1    **Nonhuman primates across sub-Saharan Africa are infected with the yaws**

2    **bacterium *Treponema pallidum* subsp. *pertenue***

3

4    Running title: Yaws bacterium in wild nonhuman primates

5

6    Sascha Knauf (sknauf@dpz.eu)[1¶], Jan F. Gogarten (jan.gogarten@gmail.com)[2,3,4¶], Verena J.

7    Schuenemann (verena.schuenemann@uni-tuebingen.de)[5¶], Hélène M. De Nys

8    (denyshelene@yahoo.com)[2,#¶], Ariane Düx (duexa@rki.de)[2,6], Michal Strouhal

9    (mstrouhal@med.muni.cz)[7], Lenka Mikalová (lmikal@med.muni.cz)[7], Kirsten I. Bos

10   (bos@shh.mpg.de)[8], Roy Armstrong (roy.armstrong@cumbria.ac.uk)[9], Emmanuel K. Batamuzi

11   (emmanuelbatamuzi@gmail.com)[10,§], Idrissa S. Chuma (chumaidr@gmail.com)[11], Bernard Davoust

12   (bernard.davoust@gmail.com)[12], Georges Diatta (georges.diatta@ird.fr)[13], Robert D. Fyumagwa

13   (rfyumagwa@yahoo.com)[14], Reuben R. Kazwala (kazwala@suanet.ac.tz)[15], Julius D. Keyyu

14   (julius.keyyu@gmail.com)[14], Inyasi A. V. Lejora (inyasi.lejora@tanzaniaparks.go.tz)[11], Anthony

15   Levasseur (anthony.levasseur@univ-amu.fr)[12], Hsi Liu (hcl6@cdc.gov)[16], Michael A. Mayhew

16   (mic1mayhew@hotmail.com)[9], Oleg Mediannikov (olegusss1@gmail.com)[12], Didier Raoult

17   (Didier.raoult@gmail.com)[12,13], Roman M. Wittig (wittig@eva.mpg.de)[3,17], Christian Roos

18   (CRoos@dpz.eu)[18], Fabian H. Leendertz (leendertzf@rki.de)[2], David Šmajs

19   (dsmajs@med.muni.cz)[7&*], Kay Nieselt (kay.nieselt@uni-tuebingen.de)[19&], Johannes Krause

20   (krause@shh.mpg.de)[5,8&*], Sébastien Calvignac-Spencer (CalvignacS@rki.de)[2,6&*]

21

22   [1]Work Group Neglected Tropical Diseases, Pathology Unit, German Primate Center, Leibniz Institute

23   for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany.

24   [2]Epidemiology of Highly Pathogenic Microorganisms, Robert Koch Institute, Seestraße 10, 13353

25   Berlin, Germany.

26   [3]Primatology Department, Max Planck Institute for Evolutionary Anthropology, Deutscher Pl. 6, 04103

27   Leipzig, Germany.

28  [4]Department of Biology, McGill University, 1205 Avenue du Docteur-Penfield, Montréal, QC H3A

29  1B1, Quebec, Canada.

30  [5]Institute for Archaeological Sciences, University of Tübingen, Ruemelinstrasse 23, 72070 Tübingen,

31  Germany.

32  [6]Viral Evolution, Robert Koch Institute, Seestraße 10, 13353 Berlin, Germany.

33  [7]Department of Biology, Faculty of Medicine, Masaryk University, Kamenice 5/A6, 625 00 Brno,

34  Czech Republic.

35  [8]Max Planck Institute for the Science of Human History, Kahlaische Strasse 10, 7745 Jena, Germany.

36  [9]Department of Science, Natural Resources and Outdoor Studies, University of Cumbria, Fusehill

37  Street, Carlisle, Cumbria, CA12HH, UK.

38  [10]Department of Veterinary Surgery and Theriogenology, College of Veterinary and Medical Sciences,

39  Sokoine University of Agriculture, Morogoro, Tanzania.

40  [11]Ecological Monitoring Department, Tanzania National Parks (TANAPA), P.O.Box 3134, Arusha,

41  Tanzania.

42  [12]Research Unit of Emerging Infectious and Tropical Diseases (URMITE) UM63, CNRS 7278, IRD

43  198, INSERM 1095, Aix-Marseille University, 27, Bd Jean Moulin, 13385 Marseille cedex 05,

44  Marseille, France.

45  [13]Research Unit of Emerging Infectious and Tropical Diseases (URMITE) IRD 198, Campus

46  IRD/UCAD, Hann Les Maristes, Dakar, Senegal.

47  [14]Tanzania Wildlife Research Institute (TAWIRI), P.O. Box 661, Arusha, Tanzania.

48  [15]Department of Veterinary Medicine and Public Health, College of Veterinary and Medical Sciences,

49  Sokoine University of Agriculture, Morogoro, Tanzania.

50  [16]National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Diseases

51  Control and Prevention, Atlanta, GA, USA.

52  [17]Taï Chimpanzee Project, Centre Suisse de Recherches Scientifiques, BP 1303, Abidjan 01, Côte

53  d'Ivoire.

54  [18]Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute

55  for Primate Research, Kellnerweg 4, 37077 Göttingen, Germany.

56    [19]Center for Bioinformatics, University of Tübingen, Sand 14, 72076 Tübingen, Germany.

57    [#]Current address: Unité Mixte Internationale 233, Institut de Recherche pour le Développement,

58    INSERM U1175, and University of Montpellier, Montpellier, France.

59

60    [*]Correspondence to: Dr. Sébastien Calvignac-Spencer, Robert Koch Institute, Seestrasse 10, 13353

61    Berlin, Germany, tel.: +49187542501, calvignacs@rki.de; Prof. David Šmajs, Faculty of Medicine,

62    Masaryk University, 625 00 Brno, Czech Republic, tel.: +420 549 497 496, dsmajs@med.muni.cz; Prof.

63    Dr. rer. nat. Johannes Krause, Max Planck Institute for the Science of Human History, Jena, Germany,

64    tel.: +49 3641 686600, krause@shh.mpg.de.

65    [¶]These authors contributed equally to this work.

66    [§]Deceased.

67    [&]These authors also contributed equally to this work.

68

69  Dear Editor,

70      The bacterium *Treponema pallidum* (*TP*) causes human syphilis (subsp. *pallidum*; *TPA*), bejel

71  (subsp. *endemicum*; *TEN*) and yaws (subsp. *pertenue*; *TPE*)[1]. While syphilis reached a world-wide

72  distribution[2], bejel and yaws are endemic diseases. Bejel is found in dry areas in Sahelian Africa and

73  Saudi Arabia, whereas yaws is present in the humid tropics[1]. Yaws is currently reported endemic in 14

74  countries and an additional 84 countries have a known history of yaws but lack recent epidemiological

75  data[3,4]. The disease was subject to global eradication efforts in the mid-20th century but reemerged in

76  West Africa, Southern Asia, and the Pacific region[5]. New large-scale treatment options triggered the

77  ongoing second eradication campaign, which aims to eradicate yaws globally by 2020[5].

78      *TPE* is usually considered as a strictly human pathogen. This perception may however partly

79  result from the lack of detailed data on nonhuman primate (NHP)-infecting treponemes. Indeed, a

80  number of African NHPs show skin ulcerations suggestive of treponemal infection and antibodies

81  against *TP* have been detected in wild NHP populations[6,7]. While genetic studies confirmed monkeys

82  and great apes are infected with *TP* strains[8-10], most of these analyses only determined short DNA

83  sequences. The small number of polymorphic sites examined largely precludes assignment of these

84  strains to a particular *TP* subspecies[9], especially considering that sporadic recombination events

85  between subspecies have been reported[11]. The only simian strain whose whole genome was sequenced

86  - Fribourg-Blanc, isolated from a Guinea baboon (*Papio papio*) in 1966[7] - unambiguously clustered

87  with human-infecting *TPE* strains[12].

88      A fundamental question with regard to yaws evolution and possibly yaws eradication is whether

89  humans and NHPs are commonly infected with the same pathogen, *TPE*, and whether transmission

90  between NHPs and humans occurs. To determine which pathogen causes treponematoses in NHPs

91  across sub-Saharan Africa, we collected samples from symptomatic wild individuals belonging to three

92  NHP species (*Cercocebus atys*, *Chlorocebus sabaeus*, and *Papio anubis*) from four independent

93  populations in West and East Africa (**Fig 1**, **Supplementary Table S1, Supplementary Materials**).

94  Samples were collected at Taï National Park (TaïNP; Côte d'Ivoire), Bijilo Forest Park (BFP, the

95  Gambia), Niokolo-Koba National Park (NKNP, Senegal), and Lake Manyara National Park (LMNP,

96    Tanzania). Monkeys presented yaws-like orofacial and limb lesions (TaïNP, BFP) or ulcerative

97    anogenital skin lesions (BFP, NKNP, LMNP)[9].

98        Using PCR, we showed the presence of *TP* in skin lesion biopsies or swabs from NHPs at

99    TaïNP (*C. atys*), BFP, and NKNP (*C. sabaeus*). *TP* infection in olive baboons (*P. anubis*) at LMNP had

100    previously been confirmed[6]. Two samples per NHP population were selected for whole genome

101    sequencing based on high *TP* copy number or the ability to amplify long PCR fragments

102    (**Supplementary Table S2**). To overcome the background of host genomic DNA, we used targeted

103    DNA capture coupled with NGS to reconstruct whole *TP* genomes[2,8]. Following quality filtering,

104    removal of PCR duplicates, merging of different sequencing runs from the same sample, and mapping

105    against the *TPE* strain Fribourg-Blanc reference genome, we obtained a range of 22,886-470,303 DNA

106    sequencing reads per sample. All samples showed at least 80% coverage of the reference genome with

107    depth coverage of three or higher; average genome coverage depth was between 6.1 and 121.0-fold

108    (**Supplementary Table S3**).

109        We generated maximum likelihood, Bayesian and maximum parsimony trees based on the

110    genomes reconstructed in our study and all available reference genomes (total sequence length:

111    1,133,379 nucleotides). In all trees, *TPE* and *TPA* strains formed reciprocally monophyletic groups,

112    with a mean *TPE*/*TPA* strain divergence of 0.099%. NHP-infecting *TP* strains all clustered with human-

113    infecting *TPE* strains (**Fig 1**; **Supplementary Figure S1**). The *TPE* clade exhibited a star-like branching

114    pattern whereby all basal branches were very short and received low statistical support. Importantly,

115    this pattern does not support a clear reciprocal monophyly of the *TPE* strains infecting humans and

116    NHPs. In line with this, the minimum divergence between strains infecting humans and NHPs was

117    lower than the maximum divergence amongst human and NHP-infecting strains (0.011% versus

118    0.015% and 0.024%). Human-infecting *TPE* strains Samoa D, CDC-2, CDC-2575, Ghana-051, and

119    Gauthier, which span a broad geographic and temporal range (at least four decades), were less divergent

120    from each other than the two strains infecting sooty mangabeys from a single social group at TaïNP

121    (0.011% versus 0.017% sequence divergence, respectively). While intra-group strain divergence was

122    low for the two African green monkey populations and the olive baboons (0.0003% and 0.0017%,

123    respectively), intra-species strain divergence for African green monkeys was relatively high when

124    compared to the divergence observed between the two most divergent human strains (0.0094% versus

125    0.015%).

126          For the sample LMNP-1, we determined the complete genome sequence and structure (average

127    depth of coverage: 169x; GenBank: CP021113; **Supplementary Tables S5-6**)[12]. The LMNP-1 genome

128    showed the same structure as published complete genomes of human-infecting *TPE* strains and the

129    simian strain Fribourg-Blanc. It was more similar to the human-infecting *TPE* Gauthier strain than the

130    simian isolate Fribourg-Blanc, showing differences at 266 and 325 chromosomal positions,

131    respectively. Most differences were single nucleotide substitutions or small indels (**Supplementary**

132    **Table S7**). The LMNP-1 and Gauthier strains exhibited the same number of the 24-bp repeats in the

133    *TP_0470* gene (n=25) and Gauthier had only one 60-bp repeat more than LMNP-1 strain in the *arp*

134    gene (LMNP-1 n=9 vs. Gauthier n=10). All 60-bp repeats in the *arp* gene of LMNP-1 were of Type II

135    and were identical to other *TPE* strains[13]. The *tprK* gene of LMNP-1 only had three variable regions,

136    V5-V7, when compared to other *TPE* strains. In addition to differences in *TP_0433*, *TP_0470*, and *tprK*

137    genes, relatively large indels were determined in *TPEGAU_0136* (33-nt long deletion; specific for

138    strains Gauthier and Samoa D), in *TPFB_0548* (42-nt long deletion; specific for strain Fribourg-Blanc),

139    in *TPEGAU_0858* (79-nt long deletion; specific for strain Gauthier), in the intergenic region (IGR)

140    between *TPEGAU_0628* and *TPEGAU_0629* (302-nt long deletion; specific for strain Gauthier), and

141    in IGR between *TPFB_0696* and *TPFB_0697* (430-nt long insertion; specific for strain Fribourg-

142    Blanc); the length of other sequence differences ranged between 1-15 nts. RNA operons structure of

143    the LMNP-1 genome (coordinates 231,180-236,139; 279,584-284,533; according to *TPE* strain

144    Gauthier: NC_016843.1) was similar to strains Gauthier, CDC-2, and Fribourg-Blanc, but different to

145    the strains Samoa D, Samoa F, and CDC-1. The LMNP-1 16S-5S-23S was identical in both operons

146    and 23S rRNA sequences were identical to other *TPE* strains except for strain Fribourg-Blanc (having

147    a single nucleotide difference at position 458). We did not find any mutations associated with macrolide

148    resistance (e.g. A2058G, A2059G)[14]. When the two NHP-infecting TPE strains, Fribourg-Blanc and

149    LMNP-1, were compared to the closest human-pathogenic TPE strains CDC-2 and Gauthier,

150    respectively, only 7.2% and 9.1% of all coding sequences (77 and 97 coding sequences out of 1065)

151 contained amino acid substitutions, suggesting limited functional divergence (**Supplementary Tables**
152 **S7-9**).

153       Our findings unambiguously indicate that at least three African NHP species (representing four
154 populations) from West and East Africa currently suffer from treponematosis caused by *TPE*. Taking
155 into account the isolation of the Fribourg–Blanc strain from Guinea baboons in 1966 and its recent
156 sequencing and identification as a member of the *TPE* clade[12], this represents four African NHP species
157 and five populations whose symptoms can be explained by *TPE* infections. Coupled with a growing
158 number of clinical and serological observations[6,7,9,10], this suggests infection of NHPs with *TPE* is
159 common throughout sub-Saharan Africa. Humans are not the exclusive host for the yaws bacterium and
160 NHPs are infected with the same bacterial agent.

161       *TPE* strains in NHPs exhibit considerable genetic diversity, which at least equals that found
162 among published human-infecting *TPE* strains. Importantly, we found no evidence for a clear sub-
163 differentiation of NHP- and human-infecting *TPE* strains, i.e. these strains did not form well-supported
164 reciprocally monophyletic groups. Rather, the star-like topology of our phylogenomic tree suggests a
165 rapid initial radiation of the ancestor of *TPE* which may have involved transmission across primate
166 species barriers in a relatively distant past (with respect to the *TPE* clade depth). These results neither
167 support, nor allow us to exclude, a possible recent transmission of *TPE* between NHPs and humans,
168 especially due to the large geographic and temporal separation between the two groups of samples being
169 compared. A major hurdle in identifying such potential transmission events is the availability of enough
170 bacterial genomes. Despite large numbers of human cases, very few genomes have been determined
171 from human-infecting *TPE* strains and only from a very limited geographic range. Generating additional
172 human-infecting *TPE* genomes represents an important area of research that, coupled with the genomes
173 of *TPE* strains infecting NHPs presented here, could now enable the detection of recent zoonotic
174 transmission events, would any exist.

175       Since yaws has not been reported for several decades in humans in countries where we find
176 NHPs to be infected with *TPE*, we expect that if transmission happens, it is only at very low frequency
177 (as is the case for many zoonotic diseases). Of course, such low frequency zoonotic transmission does
178 not explain the reemergence of yaws, which is the consequence of continued human-to-human

179  transmission. However, now that eradication of yaws seems within reach[15], the finding that *TPE* strains

180  circulate in NHPs certainly calls for more research into their diversity and zoonotic potential.

181

182  **Data availability**

183      All raw read files have been deposited in NCBI as part of the BioProject PRJNA343706.

184

185  **Competing interests**

186      The authors declare that they have no competing interests.

187

188  **Supporting information**

189  **Supplementary Materials.** This document comprises supporting methods and additional results.

190  **Supplementary Figure S1.** Phylogenetic trees of *TP* whole genome sequences.

191  **Supplementary Table S1**. Nonhuman primates anesthetized for this study.

192  **Supplementary Table S2**. Molecular analyses (PCR and sequencing) performed on blood samples,

193  skin tissue samples, and lesion swabs.

194  **Supplementary Table S3**. Read mapping and genotyping results.

195  **Supplementary Table S4**. Published genomes used for phylogenetic analyses.

196  **Supplementary Table S5**. List of primers used for long-range PCR amplification of *TP* intervals of

197  the East African baboon genome (strain LMNP-1).

198  **Supplementary Table S6**. Summary of the PSGS sequencing results of four genomic DNA (gDNA)

199  pools of the East African baboon genome (strain LMNP-1).

200  **Supplementary Table S7**. Number of nucleotide differences (i.e. indels and SNVs) of various

201  lengths between the genome of the baboon (strain LMNP-1) and the published *TPE* genome of strains

202  Gauthier and Fribourg-Blanc.

203  **Supplementary Table S8**. Proteins encoded by the *TPE* strain Fribourg-Blanc genome with 1 and

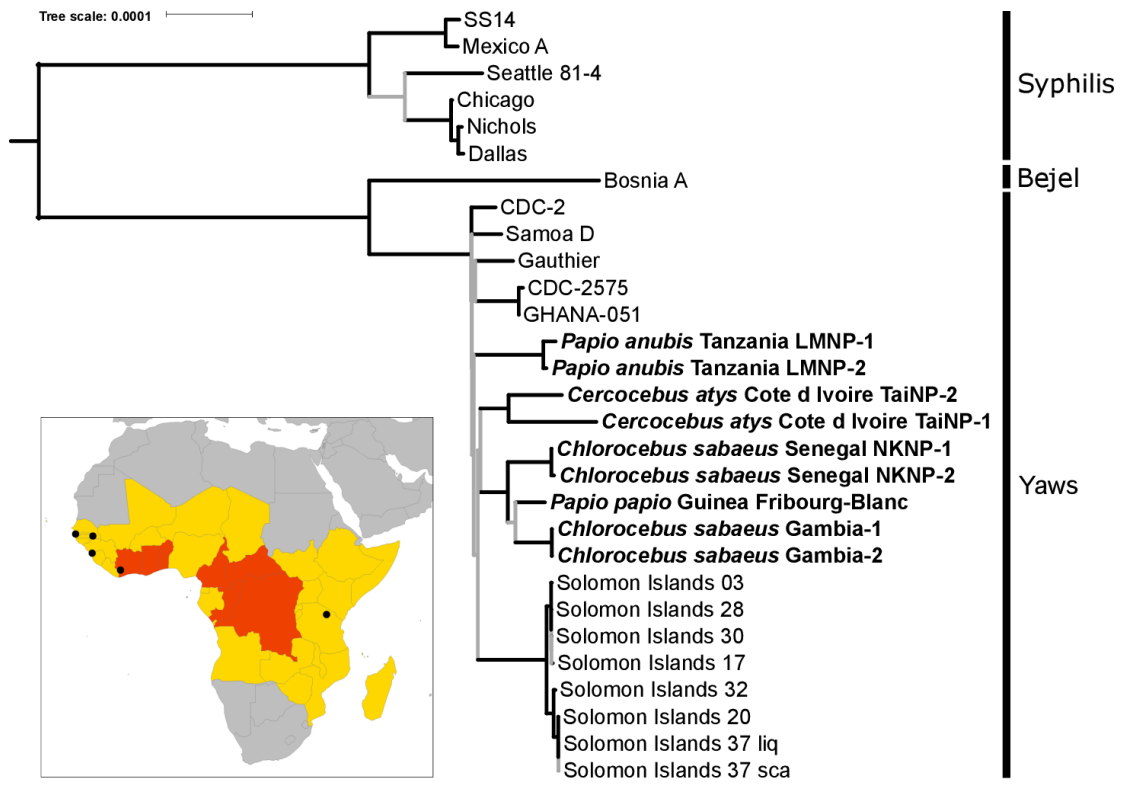204  more amino acid changes when compared to the *TPE* strain CDC-2 proteome.

205     **Supplementary Table S9**. Proteins encoded by the LMNP-1 baboon strain with one and more amino

206     acid changes when compared to the *TPE* strain Gauthier proteome.

207

**Figure Legend**

209  **Fig 1. Phylogenomic analysis of NHP- and human-infecting *Treponema pallidum* strains.** In this

210  maximum likelihood tree, nodes that had less than 95% ultrafast bootstrap approximation support are

211  indicated with grey lines. Tip labels indicate the NHP species sampled, the country of origin, and the

212  sample ID. The scale is in nucleotide substitution per site. The inset is a map of Africa where sites of

213  origin of the NHP samples from which a *TP* genome was determined are indicated with black circles.

214  A country's 2013 yaws status based on the World Health Organization's Global Health Observatory

215  (http://www.who.int/gho/en/) is indicated by its color: grey indicates no previous history of yaws

216  infections in humans, yellow indicates a country previously endemic for yaws though the current status

217  is unknown, and countries in red indicate countries which are currently considered endemic for yaws.

218

**Figures**

**Fig 1**

## Acknowledgments

## References

1    Giacani, L. & Lukehart, S. A. The endemic treponematoses. *Clinical microbiology reviews* **27**, 89-115, doi:10.1128/CMR.00070-13 (2014).

2    Arora, N. *et al.* Origin of modern syphilis and emergence of a pandemic Treponema pallidum cluster. *Nat Microbiol* **2**, 16245, doi:10.1038/nmicrobiol.2016.245 (2016).

3    Marks, M. Yaws: towards the WHO eradication target. *Trans R Soc Trop Med Hyg*, doi:10.1093/trstmh/trw032 (2016).

4    World Health Organization. Eradication of yaws: procedures for verification and certification of interruption of transmission., (Geneva, 2018).

249    5    Asiedu, K., Fitzpatrick, C. & Jannin, J. Eradication of yaws: historical efforts and achieving

250        WHO's 2020 target. *PLoS neglected tropical diseases* **8**, e3016,

251        doi:10.1371/journal.pntd.0003016 (2014).

252    6    Felsenfeld, O. & Wolf, R. H. Serological reactions with treponemal antigens in nonhuman

253        primates and the natural history of treponematosis in man. *Folia Primatol (Basel)* **16**, 294-305

254        (1971).

255    7    Fribourg-Blanc, A. & Mollaret, H. H. Natural treponematosis of the African primate. *Primates*

256        *Med* **3**, 113-121 (1969).

257    8    Gogarten, J. F. *et al.* Tools for opening new chapters in the book of Treponema pallidum

258        evolutionary history. *Clin Microbiol Infect* **22**, 916-921, doi:10.1016/j.cmi.2016.07.027 (2016).

259    9    Knauf, S. *et al.* Treponema infection associated with genital ulceration in wild baboons.

260        *Veterinary pathology* **49**, 292-303, doi:10.1177/0300985811402839 (2012).

261    10    Chuma IS, B. E., Collins DA, Fyumagwa RD, Hallmaier-Wacker LK, Kazwala RR, et al.

262        Widespread *Treponema pallidum* Infection in Nonhuman Primates, Tanzania. *Emerg Infect*

263        *Dis*. **24**, 1002-1009, doi: https://doi.org/10.3201/eid2406.180037 (2018).

264    11    Petrosova, H. *et al.* Whole genome sequence of Treponema pallidum ssp. pallidum, strain

265        Mexico A, suggests recombination between yaws and syphilis strains. *PLoS neglected tropical*

266        *diseases* **6**, e1832, doi:10.1371/journal.pntd.0001832 (2012).

267    12    Zobanikova, M. *et al.* Whole genome sequence of the Treponema Fribourg-Blanc: unspecified

268        simian isolate is highly similar to the yaws subspecies. *PLoS neglected tropical diseases* **7**,

269        e2172, doi:10.1371/journal.pntd.0002172 (2013).

270    13    Harper, K. N. *et al.* The sequence of the acidic repeat protein (arp) gene differentiates venereal

271        from nonvenereal *Treponema pallidum* subspecies, and the gene has evolved under strong

272        positive selection in the subspecies that causes syphilis. *FEMS immunology and medical*

273        *microbiology* **53**, 322-332, doi:10.1111/j.1574-695X.2008.00427.x (2008).

274    14    Lukehart, S. A. *et al.* Macrolide resistance in *Treponema pallidum* in the United States and

275        Ireland. *The New England journal of medicine* **351**, 154-158, doi:10.1056/NEJMoa040216

276        (2004).

277 15    Marks, M. *et al.* Mathematical Modeling of Programmatic Requirements for Yaws Eradication.

278    *Emerg Infect Dis* **23**, 22-28, doi:10.3201/eid2301.160487 (2017).

279