



Sequence Variation of Rare Outer Membrane Protein β -Barrel Domains in Clinical Strains Provides Insights into the Evolution of *Treponema pallidum* subsp. *pallidum*, the Syphilis Spirochete

Sanjiv Kumar,^{a*} Melissa J. Caimano,^{a,b,c} Arvind Anand,^a Abhishek Dey,^{a*} Kelly L. Hawley,^{b,d} Morgan E. LeDoyt,^a Carson J. La Vake,^b Adriana R. Cruz,^e Lady G. Ramirez,^e Lenka Paštěková,^f Irina Bezsonova,^c David Šmajš,^f Juan C. Salazar,^{b,d,e,g} Justin D. Radolf^{a,b,c,g,h}

^aDepartment of Medicine, UConn Health, Farmington, Connecticut, USA

^bDepartment of Pediatrics, UConn Health, Farmington, Connecticut, USA

^cDepartment of Molecular Biology and Biophysics, UConn Health, Farmington, Connecticut, USA

^dDivision of Pediatric Infectious Diseases, Connecticut Children's Medical Center, Hartford, Connecticut, USA

^eCentro Internacional de Entrenamiento e Investigaciones Medicas (CIDEIM), Cali, Colombia

^fDepartment of Biology, Faculty of Medicine, Masaryk University, Brno, Czech Republic

^gDepartment of Immunology, UConn Health, Farmington, Connecticut, USA

^hDepartment of Genetic and Genome Sciences, UConn Health, Farmington, Connecticut, USA

ABSTRACT In recent years, considerable progress has been made in topologically and functionally characterizing integral outer membrane proteins (OMPs) of *Treponema pallidum* subspecies *pallidum*, the syphilis spirochete, and identifying its surface-exposed β -barrel domains. Extracellular loops in OMPs of Gram-negative bacteria are known to be highly variable. We examined the sequence diversity of β -barrel-encoding regions of *tpnC*, *tpnD*, and *bamA* in 31 specimens from Cali, Colombia; San Francisco, California; and the Czech Republic and compared them to allelic variants in the 41 reference genomes in the NCBI database. To establish a phylogenetic framework, we used *T. pallidum* 0548 (*tp0548*) genotyping and *tp0558* sequences to assign strains to the Nichols or SS14 clades. We found that (i) β -barrels in clinical strains could be grouped according to allelic variants in *T. pallidum* subsp. *pallidum* reference genomes; (ii) for all three OMP loci, clinical strains within the Nichols or SS14 clades often harbored β -barrel variants that differed from the Nichols and SS14 reference strains; and (iii) OMP variable regions often reside in predicted extracellular loops containing B-cell epitopes. On the basis of structural models, nonconservative amino acid substitutions in predicted transmembrane β -strands of *T. pallidum* repeat C (TprC) and TprD2 could give rise to functional differences in their porin channels. OMP profiles of some clinical strains were mosaics of different reference strains and did not correlate with results from enhanced molecular typing. Our observations suggest that human host selection pressures drive *T. pallidum* subsp. *pallidum* OMP diversity and that genetic exchange contributes to the evolutionary biology of *T. pallidum* subsp. *pallidum*. They also set the stage for topology-based analysis of antibody responses to OMPs and help frame strategies for syphilis vaccine development.

IMPORTANCE Despite recent progress characterizing outer membrane proteins (OMPs) of *Treponema pallidum*, little is known about how their surface-exposed, β -barrel-forming domains vary among strains circulating within high-risk populations. In this study, sequences for the β -barrel-encoding regions of three OMP loci, *tpnC*, *tpnD*, and *bamA*, in *T. pallidum* subsp. *pallidum* isolates from a large number of patient specimens from geographically disparate sites were examined. Structural

Received 7 May 2018 Accepted 10 May 2018 Published 12 June 2018

Citation Kumar S, Caimano MJ, Anand A, Dey A, Hawley KL, LeDoyt ME, La Vake CJ, Cruz AR, Ramirez LG, Paštěková L, Bezsonova I, Šmajš D, Salazar JC, Radolf JD. 2018. Sequence variation of rare outer membrane protein β -barrel domains in clinical strains provides insights into the evolution of *Treponema pallidum* subsp. *pallidum*, the syphilis spirochete. mBio 9:e01006-18. <https://doi.org/10.1128/mBio.01006-18>.

Editor Steven J. Projan, MedImmune

Copyright © 2018 Kumar et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Justin D. Radolf, JRadolf@uchc.edu.

* Present address: Sanjiv Kumar, Division of Glycoscience, School of Biotechnology, Albanova University Center, Royal Institute of Technology, Stockholm, Sweden; Abhishek Dey, Biological Sciences, University of North Carolina at Charlotte, Charlotte, North Carolina, USA.

S.K. and M.J.C. contributed equally to this article.

This article is a direct contribution from a Fellow of the American Academy of Microbiology. Solicited external reviewers: Jorge Benach, Stony Brook University; Richard Marconi, Virginia Commonwealth University Medical Center.

models predict that sequence variation within β -barrel domains occurs predominantly within predicted extracellular loops. Amino acid substitutions in predicted transmembrane strands that could potentially affect porin channel function were also noted. Our findings suggest that selection pressures exerted within human populations drive *T. pallidum* subsp. *pallidum* OMP diversity and that recombination at OMP loci contributes to the evolutionary biology of syphilis spirochetes. These results also set the stage for topology-based analysis of antibody responses that promote clearance of *T. pallidum* subsp. *pallidum* and frame strategies for vaccine development based upon conserved OMP extracellular loops.

KEYWORDS *Treponema pallidum*, molecular subtyping, outer membrane proteins, spirochetes, syphilis

After years of steady decline during the 1990s, syphilis, a sexually transmitted infection caused by the uncultivable spirochete *Treponema pallidum* subsp. *pallidum*, has undergone a dramatic resurgence in the United States, particularly among men who have sex with men (1). Syphilis also poses a major threat globally, with an estimated 5.6 million new cases annually and 350,000 adverse pregnancy outcomes due to mother-to-child transmission (2). The failure of epidemiological approaches to curtail the spread of syphilis underscores the need for a vaccine capable of inducing protective antibody responses to geographically widespread and genetically diverse *T. pallidum* subsp. *pallidum* strains (3, 4). *T. pallidum* subsp. *pallidum* has been designated “the stealth pathogen” based on its ability to evade innate and adaptive immune responses for protracted periods, permitting repeated bouts of hematogenous dissemination and invasion of numerous organs, including the central nervous system and the fetal-placental barrier during pregnancy (5–7). While the appearance of opsonic antibodies is widely regarded as a turning point in the battle between host and pathogen (8, 9), the targets of antibodies that promote bacterial clearance during syphilitic infection are largely unidentified. How immune pressures in high-risk populations influence the epidemiology of syphilis and the evolutionary biology of *T. pallidum* subsp. *pallidum* is also poorly understood.

Dual-membrane bacteria have evolved a unique class of integral outer membrane protein (OMP) in which antiparallel, amphipathic β -strands circularize to form a closed barrel structure, often creating a central aqueous channel that permits uptake of nutrients and efflux of waste products (10–12). Extracellular loops bridge adjacent transmembrane strands, extending from the OM into the external milieu (13). Protective B-cell determinants reside in the extracellular loops and undergo sequence/antigenic variation to circumvent herd immunity to previously circulating strains (14–18). Extracellular loops also play critical roles in disease pathogenesis by promoting interactions with host cells and tissue components and protecting the bacterium against innate clearance mechanisms, such as complement-mediated lysis and neutrophil engulfment (19–25).

Multiple factors have impeded efforts to identify the syphilis spirochete’s integral OMPs. These include the recalcitrance of *T. pallidum* subsp. *pallidum* to *in vitro* cultivation (26, 27), the fragility of its outer membrane (28, 29), its relatively low abundance of outer membrane-spanning proteins (30, 31), and the lack of strong sequence relatedness between *T. pallidum* subsp. *pallidum* OMPs and well-characterized proteins in Gram-negative OMs (32, 33). To circumvent these obstacles, computational methods were employed to mine the *T. pallidum* subsp. *pallidum* Nichols strain genome for proteins predicted to form OM-associated β -barrels. This bioinformatics approach, combined with a battery of biophysical and cellular localization techniques, including opsonophagocytosis assays, yielded a panel of candidate OMPs (33–37). One of these, TP0326/BamA, is the central component of the molecular machine that chaperones newly exported precursor OMPs from the periplasm into the OM (37, 38). A homology model based on the solved structure of the *Neisseria gonorrhoeae* ortholog (38) predicts that the β -barrel of *T. pallidum* subsp. *pallidum* BamA contains 16 transmembrane

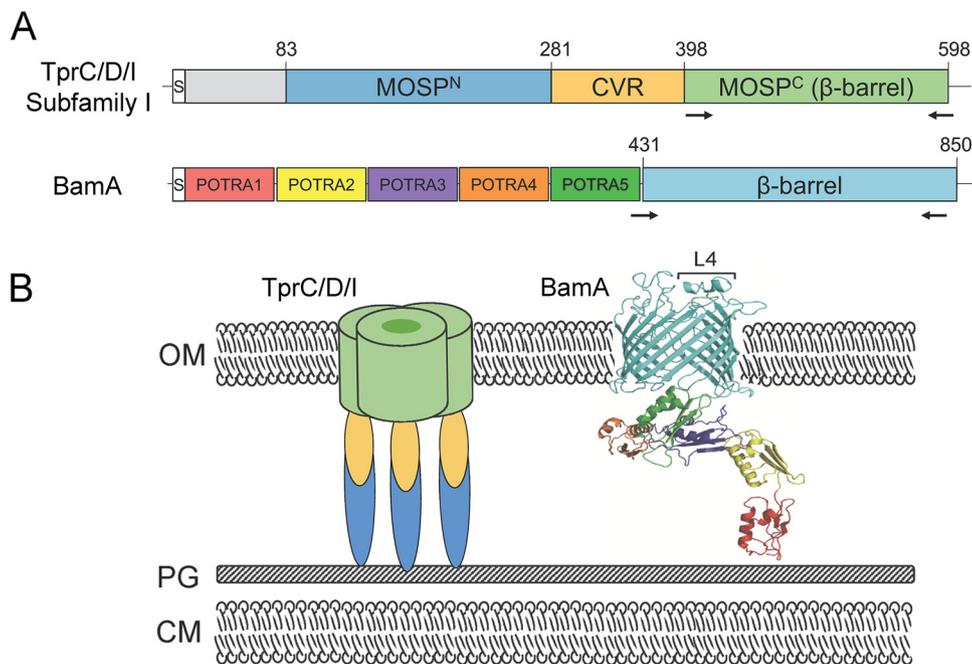


FIG 1 Domain architectures and membrane topologies of TprC (TP0117), TprD (TP0131), TprI (TP0620), and BamA (TP0326). (A) *T. pallidum* repeat (Tpr) subfamily I paralogs TprC (TP0117), TprD (TP0131), and TprI (TP0620) have identical domain architectures (35, 36). MOSP^N and MOSP^C correspond to conserved domains shared with the N and C termini of the major outer sheath protein (MOSP) of *T. denticola*, the parental Tpr ortholog, identified by the NCBI conserved domain database (CDD) server. Arrows indicate the regions that were subjected to PCR amplification for sequencing (see Table S3 for primers). The designation CVR (central variable region) denotes a sequence-variable stretch present in all Tpr orthologs (40, 77). BamA consists of a C-terminal β -barrel and five periplasmic polypeptide transport-associated (POTRA) domains (34, 37). Numbers refer to amino acid positions within the full-length proteins (signal peptides; denoted by "S," included) from *T. pallidum* subsp. *pallidum* Nichols. (B) Membrane topologies. In Tpr orthologs and MOSP, the MOSP^C domain forms the surface-exposed β -barrel (35, 36, 42). Immunofluorescence experiments in *T. pallidum* have confirmed the periplasmic location of MOSP^N and the CVR of TprC/D (Nichols) and TprI (35, 36). Moreover, the periplasmic portions of TprC/D and TprI form extended structures, as determined by small-angle X-ray scattering analysis, that anchor the β -barrels to the peptidoglycan sacculus (35, 36). A homology model based on the solved structure of the *Neisseria gonorrhoeae* ortholog (38) predicts that the β -barrel of *T. pallidum* subsp. *pallidum* BamA contains 16 transmembrane β -strands and 8 extracellular loops (37). OM, outer membrane; PG, peptidoglycan; CM, cytoplasmic membrane; L4, BamA immunodominant extracellular loop 4.

β -strands and eight extracellular loops (37) (Fig. 1). One extracellular loop, L4, was previously shown by the use of human syphilitic serum samples to contain an immunodominant epitope; antibodies against this loop also promoted opsonization of *T. pallidum* subsp. *pallidum* by rabbit peritoneal macrophages (37). A second group of candidate OMPs, *T. pallidum* repeats C and D (TprC/D) (TP0117/TP0131) and TprI (TP0620), consists of members of the paralogous Tpr family (39, 40). Analysis of recombinant TprC/D and TprI suggests that their β -barrel domains (Fig. 1) form aqueous channels in liposomes (35, 36), which is consistent with their potential functions as porins. As with porins from Gram-negative bacteria (11), TprC/D and TprI form trimers, with the β -barrel domains being essential for trimerization (35, 36). A similar bipartite topology with a periplasmic N-terminal major outer sheath protein (MOSP^N) and OM-embedded C-terminal MOSP (MOSP^C) trimeric β -barrel (Fig. 1) also has been demonstrated for the MOSP of the oral commensal *T. denticola*, the parental ortholog for the Tpr family (41, 42). Surface epitope mapping of subfamily I members based on accurate structural models has yet to be performed.

Previously, Centurion-Lara and coworkers (40) performed a detailed analysis of *tpr* genes from four subspecies of pathogenic treponemes, including a small number of *T. pallidum* subsp. *pallidum* isolates. Their predictions regarding sequence variability and immune pressure, however, were based on structural models for TprC/D and TprI that used the full-length polypeptides rather than just the OM-embedded β -barrel-forming

MOSP^C domains. Thus, no study to date has looked at sequence variation within the regions of *T. pallidum* subsp. *pallidum* OMPs known to reside at the host-pathogen interface (i.e., surface-exposed regions) in multiple *T. pallidum* subsp. *pallidum* strains circulating within at-risk populations. Here, we examined the β -barrel-encoding domains of the *tprC*, *tprD*, and *bamA* genes in DNA extracted from *T. pallidum* subsp. *pallidum* within 31 clinical specimens obtained from early syphilis patients in Cali, Colombia (43, 44); San Francisco, CA (SF) (45); and the Czech Republic (CZ) (46, 47). The resulting sequences were compared to the corresponding loci within the 41 *T. pallidum* subsp. *pallidum* reference genomes available from NCBI databases. On the basis of structural models for TprC/D and BamA, much of the sequence variability within all three OMPs is predicted to lie within extracellular loops containing B-cell epitopes. The models also identified amino acid substitutions in predicted transmembrane β -strands for TprC/D and TprD2, which could affect the selectivity of their porin channels. Lastly, OMP profiles of clinical strains at all three loci appear to be mosaics of alleles represented in the *T. pallidum* subsp. *pallidum* reference genomes and did not correlate with results from enhanced molecular typing. The findings presented here are consistent with the notion that selection pressures within human populations drive *T. pallidum* subsp. *pallidum* OMP diversity and that genetic exchange within and between the Nichols and SS14 clades contributes to the evolutionary biology of syphilis spirochetes. These results also set the stage for topology-based analyses of antibody responses that promote clearance of individual *T. pallidum* subsp. *pallidum* strains and, importantly, could be used to develop a broadly protective vaccine based on conserved extracellular loops.

(Portions of this work were presented at the STI & HIV World Congress, Rio de Janeiro, Brazil, 9 to 12 July 2017.)

RESULTS

Patients and clinical samples. Skin biopsy specimens from secondary syphilis rashes were obtained from patients seen in Cali, Colombia; swabs from exudative lesions were obtained from patients with early syphilis in San Francisco, CA (SF), and in Brno and Prague, Czech Republic (CZ). Specimens from Cali and SF were chosen for amplification and sequencing of β -barrel regions based on treponemal burdens determined by *polA* quantitative PCR (qPCR). The CZ specimens selected were PCR positive for all typing loci tested (47), reflecting high *T. pallidum* subsp. *pallidum* burdens. Table S1 in the supplemental material contains a summary of the available demographic information for the samples used in this study.

The *T. pallidum* subsp. *pallidum* strains in the clinical samples belonged to the Nichols and SS14 clades. Prior multilocus and genomic sequence analyses demonstrated that syphilis spirochetes cluster into two taxonomic groups, or clades, arbitrarily named after the Nichols and SS14 reference strains (48–50). Those studies also established that reliable clade designations for *T. pallidum* subsp. *pallidum* could be assigned on the basis of an 83-bp region of *T. pallidum* 0548 (*tp0548*) (48–50), a locus used in epidemiological studies as part of the enhanced *T. pallidum* subsp. *pallidum* typing scheme (51, 52) (see Fig. S1A in the supplemental material). At the outset, *T. pallidum* subsp. *pallidum* sequences in clinical samples from all three study sites were amplified using *tp0548*-specific primers; partial sequences were obtained from 29 samples (Fig. S1B). Among the 14 samples from Cali, 7 contained *tp0548* type f, placing these strains in the SS14 clade, while the remaining 7 contained an assortment of types belonging to the Nichols clade. The six SF strains were a mixture of types within the SS14 clade. Types for seven CZ strains also fell within the SS14 clade, while two contained *T. pallidum* subsp. *pallidum* that belonged to the Nichols clade.

Nucleotide polymorphisms within *tp0558* (encoding a NiCoT family nickel-cobalt inner membrane permease [53]) also can be used for clade discrimination (48, 54). On the basis of the *tp0558* sequences, clade assignments for *T. pallidum* subsp. *pallidum* within 30 clinical samples were determined (Fig. 2). While the majority were exact matches for either clade at all five *tp0558* “discriminator” nucleotide positions, Cali_84,

Strain	Nucleotide position								
	408	414	472	493	494	774	795	834	
Nichols	T	C	T	T	T	G	T	C	Nichols Clade
Cali_101	T	C	T	T	T	G	T	C	
Cali_127	T	C	T	T	T	G	T	C	
Cali_130	T	C	T	T	T	G	T	C	
Cali_145	T	C	T	T	T	G	T	C	
Cali_151	T	C	T	T	T	G	T	C	
Cali_153	T	C	T	T	T	G	T	C	
CZ_177zB	T	C	T	T	T	G	T	C	
CZ_178zB	T	C	T	T	T	G	T	C	
SS14	G	T	C	T	T	G	C	T	SS14 Clade
Cali_84	G	T	C	C	A	G	C	T	
Cali_103	G	T	C	T	T	G	C	T	
Cali_123	G	T	C	T	T	C	C	T	
*Cali_133	G	T	C	C	A	G	C	T	
Cali_143	G	T	C	T	T	G	C	T	
Cali_146	G	T	C	T	T	G	C	T	
Cali_156	G	T	C	T	T	G	C	T	
Cali_164	G	T	C	T	T	G	C	T	
Cali_167	G	T	C	T	T	G	C	T	
SF_6	T	T	C	T	T	G	C	T	
SF_7	G	T	C	T	T	G	C	T	
SF_40	G	T	C	T	T	G	C	T	
SF_46	T	T	C	T	T	G	C	T	
SF_50	T	T	C	T	T	G	C	T	
SF_58	ND	ND	C	T	T	G	C	T	
CZ_190Z	G	T	C	T	T	G	C	T	
CZ_192Z	G	T	C	T	T	G	C	T	
CZ_3218	G	T	C	T	T	G	C	T	
CZ_351	G	T	C	T	T	G	C	T	
CZ_4535	G	T	C	T	T	G	C	T	
CZ_PP1979B	G	T	C	T	T	G	C	T	
CZ_S1120	G	T	C	T	T	G	C	T	

FIG 2 Clade assignments of *T. pallidum* subsp. *pallidum* in clinical samples based upon *tp0558* sequences. Nucleotide polymorphisms in Nichols and SS14 *T. pallidum* subsp. *pallidum* reference and clinical strains are indicated. The five “discriminator” nucleotide positions used for clade assignment (48) are shown in red. Blue and green shading indicate discriminator nucleotides conserved in Nichols and SS14 reference strains, respectively. Gray shading indicates nucleotides conserved in both clades. “ND” indicates nucleotide positions that could not be determined by sequencing of PCR amplicons. All mutations are synonymous except for those at nucleotide positions 493 and 494, which result in the substitution of histidine for phenylalanine (Phe165) in Cali_84 and Cali_133. An asterisk is used to indicate Cali_133, which yielded a discordant (Nichols) clade assignment for *tp0548*.

Cali_123, and Cali_133 contained substitutions not found in any of the *T. pallidum* subsp. *pallidum* reference strains. Of the 27 strains for which we also obtained partial *tp0548* sequences, all but one were concordant at both loci; Cali_133 was Nichols by *tp0548*, and SS14 by *tp0558*.

In summary, based upon *tp0548* and/or *tp0558*, of the 31 clinical samples, *T. pallidum* subsp. *pallidum* strains within 8 clinical samples belonged to the Nichols clade, 21 belonged to the SS14 clade, and 2 were unassignable because of sequence discordance (Cali_133) or incomplete sequence data (Cali_77) (Table 1).

Classification of TprC β -barrel alleles of *T. pallidum* subsp. *pallidum* reference strains. Nucleotide sequence comparisons of full-length *tprC* (*tp0117*) genes harbored by the 41 *T. pallidum* subsp. *pallidum* reference genomes available, examined using *fastx_collapser*, identified five alleles (Table S2). Four of these are represented by

TABLE 1 TprC, TprD/D2, and BamA allele β -barrel variants and enhanced CDC typing of *T. pallidum* subsp. *pallidum* strains in patient specimens from Cali, Colombia; San Francisco; and the Czech Republic

Clade ^a	Patient	β -Barrel variant ^b			ECDCT ^c
		TprC	TprD/D2	BamA	
Nichols	Cali_101	Mexico A/PT_SIF	D2	Mexico A	14d/d
	Cali_127	Mexico A/PT_SIF	D2	Mexico A	21a/d
	Cali_130	Nichols	D	ND	XX/d
	Cali_145	Nichols	D	Nichols (partial)	XX/a
	Cali_151	Nichols	D	ND	XX/a
	Cali_153	Sea81-4	D2	ND	XX/d
	CZ_177zB	Sea81-4	D2	Sea81-4	14X/d
	CZ_178zB	Sea81-4	D2	Sea81-4	14d/d
SS14	Cali_84 ^d	Mexico A/PT_SIF	D2	Nichols	14d/X
	Cali_103	Mexico A/PT_SIF	D2	Mexico A	12d/f
	Cali_123	Mexico A/PT_SIF	D2	Mexico A	14d/f
	Cali_143	Mexico A/PT_SIF	D2	Mexico A	14d/f
	Cali_146	Mexico A/PT_SIF	D2	Mexico A	14b/f
	Cali_156	Mexico A/PT_SIF	D2	SS14 (partial)	XX/f
	Cali_164	Mexico A/PT_SIF	D2	Nichols	XX/f
	Cali_167	Mexico A/PT_SIF	D2	ND	XX/f
	CZ_190Z	Mexico A/PT_SIF	D2	SS14	14d/k
	CZ_192Z	Mexico A/PT_SIF	D2	SS14	14d/g
	CZ_351	Mexico A/PT_SIF	D2	SS14	14d/g
	CZ_3218	Mexico A/PT_SIF	D2	SS14	14e/g
	CZ_4535	Mexico A/PT_SIF	D2	SS14	14d/g
	CZ_PP1979B	Mexico A/PT_SIF	D2	SS14	14d/g
	CZ_S1120	Mexico A/PT_SIF	D2	SS14	14d/g
	SF_6	Mexico A/PT_SIF	D2	Mexico A	14d/f
	SF_7	Mexico A/PT_SIF	D2	Mexico A	14d/g
	SF_40	Mexico A/PT_SIF	D2	Mexico A	15a/e
	SF_46	Mexico A/PT_SIF	D2	Mexico A	14d/g
	SF_50	Mexico A/PT_SIF	D2	Mexico A	14d/f
SF_58	Mexico A/PT_SIF	D2	Mexico A	14d/f	
Nichols/SS14 ^e	Cali_133	Mexico A/PT_SIF	D2	Nichols	14d/y
ND ^f	Cali_77	Mexico A/PT_SIF	ND	Mexico A	16d/X

^aClade designations based on *tp0548* genotypes (Fig. S1A) and *tp0558* sequences (Fig. 2). See Materials and Methods and Table S2 for additional details. The *tp0548* genotype is indicated by the last letter of the ECDCT designation.

^bThe β -barrel-encoding domains of *tprC* from Mexico A and PT_SIF are identical. Nucleotide alignments for *tprC*, *tprD/D2*, and *bamA* β -barrels are presented in Fig. S2, S3, and S4, respectively. See Table S4 for NCBI accession numbers.

^cEnhanced CDC typing was performed as described in Materials and Methods. Strain type designations are as follows: the first numeral represents the number of repeats in the *arp* gene; the first letter represents the MseI restriction site profile in the *tprE/G/J* genes; and the second letter is based on sequence analysis of an 83-bp region of *tp0548*. The letter X is used to indicate missing data for a given position.

^dClade designation based on *tp0558* alone.

^eThe indicated strain carries a unique Nichols-like *tp0548* (genotype y) gene and an SS14 *tp0558* gene. See Fig. S1A for additional details.

^fND, not determined due to limited sample material.

reference strains Nichols, SS14, Mexico A, and Seattle81-4 (Sea81-4) (40, 48, 55), while the fifth is represented by a cohort of 25 clinical strains (designated "PT_SIF") from Lisbon, Portugal (56). The five *tprC* alleles encode four β -barrel variants (Fig. 3A and B); the Mexico A and PT_SIF β -barrel-forming domains are identical (Fig. S2) and therefore are together referred to as Mexico A/PT_SIF. The nucleotide polymorphisms that distinguish the four TprC β -barrel variants are distributed over four regions (Fig. 3A and B; see also Fig. S2). Region I consists of a 56-nucleotide stretch in which seven positions are unique to the Nichols allele. Region II contains the sole nucleotide substitution that differentiates the Mexico A/PT_SIF and SS14 β -barrel alleles. Region III consists of a 19-nucleotide stretch in which seven positions are unique to the Sea81-4 β -barrel allele. Region IV consists of a 49-nucleotide stretch in which seven positions are shared by

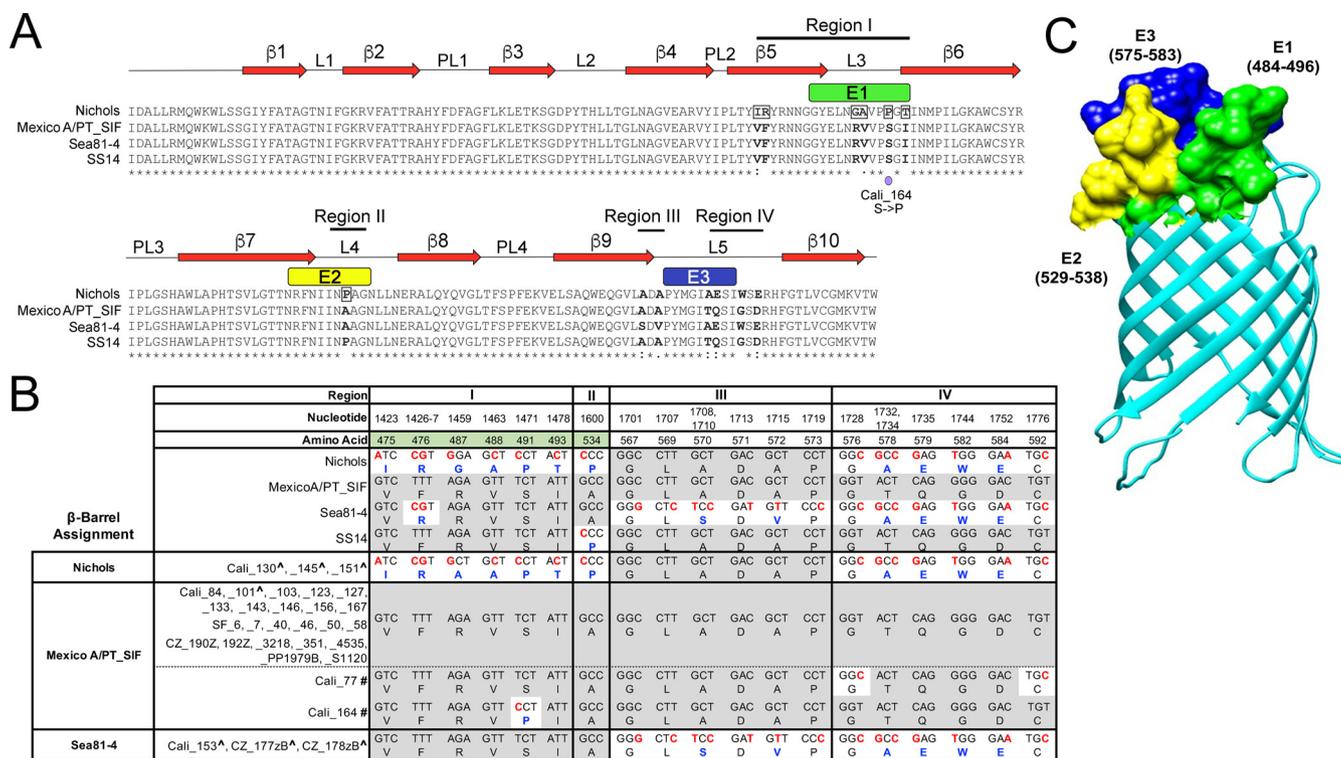


FIG 3 *T. pallidum* subsp. *pallidum* in clinical samples from Cali, San Francisco, and the Czech Republic encodes Nichols, Mexico A/PT_SIF, or Sea81-4 TprC allele β -barrel domains. (A) Multiple-sequence alignment of the four TprC allele β -barrels identified in *T. pallidum* subsp. *pallidum* reference genomes. The Mexico A and PT_SIF TprC allele β -barrel sequences are identical and are referred to here as Mexico A/PT_SIF. Identical amino acids are indicated by asterisks (*); nonsynonymous amino acid substitutions are in bold. Double dots (:) and single dots (·) indicate highly conservative and conservative changes, respectively. Amino acids that were positively selected in the branch-site model using PAML (57) are boxed in the Nichols allele sequence. Numbered red arrows and black lines indicate the positions of predicted β -strands, periplasmic loops (PL), and extracellular loops (L) in the Nichols TprC allele β -barrel structural model shown in panel C. A purple dot indicates a nonsynonymous nucleotide substitution in the Mexico A allele β -barrel variant-encoding region for Cali_164. (B) Nucleotide and amino acid differences, shown in red and blue, respectively, in the TprC allele β -barrels and variants identified in *T. pallidum* subsp. *pallidum* reference genomes and 31 clinical strains, respectively. The nucleotide and amino acid numbers for variable positions are based on full-length *T. pallidum* subsp. *pallidum* Nichols TprC (Fig. S2). Consensus positions are shaded gray. Positively selected amino acids identified by the branch-site model using PAML (57) are shaded green. Carat symbols (^) indicate clinical strains belonging to the Nichols clade (see Table 1). Hash marks (#) designate Mexico A allele TprC β -barrel variants (Cali_77 and Cali_164). (C) The TMBpro webserver (58) was used to generate a three-dimensional structural model for the β -barrel of TprC (Nichols). B-cell epitopes E1 (residues 484 to 496), E2 (residues 529 to 538), and E3 (residues 575 to 583) predicted by DiscoTope 2.0 (59) are shown in green, yellow, and dark blue.

Nichols and Sea81-4. Thirteen of the 22 polymorphisms dispersed across the β -barrel domain result in amino acid substitutions; 6 are nonconservative (Fig. 3B). The “branch-site” model in the phylogenetic analysis by maximum likelihood (PAML) package (57) identified six amino residues in full-length *tprC* as being positively selected ($P = >95\%$); all were in the β -barrel domain (Fig. 3A).

Structural modeling of the TprC β -barrel and topological mapping of the amino acid substitutions that differentiate the four β -barrel variants.

To date, there have been no solved structures for any of the Tpr proteins. As a first step toward understanding the topological and/or functional implications of the sequence data described above, a structural model for the Nichols TprC β -barrel was generated using TMBpro (58). As shown in Fig. 3C, the predicted β -barrel consists of 10 antiparallel β -strands with five connecting extracellular loops of various sizes. Overall, this model is consistent with the general principles of amphipathic β -barrel structure (10, 13) in that the external surface facing the lipid bilayer is highly hydrophobic, while charged residues line the channel (Fig. 4). Note that the strong positive charge within the channel could explain its high conductivity for the fluorophore $Tb(DPA)_3^{3-}$ used in previous studies of MOSP^C domain porin activity (35, 36). Also consistent with general OMP structures are the relatively large extracellular loops and short periplasmic turns. Depending on location, substitutions in the barrel could affect either the properties of the aqueous channel or surface interactions between the treponeme and its obligate

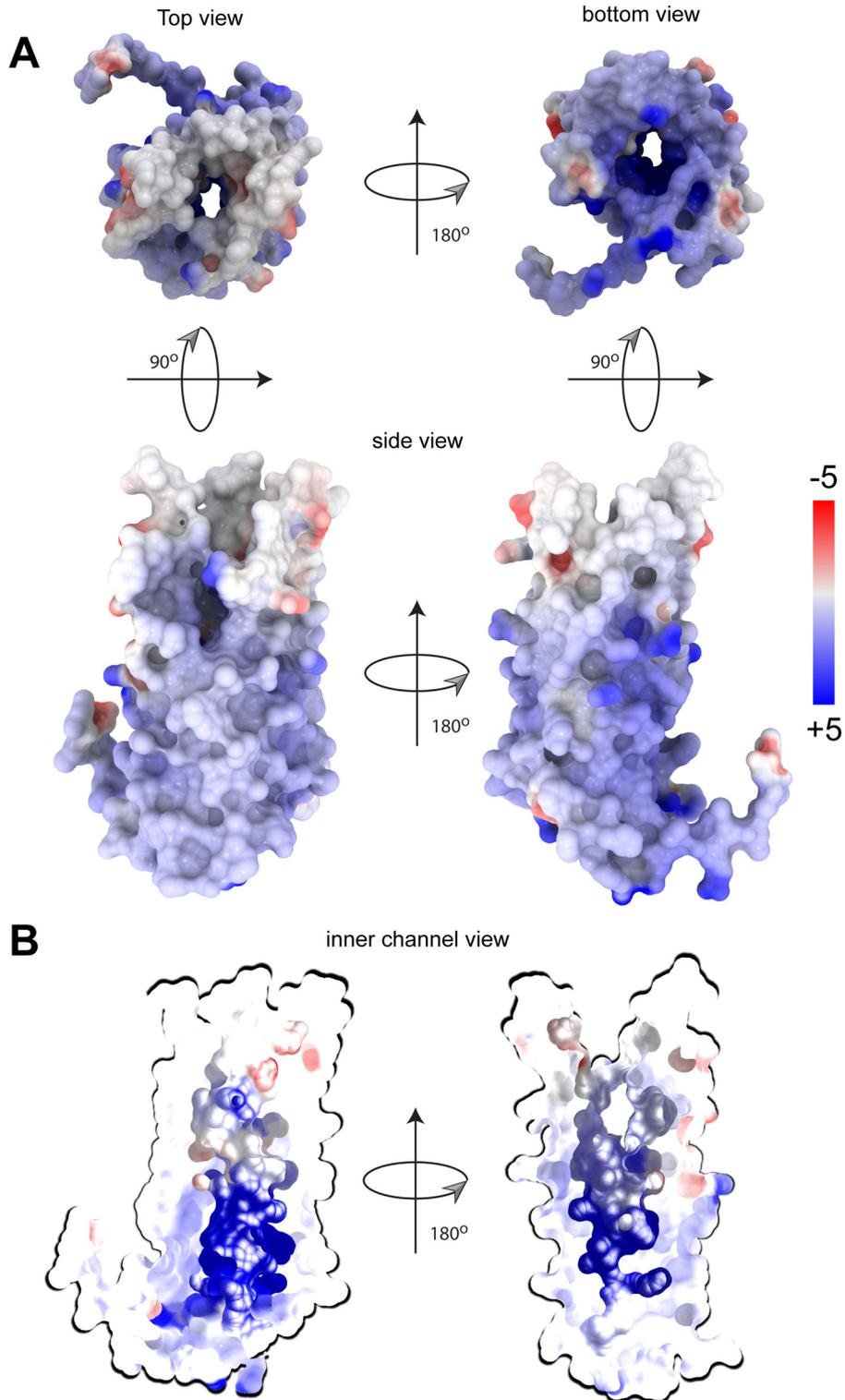


FIG 4 The Nichols TrpC/D allele β -barrel channel interior is highly positively charged. (A) Structural model of the Nichols TrpC/D allele β -barrel shown as a surface colored by electrostatic potential. Positively and negatively charged residues are shown in blue and red, respectively. (B) A slice through the Nichols TrpC/D allele β -barrel porin channel reveals a charged interior pore. The pore is colored according to its charge distribution scale shown on the right. The electrostatic potential display was generated using ICM MolBrowserPro (97).

human host. Nine of the 13 amino acid polymorphisms in the four TprC β -barrel domain variants are in extracellular loops (3 in L3, 1 in L4, and 5 in L5), while 4 are in predicted transmembrane strands (including two positively predicted residues in β 5) (Fig. 3A).

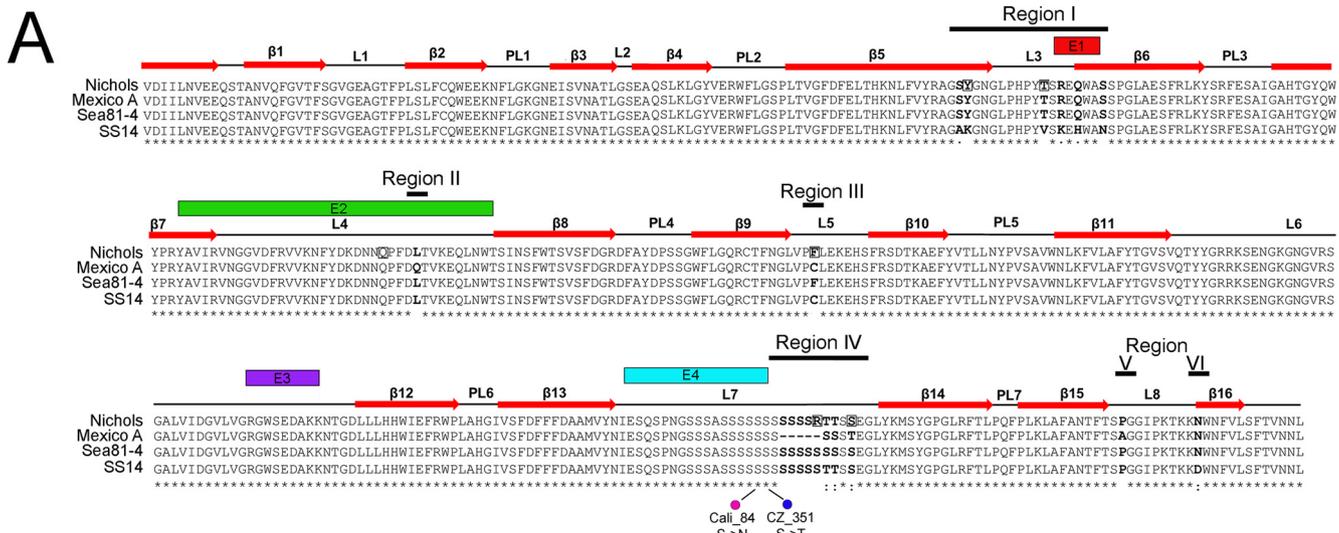
The majority of B-cell epitopes predicted for TprC localize to extracellular loops. Immunodominant B-cell epitopes of Gram-negative OMPs typically are located in extracellular loops and often are sequence variable (14–18). Analysis of the TprC β -barrel structural model using DiscoTope 2.0 (59) revealed that the predicted B-cell epitopes align well with extracellular loops L3, L4, and L5 (Fig. 3C) and closely correspond to variable regions (VRs) I, II, and III/IV, respectively (Fig. 3A). Importantly, each predicted B-cell epitope contains at least one nonconservative substitution (Fig. 3A). Of note, the single amino acid residue (Pro534) that differentiates the Mexico A/PT_SIF and SS14 alleles (region II) occurs in predicted epitope E2 (Fig. 3A).

***T. pallidum* subsp. *pallidum* strains within clinical samples from Cali, San Francisco, and the Czech Republic contain three TprC β -barrel variants, with Mexico A/PT_SIF predominating.** TprC β -barrel sequences were obtained from all 31 clinical samples (Table 1; see also Fig. 3B). Overall, the Mexico A/PT_SIF β -barrel variant predominated (25/31). While Cali_77 and Cali_164 contained nucleotide changes in regions IV and I, respectively, they were designated Mexico A/PT_SIF-like because they contain the “differentiator” G at nucleotide position 1600 (region II). Cali_164 contained a proline in place of a serine at residue 491 in region I (L3/E1; Fig. 3B), thereby creating a novel E1 epitope. The remaining TprC β -barrel variants within the clinical samples were exact matches for either the Nichols or Sea81-4 allele. Notably, of the eight *T. pallidum* subsp. *pallidum* clinical strains assigned to the Nichols clade (indicated by carat symbols in Fig. 3B), only three contained a Nichols β -barrel variant, while all of the clinical strains assigned to the SS14 clade contained a Mexico A/PT_SIF β -barrel variant (Table 1; see also Fig. 3B). Cali_133, the strain with discordant *tp0548* and *tp0558* sequences (Fig. 1A and Fig. 2, respectively), also contained a Mexico A/PT_SF β -barrel variant.

***T. pallidum* subsp. *pallidum* reference strains encode only two TprD β -barrel variants.** Centurion-Lara and coworkers (40, 60) were the first to report that the *tp0131* locus can harbor either a *tprD* allele (which is identical to *tprC* in the Nichols reference strain) or a *tprD2* allele. As noted earlier (40), the *tprD* allele occurs only in reference genomes with a Nichols *tprC* allele (Table S2). Alignment of full-length *tprD* and *tprD2* from the reference strains (Fig. S3) reveals much greater sequence divergence than has been seen with *tprC* (Fig. S2); *tprD* and *tprD2* encode identical MOSP^N domains but divergent central variable regions (CVRs) and MOSP^C domains (Fig. S3). Comparison of the TprD and TprD2 β -barrel-encoding (MOSP^C) domains revealed four regions of variability (Fig. 5A; see also Fig. S3). Region I consists of a single nucleotide change that results in substitution of arginine for lysine. Regions II to IV contain numerous nucleotide polymorphisms, many of which result in nonconservative amino acid substitutions. Region III also contains two single nucleotide indels.

Topological mapping of predicted B-cell epitopes and sequence differences used to distinguish TprD and D2 allele β -barrels. On the basis of the predicted structural model (Fig. 5B), region I lies in the extreme N terminus of the β -barrel. Region II encompasses β -strands β -2 and β -3 and the intervening PL1 periplasmic loop, while regions III and IV center on predicted extracellular loops L3 and L4, respectively. Of the five B-cell epitopes predicted by DiscoTope 2.0, four reside in extracellular loops (Fig. 5A and B). Of note, L2, which coincides with epitope E2, is conserved between TprC, TprD, and TprD2 (Fig. 3A and 5A). Whereas L5 in TprC is variable (Fig. 3A), this loop and the corresponding epitopes (E3 and E5, respectively) are identical in TprD and D2 (Fig. 5A). In addition to amino acid differences in the extracellular loops, the models for TprC/D and TprD2 identified amino acid substitutions at the entrance and exit of the channel (Fig. 5C, highlighted in red) with the potential to affect porin functionality.

TprD2 allele β -barrels predominate in the clinical strains. *T. pallidum* subsp. *pallidum* strains encoding TprD2 allele β -barrels predominated in the 30 clinical sam-



B

Region	I						II		III			IV					V			VI		
	1545	1549	1552, 1554	1579, 1580	1586	1593	1601	1778	1919	2315	2318	2325	Del-Ser			I-Ser	2350	2446	2473			
Amino acid	515	517	518	527	529	531	534	589	640	772	773	775	779	780	781	782	783	784	816	825		
β-barrel Assignment	Nichols	GCG	TCA	TAC	ACG	AGG	CAG	AGT	CTG	TTT	AGC	AGC	AGT	AGT	AGA	ACC	ACC	AGC	TCT	CCC	AAT	
	MexicoA	A	S	Y	T	R	Q	S	CAG	TGT	S	S	S	S	R	T	T	S	ACT	CCG	AAT	
	Sea81-4	GCG	TCA	TAC	ACG	AGG	CAG	AGT	CTG	TTT	AGC	AGC	AGT	AGT	AGT	AGC	AGC	AGC	TCT	CCC	AAT	
	SS14	GCA	GCA	AAA	GTG	AAG	CAC	AAT	CTG	TGT	S	S	S	S	S	T	T	S	TCT	CCC	GAT	
Nichols	Cali_133	GCG	TCA	TAC	ACG	AGG	CAG	AGT	CTG	TTT	AGC	AGC	AGT	AGT	AGA	ACC	ACC	AGC	TCT	CCC	AAT	
	Cali_84#	A	S	Y	T	R	Q	S	L	F	S	S	S	S	R	T	T	S	TCT	CCC	ND	
	Cali_145^A	ND	ND	ND	ND	ND	ND	AGT	CTG	TTT	AGC	AGC	AGT	AGT	AGA	ACC	ACC	AGC	TCT	CCC	ND	
	Cali_164	GCG	TCA	TAC	ACG	AGG	CAG	AGT	CTG	TTT	AGC	AGC	AGT	AGT	AGA	ACC	ACC	AGC	TCT	S	ND	ND
Mexico A	Cali_77_101^A_103_123_127^A_143_146	GCG	TCA	TAC	ACG	AGG	CAG	AGT	CAG	TGT	AGC	AGC	AGC	15 bp deletion				ACT	CCG	AAT		
	SF_6_7_40_46_50_58	A	S	Y	T	R	Q	S	C	C	S	S	S	15 bp deletion				T	P	N		
Sea81-4	CZ_177zB^A_178zB^A#	GCG	TCA	TAC	ACG	AGG	CAG	AGT	CTG	TTT	AGC	AGC	AGT	AGT	AGT	AGC	AGC	AGC	TCT	CCC	AAT	
		A	S	Y	T	R	Q	S	L	F	S	S	S	S	S	S	S	S	TCT	S	ND	ND
SS14	CZ_190Z_192Z_3218_4535_PP1979B_S1120	GCA	GCA	AAA	GTG	AAG	CAC	AAT	CTG	TGT	AGC	AGC	AGT	AGT	AGC	ACC	ACC	AGC	TCT	CCC	GAT	
		A	A	K	V	K	H	N	L	C	S	S	S	S	T	T	T	S	S	P	D	
	CZ_351#	GCA	GCA	AAA	GTG	AAG	CAC	AAT	CTG	TGT	AGC	ACC	AGT	AGT	AGC	ACC	ACC	AGC	TCT	CCC	GAT	
	A	A	K	V	K	H	N	L	C	S	T	S	S	S	T	T	S	TCT	S	P	D	
	Cali_156	ND	ND	ND	ND	ND	ND	ND	CTG	TGT	AGC	AGC	AGT	AGT	AGC	ACC	ACC	AGC	TCT	S	ND	ND

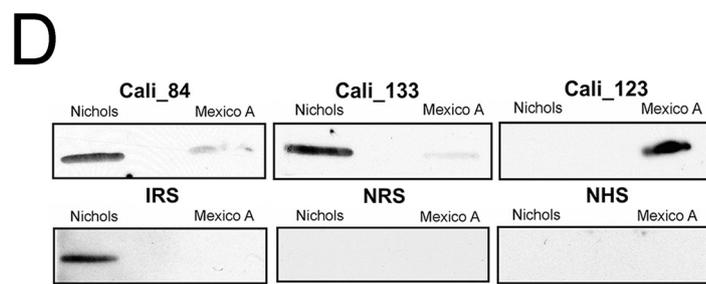
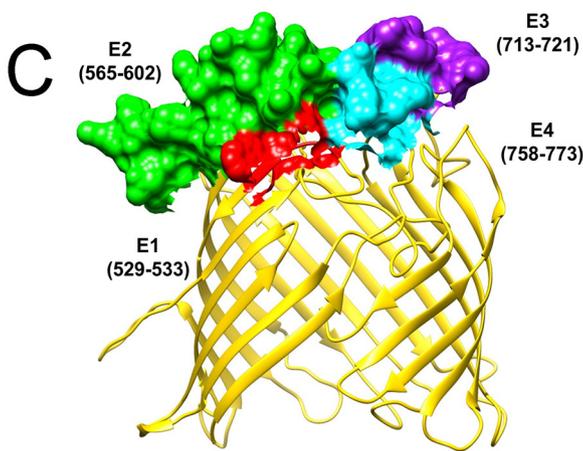


FIG 6 *T. pallidum* subsp. *pallidum* in clinical samples from Cali, San Francisco, and the Czech Republic encode Nichols, Mexico A, Sea81-4, and SS14 BamA allele β -barrel variants. (A) Multiple-sequence alignment of the four BamA allele β -barrels identified in *T. pallidum* subsp. *pallidum* reference genomes. Identical amino acids are indicated by asterisks (*); nonsynonymous amino acid substitutions are in bold. Double dots (:) and single dots (·) indicate highly conservative and conservative changes, respectively. Amino acids positively selected by the site model using PAML (57) are boxed. Numbered red arrows and black lines indicate the positions of β -strands, periplasmic loops (PL), and extracellular loops (L) in the Nichols BamA structural homology model (37) shown in panel B. Magenta

(Continued on next page)

Mexico A allele. Region III also consists of a single nucleotide with a “G” in the Mexico A and SS14 alleles and a “T” in Nichols and Sea81-4 alleles. Region IV, the most polymorphic, consists of either a single nucleotide substitution along with a 15-nucleotide deletion unique to the Mexico A allele or nonidentical stretches of 15 nucleotides in the other three alleles; of note, substitutions in this region modify the polyserine tract (Fig. 6A and B; see also Fig. S4), a unique feature of BamA orthologs in pathogenic treponemes (34, 61, 62). Regions V and VI consist of single nucleotide changes specific to the Mexico A and SS14 alleles, respectively. The “branch-site” model (57) did not identify any positively selected amino acids in the β -barrel region; however, the “site” model identified six positively selected amino acids ($P = >95\%$), distributed over regions I to IV (Fig. 6A and B).

Topological mapping of the amino acid substitutions used to distinguish between the β -barrels encoded by BamA alleles. We previously described and partially validated by immunofluorescence analysis a structural homology model for *T. pallidum* subsp. *pallidum* BamA consisting of 16 transmembrane β -strands and eight extracellular loops (37). Variable regions I through V are located entirely or largely in extracellular loops (Fig. 6A). The DiscoTope 2.0 server predicts that extracellular loops L4, L6, and L7 contain major epitopes (E2, E3, and E4, respectively; Fig. 6A and C). Note that variable region IV in L7 lies outside E4, raising the possibility that immune pressure may not be driving variability in this loop.

The clinical strains encode all four BamA β -barrel variants. β -Barrel sequences matching all four of the BamA reference alleles were identified in 27 clinical samples (Table 1; see also Fig. 6B). However, several varied from their corresponding references (Fig. 6B). Most notably, the polyserine tract of CZ_177zB and CZ_178zB, both of which encoded Sea81-4-like BamA allele β -barrels, contained a 12-nucleotide insertion that adds four additional serine residues (Fig. 6B). *bamA* β -barrel sequences were obtained from 25 *T. pallidum* subsp. *pallidum* clinical strains with clade designations (Table 1). Of the five assigned to the Nichols clade, only Cali_145 contained a Nichols BamA β -barrel variant. CZ_177zB and CZ_178zB contained the Sea81-4 variant, while Cali_101 and Cali_127 contained the Mexico A β -barrels. Of the 20 strains belonging to the SS14 clade, only 8 contained an SS14 β -barrel variant; of the remaining 12, 10 encoded the Mexico A *bamA* β -barrel variant and 2 had the Nichols variant.

A single amino acid substitution in BamA extracellular loop 4 alters immunoreactivity by patient serum samples. As noted above, the nucleotide change in variable region II results in an amino acid substitution (glutamine for leucine at residue 589) unique to the Mexico A allele β -barrel. We recently reported that this polymorphism markedly alters the reactivity of patient serum samples with a heterologous L4 peptide (37). Immunoblotting performed with serum samples from Cali patients infected with *T. pallidum* subsp. *pallidum* strains containing Nichols (Cali_84 and Cali_133) or Mexico A BamA β -barrel variants (Cali_123) confirmed this finding (Fig. 6D).

β -Barrel profiles in clinical strains can be mosaics of reference genome profiles. Clade distributions and “across-the-board” β -barrel profiles of clinical strains were

FIG 6 Legend (Continued)

and green circles indicate nonsynonymous substitutions in the Cali_84 (Nichols) and CZ_351 (SS14) BamA β -barrels, respectively. B-cell epitopes E1 (residues 529 to 533), E2 (residues 565 to 602), E3 (residues 713 to 721), and E4 (residues 758 to 773), predicted by DiscoTope 2.0, are shown in red, green, purple, and cyan boxes, respectively. (B) Nucleotide and amino acid differences, shown in red and blue, respectively, in the BamA allele β -barrels and variants found in *T. pallidum* subsp. *pallidum* reference genomes and 27 clinical strains. Nucleotide and amino acid numbers for variable positions are based on full-length Nichols BamA (Fig. S4). The column designated “Del-Ser” indicates a 15-bp deletion within the Mexico A BamA allele β -barrel. The column designated “I-Ser” indicates a 12-nucleotide insertion in CZ_177zB and CZ_178zB that extends the polyserine tract by 4 residues. Consensus positions are shaded gray. Positively selected amino acids identified by branch-site model are shaded green. “ND” indicates nucleotide and amino acid positions that could not be determined by sequencing of PCR amplicon. A caret symbol (^) indicates clinical strains assigned to the Nichols clade (see Table 1). Hash marks (#) designate BamA β -barrel variants encoded by Cali_84 (Nichols), Cz177zB (Sea81-4), Cz178zB (Sea81-4), and CZ_351 (SS14). (C) Structural homology model for the Nichols BamA allele β -barrel showing the position of B-cell epitopes using the same color scheme as in panel A. (D) A nonconservative amino acid substitution in L4 of BamA markedly alters the reactivity of syphilis patient serum samples. Serum samples from Cali_84 and Cali_133, patients infected with strains containing Nichols BamA β -barrel variants, preferentially recognize Nichols L4 whereas serum from Cali_123, infected with a strain containing a Mexico A allele BamA β -barrel, preferentially recognizes Mexico A L4. Each lane contained ~100 ng of affinity-purified recombinant histidine-tagged protein. NRS, normal rabbit serum; IRS, immune rabbit serum obtained from a rabbit infected with *T. pallidum* subsp. *pallidum* Nichols; NHS, normal human serum.

TABLE 2 Summary of β -barrel allele profiles and molecular types in clinical samples with complete OMP profiles and belonging to either the Nichols or SS14 clade

Clade	No. of samples	<i>tprC</i> β -barrel(s)	<i>tprD/D2</i> β -barrel	<i>bamA</i> β -barrel	Reference genome OMP profile ^a	ECDCT ^b
Nichols	1	Nichols	D	Nichols	Nichols	XX/a (1)
	2	Mexico A/PT_SIF	D2	Mexico A	Mexico A	14d/d (1), 21a/d (1)
	2	Sea81-4	D2	Sea81-4	Sea81-4	14d/d (1), 14X/d (1)
SS14	10	Mexico A/PT_SIF	D2	Mexico A	Mexico A	12d/f (1), 14d/f (5), 14b/f (1), 14d/g (2), 15a/e (1)
	8	Mexico A/PT_SIF	D2	SS14	Amoy	XX/f (1), 14d/k (1), 14d/g (5), 14e/g (1)
	2	Mexico A/PT_SIF	D2	Nichols	Unique	14d/X (1), XX/f (1)

^aData are based on matching the reference genome at all three OMP loci.

^bNumbers in parentheses indicate numbers of clinical samples containing *T. pallidum* subsp. *pallidum* of the corresponding ECDC type. X, undetermined genotype.

compared with those of the *T. pallidum* subsp. *pallidum* reference genomes (Table S2). Four of the six reference strains belonging to the Nichols clade (CDC A, Nichols, Chicago, and DAL-1) had Nichols *tprC*, *tprD*, and *bamA* alleles, while Sea81-4 and UW189B harbored *tprD2* in addition to Sea81-4 *tprC* and *bamA*. Although all members of the SS14 clade harbored *tprD2* alleles, variability at the other two loci, particularly *tprC*, was noted (Table S2). Parsimony analysis (63) of the *T. pallidum* subsp. *pallidum* reference genomes confirmed that the three OMP loci were poor predictors of clade assignment. *tprC*, *tprD*, and *bamA* contained only 25.9%, 0%, and 5.6% parsimony informative sites, respectively, compared to 80.6% for *tp0548* and 100% for *tp0558*.

Sequences were obtained from all three OMP loci in five clinical strains assigned to the Nichols clade (Tables 1 and 2). Of these, only Cali_145 possessed an OMP profile resembling that of the Nichols reference strain. CZ_177zB and CZ_178zB had Sea81-4-like profiles, while Cali_101 and Cali_127 had Mexico A-like profiles. Thus, two of five Nichols clade strains had β -barrel variant profiles matching those of strains of *T. pallidum* subsp. *pallidum* belonging to the SS14 clade. Complete profiles were obtained for 20 strains assigned to the SS14 clade (Tables 1 and 2). Strikingly, although all had TprD2 β -barrel variants, none had profiles matching the SS14 reference strain profile. All had the Mexico A/PT_SIF *tprC* β -barrel variant, while only eight had the SS14 strain BamA allele β -barrel variant. The other BamA β -barrel variants were either Mexico A ($n = 10$) or Nichols ($n = 2$). Stratified by geographic location (Table 1), the uniformity of the SF strains contrasted with the diversity of strains from Cali.

Molecular typing did not correlate with the β -barrel profiles. The enhanced CDC typing (ECDCT) system has been widely used to study the diversity and epidemiology of *T. pallidum* subsp. *pallidum* strains in numerous global locales (64, 65). To determine whether molecular typing is predictive of OMP profiles, ECDCT was completed on *T. pallidum* subsp. *pallidum* strains in 21 clinical samples (Tables 1 and 2). Altogether, 10 different ECDCTs were detected. 14d/g ($n = 7$) and 14d/f ($n = 5$) were the most prevalent and, along with 14d/d ($n = 2$), the only genotypes found at more than one site. Five strains had an ECDCT (14d/f) matching the SS14 reference strain but Mexico A-like OMP profiles; of the 12 strains with Mexico A allele β -barrels in all three OMP loci, none matched the Mexico A reference strain genotype (16d/e). Two ECDCTs (14d/d and 14d/g) were associated with more than one OMP profile. Conversely, several OMP profiles were associated with more than one ECDCT. Collectively, these data indicate that OMP profiles of *T. pallidum* subsp. *pallidum* strains cannot be predicted based on ECDCT data.

DISCUSSION

For decades, syphilologists have sought means to distinguish strains of *T. pallidum* subsp. *pallidum* for epidemiological, pathogenesis-related, and vaccine-related investigations. Serological analyses of live “street strains” by Turner and Hollander in the 1950s (55) yielded evidence for antigenic differences, presumably attributable to surface-exposed epitopes. The molecular typing method introduced by Pillay et al. in 1998 (52), based on numbers of repeats in the *arp* (*tp0433*) gene and sequence polymorphisms in

Tpr subfamily II genes *tprE* (*tp0313*), *tprG* (*tp0317*), and *tprJ* (*tp0621*), was a major advance, although subsequent studies revealed that this system insufficiently distinguishes common *T. pallidum* subsp. *pallidum* strains circulating globally (51, 64). The addition of subtyping based on sequences from *tp0548* markedly improved the discriminatory power of the CDC typing scheme (51, 65). In parallel, Šmajš and coworkers (48, 49) found that *tp0548* is one of several loci that phylogenetically separate *T. pallidum* subsp. *pallidum* strains into two clusters, named for the Nichols and SS14 reference strains, and that *tp0548* subtypes segregate into these two clades. Most recently, phylogenetic analysis of whole-genome sequences of geographically diverse *T. pallidum* subsp. *pallidum* strains has lent strong support to the concept of two clades, with strains belonging to the SS14 clade supplanting those belonging to the Nichols clade (50). Consistent with this notion, the majority of *T. pallidum* subsp. *pallidum* in clinical samples examined in this study grouped with the SS14 clade. They also revealed, however, that strains of *T. pallidum* subsp. *pallidum* within the Nichols clade are still actively circulating in the Eastern and Western hemispheres.

Arora et al. (50) reported that the SS14 clade contains a central dominant haplotype, designated SS14-Ω, which does not include the Mexico A strain (also classified within the SS14 clade). The observation that SS14 clade members in both Cali and San Francisco have a preponderance of Mexico A allele β-barrel at all three OMP loci suggests that Mexico A-like strains are circulating more widely than their analysis suggests. An alternative possibility, which sequence data do not rule out, is that the SS14 clade strains in the study cohort fall within SS14-Ω but some contain Mexico A-like OMP allelic remnants. Exclusion of *bamA* sequences, along with the absence of *tpr* genes in draft genomes used by Arora et al. (50), precludes a direct comparison of their data with those presented here.

Phylogenetic reconstructions, including those used to distinguish *T. pallidum* subsp. *pallidum* clades (48, 50), rely upon genes whose sequence variation recapitulates the vertical evolution of the bacterium (66). OMP-encoding genes typically are excluded from such analyses because they undergo mutations and rearrangements that give rise to phylogenetic trees in conflict with those derived from “reference” genes (66). However, sequence variation of OMPs to enhance environmental fitness, virulence, and immune evasion plays a central role in the evolution and epidemiology of pathogenic bacteria (15–17, 67, 68). Parsimony analysis revealed that the evolutionary histories of *tprC*, *tprD*, and *bamA* diverge greatly from those of genes used for clade differentiation (i.e., *tp0548* and *tp0558*). Inspection of the three OMP genes/proteins in reference and clinical strains explains this dichotomy and provides evidence that OMPs in *T. pallidum* subsp. *pallidum* are subject to host-driven, adaptive mechanisms. Except for a few nucleotides in the N-terminal periplasmic portion of Mexico A *tprC*, the variable regions for *tprC* and *bamA* reside within the surface-exposed, OM-embedded β-barrel, just as one would expect if selection pressures exerted by the host were driving variation.

Sequence variation within the β-barrel-encoding regions of *tprC*, *tprD*, and *bamA* ranged from point mutations with single amino acid substitutions to small stretches of DNA encoding multiple amino acid differences. Gray et al. (69) contended that the variable regions in *tprC* containing multiple amino acid changes are more consistent with small “site-specific” gene conversion events than with accumulated point mutations, although they were uncertain whether the source(s) of the acquired sequences is intra- or intergenomic. Regardless, there is no reason that a recombinatorial mechanism(s) would be limited to *tprC*. Indeed, the mosaic OMP profiles in some clinical strains can best be explained by recombination/conversion of larger DNA fragments within and between clades. It is worth noting that genomic sequencing identified a number of OMP loci in pathogenic treponemes, including BamA in *T. pallidum* subsp. *pallidum*, where recombination appears to have occurred (50, 70–72). The most plausible scenario for intergenomic exchange during human syphilis would be the presence of anogenital ulcers coinfecting with *T. pallidum* subsp. *pallidum* strains containing genetically divergent OMP loci.

According to structural models, the variable regions in *T. pallidum* subsp. *pallidum*

OMP β -barrels coincide with extracellular loops predicted to contain B-cell epitopes. Immunoblot results for L4 of BamA underscore the antigenic impact of even single nonsynonymous amino acid substitutions in a surface-exposed region. Substitutions and antigenic diversity in the predicted extracellular loops for Tpr β -barrels likewise could serve in a similar role in immune evasion. The cumulative effect of the OMP loop variants noted in this study, along with those in TprK (56, 65, 73) and as-yet-uncharacterized *T. pallidum* subsp. *pallidum* OMPs (33, 40), would be to foster strain diversity among *T. pallidum* subsp. *pallidum* strains and thus their ability to persist at the population level. There are numerous examples of virulence-related OMPs in bacterial pathogens playing a role in maintenance of cellular homeostasis and/or outer membrane integrity (19–21, 23, 25, 74–76). One also must consider the possibility, therefore, that extracellular loop variants impact interactions at the host-pathogen interface during syphilitic infection. The hot spot for variation in extracellular loop L7 of BamA, which lies outside the adjacent predicted major epitope (E4), might be one example.

TprD and TprD2 contain identical MOSP^N domains but highly divergent CVRs and β -barrels (40, 69, 77). Thus, although involving only two alleles, generation of diversity at the *tprD* locus was a more complex process than was seen with *tprC* and cannot be explained solely by events shaping the β -barrels. Moreover, since the CVR is periplasmic (36), nonimmunological selection pressures must also have been at work. Comparison of *tprD* and *D2* genes suggests that variation in the β -barrel likely arose from gene conversion events involving relatively small segments of DNA similar to those proposed for *tprC* (69). However, comparison of the variable regions in the *tprC* and *tprD/D2* loci (see Fig. S5 in the supplemental material) brings to light important differences. While separate conversions involving L3 appear to have occurred in the *tprC* and *tprD* genes, two exchanges occurred only in the *tprD* locus: region II, spanning transmembrane strands β -2 and β 3, and region IV, containing L4 plus flanking DNA from strands β 7 and β 8. These differences could affect not only the corresponding predicted extracellular loops of TprD2 but also its porin channel function. Differences in substrate preferences that broaden or enhance the capacity to import water-soluble nutrients across the outer membrane theoretically would be of great benefit to an extreme auxotroph and obligate pathogen such as *T. pallidum* subsp. *pallidum* (27, 77). Additional point mutations, such as those in the TprD2 β -barrels of SF_6 and SF_58, might further fine-tune channel functionality. The current worldwide predominance of strains containing *tprD2* could reflect, at least in part, the greater fitness conferred by the presence of functionally and antigenically distinct TprC and TprD OMPs.

Compared to OMPs of many dual-membrane pathogens (11, 18, 78–81), including the sexually transmitted organisms *Neisseria gonorrhoeae* (82, 83) and *Chlamydia trachomatis* (84, 85), a surprisingly small number of variants for each *T. pallidum* subsp. *pallidum* OMP were found. The high degree of similarity between the β -barrel sequences of *T. pallidum* subsp. *pallidum* in clinical samples and their reference allele counterparts argues that this limited diversity cannot be explained solely by the comparatively small number of *T. pallidum* subsp. *pallidum* genomes sequenced to date. Two possible explanations, which are not mutually exclusive, can be envisioned. One is that *T. pallidum* subsp. *pallidum* OMPs are subject to “uneven” immune pressure due to various microbiologic factors, such as low copy numbers and/or heterogeneous expression within spirochete populations (33, 77), and to variability of antibody responses in persons with different genetic backgrounds. A second possibility is that structural constraints counterbalance immunologic forces promoting loop diversity. The notion of the presence of favored loop sequences that protect the bacterium has been invoked to explain the preservation of sequence types of PorB, the dominant porin of *Neisseria meningitidis*, commonly associated with invasive disease in surveys of endemic and epidemic meningococcal strains (18).

The worldwide resurgence of syphilis has kindled a sense of urgency for vaccine development (2, 4, 86). To be efficacious globally, a syphilis vaccine must target surface-exposed (i.e., antibody-accessible) determinants expressed by geographically

disparate *T. pallidum* subsp. *pallidum* strains. Data presented here make clear that vaccine development based on integral OMPs needs to proceed along two broad fronts. One is to expand the list of candidate vaccinogens through topological and structural characterization of proteins known or predicted to form outer membrane-embedded β -barrels (33), coupled with assessment of opsonophagocytosis activity *ex vivo* and protection in the experimental rabbit model (86). The second is to refine methods for genomic sequencing of *T. pallidum* subsp. *pallidum* strains in clinical samples (56) to catalog sequence diversity among individual OMPs on a global level. From a vaccine standpoint, the objective would be to develop a mono- or multivalent vaccine based on OMP alleles from *T. pallidum* subsp. *pallidum* strains that are circulating in at-risk populations worldwide. Similar vaccine strategies have been proposed for *Borrelia burgdorferi*, the Lyme disease spirochete (87–89). The ability to use conserved extracellular loops (e.g., L2 in TprC/D/D2, L5 in TprD/D2, and L3 in BamA) would circumvent the difficulties associated with expression and purification of full-length OMPs on a mass scale (70).

MATERIALS AND METHODS

Clinical samples. Blood and skin biopsy specimens were obtained during 2009 to 2014 from patients with untreated secondary syphilis identified and referred for enrollment through a previously described network of health care professionals in Cali, Colombia (43, 44), according to protocols approved at Centro Internacional de Entrenamiento e Investigaciones Médicas (CIDEIM). Swabs from primary syphilis lesions were obtained during 2004 to 2007 at the San Francisco Municipal Sexually Transmitted Disease (STD) Clinic according to protocols approved by the University of California, San Francisco, and the Centers for Disease Control and Prevention (45). Swabs from primary and secondary syphilis lesions were obtained during 2012 to 2013 in the Czech Republic at Brno (Department of Dermatovenereology, St. Anne's Faculty Hospital, Masaryk University) and Prague (National Reference Laboratory for the Diagnostics of Syphilis, the National Institute of Public Health) (48) in accordance with protocols approved by the Ethics Committee of the Faculty of Medicine, Masaryk University. Samples and clinical data were deidentified at the sites of origin prior to transmission to UConn Health.

Propagation of *T. pallidum* subsp. *pallidum* and generation of immune rabbit serum (IRS). Animal protocols were approved by the UConn Health Institutional Animal Care and Use Committee under the auspices of Animal Welfare Assurance A347-01. *T. pallidum* subsp. *pallidum* Nichols was propagated by intratesticular inoculation of New Zealand White rabbits (90). To generate IRS, animals were challenged intradermally with 1×10^3 treponemes at each of six dorsal sites ~60 days postinoculation and monitored for 30 days.

Quantitation of treponemal burdens. DNAs were extracted using a DNeasy blood and tissue kit (Qiagen, Valencia, CA) and were eluted in 100 to 200 μ l AE buffer (10 mM Tris-Cl, 0.5 mM EDTA; pH 9.0). DNA concentrations were determined by the use of a NanoDrop instrument (Thermo Fisher, Pittsburgh, PA) or by analysis of absorbance at 260/280 nm. qPCR of *poA* (*tp0105*) was performed as described previously (91).

Molecular typing. Subtyping based on *arp* repeats and MseI polymorphisms in *tprE*, *tprG*, and *tprJ* was performed as described previously (52). Strain typing was based upon sequence variability in *tp0548* as described previously by Marra et al. (51) using primers listed in Table S4 in the supplemental material. Strain typing was based upon sequence variability in *tp0548* partial sequences as described previously.

Nested PCR and sequencing of the β -barrel-encoding regions of *tp0558*, *tprC* (*tp0117*), *tprD* (*tp0131*), and *bamA* (*tp0326*). Table S3 lists unpublished primers used in this study. Nested PCR of *tp0558* was performed as described previously (48). First-round amplifications of the *tprC* and *tprD* β -barrel-encoding regions were performed using GoTaq Flexi DNA polymerase (Promega, Madison, WI) according to the manufacturer's instructions. The resulting amplicons were subjected to gel purification using a QIAquick gel extraction kit (Qiagen) and then reamplified using internal (second-round) primers and *Ex Taq* (Clontech, Mountain View, CA). Nested PCR was carried out for *bamA* using GoTaq Flexi according to the manufacturer's instructions. Second-round amplicons for *tprC*, *tprD*, and *bamA* were subjected to gel purification and sequenced in both forward and reverse orientations. For patients Cali_84 and Cali_133, *bamA* second-round amplicons were also cloned into pCR2.1-TOPO vector (Invitrogen) according to the manufacturer's instructions and 10 individual clones were sequenced.

Cloning, expression, and purification of BamA L4 loops. Cloning, expression, and purification of the L4 loops of *T. pallidum* subsp. *pallidum* Nichols and Mexico A were described previously (37).

SDS-PAGE and immunoblotting. Recombinant His-tagged proteins were resolved by the use of AnykD Mini-Protean TGX gels (Bio-Rad) and transferred to nylon-supported nitrocellulose. Membranes were blocked and then probed overnight at 4°C with normal or immune rabbit serum at a dilution of 1:500 or with normal human or human syphilitic (Cali_84, Cali_123, and Cali_133) serum samples at a dilution of 1:250. Bound antibody was detected with horseradish peroxidase (HRP)-conjugated goat anti-rabbit antibody (Southern Biotech, Birmingham, AL) or HRP-conjugated goat anti-human IgG antibody (Pierce, Rockford, IL) at a dilution of 1:30,000. Immunoblots were developed using SuperSignal West Pico chemiluminescent substrate (Thermo Fisher Scientific, Waltham, MA).

Phylogenetic analysis. Sequence alignments were performed using MacVector v.16.0.8 (Apex, NC). Multiple-sequence alignments were performed using either MacVector or fastx_collapser from the Fastx-toolkit (version 0.0.14) (http://hannonlab.cshl.edu/fastx_toolkit/). Genes evolving under conditions of positive selection were identified with the maximum likelihood method (92) implemented in PAML version 4 (93) and its user interface PAMLX (94). Site models of PAML allow the ratio of nonsynonymous/synonymous mutations (ω) to vary in each codon (site) in the gene. Branch-site models search for positive selection in lineages where different rates of ω may occur (95). Two site models and one branch-site model of PAML were used.

Structural modeling and epitope prediction. Algorithms from the TMBpro web server (58) were used to generate the three-dimensional structures for the β -barrels of TprC (Nichols) and TprD2. The structural model of BamA (Nichols) was generated as described by Luthra et al. (37) using the solved structure of full-length BamA (PDB ID: 3KGP) from *Neisseria gonorrhoeae* as the template and the ModWeb server (<https://modbase.compbio.ucsf.edu/modweb/>). Discontinuous epitopes were predicted using the DiscoTope 2.0 server (59), and the calculated epitopes were projected onto their respective structural models using Chimera (96). The electrostatic potential display was generated using ICM MolBrowserPro (97).

Accession number(s). Table S4 contains GenBank accession numbers for the *tp0558*, *tprC*, *tprD/D2*, and *bamA* β -barrel sequences.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.01006-18>.

TEXT S1, DOCX file, 0.01 MB.

FIG S1, TIF file, 1 MB.

FIG S2, DOCX file, 0.1 MB.

FIG S3, DOCX file, 0.1 MB.

FIG S4, DOCX file, 0.1 MB.

FIG S5, TIF file, 0.5 MB.

TABLE S1, DOCX file, 0.02 MB.

TABLE S2, DOCX file, 0.04 MB.

TABLE S3, DOCX file, 0.02 MB.

TABLE S4, DOCX file, 0.02 MB.

ACKNOWLEDGMENTS

We express our gratitude to Susan Philip, San Francisco City Clinic, Population Health Division, San Francisco Department of Public Health, and Allan Pillay, Centers for Disease Control and Prevention, for providing extracted DNAs from SF patients. We are indebted to the unwavering enthusiasm and support of Nancy Saravia; CIDEIM staff members Maria Alejandra Castrillón, Juan Pablo Garcés, Luisa Rubiano, and Laura Potes; the city of Cali Secretariat of Public Health and the health care providers from Cali's public health hospital network who participated in patient recruitment.

This work was partially supported by NIH grants R01 AI26756 (J.D.R.) and R03 TW009172 (A.R.C. and J.C.S.); by Colciencias grant number 222956933229 (A.R.C.); by Grantová Agentura České Republiky (GACR) GA17-25455S (D.S.); by Ministry of Health, Czech Republic 17-31333A (D.S.), and by the generous support of the Department of Research, Connecticut Children's Children Medical Center.

We report no commercial or other associations that might pose a conflict of interest.

REFERENCES

- Patton ME, Su JR, Nelson R, Weinstock H, Centers for Disease Control and Prevention. 2014. Primary and secondary syphilis—United States, 2005–2013. *MMWR Morb Mortal Wkly Rep* 63:402–406.
- Peeling RW, Mabey D, Kamb ML, Chen XS, Radolf JD, Benzaken AS. 2017. Syphilis. *Nat Rev Dis Prim* 3:17073. <https://doi.org/10.1038/nrdp.2017.73>.
- Gottlieb SL, Deal CD, Giersing B, Rees H, Bolan G, Johnston C, Timms P, Gray-Owen SD, Jerse AE, Cameron CE, Moorthy VS, Kiarie J, Broutet N. 2016. The global roadmap for advancing development of vaccines against sexually transmitted infections: update and next steps. *Vaccine* 34:2939–2947. <https://doi.org/10.1016/j.vaccine.2016.03.111>.
- Lithgow KV, Cameron CE. 2017. Vaccine development for syphilis. *Expert Rev Vaccines* 16:37–44. <https://doi.org/10.1080/14760584.2016.1203262>.
- Salazar JC, Hazlett KR, Radolf JD. 2002. The immune response to infection with *Treponema pallidum*, the stealth pathogen. *Microbes Infect* 4:1133–1140. [https://doi.org/10.1016/S1286-4579\(02\)01638-6](https://doi.org/10.1016/S1286-4579(02)01638-6).
- Lafond RE, Lukehart SA. 2006. Biological basis for syphilis. *Clin Microbiol Rev* 19:29–49. <https://doi.org/10.1128/CMR.19.1.29-49.2006>.
- Radolf JD, Tramont EC, Salazar JC. 2014. Syphilis (*Treponema pallidum*), p 2684–2709. In Bennett JE, Dolin R, Blaser MJ (ed), *Mandell, Douglas and Bennett's principles and practice of infectious diseases*, 8th ed. Churchill Livingstone Elsevier, Philadelphia, PA.
- Lukehart SA. 2008. Scientific monogamy: thirty years dancing with the same bug: 2007 Thomas Parran Award Lecture. *Sex Transm Dis* 35:2–7. <https://doi.org/10.1097/OLQ.0b013e318162c4f2>.
- Hawley KL, Cruz AR, Benjamin SJ, La Vake CJ, Cervantes JL, LeDoyt M, Ramirez LG, Mandich D, Fiel-Gan M, Caimano MJ, Radolf JD, Salazar JC.

2017. IFN γ enhances CD64-potentiated phagocytosis of *Treponema pallidum* opsonized with human syphilitic serum by human macrophages. *Front Immunol* 8:1227. <https://doi.org/10.3389/fimmu.2017.011227>.
10. Koebnik R, Locher KP, Van Gelder P. 2000. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol* 37:239–253. <https://doi.org/10.1046/j.1365-2958.2000.01983.x>.
 11. Nikaido H. 2003. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* 67:593–656. <https://doi.org/10.1128/MMBR.67.4.593-656.2003>.
 12. Chevalier S, Bouffartigues E, Bodilis J, Maillot O, Lesouhaitier O, Feuilloley MGJ, Orange N, Dufour A, Cornelis P. 2017. Structure, function and regulation of *Pseudomonas aeruginosa* porins. *FEMS Microbiol Rev* 41: 698–722. <https://doi.org/10.1093/femsre/fux020>.
 13. Wimley WC. 2003. The versatile β -barrel membrane protein. *Curr Opin Struct Biol* 13:404–411. [https://doi.org/10.1016/S0959-440X\(03\)00099-X](https://doi.org/10.1016/S0959-440X(03)00099-X).
 14. Hobbs MM, Alcorn TM, Davis RH, Fischer W, Thomas JC, Martin I, Ison C, Sparling PF, Cohen MS. 1999. Molecular typing of *Neisseria gonorrhoeae* causing repeated infections: evolution of porin during passage within a community. *J Infect Dis* 179:371–381. <https://doi.org/10.1086/314608>.
 15. Haake DA, Suchard MA, Kelley MM, Dundoo M, Alt DP, Zuerner RL. 2004. Molecular evolution and mosaicism of leptospiral outer membrane proteins involves horizontal DNA transfer. *J Bacteriol* 186:2818–2828. <https://doi.org/10.1128/JB.186.9.2818-2828.2004>.
 16. Cody AJ, Maiden MJ, Dingle KE. 2009. Genetic diversity and stability of the *porA* allele as a genetic marker in human *Campylobacter* infection. *Microbiology* 155:4145–4154. <https://doi.org/10.1099/mic.0.031047-0>.
 17. Stenkova AM, Isaeva MP, Shubin FN, Rasskazov VA, Rakin AV. 2011. Trends of the major porin gene (*ompF*) evolution: insight from the genus *Yersinia*. *PLoS One* 6:e20546. <https://doi.org/10.1371/journal.pone.0020546>.
 18. Matthias KA, Strader MB, Nawar HF, Gao YS, Lee J, Patel DS, Im W, Bash MC. 2017. Heterogeneity in non-epitope loop sequence and outer membrane protein complexes alters antibody binding to the major porin protein PorB in serogroup B *Neisseria meningitidis*. *Mol Microbiol* 105: 934–953. <https://doi.org/10.1111/mmi.13747>.
 19. Lin J, Huang S, Zhang Q. 2002. Outer membrane proteins: key players for bacterial adaptation in host niches. *Microbes Infect* 4:325–331. [https://doi.org/10.1016/S1286-4579\(02\)01545-9](https://doi.org/10.1016/S1286-4579(02)01545-9).
 20. McClean S. 2012. Eight stranded beta-barrel and related outer membrane proteins: role in bacterial pathogenesis. *Protein Pept Lett* 19: 1013–1025.
 21. Mikula KM, Kolodziejczyk R, Goldman A. 2012. *Yersinia* infection tools-characterization of structure and function of adhesins. *Front Cell Infect Microbiol* 2:169. <https://doi.org/10.3389/fcimb.2012.00169>.
 22. Confer AW, Ayalew S. 2013. The OmpA family of proteins: roles in bacterial pathogenesis and immunity. *Vet Microbiol* 163:207–222. <https://doi.org/10.1016/j.vetmic.2012.08.019>.
 23. Haake DA, Zückert WR. 2015. The leptospiral outer membrane. *Curr Top Microbiol Immunol* 387:187–221. https://doi.org/10.1007/978-3-662-45059-8_8.
 24. Seidman D, Hebert KS, Truchan HK, Miller DP, Tegels BK, Marconi RT, Carlyon JA. 2015. Essential domains of *Anaplasma phagocytophilum* invasins utilized to infect mammalian host cells. *PLoS Pathog* 11: e1004669. <https://doi.org/10.1371/journal.ppat.1004669>.
 25. Wu Z, Periaswamy B, Sahin O, Yaeger M, Plummer P, Zhai W, Shen Z, Dai L, Chen SL, Zhang Q. 2016. Point mutations in the major outer membrane protein drive hypervirulence of a rapidly expanding clone of *Campylobacter jejuni*. *Proc Natl Acad Sci U S A* 113:10690–10695. <https://doi.org/10.1073/pnas.1605869113>.
 26. Cox DL. 1994. Culture of *Treponema pallidum*. *Methods Enzymol* 236: 390–405. [https://doi.org/10.1016/0076-6879\(94\)36029-4](https://doi.org/10.1016/0076-6879(94)36029-4).
 27. Norris SJ, Cox DL, Weinstock GM. 2001. Biology of *Treponema pallidum*: correlation of functional activities with genome sequence data. *J Mol Microbiol Biotechnol* 3:37–62.
 28. Penn CW, Cockayne A, Bailey MJ. 1985. The outer membrane of *Treponema pallidum*: biological significance and biochemical properties. *J Gen Microbiol* 131:2349–2357. <https://doi.org/10.1099/00221287-131-9-2349>.
 29. Cox DL, Chang P, McDowall AW, Radolf JD. 1992. The outer membrane, not a coat of host proteins, limits antigenicity of virulent *Treponema pallidum*. *Infect Immun* 60:1076–1083.
 30. Radolf JD, Norgard MV, Schulz WW. 1989. Outer membrane ultrastructure explains the limited antigenicity of virulent *Treponema pallidum*. *Proc Natl Acad Sci U S A* 86:2051–2055. <https://doi.org/10.1073/pnas.86.6.2051>.
 31. Walker EM, Zampighi GA, Blanco DR, Miller JN, Lovett MA. 1989. Demonstration of rare protein in the outer membrane of *Treponema pallidum* subsp. *pallidum* by freeze-fracture analysis. *J Bacteriol* 171:5005–5011. <https://doi.org/10.1128/jb.171.9.5005-5011.1989>.
 32. Fraser CM, Norris SJ, Weinstock GM, White O, Sutton GG, Dodson R, Gwinn M, Hickey EK, Clayton R, Ketchum KA, Sodergren E, Hardham JM, McLeod MP, Salzberg S, Peterson J, Khalak H, Richardson D, Howell JK, Chidambaram M, Utterback T, McDonald L, Artiach P, Bowman C, Cotton MD, Fujii C, Garland S, Hatch B, Horst K, Roberts K, Sandusky M, Weidman J, Smith HO, Venter JC. 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* 281:375–388. <https://doi.org/10.1126/science.281.5375.375>.
 33. Radolf JD, Kumar S. 29 August 2017. The *Treponema pallidum* outer membrane. *Curr Top Microbiol Immunol* https://doi.org/10.1007/82_2017_44.
 34. Desrosiers DC, Anand A, Luthra A, Dunham-Ems SM, LeDoyt M, Cummings MA, Eshghi A, Cameron CE, Cruz AR, Salazar JC, Caimano MJ, Radolf JD. 2011. TP0326, a *Treponema pallidum* β -barrel assembly machinery A (BamA) orthologue and rare outer membrane protein. *Mol Microbiol* 80: 1496–1515. <https://doi.org/10.1111/j.1365-2958.2011.07662.x>.
 35. Anand A, Luthra A, Dunham-Ems S, Caimano MJ, Karanian C, LeDoyt M, Cruz AR, Salazar JC, Radolf JD. 2012. TprC/D (Tp0117/131), a trimeric, pore-forming rare outer membrane protein of *Treponema pallidum*, has a bipartite domain structure. *J Bacteriol* 194:2321–2333. <https://doi.org/10.1128/JB.00101-12>.
 36. Anand A, LeDoyt M, Karanian C, Luthra A, Koszelak-Rosenblum M, Malkowski MG, Puthenveetil R, Vinogradova O, Radolf JD. 2015. Bipartite topology of *Treponema pallidum* repeat proteins C/D and I: outer membrane insertion, trimerization, and porin function require a C-terminal β -barrel domain. *J Biol Chem* 290:12313–12331. <https://doi.org/10.1074/jbc.M114.629188>.
 37. Luthra A, Anand A, Hawley KL, LeDoyt M, La Vake CJ, Caimano MJ, Cruz AR, Salazar JC, Radolf JD. 2015. A homology model reveals novel structural features and an immunodominant surface loop/opsonic target in the *Treponema pallidum* BamA ortholog TP_0326. *J Bacteriol* 197: 1906–1920. <https://doi.org/10.1128/JB.00086-15>.
 38. Noinaj N, Kuszak AJ, Gumbart JC, Lukacik P, Chang H, Easley NC, Lithgow T, Buchanan SK. 2013. Structural insight into the biogenesis of β -barrel membrane proteins. *Nature* 501:385–390. <https://doi.org/10.1038/nature12521>.
 39. Centurion-Lara A, Castro C, Barrett L, Cameron C, Mostowfi M, Van Voorhis WC, Lukehart SA. 1999. *Treponema pallidum* major sheath protein homologue TprK is a target of opsonic antibody and the protective immune response. *J Exp Med* 189:647–656. <https://doi.org/10.1084/jem.189.4.647>.
 40. Centurion-Lara A, Giacani L, Godornes C, Molini BJ, Brinck Reid T, Lukehart SA. 2013. Fine analysis of genetic diversity of the *tpr* gene family among treponemal species, subspecies and strains. *PLoS Negl Trop Dis* 7:e2222. <https://doi.org/10.1371/journal.pntd.0002222>.
 41. Anand A, Luthra A, Edmond ME, LeDoyt M, Caimano MJ, Radolf JD. 2013. The major outer sheath protein (Msp) of *Treponema denticola* has a bipartite domain architecture and exists as periplasmic and outer membrane-spanning conformers. *J Bacteriol* 195:2060–2071. <https://doi.org/10.1128/JB.00078-13>.
 42. Puthenveetil R, Kumar S, Caimano MJ, Dey A, Anand A, Vinogradova O, Radolf JD. 2017. The major outer sheath protein forms distinct conformers and multimeric complexes in the outer membrane and periplasm of *Treponema denticola*. *Sci Rep* 7:13260. <https://doi.org/10.1038/s41598-017-13550-6>.
 43. Cruz AR, Pillay A, Zuluaga AV, Ramirez LG, Duque JE, Aristizabal GE, Fiel-Gan MD, Jaramillo R, Trujillo R, Valencia C, Jagodzinski L, Cox DL, Radolf JD, Salazar JC. 2010. Secondary syphilis in Cali, Colombia: new concepts in disease pathogenesis. *PLoS Negl Trop Dis* 4:e690. <https://doi.org/10.1371/journal.pntd.0000690>.
 44. Cruz AR, Ramirez LG, Zuluaga AV, Pillay A, Abreu C, Valencia CA, La Vake C, Cervantes JL, Dunham-Ems S, Cartun R, Mavilio D, Radolf JD, Salazar JC. 2012. Immune evasion and recognition of the syphilis spirochete in blood and skin of secondary syphilis patients: two immunologically distinct compartments. *PLoS Negl Trop Dis* 6:e1717. <https://doi.org/10.1371/journal.pntd.0001717>.
 45. Katz KA, Pillay A, Ahrens K, Kohn RP, Hermanstynne K, Bernstein KT, Ballard RC, Klausner JD. 2010. Molecular epidemiology of syphilis—San Francisco, 2004–2007. *Sex Transm Dis* 37:660–663. <https://doi.org/10.1097/OLQ.0b013e3181e1a77a>.

46. Flasarová M, Pospíšilová P, Mikalová L, Vališová Z, Dastychová E, Strnadel R, Kuklová I, Woznicová V, Zákoucká H, Šmajš D. 2012. Sequencing-based molecular typing of *Treponema pallidum* strains in the Czech Republic: all identified genotypes are related to the sequence of the SS14 strain. *Acta Derm Venereol* 92:669–674. <https://doi.org/10.2340/00015555-1335>.
47. Grillová L, Pětrošová H, Mikalová L, Strnadel R, Dastychová E, Kuklová I, Kojanová M, Kreidlová M, Vaňousová D, Hercogová J, Procházka P, Zákoucká H, Krchňáková A, Vašků V, Šmajš D. 2014. Molecular typing of *Treponema pallidum* in the Czech Republic during 2011 to 2013: increased prevalence of identified genotypes and of isolates with macrolide resistance. *J Clin Microbiol* 52:3693–3700. <https://doi.org/10.1128/JCM.01292-14>.
48. Nechvátal L, Pětrošová H, Grillová L, Pospíšilová P, Mikalová L, Strnadel R, Kuklová I, Kojanová M, Kreidlová M, Vaňousová D, Procházka P, Zákoucká H, Krchňáková A, Šmajš D. 2014. Syphilis-causing strains belong to separate SS14-like or Nichols-like groups as defined by multilocus analysis of 19 *Treponema pallidum* strains. *Int J Med Microbiol* 304:645–653. <https://doi.org/10.1016/j.ijmm.2014.04.007>.
49. Mikalová L, Strouhal M, Čejková D, Zobaníková M, Pospíšilová P, Norris SJ, Sodergren E, Weinstock GM, Šmajš D. 2010. Genome analysis of *Treponema pallidum* subsp. *pallidum* and subsp. *pertenue* strains: most of the genetic differences are localized in six regions. *PLoS One* 5:e15713. <https://doi.org/10.1371/journal.pone.0015713>.
50. Arora N, Schuenemann VJ, Jäger G, Peltzer A, Seitz A, Herbig A, Strouhal M, Grillová L, Sánchez-Busó L, Kühnert D, Bos KI, Davis LR, Mikalová L, Bruisten S, Komericki P, French P, Grant PR, Pando MA, Vaulet LG, Fermepin MR, Martínez A, Centurion Lara A, Giacani L, Norris SJ, Šmajš D, Boshard PP, González-Candelas F, Nieselt K, Krause J, Bagheri HC. 2016. Origin of modern syphilis and emergence of a pandemic *Treponema pallidum* cluster. *Nat Microbiol* 2:16245. <https://doi.org/10.1038/nmicrobiol.2016.245>.
51. Marra C, Sahi S, Tantaló L, Godornes C, Reid T, Behets F, Rompalo A, Klausner JD, Yin Y, Mulcahy F, Golden MR, Centurion-Lara A, Lukehart SA. 2010. Enhanced molecular typing of *Treponema pallidum*: geographical distribution of strain types and association with neurosyphilis. *J Infect Dis* 202:1380–1388. <https://doi.org/10.1093/infdis/jiq138>.
52. Pillay A, Liu H, Chen CY, Holloway B, Sturm AW, Steiner B, Morse SA. 1998. Molecular subtyping of *Treponema pallidum* subspecies *pallidum*. *Sex Transm Dis* 25:408–414. <https://doi.org/10.1097/00007435-199809000-00004>.
53. Eitinger T, Suhr J, Moore L, Smith JA. 2005. Secondary transporters for nickel and cobalt ions: theme and variations. *Biometals* 18:399–405. <https://doi.org/10.1007/s10534-005-3714-x>.
54. Gallo Vaulet L, Grillová L, Mikalová L, Casco R, Rodríguez Fermepin M, Pando MA, Šmajš D. 2017. Molecular typing of *Treponema pallidum* isolates from Buenos Aires, Argentina: frequent Nichols-like isolates and low levels of macrolide resistance. *PLoS One* 12:e0172905. <https://doi.org/10.1371/journal.pone.0172905>.
55. Turner TB, Hollander DH. 1957. Biology of the treponematoses. World Health Organization, Geneva, Switzerland.
56. Pinto M, Borges V, Antelo M, Pinheiro M, Nunes A, Azevedo J, Borrego MJ, Mendonça J, Carpinteiro D, Vieira L, Gomes JP. 2016. Genome-scale analysis of the non-cultivable *Treponema pallidum* reveals extensive within-patient genetic variation. *Nat Microbiol* 2:16190. <https://doi.org/10.1038/nmicrobiol.2016.190>.
57. Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* 22:2472–2479. <https://doi.org/10.1093/molbev/msi237>.
58. Randall A, Cheng J, Sweredoski M, Baldi P. 2008. TMBpro: secondary structure, β -contact and tertiary structure prediction of transmembrane beta-barrel proteins. *Bioinformatics* 24:513–520. <https://doi.org/10.1093/bioinformatics/btm548>.
59. Kringelum JV, Lundegaard C, Lund O, Nielsen M. 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol* 8:e1002829. <https://doi.org/10.1371/journal.pcbi.1002829>.
60. Centurion-Lara A, Sun ES, Barrett LK, Castro C, Lukehart SA, Van Voorhis WC. 2000. Multiple alleles of *Treponema pallidum* repeat gene D in *Treponema pallidum* isolates. *J Bacteriol* 182:2332–2335. <https://doi.org/10.1128/JB.182.8.2332-2335.2000>.
61. Cameron CE, Lukehart SA, Castro C, Molini B, Godornes C, Van Voorhis WC. 2000. Opsonic potential, protective capacity, and sequence conservation of the *Treponema pallidum* subspecies *pallidum* Tp92. *J Infect Dis* 181:1401–1413. <https://doi.org/10.1086/315399>.
62. Pětrošová H, Zobaníková M, Čejková D, Mikalová L, Pospíšilová P, Strouhal M, Chen L, Qin X, Muzny DM, Weinstock GM, Šmajš D. 2012. Whole genome sequence of *Treponema pallidum* ssp. *pallidum*, strain Mexico A, suggests recombination between yaws and syphilis strains. *PLoS Negl Trop Dis* 6:e1832. <https://doi.org/10.1371/journal.pntd.0001832>.
63. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729. <https://doi.org/10.1093/molbev/mst197>.
64. Peng RR, Wang AL, Li J, Tucker JD, Yin YP, Chen XS. 2011. Molecular typing of *Treponema pallidum*: a systematic review and meta-analysis. *PLoS Negl Trop Dis* 5:e1273. <https://doi.org/10.1371/journal.pntd.0001273>.
65. Ho EL, Lukehart SA. 2011. Syphilis: using modern approaches to understand an old disease. *J Clin Invest* 121:4584–4592. <https://doi.org/10.1172/JCI57173>.
66. Soucy SM, Huang J, Gogarten JP. 2015. Horizontal gene transfer: building the web of life. *Nat Rev Genet* 16:472–482. <https://doi.org/10.1038/nrg3962>.
67. van der Ley P, Heckels JE, Virji M, Hoogerhout P, Poolman JT. 1991. Topology of outer membrane porins in pathogenic *Neisseria* spp. *Infect Immun* 59:2963–2971.
68. Azghani AO, Idell S, Bains M, Hancock RE. 2002. *Pseudomonas aeruginosa* outer membrane protein F is an adhesin in bacterial binding to lung epithelial cells in culture. *Microb Pathog* 33:109–114. <https://doi.org/10.1006/mpat.2002.0514>.
69. Gray RR, Mulligan CJ, Molini BJ, Sun ES, Giacani L, Godornes C, Kitchen A, Lukehart SA, Centurion-Lara A. 2006. Molecular evolution of the *tpcC*, *D*, *I*, *K*, *G*, and *J* genes in the pathogenic genus *Treponema*. *Mol Biol Evol* 23:2220–2233. <https://doi.org/10.1093/molbev/msl092>.
70. Hollingshead S, Jongerius I, Exley RM, Johnson S, Lea SM, Tang CM. 2018. Structure-based design of chimeric antigens for multivalent protein vaccines. *Nat Commun* 9:1051. <https://doi.org/10.1038/s41467-018-03146-7>.
71. Staudová B, Strouhal M, Zobaníková M, Čejková D, Fulton LL, Chen L, Giacani L, Centurion-Lara A, Bruisten SM, Sodergren E, Weinstock GM, Šmajš D. 2014. Whole genome sequence of the *Treponema pallidum* subsp. *endemicum* strain Bosnia A: the genome is related to yaws treponemes but contains few loci similar to syphilis treponemes. *PLoS Negl Trop Dis* 8:e3261. <https://doi.org/10.1371/journal.pntd.0003261>.
72. Mikalová L, Strouhal M, Oppelt J, Grange PA, Janier M, Benhaddou N, Dupin N, Šmajš D. 2017. Human *Treponema pallidum* 11q/j isolate belongs to subsp. *endemicum* but contains two loci with a sequence in TP0548 and TP0488 similar to subsp. *pertenue* and subsp. *pallidum*, respectively. *PLoS Negl Trop Dis* 11:e0005434. <https://doi.org/10.1371/journal.pntd.0005434>.
73. Centurion-Lara A, LaFond RE, Hevner K, Godornes C, Molini BJ, Van Voorhis WC, Lukehart SA. 2004. Gene conversion: a mechanism for generation of heterogeneity in the *tpkK* gene of *Treponema pallidum* during infection. *Mol Microbiol* 52:1579–1596. <https://doi.org/10.1111/j.1365-2958.2004.04086.x>.
74. Smith SG, Mahon V, Lambert MA, Fagan RP. 2007. A molecular Swiss army knife: OmpA structure, function and expression. *FEMS Microbiol Lett* 273:1–11. <https://doi.org/10.1111/j.1574-6968.2007.00778.x>.
75. Maruvada R, Kim KS. 2011. Extracellular loops of the *Escherichia coli* outer membrane protein A contribute to the pathogenesis of meningitis. *J Infect Dis* 203:131–140. <https://doi.org/10.1093/infdis/jiq009>.
76. Fairman JW, Noinaj N, Buchanan SK. 2011. The structural biology of β -barrel membrane proteins: a summary of recent reports. *Curr Opin Struct Biol* 21:523–531. <https://doi.org/10.1016/j.sbi.2011.05.005>.
77. Radolf JD, Deka RK, Anand A, Šmajš D, Norgard MV, Yang XF. 2016. *Treponema pallidum*, the syphilis spirochete: making a living as a stealth pathogen. *Nat Rev Microbiol* 14:744–759. <https://doi.org/10.1038/nrmicro.2016.141>.
78. Zhang Q, Meitzler JC, Huang S, Morishita T. 2000. Sequence polymorphism, predicted secondary structures, and surface-exposed conformational epitopes of *Campylobacter* major outer membrane protein. *Infect Immun* 68:5679–5689. <https://doi.org/10.1128/IAI.68.10.5679-5689.2000>.
79. Alcalá B, Salcedo C, Arreaza L, Abad R, Enríquez R, De La Fuente L, Uría MJ, Vázquez JA. 2004. Antigenic and/or phase variation of PorA protein in non-subtypable *Neisseria meningitidis* strains isolated in Spain. *J Med Microbiol* 53:515–518. <https://doi.org/10.1099/jmm.0.05517-0>.
80. Isaeva MP, Stenková AM, Guzev KV, Bystritskaya EP, Shubin FN, Rasskazov VA, Rakin A. 2012. Diversity and adaptive evolution of a major porin

- gene (*ompF*) in *Yersinia pseudotuberculosis*. *Adv Exp Med Biol* 954:39–43. https://doi.org/10.1007/978-1-4614-3561-7_5.
81. Davies RL, Lee I. 2004. Sequence diversity and molecular evolution of the heat-modifiable outer membrane protein gene (*ompA*) of *Mannheimia (Pasteurella) haemolytica*, *Mannheimia glucosida*, and *Pasteurella trehalosi*. *J Bacteriol* 186:5741–5752. <https://doi.org/10.1128/JB.186.17.5741-5752.2004>.
 82. McKnew DL, Lynn F, Zenilman JM, Bash MC. 2003. Porin variation among clinical isolates of *Neisseria gonorrhoeae* over a 10-year period, as determined by Por variable region typing. *J Infect Dis* 187:1213–1222. <https://doi.org/10.1086/374563>.
 83. Unemo M, Vorobieva V, Firsova N, Ababkova T, Leniv I, Haldorsen BC, Fredlund H, Skogen V. 2007. *Neisseria gonorrhoeae* population in Arkhangel'sk, Russia: phenotypic and genotypic heterogeneity. *Clin Microbiol Infect* 13:873–878. <https://doi.org/10.1111/j.1469-0691.2007.01764.x>.
 84. Nunes A, Borrego MJ, Nunes B, Florindo C, Gomes JP. 2009. Evolutionary dynamics of *ompA*, the gene encoding the *Chlamydia trachomatis* key antigen. *J Bacteriol* 191:7182–7192. <https://doi.org/10.1128/JB.00895-09>.
 85. Gomes JP, Bruno WJ, Nunes A, Santos N, Florindo C, Borrego MJ, Dean D. 2007. Evolution of *Chlamydia trachomatis* diversity occurs by widespread interstrain recombination involving hotspots. *Genome Res* 17:50–60. <https://doi.org/10.1101/gr.5674706>.
 86. Cameron CE, Lukehart SA. 2014. Current status of syphilis vaccine development: need, challenges, prospects. *Vaccine* 32:1602–1609. <https://doi.org/10.1016/j.vaccine.2013.09.053>.
 87. Izac JR, Oliver LD, Jr, Earnhart CG, Marconi RT. 2017. Identification of a defined linear epitope in the OspA protein of the Lyme disease spirochetes that elicits bactericidal antibody responses: implications for vaccine development. *Vaccine* 35:3178–3185. <https://doi.org/10.1016/j.vaccine.2017.04.079>.
 88. Oliver LD, Jr, Earnhart CG, Virginia-Rhodes D, Theisen M, Marconi RT. 2016. Antibody profiling of canine IgG responses to the OspC protein of the Lyme disease spirochetes supports a multivalent approach in vaccine and diagnostic assay development. *Vet J* 218:27–33. <https://doi.org/10.1016/j.tvjl.2016.11.001>.
 89. Earnhart CG, Marconi RT. 2007. Construction and analysis of variants of a polyvalent Lyme disease vaccine: approaches for improving the immune response to chimeric vaccinogens. *Vaccine* 25:3419–3427. <https://doi.org/10.1016/j.vaccine.2006.12.051>.
 90. Cox DL, Luthra A, Dunham-Ems S, Desrosiers DC, Salazar JC, Caimano MJ, Radolf JD. 2010. Surface immunolabeling and consensus computational framework to identify candidate rare outer membrane proteins of *Treponema pallidum*. *Infect Immun* 78:5178–5194. <https://doi.org/10.1128/IAI.00834-10>.
 91. Chen CY, Chi KH, George RW, Cox DL, Srivastava A, Rui Silva M, Carneiro F, Lauwers GY, Ballard RC. 2006. Diagnosis of gastric syphilis by direct immunofluorescence staining and real-time PCR testing. *J Clin Microbiol* 44:3452–3456. <https://doi.org/10.1128/JCM.00721-06>.
 92. Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13:555–556. <https://doi.org/10.1093/bioinformatics/13.5.555>.
 93. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591. <https://doi.org/10.1093/molbev/msm088>.
 94. Xu B, Yang Z. 2013. PAMLX: a graphical user interface for PAML. *Mol Biol Evol* 30:2723–2724. <https://doi.org/10.1093/molbev/mst179>.
 95. Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* 19:908–917. <https://doi.org/10.1093/oxfordjournals.molbev.a004148>.
 96. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comp Chem* 25:1605–1612. <https://doi.org/10.1002/jcc.20084>.
 97. Abagyan R, Totrov M. 1994. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J Mol Biol* 235:983–1002. <https://doi.org/10.1006/jmbi.1994.1052>.