in the *tprD* gene locus, (5) the presence of a 51-bp insertion between *tp0126* and *tp0127*, and (6) sequence analysis of an 84-bp region of *tp0548*. The combination of targets 1 and 2 comprises the Centers for Disease Control and Prevention (CDC) *T. pallidum* subtyping method.

Results. Adding sequence analysis of *tp0548* to the CDC method yielded the most discriminating typing system. Twenty-five strain types were identified and designated as "CDC subtype/*tp0548* sequence type." Type 14d/f was found in samples from 5 of 6 locations. In Seattle, Washington, strain types changed from 1999 through 2008 ($P < .001$). Twenty-one (50%) of 42 patients infected with type 14d/f had neurosyphilis compared with 10 (24%) of 41 patients infected with any of the other types combined ($P = .02$).

Conclusion. We describe an enhanced *T. pallidum* strain typing system that shows biological and clinical relevance.

Strain typing is a powerful tool for determining the diversity of pathogens and the epidemiology of infections. In 1998, investigators at the Centers for Disease Control and Prevention (CDC) published a molecular method to distinguish among subtypes of *Treponema pallidum* subspecies *pallidum* (hereafter *T. pallidum*), the causative agent of syphilis [1]. The method is based on (1) determination of the number of 60-bp repeats in the acidic repeat protein (*arp*) gene and (2) sequence differences in the *T. pallidum* repeat (*tpr*) subfamily II genes (*tprE [tp0313]*, *tprG [tp0317]*, and *tprJ [tp0621]*) determined by restriction fragment length polymorphism (RFLP) analysis. Subtype designation is based on the number of repeats and the RFLP pattern—for example, subtype 14d. This subtyping method has been applied to bacterial DNA recovered from chancres, condyloma lata, mouth scrapings, ear scrapings, blood samples, and cerebrospinal fluid (CSF) samples and to laboratory-passaged *T. pallidum* isolates from diverse

Received 15 April 2010; accepted 26 May 2010; electronically published 24 September 2010.

Potential conflicts of interest: none reported.

Presented in part: 17th Conference on Retroviruses and Opportunistic Infections, San Francisco, CA, 16–19 February 2010 (abstract 177).

Financial support: National Institutes of Health (grants NS34235, AI42143, and AI63940); Sexually Transmitted Diseases Clinical Trials Unit (grant N01-AI-75329); Sexually Transmitted Infection Clinical Trials Group (grant N01-AI-40073).

Reprints or correspondence: Dr Christina M. Marra, University of Washington School of Medicine, Dept of Neurology, Harborview Medical Center, Box 359775, 325 9th Ave, Seattle, WA 98104-2499 (cmarra@u.washington.edu).

The Journal of Infectious Diseases 2010; 202(9):1380–1388

0022-1899/2010/20209-0012$15.00
DOI: 10.1086/656533

geographic areas in the hopes that molecular typing could be used to characterize an epidemic and assist in control efforts [1–8]. However, a particular drawback of the published subtyping system is its inability to discriminate among the most common subtypes.

We performed this study to determine whether examination of additional gene targets might lead to a more discriminating typing system than the published CDC system. We reported that identification of a 51-bp insertion in the intergenic spacer between the *tp0126* and *tp0127* open reading frames could be used to discriminate among *T. pallidum* organisms [9]. Strain-to-strain sequence variations in the *T. pallidum tprC* (*tp0117*) and *tprD* (*tp0131*) genes have been noted [10, 11], and sequence analysis of the *tp0548* gene has been used to distinguish among *T. pallidum* from patient-derived samples [12, 13]. We examined each of these targets in samples from different areas of the world and determined that sequence analysis of a short region of the *tp0548* gene, in addition to analysis of the *arp* and *tpr* subfamily II genes, greatly improved discrimination among strains, allowing us to divide 14 CDC subtypes into 25 strain types. Using this improved strain typing system, we demonstrate geographic distributions of strain types, changes in strain types over time in a single geographic region, and an association between strain type and neurosyphilis.

METHODS

Sample sources and DNA extraction. Samples from 173 individual patients from different geographical sites were examined in this study: (1) 72 *T. pallidum* isolates derived from blood samples, CSF samples, or lesion exudates in Seattle, Washington, collected in 1981, 1983, 1987, and from 1999 through 2008; (2) 11 historical *T. pallidum* isolates, including the reference Nichols strain, which was derived from a CSF sample in 1912; (3) 16 blood samples collected in Seattle from 1999 through 2008; and (4) from randomly selected swab samples from primary or secondary lesion exudates collected at the following sites: Antananarivo, Toamasina, and Mahajanga in Madagascar from 2003 through 2008 ($n = 20$) [14]; San Francisco, California, from 2001 through 2007 ($n = 19$) [15]; Baltimore, Maryland, from 1999 through 2001 ($n = 15$); Nanning, Guangxi Autonomous Region, China, from 2006 through 2007 ($n = 10$), and Dublin, Ireland, in 2002 ($n = 10$) [16]. The study protocol was reviewed and approved by local institutional review boards, and human experimentation guidelines were followed in the conduct of this research. Written informed consent was obtained from all participants.

The relationship between strain types and neurosyphilis was examined in a subset of 84 patients enrolled in an ongoing study of CSF abnormalities in patients with syphilis conducted at the University of Washington (Seattle, WA) from 1999 through 2008 [17]. Briefly, individuals were eligible for enrollment if they had clinical or serological evidence of syphilis and were deemed by the referring health care provider as possibly having neurosyphilis. All participants underwent a structured history and neurological examination, lumbar puncture, and venipuncture.

DNA was extracted from swabs and *T. pallidum* suspensions that were stored in 10 mmol/L Tris-hydrochloride, 0.1 mol/L ethylenediaminetetraacetic acid (EDTA), and 0.5% sodium dodecyl sulfate with the use of a QIAamp DNA Mini kit (Qiagen) according to the manufacturer's instructions. DNA was extracted from blood samples stored in 20 mmol/L Tris-hydrochloride, 0.2 mol/L EDTA, and 1% sodium dodecyl sulfate with the use of a QIAamp DNA Blood Midi kit (Qiagen) according to the manufacturer's instructions.

Determination of the number of 60-bp repeats in the* arp *gene. Ten microliters of DNA was amplified using 2.5 U of GoTaq Flexi DNA polymerase (Promega) in a 50-μL reaction with 200 nmol/L dNTP, 0.6 μmol/L primers, 1.5 mmol/L magnesium chloride ($MgCl_2$), and 1× GoTaq Flexi buffer. The following primer sequences were used: sense, 5′AGCGTGATCCTCTGTCATCC3′; antisense, 5′TATGCTGAGAAGCGACCTCA3′. The cycling conditions were as follows: 94°C for 4 min; then 40 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min; followed by 72°C for 7 min. The polymerase chain reaction (PCR) products were separated by agarose gel electrophoresis. The size of all PCR products was determined by comparison with a 1-kb DNA ladder (Invitrogen) with the use of Quantity One software (version 4.6.7; Biorad), and the number of repeats was established by comparing the size of the amplified product with that of the product amplified from the Nichols strain, which contains 14 repeats [1].

RFLP analysis of the* tprE, tprG, *and* tprJ *genes. Five microliters of DNA was amplified using 2.5 U of GoTaq Flexi DNA polymerase in a 50-μL reaction with 200 nmol/L dNTP, 0.6 μmol/L primers, 2.5 mmol/L $MgCl_2$, and 1× GoTaq Flexi buffer. The following primer sequences were used: sense, 5′CAGGTTTTGCCGTTAAGC3′; antisense, 5′AATCAAGGGAGAATACCGTC3′. The cycling conditions were as follows: 94°C for 4 min; then 40 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min; followed by 72°C for 7 min.

The *tprE, tprG,* and *tprJ* amplicons were digested according to published methods with minor modifications. Digestion products were separated by agarose gel electrophoresis, and banding patterns were compared to published data [1].

RFLP analysis of the* tprC *gene. Five microliters of DNA was amplified using 1.25 U of GoTaq Flexi DNA polymerase in a 50-μL reaction with 200 nmol/L dNTP, 2.5 mmol/L $MgCl_2$, 0.6 μmol/L primers, and 1× GoTaq Flexi buffer. The following primer sequences were used: sense, 5′GTACATCCCCCTCACCTAC3′; antisense, 5′AAAAATGGTACAGGCGTTGC3′. The cycling conditions were as follows: 95°C for 2 min; then

Table 1. Primers Used to Amplify the *tp0548* Gene of *Treponema pallidum*

Primer designation	Primer sequence
Sense 1	5′GCGTGGTGGTGAGTTCTTCT3′
Antisense 1	5′ACGGCAGGCTAGTTGAGAAT3′
Sense 2	5′GGTCCCTATGATATCGTGTTCG3′
Antisense 2	5′GTCATGGATCTGCGAGTGG3′
Antisense 3	5′CGTTTCGGTGTGTGAGTCAT3′

40 cycles of 95°C for 1 min, 60°C for 30 s, and 72°C for 1 min; followed by 72°C for 7 min.

Amplified products were digested using 5 U each of BsiEI and BsiHKAI enzymes (New England Biolabs) for 2 h at 60°C and then 2 h at 65°C. The products of the restriction digestion were separated by agarose gel electrophoresis. The size of the largest band (276, 249, or 203 bp) was used to distinguish between organisms.

Analysis of the* tprD *gene locus. Five microliters of DNA was amplified using 1.25 U of GoTaq Flexi DNA polymerase in a 50-μL reaction with 200 nmol/L dNTP, 3.0 mmol/L $MgCl_2$, 0.6 μmol/L primers, and 1× GoTaq Flexi buffer. The following primer sequences were used: sense, 5′AGTACCAGGTGGGACTGACG3′; antisense, 5′GAACGGGTCTCCACACTCAC3′. The cycling conditions were as follows: 95°C for 2 min; then 40 cycles of 95°C for 1 min, 65°C for 30 s, and 72°C for 1 min; followed by 72°C for 5 min.

Amplified products were digested using 10 U of MluI enzyme (New England Biolabs) at 37°C for 2 h. Products of the restriction digestion were separated by agarose gel electrophoresis. The size of the products identified the *tprC/D* gene (232 and 167 bp) or the *tprD2* gene (399 bp).

Sequence analysis of the* tp0548 *gene. Two sense and 3 antisense primers were used to amplify DNA from bacterial isolates (sense 1 and antisense 1), lesion exudates (sense 2 and antisense 2), and blood samples (sense 2 and antisense 3) (Table 1). Different primer sets were used for the 3 sample types to optimize analytic sensitivity. For isolates, 5 μL of DNA was amplified using 2.5 U of GoTaq Flexi DNA polymerase in a 50-μL reaction with 200 nmol/L dNTP, 0.8 μmol/L primers, 1.5 mmol/L $MgCl_2$, and 1× GoTaq Flexi buffer. The cycling conditions were as follows: 95°C for 2 min; then 40 cycles of 95°C for 1 min, 60°C for 2 min, and 72°C for 1 min; followed by 72°C for 10 min. For amplification of DNA from exudate swabs or blood samples, the methods were identical except that 0.6 μmol/L of each primer was used and the annealing step was performed at 62°C for 75 s. PCR products were purified using a QIAquick PCR purification kit (Qiagen) and sequenced using standard methods. Variability in a portion of the gene 131 bp downstream from the start codon was noted. Nine distinct sequence groups were observed and designated as a–i (Figure 1).

Cloning. Amplicons were ligated into the TOPO TA vector (Invitrogen) according to the manufacturer's instructions, and clones were sequenced according to standard methods.

Clinical laboratory methods. CSF Venereal Disease Research Laboratory (CSF-VDRL) tests were performed and CSF white blood cell (WBC) counts were measured in laboratories approved by the Clinical Laboratory Improvement Amendments. The serum rapid plasma reagin (RPR) test was performed in a single research laboratory according to standard methods [18].

Statistical methods. Associations between categorical variables were determined by the χ^2 test or the Fisher exact test. Results for which $P < .05$ were considered to be statistically significant.

RESULTS

Determination of the optimal typing system. We first analyzed 84 samples with the use of all 6 targets to determine which combination provided the highest level of discrimination. Among these samples, the CDC method identified 12 subtypes, most frequently 14a ($n = 10$), 14d ($n = 51$), and 15d ($n = 9$). Compared with the CDC method, determination of the presence or absence of the insertion between *tp0126* and *tp0127* and RFLP analysis of *tprC* did not improve discrimination. Sequence analysis of *tp0548* divided the most common

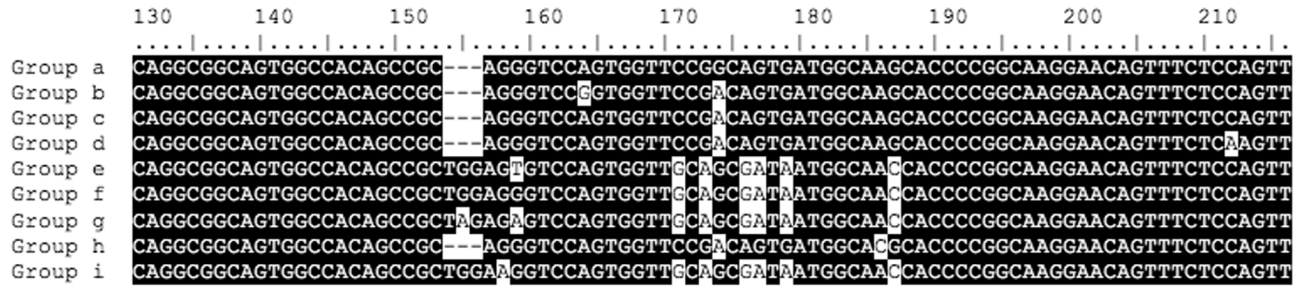

Figure 1. Typing groups determined by the sequence of base pairs 131–215 in the *tp0548* open reading frame of *Treponema pallidum*.

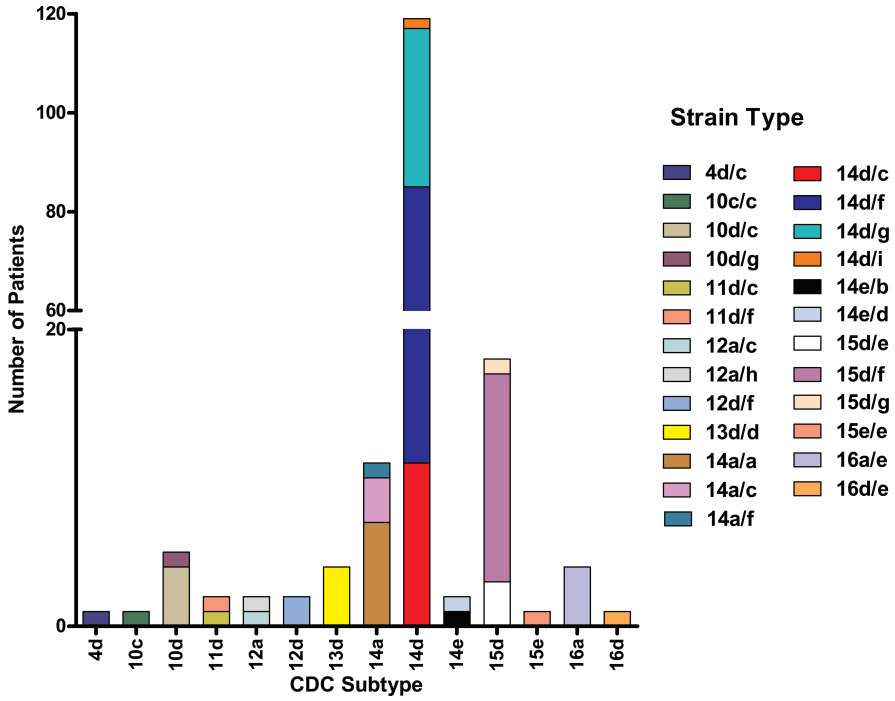

Figure 2. Comparison of Centers for Disease Control and Prevention (CDC) subtypes of *Treponema pallidum* with strain types that are defined by the addition of sequence analysis of the *tp0548* gene to the CDC method of subtyping. Fourteen CDC subtypes could be divided into 25 strain types. Importantly, the 3 subtypes that contain 14 repeats of the acidic repeat protein gene could be separated into 9 separate types by use of the enhanced typing scheme. The lower portion of the *Y*-axis has been expanded to clearly show the low-frequency molecular types.

CDC subtype, 14d, into 4 separate groups, and subtypes 12a and 14a could each be divided into 2 groups. Addition of the *tprD* analysis to that provided by *tp0548* only incrementally improved discriminatory ability, allowing us to divide 1 additional CDC subtype, 10d, into 2 groups.

On the basis of practical considerations of cost versus increased discriminating ability, we propose a new *T. pallidum* typing method that includes the published CDC method plus sequence analysis of base pairs 131–215 of the *tp0548* open reading frame. From here forward, we refer to the type found by the revised method as the "strain type," to distinguish it from that found by the CDC method, which has been termed "subtype." We express the strain type as "CDC subtype/*tp0548* sequence type"; for example, 14d/f. Using this method, we subsequently determined strain type of all 173 samples.

Figure 2 shows the relationships between the CDC subtypes and the strain types. Analysis of all 173 samples identified 14 CDC subtypes, which could be further divided into 25 strain types. Specifically, subtypes 10d, 11d, 12a, 14a, 14d, 14e, and 15d could all be subdivided using the revised strain typing system, which attests to its improved discrimination. Types 14d/f, 14d/g, and 15d/f were the most common strain types found. Table 2 shows a summary of the newly defined strain types and tested samples.

Stability of strain type within isolates. Pillay and colleagues [1] showed that the CDC subtype of the Nichols strain did not change with repeated rabbit passages. We confirmed this finding. Five clones of the Nichols strain at 22 passages after infection and 5 clones of the original inoculum showed identical strain type (14a/a). We examined strain type stability for 2 other *T. pallidum* isolates. The strain type of the Sea 81-4 isolate (14e/b) remained the same at rabbit passages 12 and 22. Previous work showed that the Chicago C isolate rapidly acquires diversity of the *tprK* gene sequence with serial rabbit passage [19]. Even after 12 passes of the Chicago C isolate in rabbits, this isolate demonstrated strain identity with the original inoculum (14a/a).

Geographic distribution of strains. Twenty of the 25 strain types were identified in samples from Baltimore, San Francisco, China, Dublin, and Madagascar and among the participants enrolled in the study at the University of Washington in Seattle (Figure 3 and Table 2). More than 1 strain type was found in samples from each of the 6 locations. Types 14d/c, 14d/f, 14d/g, 15d/f, and 16a/e were found in samples from at least 2 different locations, with type 14d/f found in samples from everywhere except in Madagascar. Three of the 4 California strain types were also represented in Seattle, perhaps suggesting a sexual connection between the San Francisco and Seattle cohorts.

Characteristics of the 84 patients enrolled in the study at the University of Washington, Seattle. The characteristics of

Table 2. Strain Types and Sample Sources of *Treponema pallidum*

Strain type, geographic site	No. of samples	Representative isolate
4d/c		NA[a]
Madagascar	1	
10c/c		NA[a]
Madagascar	1	
10d/c		NA[a]
Madagascar	4	
10d/g		NA[a]
Dublin	1	
11d/c		NA[a]
Madagascar	1	
11d/f		NA[a]
Baltimore	1	
12a/c[b]		Bal 8
12a/h[b]		Bal 9
12d/f		NA[a]
Baltimore	2	
13d/d		UW279
Seattle	4	
14a/a[b]		Nichols
14a/c		NA[a]
Madagascar	3	
14a/f		NA[a]
China	1	
14d/c		NA[a]
San Francisco	1	
Madagascar	10	
14d/f		Street 14
San Francisco	10	
China	9	
Dublin	1	
Baltimore	7	
Seattle	43	
14d/g		UW249
San Francisco	4	
Dublin	7	
Seattle	21	
14d/i		UW149
Seattle	2	
14e/b[b]		Sea 81-4
14e/d		NA[a]
Dublin	1	
15d/e		NA[a]
Baltimore	3	
15d/f		UW116
San Francisco	4	
Seattle	10	
15d/g		NA[a]
Seattle	1	
15e/e		NA[a]
Seattle	1	
16a/e		UW284
Baltimore	2	
Seattle	2	
16d/e[b]		Mexico

[a] *T. pallidum* isolate not available because these strain types were identified from patient samples only.

[b] Strain type not found in samples from any of the 6 geographic sites.

the 84 study participants in Seattle are shown in Table 3. Most patients were white men who have sex with men (MSM). There was a trend toward a difference in strain type by human immunodeficiency virus (HIV) status ($P = .08$).

The strain type was identified in samples from 3 subjects who returned during our study with a second episode of syphilis. Subject 1's first episode of syphilis was in July 2003, when he received treatment for secondary syphilis. He returned with secondary syphilis in January 2004. Strain type 13d/d was identified in samples collected during each episode, which suggests reinfection from a member of the same sexual network or treatment failure. Subject 2's first episode of syphilis was in April 2006, when he received treatment for secondary syphilis. He returned with early latent syphilis in July 2008. Strain type 14d/f was identified in the sample collected during the first episode and type 14d/g in the sample collected during the second episode, which indicates reinfection with a new strain. Subject 3 presented in January 2003, when he received treatment for secondary syphilis and neurosyphilis. Seven months after treatment, the serum RPR titer decreased from 1:512 to 1:64. He returned with early latent syphilis in January 2009. Strain type 14d/f was identified in the sample collected during the first episode and type 14d/g in the sample collected during the second episode, which again demonstrates reinfection with a new strain.

Temporal distribution of strain types in Seattle. Typing of a large number of samples collected over 10 years in Seattle enabled us to examine the temporal distribution of strain types in a single community. Figure 4 shows the 8 Seattle strain types identified by year of collection. Type 14d/f predominated from 1999 through 2003 and remained as a minor strain through 2007. Type 15d/f appeared in 2002 and was not seen after 2005. Type 14d/g was first seen in 2004 and was identified through 2008. It is notable that the introduction of type 14d/g coincided with an increasing incidence of syphilis in Seattle [20]. Type 16a/e was found in samples from 2 African American heterosexual partners in 2004. A single patient with type 15d/g was seen in 2004, single patients with type 14d/i were seen in 2003 and 2004, and a single patient with type 15e/e was seen in 2006.

For purposes of analysis, the years of sample collection were divided into 3 epochs (1999–2002, 2003–2005, and 2006–2008). There was a significant change in strain type over the 3 time periods ($P < .001$). We also determined the strain type of 4 isolates collected in Seattle in 1981, 1983, and 1987. Two isolates (collected in 1983 and 1987) were type 14d/f, the most common strain type identified in samples collected in Seattle from 1999 through 2008. In contrast, 1 isolate collected in 1981 was type 14e/b and another collected in 1983 was type 14a/a. Neither of these types was found in samples collected in Seattle from 1999 through 2008.

Association of strain type with neurosyphilis. A sensitive

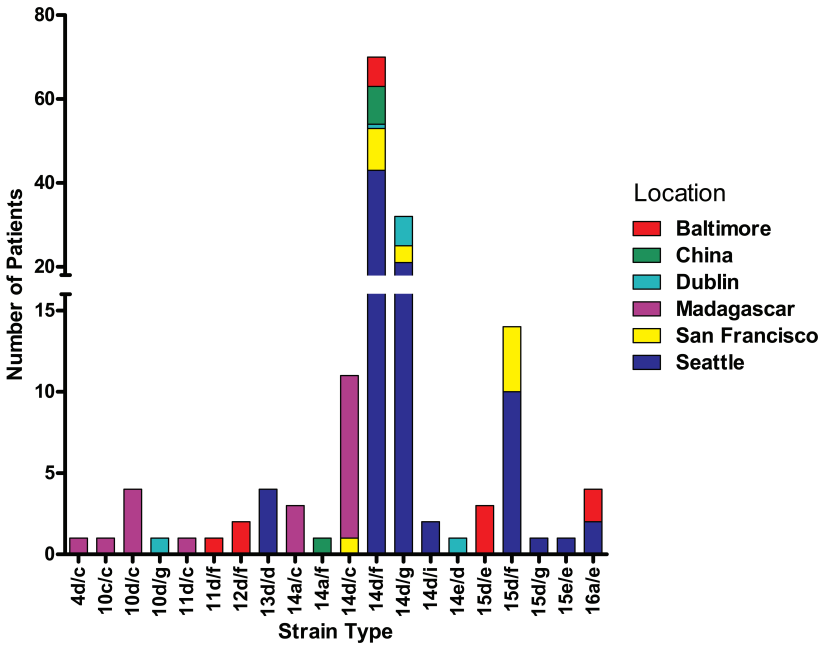

Figure 3. Strain types of *Treponema pallidum* identified in samples from 6 geographic regions. More than 1 strain type was found in samples from all locations, and type 14d/f was identified in samples from all sites other than Madagascar. Several strain types were found in samples from only a single location. The lower portion of the *Y*-axis has been expanded to clearly show the low-frequency molecular types.

and discriminating typing method would be particularly useful if strain type could be associated with clinical findings. The samples from patients in Seattle were obtained in an ongoing study of neurosyphilis, and we examined whether any particular strain type was more commonly associated with this disease. Twenty-one (50%) of 42 patients infected with type 14d/f had neurosyphilis (defined as a CSF WBC count of >20 cells/μL or a reactive CSF-VDRL test result) compared with 10 (24%) of 41 patients infected with any of the other 7 strains combined ($P = .02$). Similarly, 15 (36%) of 42 patients infected with type 14d/f had a reactive CSF-VDRL test result, compared with 5 (12%) of 41 patients infected with any of the other 7 strains combined ($P = .01$). There was no difference in the proportion of patients who had neurosyphilis over the 10-year period of observation, and type 14d/f was equally common in HIV-infected patients and HIV-uninfected patients.

DISCUSSION

Several previous studies have used the CDC method to investigate *T. pallidum* subtype prevalence among communities. For example, Sutton and colleagues [2] determined CDC subtypes from *T. pallidum* DNA isolated from blood samples and genital ulcer swabs during an outbreak of syphilis in Phoenix, Arizona. About half of the samples were CDC subtype 14f. Nine other subtypes were identified, including 14d. Subtypes 14d and 14f have been identified in samples from diverse geographic settings, including North and South Carolina [4], Lisbon [7], Scotland [8], South Africa [1, 3], and China [21]. Among the samples examined in our study, CDC subtype 14d was most common, and we did not identify any subtype 14f samples. By enhancing the CDC subtyping method with sequence analysis of a small region of the *tp0548* gene, we were able to separate 14 CDC subtypes into 25 different strain types. Importantly, we were able to separate the subtypes with 14 *arp* repeats into 9 individual strain types.

We found a change in the circulating strain types in Seattle, Washington, from 1999 through 2008, with the introduction and disappearance of strain types in the region during that period. Retrospectively, we also saw the disappearance of 2 strains that were evident in Seattle in the 1980s, compared with the strains identified in samples collected from 1999 through 2008. The expansion of the introduced types within our recent cohort of patients who were predominantly MSM is consistent with 1 or more overlapping sexual networks in the community of MSM in Seattle. Moreover, we saw a trend toward an association between strain type and HIV status. This finding further suggests that there are separate networks among MSM in our community and is consistent with serosorting such that patients choose sexual partners based on their HIV status. We identified an unusual type (type 16a/e) in only 2 individuals in 2004, without subsequent detection in other patients. These 2 individuals were African American heterosexual partners. In addition, a single African American patient was infected with a type 15e/e strain. The finding of these unique strains in Af-

Table 3. Characteristics of Participants Enrolled in the Study at the University of Washington, Seattle, Washington

Characteristic	No. (%) of participants (N = 84)
Male	82 (98)
Race	
White	72 (86)
African American	8 (10)
Other	4 (4)
Sexual identification	
MSM	73 (87)
Bisexual	6 (7)
Heterosexual	4 (5)
Transgender	1 (1)
HIV-infected	70 (83)
Injection drug use, ever	19 (23)
Syphilis	
Primary, secondary, or early latent	78 (93)
Late latent or unknown duration	6 (7)
Serum RPR titer, median (IQR)	128 (64–256)
CSF WBC count of >20 cells/μL or reactive CSF-VDRL test result	31 (37)[a]
Reactive CSF-VDRL test result	20 (24)[a]

NOTE. Data are no. (%) of participants, unless otherwise indicated. CSF-VDRL, cerebrospinal fluid Venereal Disease Research Laboratory; HIV, human immunodeficiency virus; MSM, men who have sex with men; RPR, rapid plasma reagin.

[a] The CSF-VDRL test result was missing for 1 participant

rican American patients in the setting of an epidemic among predominantly white MSM provides additional evidence for separate sexual networks.

Many of the samples examined in this study have also been tested for 23S ribosomal DNA mutations associated with macrolide resistance in *T. pallidum* [14, 22, 23], and we considered including either or both of these mutations in our typing scheme. Although these markers were able to distinguish among strains within some types, we were uncertain about the long-term stability of this marker within a strain, particularly in light of studies that have shown the association of the A2058G mutation with prior exposure to macrolides [9, 15]. Similarly, in a small number of samples, we examined 2 other molecular targets that proved to be unsuitable for strain typing: sequence analysis of the *tp0136* gene [12] and determination of the number of 24-bp repeats in the *tp0470* gene. Sequence analysis of the *tp0136* gene lacked discriminatory ability, and the number of repeats in the *tp0470* gene varied within a single bacterial isolate (data not shown).

The association between strain type and neurosyphilis is particularly notable. Although literature from the first part of the 20th century discusses the existence of neuroinvasive or neurotropic strains of *T. pallidum*, this concept has not been rigorously studied in humans. In a rabbit model, we showed that the clinical phenotype of disease differed depending on the inoculating strain [24]. In that study, rabbits were infected with 3 different strain types: 14a/a, 14e/b, and 14d/f. Animals infected with 1 type 14a/a strain and 1 type 14d/f strain had the greatest degree of neuroinvasion. In the study reported here, we found that patients with neurosyphilis were most commonly

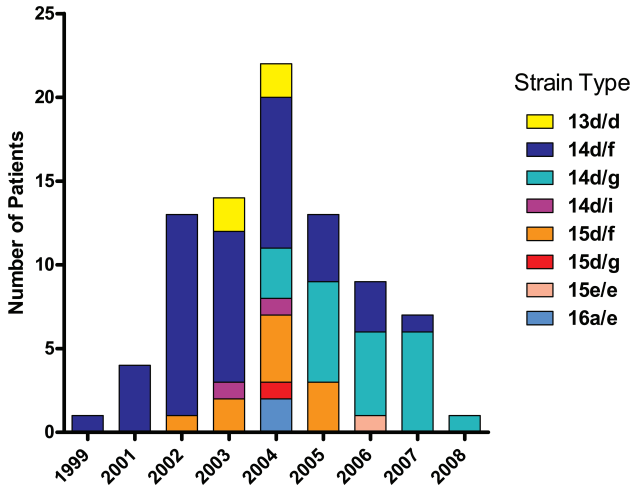

Figure 4. Strain types of *Treponema pallidum* identified in samples collected each year from 1999 through 2008 from patients from Seattle, Washington. The graph shows the introduction and loss of some strain types during that period. There was significant change in type over time ($P < .001$).

infected with type 14d/f, which is consistent with the findings of the rabbit studies. We cannot comment on the neuroinvasive potential of type 14e/b in our study of syphilis in humans, as none of the patients in Seattle was infected with this strain. *T. pallidum* type 14d/f strains may be more neuroinvasive, or they may be better able to evade host immune responses in the central nervous system (CNS) compared with other strain types.

The ability to identify the infecting strain types in patients with syphilis could have important clinical implications. Controversy abounds regarding which patients with syphilis should undergo lumbar puncture to evaluate the possibility of neurosyphilis. Currently, the best indicator of risk is a serum RPR titer of $\geqslant$1:32, but that criterion leads to some false positives [17, 25]. If the strain type could be determined in a blood sample, then it could help to identify those patients who are at greatest risk for neurosyphilis. These individuals, particularly those with RPR titers of $\geqslant$1:32, could be then targeted for lumbar puncture or empiric neurosyphilis treatment. Our data suggest that future larger investigations of the correlation between the strain type and the clinical or laboratory markers that indicate increased risk of neurosyphilis are warranted. In addition, the association between specific strain type and neurosyphilis could lead a new understanding of the molecular mechanisms underlying neuroinvasion.

The enhanced typing method that we describe shows biological and clinical relevance, as well as epidemiological utility. It represents a significant advance in our ability to study the molecular epidemiology of syphilis, and it offers the potential to learn more about the pathogenesis of CNS disease. Future epidemiological studies that combine social network analysis and strain typing data are required to determine the ultimate utility of this new typing method for syphilis investigation and control. Similarly, continued study of risks for neurosyphilis is required to determine the ultimate role of strain typing as part of risk assessment.

Acknowledgments

We thank Dr Thanh Ton of the University of Washington School of Medicine Department of Neurology (Seattle, WA) for valuable statistical support. We gratefully acknowledge the contributions of Trudy Jones of the University of Washington School of Medicine Department of Neurology, Rolf Pederson and the staff of the Seattle King County Sexually Transmitted Diseases Clinic (Seattle, WA), the members of the University of North Carolina–Madagascar study team in Madagascar, the staff of the San Francisco City Clinic (San Francisco, CA), and Dr Bang-Yong Zhu of the Guangxi Autonomous Region Institute of Dermatology (Nanning, China).

References

1. Pillay A, Liu H, Chen CY, et al. Molecular subtyping of *Treponema pallidum* subspecies *pallidum*. Sex Transm Dis **1998**; 25:408–414.
2. Sutton MY, Liu H, Steiner B, et al. Molecular subtyping of *Treponema pallidum* in an Arizona county with increasing syphilis morbidity: use of specimens from ulcers and blood. J Infect Dis **2001**; 183:1601–1606.
3. Pillay A, Liu H, Ebrahim S, et al. Molecular typing of *Treponema pallidum* in South Africa: cross-sectional studies. J Clin Microbiol **2002**; 40:256–258.
4. Pope V, Fox K, Liu H, et al. Molecular subtyping of *Treponema pallidum* from North and South Carolina. J Clin Microbiol **2005**; 43:3743–3746.
5. Molepo J, Pillay A, Weber B, Morse SA, Hoosen AA. Molecular typing of *Treponema pallidum* strains from patients with neurosyphilis in Pretoria, South Africa. Sex Transm Infect **2007**; 83:189–192.
6. Florindo C, Reigado V, Gomes JP, Azevedo J, Santo I, Borrego MJ. Molecular typing of *Treponema pallidum* clinical strains from Lisbon, Portugal. J Clin Microbiol **2008**; 46:3802–3803.
7. Castro R, Prieto E, Aguas MJ, Manata MJ, Botas J, Pereira FM. Molecular subtyping of *Treponema pallidum* subsp. *pallidum* in Lisbon, Portugal. J Clin Microbiol **2009**; 47:2510–2512.
8. Cole MJ, Chisholm S, Palmer HM, Wallace LA, Ison CA. Molecular epidemiology of syphilis in Scotland. Sex Transm Infect **2009**; 85:447–451.
9. Marra CM, Colina AP, Godornes C, et al. Antibiotic selection may contribute to increases in macrolide-resistant *Treponema pallidum*. J Infect Dis **2006**; 194:1771–1773.
10. Centurion-Lara A, Sun ES, Barrett LK, Castro C, Lukehart SA, Van Voorhis WC. Multiple alleles of *Treponema pallidum* repeat gene D in *Treponema pallidum* isolates. J Bacteriol **2000**; 182:2332–2335.
11. Sun ES, Molini BJ, Barrett LK, Centurion-Lara A, Lukehart SA, Van Voorhis WC. Subfamily I *Treponema pallidum* repeat protein family: sequence variation and immunity. Microbes Infect **2004**; 6:725–737.
12. Woznicova V, Smajs D, Wechsler D, Matejkova P, Flasarova M. Detection of *Treponema pallidum* subsp. *pallidum* from skin lesions, serum, and cerebrospinal fluid in an infant with congenital syphilis after clindamycin treatment of the mother during pregnancy. J Clin Microbiol **2007**; 45:659–661.
13. Flasarova M, Smajs D, Matejkova P, Woznicova V, Heroldova-Dvorakova M, Votava M. Molecular detection and subtyping of *Treponema pallidum* subsp. *pallidum* in clinical specimens [in Czech]. Epidemiol Mikrobiol Imunol **2006**; 55:105–111.
14. Van Damme K, Behets F, Ravelomanana N, et al. Evaluation of azithromycin resistance in *Treponema pallidum* specimens from Madagascar. Sex Transm Dis **2009**; 36:775–776.
15. Mitchell SJ, Engelman J, Kent CK, Lukehart SA, Godornes C, Klausner JD. Azithromycin-resistant syphilis infection: San Francisco, California, 2000–2004. Clin Infect Dis **2006**; 42:337–345.
16. Hopkins S, Lyons F, Coleman C, Courtney G, Bergin C, Mulcahy F. Resurgence in infectious syphilis in Ireland: an epidemiological study. Sex Transm Dis **2004**; 31:317–321.
17. Marra CM, Maxwell CL, Smith SL, et al. Cerebrospinal fluid abnormalities in patients with syphilis: association with clinical and laboratory features. J Infect Dis **2004**; 189:369–376.
18. Larsen SA, Pope V, Johnson RE, Kennedy EJ Jr. A manual of tests for syphilis. 9th ed. Washington, DC: American Public Health Association, **1998**:193–207.
19. LaFond RE, Centurion-Lara A, Godornes C, Van Voorhis WC, Lukehart SA. TprK sequence diversity accumulates during infection of rabbits with *Treponema pallidum* subsp. *pallidum* Nichols strain. Infect Immun **2006**; 74:1896–1906.
20. Seattle and King County Department of Public Health. Sexually transmitted diseases epidemiology report 2008. Seattle and King County Department of Public Health Web site. http://www.kingcounty.gov/healthservices/health/communicable/std/statistics.aspx. Accessed 7 April 2010.
21. Martin IE, Gu W, Yang Y, Tsang RS. Macrolide resistance and molecular types of *Treponema pallidum* causing primary syphilis in Shanghai, China. Clin Infect Dis **2009**; 49:515–521.
22. Lukehart SA, Godornes C, Molini BJ, et al. Macrolide resistance in *Treponema pallidum* in the United States and Ireland. N Engl J Med **2004**; 351:154–158.
23. Matejkova P, Flasarova M, Zakoucka H, et al. Macrolide treatment

failure in a case of secondary syphilis: a novel A2059G mutation in the 23S rRNA gene of *Treponema pallidum* subsp. *pallidum*. J Med Microbiol **2009**; 58:832–836.
24. Tantalo LC, Lukehart SA, Marra CM. *Treponema pallidum* strain-specific differences in neuroinvasion and clinical phenotype in a rabbit model. J Infect Dis **2005**; 191:75–80.
25. Libois A, De Wit S, Poll B, et al. HIV and syphilis: when to perform a lumbar puncture. Sex Transm Dis **2007**; 34:141–144.