1 Technical appendix

2

3 Strain diversity of *Treponema pallidum* subsp. *pertenue* suggests rare interspecies

4 transmission in African nonhuman primates

5

6	Authors: Idrissa S. Chuma ^{1,2,3*} , Christian Roos ^{2,4*} , Anagaw Atickem ⁵ , Torsten Bohm ⁶ , D.
7	Anthony Collins ⁷ , Linda Grillová ^{8,9} , Luisa K. Hallmaier-Wacker ^{1,2} , Rudovick R. Kazwala ³ ,
8	Julius D. Keyyu ⁹ , Simone Lüert ^{1,2} , Ulrich Maloueki ^{6,11} , Jan Oppelt ^{9,12} , Klára J.
9	Petrželková ^{13,14,15} , Alexander Piel ^{16,17} , Fiona A. Stewart ^{16,17} , David Šmajs ⁹ , Sascha Knauf ^{1,18§}
10	
11	Design of the Multi-Locus Sequence Typing system
12	Identification of most variable genes in TPE

We identified six candidate genes with the SNVs frequency per kbp ranging from 8-29, which
code outer membrane proteins or hypothetical proteins (*TP0136*, *TP0326*, *TP0488*, *TP0548*, *TP0858*, *TP0865*; Table 1). Most of the loci were identified as putative intra- or inter-strain
recombination genes (*TP0136*, *TP0326*, *TP0488*, *TP548* and *TP0865*)^{30,32,37-38}.

17

18 Resolution power of genome-wide data and individual loci

Genome-wide maximum-likelihood (ML) trees of 23 available complete and draft genome sequences were based on 1,207 variable sites and allowed to distinguish 22 haplotypes (data not showed). Given the fact that strains CDC 2575 and Ghana-051 are completely identical²⁰, the genome-wide tree had 100% resolution of whole genome sequences. Among six candidate loci, the highest resolution power was observed in *TP0488* (70%) followed by *TP0326*,

Chuma et al.

TP0548, *TP0858* (all 57%), *TP0136* (44%) and *TP0865* (31.8%). Interestingly, the concatenated sequences of all candidate loci did not reveal higher resolution then the resolution observed in the *TP0488* gene (70%).

27

Identification of the most suitable typing loci for TPE MLST among and within the candidate
genes

Since the amplification efficiency of *Treponema* DNA from clinical samples has been
shown to be dependent on the length of the PCR products^{10,39}, we selected loci with the
highest occurrence of variable sites accumulated in as short DNA regions as possible (Table
S2). Furthermore, we preferred loci with the highest percentage of genome-wide resolution
and last, we selected loci that were able to clearly distinguish treponematoses caused by *TPE*from the *TPA/TEN* infections (Table S2).

36

37 DNA extraction

DNA extraction for tissue samples was performed as described elsewhere². Swab materials 38 39 followed the same protocol as described for the skin tissues using the QIAamp DNA Mini Kit 40 (Qiagen, Hilden, Germany) with some modifications. Briefly, each swab was digested in 450 41 µl custom-made lysis buffer (10 mM Tris [pH 8.0], 0.1 M EDTA [pH 8.0], and 0.5% sodium 42 dodecyl sulfate), the same buffer in which the swab was collected. In addition, the remaining 43 lysis buffer was also digested. In both reactions, care was taken to keep the ratio of buffer and proteinase K as recommended by the manufacturer. The samples were digested overnight at 44 45 56°C and 900 rpm (Thermomixer Comfort, Eppendorf, Hamburg, Germany). On the next day, reaction tubes that contained swabs were placed on top of a new PCR clean reaction tube of 46 47 2.0 ml volume. A sterile 20G needle was used to penetrate the bottom of the reaction tube to 48 allow for subsequent separation of DNA containing fluid from the swab through vigorous 49 centrifugation at 6,000 xg for 5 min at room temperature. All subsequent steps followed the Technical appendix - 2 -

50 manufactures protocol. The DNA was eluted twice with 100 µL AE buffer and was further purified using glycogen precipitation according to the protocol published in Knauf et al.²⁴. 51 52 Ethanol (98%)-preserved fecal samples from western lowland gorillas (Gorilla gorilla gorilla) were extracted using the First DNA All-tissue extraction kit (Gen-ial, Troisdorf, 53 54 Germany). Briefly, the procedure followed the manufacturer's guidance with some minor 55 modifications. Feces were dried from ethanol overnight and subsequently resolved in 1 ml of the kit's containing lysis buffers #1 and 100 µl of lysis buffer #2. Samples where then 56 57 incubated with 20 µl proteinase K at 65°C and 1000 rpm (Thermomixer Comfort, Eppendorf, 58 Hamburg, Germany) for 50 min. All subsequent steps followed those of the manufacture's 59 protocol.

Chuma et al.

61 Table S1. *TPE* genomes used to identify most suitable gene candidates for the MLST 62 design. ^{\$} In order to determine the whole genome sequence, the *TPE* strain was multiplied in 63 experimental animals prior to the NGS. [§] Culture-independent enrichments (hybridization 64 captures) were used prior to the NGS in order to separate the *TP* genetic material from the 65 host DNA. Compl.=complete

Strains ID	Genome	Source	Host	Year of isolation	Geographic area	References
CDC-1	Compl.	Rabbit inoculation [§]	Human	1980	Ghana	unpublished
CDC-2	Compl.	Rabbit inoculation ^{\$}	Human	1980	Ghana	40
Samoa D	Compl.	Rabbit inoculation ^{\$}	Human	1953	Western Samoa	40
Gauthier	Compl.	Rabbit inoculation ^{\$}	Human	1960	Congo	40
Fribourg- Blanc	Compl.	Rabbit inoculation [§]	NHPs	1966	Ghana	41
CDC 2575	Compl.	Rabbit inoculation [§]	Human	1980	Ghana	21
Ghana-051	Compl.	Rabbit inoculation ^{\$}	Human	1988	Ghana	21
Sei Geringging K403	Compl.	Rabbit inoculation [§]	Human	1990	Indonesia	42
Kampung Dalan K363	Compl.	Rabbit inoculation ^{\$}	Human	1990	Indonesia	42
LMNP-1	Compl.	Clinical§	NHPs	2007	Tanzania	1
LMNP-2	Draft	Clinical§	NHPs	2007	Tanzania	1
Gambia-1	Draft	Clinical§	NHPs	unknown	Gambia	1
Gambia-2	Draft	Clinical [§]	NHPs	unknown	Gambia	1
Senegal NKNP-1	Draft	Clinical§	NHPs	unknown	Senegal	1
Senegal NKNP-2	Draft	Clinical§	NHPs	unknown	Senegal	1
Cote d'Ivoire TaiNP-1	Draft	Clinical§	NHP	unknown	Ivory Coast	1
Cote d'Ivoire TaiNP-2	Draft	Clinical§	NHP	unknown	Ivory Coast	1

ERR1470330	Draft	Clinical [§]	Human	2013	Solomon Islands	43
ERR1470331	Draft	Clinical [§]	Human	2013	Solomon Islands	43
ERR1470334	Draft	Clinical§	Human	2013	Solomon Islands	43
ERR1470338	Draft	Clinical [§]	Human	2013	Solomon Islands	43
ERR1470343	Draft	Clinical§	Human	2013	Solomon Islands	43
ERR1470344	Draft	Clinical§	Human	2013	Solomon Islands	43

68 **Table S2: Characteristics of candidate genes for** *TPE* **typing.** ^{\$} According to the yaws

- 69 reference genome Samoa D (GenBank accession number CP002374.1). § The genomes listed
- 70 in Table 1 including representatives of TPE strains, representative of TEN (Bosnia A,
- 71 GenBank accession number CP007548.1) and representatives of TPA (Nichols, GenBank
- 72 accession number CP004010.2; SS14, GenBank accession number CP004011.1; Mexico,
- 73 GenBank accession number CP003064.1) were used to test whether the corresponding
- 74 sequences can be used for subspecies classification (*TPA/TPE/TEN*).

Gene ^{\$}	Variable region length (bp)	Variable region coordinates ^{\$}	% of genome- wide data resolution	Differen- tiating TPE from TPA/TEN [§]	Differen- tiating TPA/TPE/TE N [§]
TP0488	782	522,942 - 523,723	70.0	Yes	No
<i>TP0548</i>	755	593,318 - 594,072	57.0	No	No
TP0858	824	936,118 - 936,941	57.0	Yes	Yes
TP0326	2,086	346,066 - 348,151	57.0	Yes	No
TP0136	910	157,823 - 158,733	44.0	Yes	Yes
TP0865	897	945,224 - 946,121	31.8	Yes	Yes

75

76

Chuma et al.

78 Table S3. Summary of TP sequence data of NHP origin included into this study. *multi-

- 79 strain infection present, NP=National Park, CA=Conservation Area, TZ=Tanzania,
- 80 ET=Ethiopia, RC=Republic of the Congo. Details on NHP species composition can be found
- 81 in Table S5.

Country	Sample location	n NHPs	n <i>TP0619</i>	n <i>TP0548</i>	n <i>TP0488</i>	n concatenated
			sequences	sequences	sequences	sequences
TZ	Gombe NP	3	1	3	3	3
ΤZ	Lake Manyara	46	41	44*	37	33
	NP					
ΤZ	Katavi NP	1	1	0	1	0
ΤZ	Mahale NP	1	1	1	0	0
ΤZ	Mikumi NP	1	0	1	0	0
ΤZ	Ngorongoro CA	9	9	9	5	5
ΤZ	Ruaha NP	6	4	5	5	4
ΤZ	Serengeti NP	7	5	7	8*	8
ΤZ	Tarangire NP	2	1	2	1	1
ΤZ	Issa Valley	2	2	2	3*	3
ΤZ	Udzungwa NP	1	1	1	0	0
ET	Awash NP	2	1	2	1	1
RC	Odzala-Kokoua	4	4	1	4	1
	NP					
	Total	85	71	78	67	59

82

Table S4. Strains diversity and year of sampling based on the concatenated sequences used for MLST (*TP0548* and *TP0488*) in NHP infecting *TP*. Strain classification letters (uppercase) shall not be confused with lower case letters that are used in the enhanced typing system in human syphilis and yaws-causing strains. The letters used here are only used to visualize and discuss the differences in *TP* strains of NHP origin.

Strain Typ	NHP sample ID	Sampling Area	Sampling Year	Species	
А	21LMF2290815	LMNP	2015	Papio anubis	
В	26F8060407	LMNP	2007	Papio anubis	
С	47M2180407	LMNP	2007	Papio anubis	
D	11LMF5200815	LMNP	2015	Papio anubis	
D	13LMM8210815	LMNP	2015	Papio anubis	
Е	4LMF8160815	LMNP	2015	Papio anubis	
Е	6LMF5170815	LMNP	2015	Papio anubis	
Е	7F5250307	LMNP	2007	Papio anubis	
Е	15F8270307	LMNP	2007	Papio anubis	
Е	16M8280307	LMNP	2007	Papio anubis	
Е	21F8040407	LMNP	2007	Papio anubis	
Е	22LMF5290815	LMNP	2015	Papio anubis	
Е	32F2110407	LMNP	2007	Papio anubis	
E	33M8120407	LMNP	2007	Papio anubis	
Е	49F8190407	LMNP	2007	Papio anubis	
Е	52F8210407	LMNP	2007	Papio anubis	
E	54M8210407	LMNP	2007	Papio anubis	
Е	55M2230407	LMNP	2007	Papio anubis	
E	60M5250407	LMNP	2007	Papio anubis	
E	63M8270407	LMNP	2007	Papio anubis	
E	67M8000507	LMNP	2007	Papio anubis	
E	69F5090507	LMNP	2007	Papio anubis	
E	74M8160507	LMNP	2007	Papio anubis	
F	34F2130407	LMNP	2007	Papio anubis	
G	70M5100507	LMNP	2007	Papio anubis	
Н	12LMF2210815	LMNP	2015	Papio anubis	
Н	19LMF8280815	LMNP	2015	Papio anubis	
Н	30LMF5190416	LMNP	2015	Papio anubis	
Н	50F2190407	LMNP	2015	Papio anubis	
Ι	5TNF11241215	TNP	2015	Papio anubis	
J	24SNM5151115	SNP	2015	Papio anubis	
Κ	29SNF2191115	SNP	2015	Papio anubis	
L	7SNM5081115	SNP	2015	Papio anubis	

М	24SNM5151115	SNP	2015	Papio anubis
Ν	20GNF5220916	GNP	2016	Papio anubis
0	19GNM2220916	GNP	2016	Papio anubis
0	3GNF8200916	GNP	2016	Papio anubis
Р	10NCM8170516	NCA	2016	Papio anubis
Р	12NCF5210317	NCA	2017	Papio anubis
Р	14NCM5210317	NCA	2017	Papio anubis
Q	17NCF2220317	NCA	2017	Papio anubis
R	18NCF8220317	NCA	2017	Papio anubis
S	2SNF2130815	SNP	2015	Papio anubis
S	6SNF2081115	SNP	2015	Papio anubis
Т	LMNP-1	LMNP	2007	Papio anubis
U	24LMM2081215	LMNP	2015	Papio anubis
V	34LMM2190317	LMNP	2017	Chlorocebus pygerythrus
W	41SNM1231115	SNP	2015	Chlorocebus pygerythrus
Х	9LMM2180815	LMNP	2015	Cercopithecus mitis
Y	5RUM2080716	RNP	2016	Chlorocebus pygerythrus
Ζ	14RUF5130716	RNP	2016	Papio cynocephalus
Ζ	16RUF8140716	RNP	2016	Papio cynocephalus
AA	6RUM2090716	RNP	2016	Chlorocebus pygerythrus
AB	32LMM2190317	LMNP	2017	Chlorocebus pygerythrus
AC	1UGF5101016	Issa Valley	2016	Papio cynocephalus
AC	2UGF8101016	Issa Valley	2016	Papio cynocephalus
AD	2UGF8101016	Issa Valley	2016	Papio cynocephalus

90 Table S5. GenBank accession numbers for the sequences included into this study. (Excel

91 Sheet)

92

	Consensus	1 100	200	300 4	po	500	600 700	800	900	1,000 1,035
	1. 1SNM5120815 (P. anubis)									
	2. 1UGF5101016 (P. cynocephalus) 3. 2F8220307 (P. anubis)		1							
	4. 2SNF2130815 1 (P. anubis) 5. 2UGF8101016 (P. cynocephalus)		1				1			
	6. 3GN F8200916 (P. anubis) 7. 3LMF5300415 (P. anubis)					-				
	8. 4AWM2051017 (C. aetiops) 9. 4LMF8160815 1 (P. anubis)	-	1					I		
	10. SRUM2080716 (C. pygerythrus) 11. STNF11241215 1 (P. anubis)					1				
	12. 6LMF5170815 1 (P. anubis) 13. 6RUM2090716 (C. pygerythrus)		1							
	14. 6SNF2081115 (P. anubis) 15. 7F5250307 (P. anubis)						1			
	16. 7MN M1300916 (C. pygerythrus) 17. 7SN M5081115 (P. anubis)		1	1						
	18. 8UN M2010716 (C. pygerythrus) 19. 9LMM2180815 (C. mitis)		1				1			
	20. 9NCF8170516 (P. anubis) 21. 10LMF8190815 (P. anubis)						1			
	22. 10NCM8170516 (P. anubis) 23. 11LMF5200815 1 (P. anubis)						1			
	24. 11LMF5200815 2 (P. anubis) 25. 12LMF2210815 1 (P. anubis)				-					
	26. 12NCF5210317 (P. anubis) 27. 12TNM9251215 (P. anubis)		1							
	28. 13LMM8210815 1 (P. anubis) 29. 13NCM2210317 (P. anubis)						1			
	30. 14AWM2051017 1 (C. aethiops) 31. 14NCM5210317 (P. anubis)	1	1					1		
	32. 14RUF5130716 (P. cynocephalus) 33. 15F8270307 (P. anubis)		1							
	34. 15NCF2220317 (P. anubis) 35. 16LMM5250815 1 (P. anubis)						1			
	36. 16M8280307 (P. anubis) 37. 16N CF2220317 (P. anubis)						I			
	38. 16RUF8140716 (P. cynocephalus) 39. 17LMF5270815 1 (P. anubis)		1							
	40. 17NCF2220317 (P. anubis) 41. 17RUM5140716 (P. cynocephalus)		1				1			
	42. 18LMF5270815 1 (P. anubis) 43. 18N CF8220317 (P. anubis)						i			
	44. 19GN M2220916 (P. anubis) 45. 19LME8280815 1 (P. anubis)				1-	-				
	46. 20GN F5220916 (P. anubis) 47. 20M5030407 (P. anubis)				1-		1			
	48. 21F8040407 (P. anubis) 49. 211 ME2290815 1 (P. anubis)									
	50. 22LMF5290815 1 (P. anubis) 51. 24LMM2081215 (P. anubis)									
52 73.000000000000000000000000000000000000	52. 24SNM5151115 (P. anubis)				- 11					
	54. 27MKM5180715 (P. cynocephalus)		1				1			
	56. 30LMF5190415 (P. anubis) 57. 32E2110407 (P. anubis)									
	58. 32LMM2190317 (P. anubis) 59. 33M8120407 (P. anubis)		1							
C:: 2.: 2:: 2:: 2:: 2:: 2:: 2:: 2:: 2:: 2:: 2:: 2:: 2:: <td>60. 34F2130407 (P. anubis) 61. 34I MM2190217 (C. pygenthrus)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	60. 34F2130407 (P. anubis) 61. 34I MM2190217 (C. pygenthrus)									
GA (15)(11)S) (C. pugenthus) III IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	62. 39M8150407 (P. anubis)									
62 8250007 (F, anubis) 63 925017 (Soff, anubis) 65 925017 (Soff, anubis) 65 925017 (Soff, anubis) 75 925027 (Soff, anubis) 76 925027 (Soff, anubis) 77 925027 (Soff, anubis) 78 925027 (Soff, anubis) 77 925027 (Soff, anubis) 78 92502 (Soff, anubis) 78 92502 (Soff, anubis) 78 925	64. 41SNM1231115 (C. pygerythrus)				1-	-				
62 555 (1007) [2, anubis]	66. 49F8190407 (P. anubis) 67. 502190407 (P. anubis)									
	68. 52F8210407 (P. anubis) 69. 54K9210407 (P. anubis)									
72 CARGEORD (F. anubis) 72 CARGEORD (F. anubis) 74 CARGEORD (F. anubis) 75 CARGEORD (F. anubis) 76 CARGEORD (F. anubis) 77 CARGEORD (F. anubis) 78 CARGEORD (F. anubis) 77 CARGEORD (F. anubis) 78 CARGEORD (F. anubis) 79 CARGEORD (F. anubis) 70 CARGEORD (F. anubis) 71 Tartis 72 CARGEORD (F. anubis) 73 CARGEORD (F. anubis) 74 Tartis 75 CARGEORD (F. anubis) 76 CARGEORD (F. anubis) 77 CARGEORD (F. anubis) 78 CARGEORD (F. anubis) 78 CARGEORD (F. anubis) 78 CARGEORD (F. anubis) 78 CARGEORD (F. anubis) <td< td=""><td>70. 55M2230407 (P. anubis)</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>	70. 55M2230407 (P. anubis)									
24 Decomposition (Control Control Contrel Control Control Control Control Contrel Control Control Contro	72. 60M5250407 (P. anubis)									
20 20 <td< td=""><td>73. 67M8070507 (P. anubis)</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>	73. 67M8070507 (P. anubis)									
72 A202301401/20.7 1 1 1 1 1 87 Photourgalian 2 (C PO0230) 1 1 1 1 1 87 Photourgalian 2 (C PO0230) 1 1 1 1 1 88 Photourgalian 2 (C PO0230) 1 1 1 1 1 88 Photourgalian 2 (C PO0230) 1 1 1 1 1 88 Photourgalian 2 (C PO0230) 1 1 1 1 1 88 Photourgalian 2 (C PO0230) 1 1 1 1 1 88 Photourgalian 2 (C PO0230) 1 1 1 1 1 80 Photourgalian 2 (C PO0230) 1 1 1 1 1 80 Photourgalian 2 (C PO0230) 1 1 1 1 1 80 Photourgalian 2 (C PO0230) 1 1 1 1 1 80 Photourgalian 2 (C PO0230) 1 1 1 1 1 80 Photourgalian 2 (C PO0230) 1 1 1 1 1 91 C C C 2 (PO0230) 1 1 1 1 1 92 <td>75. 69=5090507 (P. anubis) 76. 70M5100507 (P. anubis)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	75. 69=5090507 (P. anubis) 76. 70M5100507 (P. anubis)									
C/2 PID2002000 Image: 1 (Procession) Image: 1 (Procession) S2 CCC 2275 (Procession) Image: 1 (Procession) Image: 1 (Procession) S2 CCC 2275 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession) S3 Samous D1 (Procession) Image: 1 (Procession) Image: 1 (Procession)	77. 74M8160507 (P. anubis) 78. 3DZAKM13280917 (Gorilla gorilla)	1	.1				1	1		
b) Controls (C)	79. Fribourg-Blanc 1 (HMS85227) 80. Fribourg-Blanc 2 (CP003902)	_								
B3. 2017.0200 / LP120267 I </td <td>81. Gnana-051 (CP020365) 82. CDC 2575 (CP020366)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>1</td> <td></td> <td></td> <td></td>	81. Gnana-051 (CP020365) 82. CDC 2575 (CP020366)						1			
Bit Ministry (1700/92/1/) Image: State (1700/92/1/) Im	83. SamoaD 1 (CP002374) 84. CDC-2 1 (HM243495)		1							
82, UAX, 1 (MS2502) N N N N N N N N 90, Chicago (CP00752) N N N N N N N 90, Chicago (CP00752) N N N N N N N 90, Chicago (CP00752) N N N N N N N 92, Mexico A (CP00357) N N N N N N N 92, Mexico A (CP00357) N N N N N N N 92, Mexico A (CP00357) N N N N N N N 95, Sita (CP00357) N N N N N N N N 97, Seattle Nichols 2 (CP00558) N N N N N N N N N 100, Nichols Houston Clone E (CP005560) N N N N N N N N 101, Nichols Houston Clone E (CP005560) N N N N N N	85. SamoaD 2 (HM245777) 86. Bosnia A (CP007548)			1						1
88) 511 (1) 1 </td <td>87. DAL-11 (HM585229) 88. Nichols 1 (AE000520)</td> <td>11-</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	87. DAL-11 (HM585229) 88. Nichols 1 (AE000520)	11-								
91. UL-22 (CV020215) 1	89. SS14 1 (CP000805) 90. Chicago (CP001752)			1 1				101		
93. DAL-1 (CPU0313) 1 1 1 1 1 95. DAL-1 (CPU0313) 1 1 1 1 1 1 96. State 1, CPU0350 1 1 1 1 1 1 1 97. Sattle (CPU0422) 1 1 1 1 1 1 1 1 98. CDC A (CPU0559) 1 1 1 1 1 1 1 1 98. CDC A (CPU0559) 1 1 1 1 1 1 1 1 100. Nichols Houston clone E (CPU0560) 1	91. CDC-2 2 (CP002375) 92. Mexico A (CP003064)			1 1						
Sp. Microls 2 (CP004010) III IIII III IIII IIII IIII IIII IIII IIII IIII IIIIIII IIIIIIII IIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	93. DAL-1 2 (CP003115) 94. Sea 81-4 (CP003679)	11-								
97. Seattle Nichols (CP010422) 1 1 1 1 1 1 97. Seattle Nichols (CP010558) 1 1 1 1 1 98. Chicage Propulation genome (CP010558) 1 1 1 1 1 99. CDC-A (CP010559) 1 1 1 1 1 1 100. Nichols Houston (Cone E (CP010560) 1 1 1 1 1 1 101. Witchis (CP010562) 1 1 1 1 1 1 1 103. Witchis (CP010563) 1<	95. Nichols 2 (CP004010) 96. SS14 2 (CP004011)			1 1					- 'I II 'I	
99. CDC-A (CP010559) III III IIII III IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	97. Seattle Nichols (CP010422) 98. Chicago Population genome (CP010558)	11-								
101. Nichols Houston clone J (CP010561) IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	99. CDC-A (CP010559) 100. Nichols Houston clone E (CP010560)	11-							1 1	
103. UW2288 1 (CP010563) IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	101. Nichols Houston clone J (CP010561) 102. UW074B (CP010562)	11-		1 1						
105. UW 2548 (CP010565) 111 111 111 111 107. Amoy (CP015162) 111 111 111 111 108. PT_SIF0650 1111 111 111 111 109. PT_SIF0571 (CP016045) 1111 111 111 111 110. PT_SIF0571 (CP016045) 1111 111 111 111 111. PT_SIF0877 (CP016047) 1111 111 111 111 112. PT_SIF0954 (CP016050) 1111 111 111 111 113. PT_SIF0954 (CP016050) 1111 111 111 111 115. PT_SIF1020 (CP016051) 1111 111 111 111 115. PT_SIF1020 (CP016052) 1111 111 111 111 117. PT_SIF127 (CP016054) 1111 111 111 111 118. PT_SIF136 (CP016055) 1111 111 111 111 119. PT_SIF140 (CP016057) 1111 111 111 111 121. PT_SIF1516 (CP016055) 1111 111 111 111 122. PT_SIF1140 (CP016057) 1111 111 111 111	103. UW228B 1 (CP010563) 104. UW228B 2 (CP010564)	-		1 1						
107. Amoy (CP01562) 111 11 11 11 108. PT SFR057 (CP016045) 111 111 11 111 109. PT SFR057 (CP016045) 111 111 111 111 111. PT SFR0877 (CP016045) 111 111 111 111 112. PT SFR0877 (CP016047) 111 111 111 111 113. PT SFR0954 (CP016050) 1111 111 111 111 114. PT SFR0877 (CP016048) 1111 111 111 111 115. PT SFR020 (CP016051) 1111 111 111 111 115. PT SFR020 (CP016052) 1111 111 111 111 117. PT SFR020 (CP016053) 1111 111 111 111 118. PT SFR127 (CP016054) 1111 111 111 111 119. PT SFR140 (CP016055) 1111 111 111 111 119. PT SFR142 (CP016055) 1111 111 111 111 121. PT SFR142 (CP016055) 1111 111 111 111 122. PT SFR142 (CP016057) 1111 111 111 111	105. UW254B (CP010565) 106. UW391B (CP010566)			 						
109. PT StrO251 (CP016046) 111. PT Str0877 (CP016047) 111. PT Str0877 (CP016048) 111. PT Str0877 (CP016048) 111. PT Str0877 (CP016048) 111. PT Str0877 (CP016048) 113. PT Str0954 (CP016050) 111. PT Str0877 (CP016048) 111. PT Str0877 (CP016049) 114. PT Str002 (CP016050) 111. PT Str0877 (CP016048) 111. PT Str0877 (CP016051) 115. PT Str002 (CP016052) 111. PT Str0877 (CP016054) 111. PT Str1717 (CP016054) 117. PT Str127 (CP016054) 111. PT Str142 (CP016055) 111. PT Str142 (CP016055) 118. PT Str142 (CP016056) 111. PT Str142 (CP016056) 111. PT Str142 (CP016056) 119. PT Str142 (CP016056) 111. PT Str142 (CP016056) 111. PT Str142 (CP016056) 121. PT Str195 (CP016058) 111. PT Str142 (CP016057) 111. PT Str142 (CP016057) 122. PT Str1156 (CP016058) 111. PT Str142 (CP016059) 111. PT Str142 (CP016059) 123. PT Str185 (CP016060) 111. PT Str142 (CP016059) 111. PT Str142 (CP016059) 124. PT Str156 (CP016059) 111. PT Str142 (CP016050) 111. PT Str142 (CP016050) 125. PT Str1280 (CP016069) 111. PT Str142 (CP016059) 111. PT Str142 (CP016059) 125. PT Str1280 (CP016069) 111. PT Str142 (CP016050) 111. PT Str142 (CP016050) 126. PT Str	107. Amoy (CP015162) 108. PT S(F0697 (CP016045)									
111. PT StP0877 (CP016048) 111. 1 111. 1 111. 1 112. PT StP098 (CP016049) 111. 1 111. 1 111. 1 113. PT StP098 (CP016050) 111. 1 111. 1 111. 1 114. PT StP098 (CP016050) 111. 1 111. 1 111. 1 115. PT StP020 (CP016051) 111. 1 111. 1 111. 1 116. PT StP1020 (CP016052) 111. 1 111. 1 111. 1 117. PT StP127 (CP016054) 111. 1 111. 1 111. 1 118. PT StP135 (CP016053) 111. 1 111. 1 111. 1 119. PT StP140 (CP016056) 111. 1 111. 1 111. 1 119. PT StP140 (CP016056) 111. 1 111. 1 111. 1 120. PT StP140 (CP016056) 111. 1 111. 1 111. 1 121. PT StP156 (CP016057) 111. 1 111. 1 111. 1 122. PT StP156 (CP016057) 111. 1 111. 1 111. 1 123. PT StP183 (CP016060) 111. 1 111. 1 111. 1 124. PT StP182 (CP0160651) 111. 1 111. 1 111. 1 125. PT StP120 (CP016062) 111. 1 111. 1 111. 1	109. PT_SIF0751 (CP016046) 110. PT_SIF0857 (CP016047)									
113. PT Str0954 (2P016050) 111. 1 111. 1 111. 1 114. PT Str002 (2P016051) 111. 1 111. 1 111. 1 115. PT Str002 (2P016052) 111. 1 111. 1 111. 1 116. PT Str002 (2P016052) 111. 1 111. 1 111. 1 117. PT Str127 (2P016054) 111. 1 111. 1 111. 1 118. PT Str127 (2P016054) 111. 1 111. 1 111. 1 119. PT Str140 (2P016056) 111. 1 111. 1 111. 1 119. PT Str140 (2P016056) 111. 1 111. 1 111. 1 121. PT Str140 (2P016056) 111. 1 111. 1 111. 1 122. PT Str140 (2P016057) 111. 1 111. 1 111. 1 121. PT Str156 (2P016057) 111. 1 111. 1 111. 1 122. PT Str156 (2P016057) 111. 1 111. 1 111. 1 123. PT Str156 (2P016057) 111. 1 111. 1 111. 1 125. PT Str120 (2P016051) 111. 1 111. 1 111. 1 125. PT Str1220 (2P016051) 111. 1 111. 1 111. 1 125. PT Str1220 (2P016052) 111. 1 111. 1 111. 1	111. PT_SIF0877 (CP016048) 112. PT_SIF0908 (CP016049)									
115. PT_SIF1020 (CP016052) 111. PT_SIF127 (CP016054) 117. PT_SIF127 (CP016055) 111. PT_SIF127 (CP016055) 118. PT_SIF140 (CP016056) 111. PT_SIF140 (CP016057) 120. PT_SIF140 (CP016057) 111. PT_SIF140 (CP016057) 121. PT_SIF156 (CP016058) 111. PT_SIF156 (CP016059) 122. PT_SIF156 (CP016057) 111. PT_SIF160 (CP016059) 123. PT_SIF183 (CP016060) 111. PT_SIF183 (CP016061) 125. PT_SIF200 (CP016052) 111. PT_SIF1220 (CP016054) 126. PT_SIF220 (CP016054) 111. PT_SIF1220 (CP016054) 127. PT_SIF220 (CP016054) 111. PT_SIF1220 (CP016054) 128. PT_SIF220 (CP016056) 111. PT_SIF1220 (CP016056) 129. PT_SIF1220 (CP016056) 111. PT_SIF1220 (CP016056) 129. PT_SIF1230 (CP016056) 111. PT_SIF1230 (CP016056) 129. PT_SIF1230 (CP016056) 111. PT_SIF1230 (CP016056) 129. PT_SIF1230 (CP016056) 111. PT_SIF1230 (CP016057) 129. PT_SIF1230 (CP016056) 111. PT_SIF1230 (CP016057) 129. PT_SIF1230 (CP016056) 111. PT_SIF1230 (CP016056) 131. PT_SIF1230 (CP016056) 111. PT_SIF1230 (CP0160567) 131. PT_SIF1230 (CP0160567) 111. PT_SIF1230 (CP0160567)	113. PT_SIF0954 (CP016050) 114. PT_SIF1002 (CP016051)									
117. PT_SIF127 (CP016055) 119. PT_SIF140 (CP016055) 120. PT_SIF140 (CP016057) 121. PT_SIF157 (CP016057) 122. PT_SIF157 (CP016057) 123. PT_SIF157 (CP016057) 124. PT_SIF157 (CP016057) 125. PT_SIF120 (CP016051) 125. PT_SIF120 (CP016051) 126. PT_SIF1220 (CP016051) 127. PT_SIF1220 (CP016053) 128. PT_SIF1220 (CP016053) 129. PT_SIF1220 (CP016056) 129. PT_SIF1220 (CP016056) 120. PT_SIF	115. PT_SIF1020 (CP016052) 116. PT_SIF1063 (CP016053)									
119 pr_Siria0 (2p0)6555 120 pr_Siria0 (2p0)6557 121 pr_Siria0 (2p0)6557 122 pr_Siria0 (2p0)6557 123 pr_Siria0 (2p0)6559 124 pr_Siria0 (2p0)6559 125 pr_Siria0 (2p0)6561 126 pr_Siria0 (2p0)6561 127 pr_Siria0 (2p0)6563 128 pr_Siria0 (2p0)6563 129 pr_Siria0 (2p0)6563 131 pr_Siria0 (2p0)6669 131 pr_Siria0 (2p0)6669	117. PT_SIF1127 (CP016054) 118. PT_SIF1135 (CP016055)	in the second seco								
121 pT_SIF1462 (Ep016053) 122 pT_SIF145 (Ep016059) 123 pT_SIF167 (Ep016050) 124 pT_SIF167 (Ep016050) 125 pT_SIF120 (Ep016051) 127 pT_SIF120 (Ep016053) 128 pT_SIF222 (Ep016054) 129 pT_SIF1220 (Ep016055) 129 pT_SIF1230 (Ep016055)	119. PT_SIF1140 (CP016056) 120. PT_SIF1142 (CP016057)	iiii								
122 pT_SIF183 (Ep016667) 123 pT_SIF128 (CP016663) 124 pT_SIF128 (CP016663) 125 pT_SIF128 (CP016663) 126 pT_SIF128 (CP016663) 127 pT_SIF128 (CP016665) 128 pT_SIF1278 (CP016665) 129 pT_SIF1278 (CP016665) 129 pT_SIF1278 (CP016665) 129 pT_SIF1278 (CP016665) 129 pT_SIF128 (CP016665) 129 pT_SIF128 (CP016665) 129 pT_SIF128 (CP016665) 129 pT_SIF128 (CP016665) 129 pT_SIF128 (CP016667) 129 pT_SIF128 (CP01667) 129 pT_SIF128 (CP01667) 129 pT_SIF128 (CP01667) 129 pT_SIF128 (CP01667) 129 pT_SIF128 (CP0	121. PT_SIF1156 (CP016058)									
125 PT_SIF1200 (CP0160652) 126 PT_SIF1220 (CP016063) 127 PT_SIF1220 (CP016064) 129 PT_SIF1200 (CP0160664) 129 PT_SIF1200 (CP0160666) 130 PT_SIF1220 (CP0160667) 131 PT_SIF1220 (CP016067) 131	123, PT_SIF1183 (CP016060) 124, PT_SIF1186 (CP016060)									
127 PT-SIE1252 (CP016062) 128 PT-SIE1253 (CP016062) 129 PT-SIE1263 (CP0160667) 131 PT-SIE1289 (CP0160667) 131 PT-SIE1289 (CP0160687) 131 PT-SIE1289 (CP0160688) 131	125. PT_SIF1200 (CP016062) 126. PT_SIF1202 (CP016062)									
120. PT_SIF128 (CP016066) IIIIIII IIIIIIIIIIIIIIIIIIIIIIIIII	127. PT_SIF1252 (CP016064)									
130. PT_SIF1290 (CP016068) III III III III III III III III III I	120. PT_SIF1201 (CP016066) 129. PT_SIF1278 (CP016066)	1111								
	131. PT_SIF1299 (CP016068) 132. PT_SIF1299 (CP016069)									

95 Figure S1. Alignment of NHP and human TPA, TEN, and TPE sequence data of TP0548

- 96 locus. Sequences of *TP* of NHP origin show only one part of the sequence where most of the
- 97 nucleotide variation is found (light green box). This distinguishes them from human TPA
- 98 (blue box) and TEN (red box) as well as human yaws-causing strains (TPE; dark green box),
- 99 where we have three and two variable regions, respectively. GenBank accession numbers for
- all NHP samples are listed in Table S5.





104 bootstrap replicates were performed and bootstrap values from 80-100% are highlighted as

105 light blue circles of respective size. The bar indicates substitutions per site.



107 Figure S3. MP tree for the *TP0548* locus of the Tanzanian strains of NHP origin. 1,000

- 108 bootstrap replicates were performed and bootstrap values from 80-100% are highlighted as
- 109 light blue circles of respective size. Gaps were coded as fifth character.



111 Figure S4. ML tree for the *TP0488* locus of the Tanzanian strains of NHP origin. 1,000

- bootstrap replicates were performed and bootstrap values from 80-100% are highlighted as
- 113 light blue circles of respective size. The bar indicates substitutions per site.



114 Figure S5. MP tree for the *TP0488* locus of the Tanzanian strains of NHP origin. 1,000

- bootstrap replicates were performed and bootstrap values from 80-100% are highlighted as
- 116 light blue circles of respective size. Gaps were coded as fifth character.



118 Figure S6. Rooted MP tree based on the concatenated sequences used for MLST

119 (*TP0548* and *TP0488*). The tree is based on 1,773 nts and 1,000 bootstrap replicates.

Technical appendix - 16 -

Bootstrap values from 80-100% are highlighted as light blue circles of respective size. NHP
species and/or GenBank accession numbers of published strains are provided in parentheses
following the name of the strain. In all cases were the species is not mentioned, sequences are
from *TP* of human origin. Blue=subsp. *pallidum*, green=subsp. *pertenue*, red=subsp. *endemicum*. The pathogen causing rabbit syphilis, *Treponema paraluisleporidarum* ecovar *Cuniculus* strain Cuniculi A, is used as an outgroup. The bar refers to substitutions per site.
Gaps were coded as fifth character.

127



128 129 Figure S7. Camera trap photos of Western lowland gorillas with severe ulcerative skin 130 lesions at Odzala-Kokoua National Park, Republic of the Congo, in 2017. Fecal samples 131 analyzed in this study were collected in the same area and originated from animals with 132 severe facial lesions. The images are not covered by the CC BY license. Image credits to 133 African Parks (www.african-parks.org). All rights reserved, used with permission. When 134 using the images online and across all platforms, African Parks and the photographer must be 135 named appropriately, and African Parks and Odzala--Kokoua National Park should be 136 mentioned using the following social media handles: on Twitter: @AfricanParks, Instagram: 137 (a)AfricanParksNetwork and Facebook: (a)AfricanParks.Odzala--Kokoua National Park.

Chuma et al.

Technical appendix

139 Supplementary References

- 140 37. Grillová, L., Noda, A. A., Lienhard, R., Blanco, O., Rodríguez, I. & Šmajs, D.
- 141 Multilocus sequence typing of *Treponema pallidum* subsp. *pallidum* in Cuba from 2012
- to 2017. J Infect Dis, jiy604 (2018).
- 143 38. Arora, N., Schuenemann, V., Jager, G., Peltzer, A., Seitz, A., Herbig, A., et al. Origin of
 144 modern syphilis and emergence of a pandemic *Treponema pallidum* cluster. *Nat*145 *Microbiol* 2, 16245 (2016).
- 146 39. Flasarová, M., Pospíšilová, P., Mikalová, L., Vališová, Z., Dastychová, E., Strnadel, R.,
- 147 et al. Sequencing-based molecular typing of treponema pallidum strains in the Czech
- 148 Republic: all identified genotypes are related to the sequence of the SS14 strain. *Acta*
- 149 *Derm Venereol* **92**, 669-674 (2012).
- 40. Čejkova, D., Zobaníková, M., Chen, L., Pospíšilová, P., Strouhal, M., Qin, X., et al.
 Whole genome sequences of three *Treponema pallidum* ssp. *pertenue* strains: yaws and
 syphilis treponemes differ in less than 0.2% of the genome sequence. *PLoS Negl Trop*
- 153 *Dis* **6**, e1471 (2012).
- 154 41. Zobaníková, M., Strouhal, M., Mikalová, L., Čejkova, D., Ambrožová, L., Pospíšilová,
- P., et al. Whole genome sequence of the *Treponema* Fribourg-Blanc: unspecified simian
 isolate is highly similar to the yaws subspecies. *PLoS Negl Trop Dis* 7, e2172 (2013).
- 157 42. Strouhal, M., Mikalová, L., Haviernik, J., Knauf, S., Bruisten, S., Noordhoek, G. T., et
- al. Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue*
- 159 from Indonesia: Modular structure of several treponemal genes. *PLoS Negl Trop Dis* **12**,
- 160 e0006867 (2018).
- 161 43. Marks, M., Fookes, M., Wagner, J., Butcher, R., Ghinai, R., Sokana, O., et al.
- 162 Diagnostics for yaws eradication: Insights from direct next-generation sequencing of
- 163 cutaneous strains of *Treponema pallidum*. *Clin Infect Dis* **66**, 818-824 (2018).