Doxycycline Resistance and 16S rRNA Mutations in *Treponema pallidum*

Mathew A. Beale, Michael Marks, Annie Luetkemeyer, Connie Celum, Matthew R. Golden, Lorenzo Giacani, Nicole A.P. Lieberman

Author affiliations: Wellcome Sanger Institute, Hinxton, UK (M.A. Beale); London School of Hygiene and Tropical Medicine, London, UK (M. Marks); University of California San Francisco, San Francisco, California, USA (A. Luetkemeyer); University of Washington, Seattle, Washington, USA (C. Celum, M.R. Golden, L. Giacani, N.A.P. Lieberman)

DOI: https://doi.org/10.3201/eid3208.260433

To the Editor: We read with interest but also concern the article by Long et al. (*1*) describing putative doxycycline resistance–associated variants of the *Treponema pallidum* ribosomal RNA 16S gene. Determining if *T. pallidum* can become resistant to doxycycline is urgent, given its use for both prevention (i.e., doxy-PEP) and treatment of syphilis, especially amid ongoing shortages of benzathine penicillin G.

As the groups that generated most of the sequencing data reanalyzed in Long et al. (*1*), our analyses were unable to replicate the primary finding of a heterozygous G966T 16S mutation (*Escherichia coli* numbering) in 9 samples. Using standard pathogen genomics methods (Figure), we did not find the variant in any reanalyzed sample at an allele frequency exceeding background technical noise.

Of note, >98% of publicly available *T. pallidum* genomes were generated from clinical specimens by metagenomic sequencing using hybrid capture probes. That method retains conserved non–*T. pallidum* DNA, such as ribosomal operons from other bacteria, in the sample (*2*,*3*), and can introduce errors when merging fragmented DNA reads into consensus sequences. Such artifacts may be incorporated into consensus *T. pallidum* genomes available from public resources such as pubMLST (https://pubmlst.org) or GenBank and misinterpreted as real mutations. Although Long et al. (*1*) reported filtering for reads arising from *Treponema*, we and others have shown the necessity of more stringent methods, such as competitive mapping versus related ribosomal sequences (*2*) or requiring near-identity to known *T. pallidum* rRNA sequences (*4*), to avoid inadvertent reporting errors.

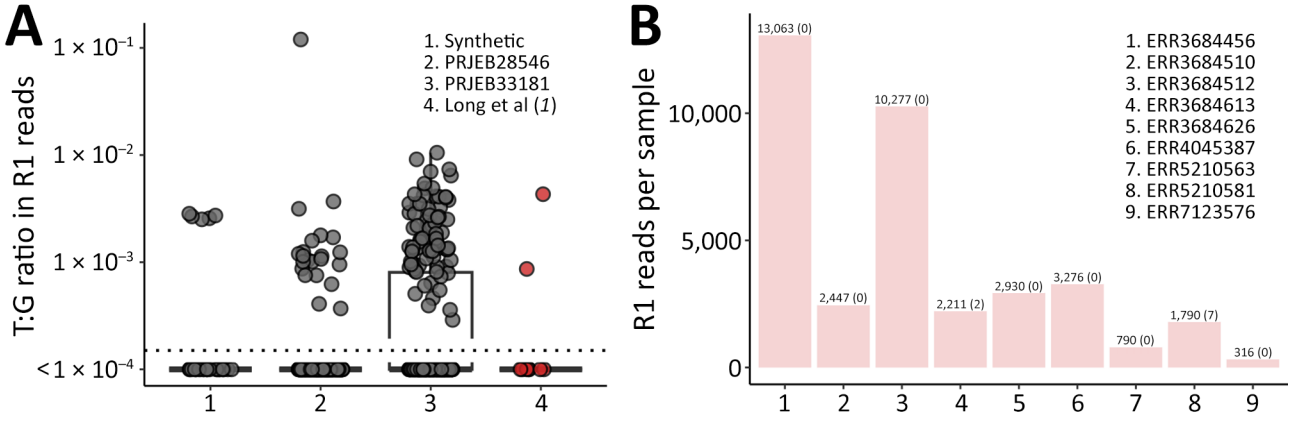

Figure. Ratio of G966T to wild-type G alleles among R1 reads of *Treponema pallidum*. We downloaded raw reads from the European Nucleotide Archive (BioProjects PRJEB28546 and PRJEB33181), which were re-assembled in Long et al. (*1*) and generated synthetic reads with the error profile of the Illumina HiSeq 2500 from 29 high-quality *T. pallidum* assemblies from GenBank using Illumina ART (Illumina ART; https://www.niehs.nih.gov/research/resources/software/biostatistics/art). We removed reads containing host sequences by using kraken 2 (https://github.com/DerrickWood/kraken2), performed quality and adaptor trimming with Trimmomatic version 0.39 (https://github.com/usadellab/trimmomatic), and filtered remaining read pairs to include only reads unambiguously arising from *T. pallidum* (taxid 160) using the standard kraken 2 16GB database with the default kmer length of 35. We isolated reads containing sequences from either of the rRNA loci by using bbduk (https://sourceforge.net/projects/bbmap) requiring a 21-mer match to the *T. pallidum* 16S locus sequence with 1 or 0 mismatches. For unambiguous identification and enumeration of the wild-type or G966T variant, R1 reads were grepped for perfect matches to the 51-base sequence centered on nt 966 (bold), equivalent to positions 232265 and 280700 in the SS14 reference sequence chromosome (CP004011): GGTGGAGCATGTGGTTTAATTCGAT**G**ATACGCGAGGAACCTTACCCGGGTT (wild-type) and GGTGGAGCATGTGGTTTAATTCGAT**T**ATACGCGAGGAACCTTACCCGGGTT (G966T), and the reverse complement of each. A) T:G ratio at rRNA position 966 in samples with >100 R1 reads (n = 622). Each point represents a sample; points below the dotted line contain 0 G966T reads. Boxplots represent medians and interquartile ranges. Samples identified in Long et al. do not have a G966T allele frequency exceeding the technical noise from PCR and sequencing. B) Total R1 reads containing wild-type and mutant sequence in 9 samples identified by Long et al. The total reads are shown for each sample and the number of reads supporting the G966T mutation is shown in parentheses.

The development of doxycycline resistance by *T. pallidum* could have devastating consequences for the control of syphilis and is being closely monitored by clinicians and scientists. We are encouraged by increasing genomic surveillance of *T. pallidum*, providing a mechanism for early detection of the emergence of resistance-associated variants.

This research was funded in whole, or in part, by the Wellcome Trust (220540/Z/20/A). M.M. is funded by the National Institute of Health and Social Care Research (NIHR304258); the views expressed are those of the author(s) and not necessarily those of the Institute or the Department of Health and Social Care. A.L. and C.C. are funded by the US National Institute of Allergy and Infectious Diseases (R01 AI143439). N.L. is supported by Gates Foundation (INV-072205). M.R.G. has received research support from Hologic and Speedx.

References

1. Long GS, Neale M, Braukmann T, Tran V, Singh N, Allen V, et al. Genomic analysis of doxycycline resistance–associated 16S rRNA mutations in *Treponema pallidum* subspecies *pallidum*. Emerg Infect Dis. 2026;32:242–5. https://doi.org/10.3201/eid3202.251060
2. Beale MA, Marks M, Sahi SK, Tantalo LC, Nori AV, French P, et al. Genomic epidemiology of syphilis reveals independent emergence of macrolide resistance across multiple circulating lineages. Nat Commun. 2019;10:3255. https://doi.org/10.1038/s41467-019-11216-7
3. Chen W, Šmajs D, Hu Y, Ke W, Pospíšilová P, Hawley KL, et al. Analysis of *Treponema pallidum* strains from China using improved methods for whole-genome sequencing from primary syphilis chancres. J Infect Dis. 2021;223:848–53. https://doi.org/10.1093/infdis/jiaa449
4. Lieberman NAP, Lin MJ, Xie H, Shrestha L, Nguyen T, Huang ML, et al. *Treponema pallidum* genome sequencing from six continents reveals variability in vaccine candidate genes and dominance of Nichols clade strains in Madagascar. PLoS Negl Trop Dis. 2021;15:e0010063. https://doi.org/10.1371/journal.pntd.0010063

Address for correspondence: Nicole Lieberman, University of Washington, 850 Republican St, Seattle, WA 98109, USA; email: naliebe@uw.edu

George S. Long, Venkata R. Duvvuri

Author affiliations: Public Health Ontario, Toronto, Ontario, Canada (G.S. Long, V.R. Duvvuri); University of Toronto, Toronto (V.R. Duvvuri)

DOI: https://doi.org/10.3201/eid3208.260943

In Response: On March 5, 2026, Drs. Lieberman and Beale contacted us concerning the identification of a heterozygous allele at position G966T (by *E. coli* reference numbering, G969T by *T. pallidum* reference numbering) of the 16S rRNA gene of *Treponema pallidum* subspecies *pallidum* (*1*). They surmised that our identification of the allele may have been caused by either genome misassemblies or contamination with sequences from other *Trepomena* species (*2*). In response to these reproducibility concerns, we undertook a separate analysis using an independent analytical process from our original work (*1*).

Our reanalysis focused on the sequencing libraries of Beale et al. (*3*–*5*) and applied a mapping-based approach, aligning reads to the 16S rRNA gene of the *T. pallidum* reference genome (NC_021490.2:231255–232803). That reanalysis concluded that the previously identified heterozygous allele was indeed the result of contamination. Accordingly, Table 4 and text associated with the heterozygous single-nucleotide polymorphism have been removed from the original article. In addition, an extra paragraph has been added to the Appendix detailing the additional analyses.

Of note, the overall findings of our study remain unaffected. We still report the sequencing and assembly of 15 *T. pallidum* subsp. *pallidum* genomes from Canada. This revision strengthens the clarity of our comparative genomics analysis and simplifies the identification of single-nucleotide polymorphisms related to putative doxycycline resistance and underscore the importance of accounting for the presence of two 16S rRNA gene copies in *T. pallidum*. This study continues to represent a vital step in establishing a genomic baseline for monitoring circulating *T. pallidum* subsp. *pallidum* lineages in Canada.

References

1. Long GS, Neale M, Braukmann T, Tran V, Singh N, Allen V, et al. Genomic analysis of doxycycline resistance–associated 16S rRNA mutations in *Treponema pallidum* subspecies *pallidum*. Emerg Infect Dis. 2026;32:242–5. https://doi.org/10.3201/eid3202.251060
2. Beale MA, Marks M, Luetkemeyer A, Celum C, Golden MR, Giacani L, et al. Doxycycline resistance and 16S rRNA mutations in *Treponema pallidum*. Emerg Infect Dis. 2026;32:1383-84. https://doi.org/10.3201/eid3208.260433
3. Beale MA, Marks M, Sahi SK, Tantalo LC, Nori AV, French P, et al. Genomic epidemiology of syphilis reveals independent emergence of macrolide resistance across multiple circulating lineages. Nat Commun. 2019;10:3255. https://doi.org/10.1038/s41467-019-11216-7
4. Beale MA, Marks M, Cole MJ, Lee MK, Pitt R, Ruis C, et al. Global phylogeny of *Treponema pallidum* lineages reveals recent expansion and spread of contemporary syphilis. Nat Microbiol. 2021;6:1549–60. https://doi.org/10.1038/s41564-021-01000-z
5. Beale MA, Thorn L, Cole MJ, Pitt R, Charles H, Ewens M, et al. Genomic epidemiology of syphilis in England: a population-based study. Lancet Microbe. 2023;4:e770–80. https://doi.org/10.1016/S2666-5247(23)00154-4

Address for correspondence: Venkata R. Duvvuri, Public Health Ontario, 661 University Ave, Ste 1701, Toronto, ON M5G 1M1, Canada; email: venkata.duvvuri@oahpp.ca