MLST typing of *Treponema pallidum* subsp. *pallidum* in the Czech Republic during 2004-2017: Clinical isolates belonged to 25 allelic profiles and harbored 8 novel allelic variants

Eliška Vrbová[1], Linda Grillová[1¤a], Lenka Mikalová[1], Petra Pospíšilová[1], Radim Strnadel[2], Eliška Dastychová[3], Martina Kojanová[4], Miluše Kreidlová[5], Daniela Vaňousová[6], Filip Rob[6], Přemysl Procházka[7], Alena Krchňáková[3], Vladimír Vašků[3], Vladana Woznicová[3], Monika Dvořáková Heroldová[3], Ivana Kuklová[4], Hana Zákoucká[8], David Šmajs[1]*

1 Department of Biology, Faculty of Medicine, Masaryk University, Brno, Czech Republic, 2 Department of Dermatovenerology, Faculty Hospital Brno, Brno, Czech Republic, 3 Department of Medical Microbiology, Faculty of Medicine, St. Anne's Hospital and Masaryk University, Brno, Czech Republic, 4 Department of Dermatology, 1st Faculty of Medicine, Charles University in Prague, Prague, Czech Republic, 5 Institute of Medical Biochemistry and Laboratory Diagnostics of the General University Hospital, The First Faculty of Medicine of Charles University in Prague, Prague, Czech Republic, 6 Department of Dermatovenerology, 2nd Faculty of Medicine, Charles University in Prague, Prague, Czech Republic, 7 Outpatient STI Clinic Medicentrum, Prague, Czech Republic, 8 National Reference Laboratory for Diagnostics of the Syphilis, National Institute for Public Health, Prague, Czech Republic

¤a Current address: Biology of Spirochetes Unit, Institut Pasteur, Paris, France
* dsmajs@med.muni.cz

OPEN ACCESS

Abstract

A recently introduced Multilocus Sequence Typing scheme for *Treponema pallidum* subsp. *pallidum* was applied to clinical samples collected from 2004 to 2017 from the two largest cities (Prague and Brno) in the Czech Republic. Altogether, a total of 675 samples were tested in this study and 281 of them were found PCR-positive for treponemal DNA and typeable. Most of the typed samples (n = 281) were swabs from primary or secondary syphilis lesions (n = 231), and only a minority were whole blood or tissue samples (n = 50). Swab samples from patients with rapid plasma regain (RPR) values of 1–1024 were more frequently PCR-positive (84.6%) compared to samples from patients with non-reactive RPR test (46.5%; p-value = 0.0001). Out of 281 typeable samples, 136 were fully-typed at all TP0136, TP0548, and TP0705 loci. Among the fully and partially typed samples, 25 different allelic profiles were identified. Altogether, eight novel allelic variants were found among fully (n = 5) and partially (n = 3) typed samples. The distribution of TPA allelic profiles identified in the Czech Republic from 2004 to 2017 revealed a dynamic character with allelic profiles disappearing and emerging over time. While the number of samples with the A2058G mutation was seen to increase (86.7% in 2016/2017), the number of samples harboring the A2059G mutation was found to have decreased over time (3.3% in 2016/2017). In addition, we found several allelic profile associations with macrolide resistance or susceptibility, the gender of patients, as well as patient residence.

Introduction

Treponema pallidum subsp. *pallidum* (TPA) is an unusual bacterial pathogen [1] that causes syphilis, a chronic venereal disease in humans. Each year, there are about 5.6 million new cases worldwide [2–3]. In recent years, there have been approximately 700–800 cases of syphilis per year in the Czech Republic, a country having about 10.5 million citizens (data provided by the Institute of Health Information and Statistics of the Czech Republic). Most of the cases were among men having sex with men (MSM), promiscuous individuals, or sex-workers [4].

While a syphilis diagnosis is mainly based on anamnestic data, clinical findings, and results of serological tests, PCR detection and molecular typing can also be used to diagnose syphilis, especially in the cases with negative serology [5]. Moreover, PCR and molecular typing have proved to be useful in cases where the infection was caused by *T. pallidum* subsp. *endemicum* [6–10].

In the last two decades, molecular typing of TPA isolates has mapped several thousand clinical isolates from different countries all over the world. During that time, the original typing technique [11] was continually improved for better resolution [12–13]. Sequencing-based molecular typing (SBMT) was introduced in 2006 [14] and was recently enhanced [15]. The recently reported treponemal MLST typing system analyses the TP0136, TP0548, and TP0705 loci, and has shown a resolution power of about 30% of whole genome sequences [15]. While MLST typing is a simple method revealing a high portion of genetic variability, whole genome sequencing is a more complicated method where just a subset of samples is usually characterized. In addition, 23S rDNA can be used to find mutations that cause macrolide-resistance, however, this analysis is not part of MLST. MLST typing has already been used to examine several hundred clinical samples from Switzerland [15], France [15, 16], and Cuba [17]. Moreover, a public treponemal MLST database for storage and analyses of typing data has been established [18].

In this communication, we performed molecular typing of TPA, using a newly introduced MLST typing system, on samples collected in the Czech Republic from 2004–2017. Some of the samples had been previously typed with the SBMT typing scheme [4, 19].

Materials and methods

Clinical material

Samples were collected during years 2004–2017 from two clinical departments in Brno (the Department of Dermatovenereology, St. Anne´s Faculty Hospital and the Department of Medical Microbiology, Faculty of Medicine, St. Anne´s Hospital and Masaryk University) and from two clinical departments in Prague (the Department of Dermatovenereology, 1st Faculty of Medicine, Charles University and the National Reference Laboratory for Diagnostics of Syphilis, National Institute for Public Health). Clinical data included patient age, gender, type of clinical material, results of serology, primary diagnosis, sexual orientation, and HIV status. Serological tests included *T. pallidum* particle agglutination (TPPA) or *T. pallidum* hemagglutination (TPHA) tests, the rapid plasma regain (RPR) test, and enzyme-linked immunosorbent assay (ELISA) or Western blot for IgM and IgG depending on the source hospital. Serological tests were provided by OMEGA Diagnostics (Reinbek, Germany), TEST-LINE (Brno, Czech Republic), and MARDX (Carlsbad, CA, USA). In total, 675 samples were examined by PCR. Patient characteristics for those with typeable samples (i.e., samples positive for at least one of typing loci) are given in Table 1.

Isolation of DNA

DNA was isolated as described previously [4] using QIAamp DNA Blood Mini Kits and a DNeasy Blood & Tissue Kits (Qiagen, Hilden, Germany).

Table 1. Clinical characteristics of patients with typed TPA samples.

Clinical characteristics of patients	Patients (n = 269)
Mean age (men/women)	44.2 (0–71)/24.9 (0–38)
Sex, n (%)	M 244 (90.71); W 25 (9.29)
Serology[a]	
TPPA/TPHA (%)	242 P (89.96); 7 N (2.6); 20 n.d. (7.43)
RPR	216 P (80.3); 47 N (17.47); 6 n.d. (2.23)
≤ 1:16	118
≥ 1:32	91
Positive without titer value	7
Diagnosis	
Primary syphilis stage	165
Secondary syphilis stage	36
Congenital syphilis	2
Undetermined syphilis stage	66
Material	**Samples (n = 281)[b]**
No. of swabs	231 (82.2%)
No. of whole blood samples	47 (16.7%)
No. of tissue samples	3 (1.1%)

P, positive; N, negative; n.d, not determined; M, men; W, women

[a]Serology is presented for the most frequently used tests.

[b]There were no differences found in multiple samples collected from one patient.

https://doi.org/10.1371/journal.pone.0217611.t001

PCR amplification

Typing loci and the 23S rDNA locus were amplified as described previously [4, 15, 16, 20]. The PCR mixture (final volume of 25 μl) contained in the first step 16.3 μl of water, 2 μl of a 2.5 mM deoxynucleotide triphosphate (dNTP) mixture, 5 μl of 5x PS GXL buffer, 0.095 μl of each primer (100 pmol/μl), 0.5 μl of PrimeSTAR GXL polymerase (Takara Bio Europe, France), and 1 μl of isolated DNA. When PCR results were negative, the amount of isolated DNA per sample was increased: 11.5 μl of water, 2.5 μl of ThermoPol Reaction buffer, 0.5 μl of a 10 mM dNTP mixture, 0.25 μl of each primer (100 pmol/μl), 0.1 μl Taq polymerase (5,000 U/ml; New England BioLabs, Ipswich, MA) and 10 μl of DNA. In the first step, using 1 μl of DNA and GXL polymerase, PCR amplification was performed under the following condition: 94˚C (1 min); 98˚C (10 s), 68˚C (15 s; −1.0˚C per cycle), 68˚C (1 min and 45 s) for 8 cycles; 98˚C (10 s), 61˚C (15 s), 68˚C (1 min and 45 s) for 35 cycles; and 68˚C (7 min). When using 10 μl of DNA and *Taq* polymerase, these following conditions were used: 94˚C (1 min); 94˚C (20 s), 55˚C (20 s; −1.0˚C per cycle), 72˚C (1 min and 45 s) for 8 cycles; 94˚C (20 s), 48˚C (20 s), 72˚C (1 min 45 s) for 35 cycles; and 72˚C (7 min). The mixture for the second step was the same for both versions of the first step and the final volume (25 μl) for one reaction contained: 20.5 μl of water, 2.5 μl of ThermoPol Reaction buffer, 0.5 μl of a 10 mM dNTP mixture, 0.25 μl of each primer (100 pmol/μl), 0.1 μl *Taq* polymerase (5,000 U/ml; New England BioLabs, Ipswich, MA) and 1 μl of PCR product from the first step. PCR were performed under following conditions: 94˚C (1 min); 94˚C (30 s), 48˚C (30 s), 72˚C (1 min and 15 s) for 40 cycles; and 72˚C (7 min). DNA of TPA strain Nichols (5 pg/μl) was used as a positive control; distilled water was used as a negative control. A list of all primers used for nested PCR is shown in S1 Table [15]. PCR products were purified using QIAquick PCR Purification Kits (Qiagen, Hilden, Germany) according to the manufacturer´s instruction.

Sequencing and sequence analysis

We used Sanger sequencing performed at GATC Biotech AG (Constance, Germany; Eurofins Genomics Company). Analyses of the sequences were performed using Lasergene software (DNASTAR v.7.1.0; DNASTAR, Madison, WI, USA). Sequences were uploaded to the PubMLST database of *Treponema pallidum* subsp. *pallidum* [18] and allelic profiles were automatically assigned. Sequences of 23S rRNA genes encoding macrolide resistance or susceptibility were evaluated at positions corresponding to positions 2058 and 2059 in the 23S rRNA gene of *Escherichia coli* (accession no. V00331), where the A for G substitution has been shown to cause macrolide resistance [20–22]. These positions were carefully analysed to check for possible mix of wild-type and mutant sequences. Sequences were obtained by Sanger sequencing. Alleles encoding resistance were marked A2058G or A2059G depending on the site of substitution.

Statistical methods

Correlations of characteristics of clinical samples with allelic profiles were performed using the two-sided Fisher´s exact test, and statistical significance was set at $p < 0.05$. Statistical analyses were performed using STATISTICA software v.12 (StatSoft, Tulsa, OK, USA).

Ethics statement

This study was approved by the ethics committee of the Faculty of Medicine, Masaryk University (5G/2017). All patients provided written informed consent.

Results

We examined 675 clinical samples collected from 2004–2017 from four hospitals in the two largest cities in the Czech Republic (2 hospitals in Brno and 2 hospitals in Prague). While all samples were tested against TP0705, TP0136, and TP0548 between 2014–2017, data for TP0136 and TP0548 came from clinical samples collected between 2004 and 2013 as part of previous studies [4, 19] and were retested in locus TP0705. We found 281 samples to be positive and typeable, i.e., at least one locus TP0136, TP0548, or TP0705, was amplified and sequenced. The majority of the typed samples were swabs from primary or secondary syphilis lesions (n = 231) and the rest were from whole blood samples (n = 47) and tissue samples (n = 3) taken *post mortem*. Most of the samples (86.48%) belonged to the SS14-like genetic group, while only 2.13% belonged to the Nichols-like genetic group. The remaining 32 samples (11.39%) were not classified as SS14-like or Nichols-like because of selective positivity for locus TP0705 which does not contain informative sites for discriminating between these two genetic groups. Clinical characteristics of patients are summarized in Table 1.

Swab samples from patients with RPR values from 1–1024 were more frequently PCR-positive (84.6%) compared to samples from patients with non-reactive RPR test (46.5%; p-value = 0.0001). No such difference was observed among whole blood samples, however, whole blood samples represented a minority of samples used in this study (16.7%; 47 out of 281).

Out of 281 typeable samples, 136 were fully-typed at the TP0136, TP0548, and TP0705 loci. Among the fully typed samples, 16 different allelic profiles were found, and the partially typed samples (n = 145) revealed 18 different allelic profiles (listed in S2 Table). Since nine allelic profiles were identified in both fully and partially typed samples, the total number of different allelic profiles identified in this study was 25. In fully-typed samples, five novel allelic variants were identified including four in TP0548 and one in TP0705. Novel allele variants (n = 3) were

Table 2. Allelic profiles identified among fully-typed samples (n = 136).

Sequence type[1]	Allelic profile[2]	23S rDNA[3] (no. of samples)	Genetic group [23].	No. of samples (%)
1	1.3.1.	S(1)/R8(44)/X(11)	SS14-like	56 (41.2)
3	1.1.8.	S(26)/R8(2)	SS14-like	28 (20.6)
2	1.1.1.	S(8)/R8(11)/X(1)	SS14-like	20 (14.7)
25	1.26.1.[4]	R8(10)/X(2)	SS14-like	12 (8.8)
11	1.1.3.	R9(5)	SS14-like	5 (3.7)
44	1.36.1.[4]	S(4)	SS14-like	4 (2.9)
26	9.7.3.	R8(1)/X(1)	Nichols-like	2 (1.5)
48	17.1.1.[4]	S(1)	SS14-like	1 (0.7)
46	1.31.1.[5]	R8(1)	SS14-like	1 (0.7)
43	1.28.1.[5]	R8(1)	SS14-like	1 (0.7)
45	4.1.1.	R8(1)	SS14-like	1 (0.7)
41	1.29.1.[5]	R8(1)	SS14-like	1 (0.7)
50	1.32.1.[5]	S(1)	SS14-like	1 (0.7)
42	1.1.16.[5]	R9(1)	SS14-like	1 (0.7)
7	1.4.1.	R8(1)	SS14-like	1 (0.7)
49	18.1.1.[4]	S(1)	SS14-like	1 (0.7)

[1] According PubMLST database of *Treponema pallidum* subsp. *pallidum* [18].

[2] Allelic profiles based on sequences of TP0136, TP0548, and TP0705 [15].

[3] Locus encoding resistance to macrolide antibiotics: S = sensitive, R8 = A2058G mutation, R9 = A2059G mutation, X = undetermined. Both A2058G and A2059G mutations result in resistance to macrolide antibiotics.

[4] Newly identified profiles, with known alleles from previous studies [4, 19]. Original description: 1.26. = SU5, 1.36. = SU7, 17.1. = U2S, 18.1. = U1S.

[5] Newly identified allelic profiles, with novel alleles.

https://doi.org/10.1371/journal.pone.0217611.t002

also identified among partially typed samples and included one in TP0548 and two in TP0705. Allelic profiles identified among fully-typed samples are summarized in Table 2, rest of identified variants are presented in S2 Table. The new allelic variants identified in this study are shown in Fig 1. Among both fully and partially typed samples, the highest amplification efficiency was found for locus 23S rDNA, which was not used for typing (positive in 233 samples), followed by TP0705 (typed in 224 samples), TP0548 (typed in 208 samples), and TP0136 (typed in 204 samples).

A phylogenetic analysis of allelic profiles found among fully typed samples is shown in Fig 2. The allelic profile 9.7.3. corresponds to the allelic profile of the Nichols-like strain, all other allelic profiles corresponded to the allelic profiles of the SS14-like strains. Except for the allelic profile 18.1.1., all identified allelic profiles of the SS14-like strains were highly related.

The distribution of TPA allelic profiles identified in the Czech Republic from 2004 to 2017 is shown in Fig 3. Whereas allelic profile 1.1.8. (which was found to be associated with susceptibility to macrolides) reached its highest proportion in 2010 and 2011 and then disappeared in 2017, there were other allelic profiles that circulated during all tested years (e.g., 1.1.1.) or even increased (e.g., 1.3.1.). In 2012, a new allelic profile (1.26.1.) emerged and persisted until the end of the test period in 2017. Other allelic profiles seem to have random occurrence indicating the dynamic nature of circulating allelic profiles in particular populations. The most frequent was allelic profile 1.3.1. (41.2%) followed by allelic profiles 1.1.8. (20.6%), and 1.1.1. (14.7%). Moreover, as the number of samples increased (13 in 2004/2007, 24 in 2008/2009, 56 in 2010/2011, 63 in 2012/2013, 71 in 2014/2015 and 54 in 2016/2017), the number of allelic profiles also increased (11 in 2004/2007, 7 in 2008/2009, 8 in 2010/2011, 16 in 2012/2013, 17 in

A. New alleles of TP0548

A. Coordinates according to TPA SS14 (CP004011.1)	154	158	167	503	772	773	775	965
TPA_SS14	G	G	G	A	A	A	G	A
TP_0548_27	A	A	G	A	G	A	G	G
TP_0548_28	A	A	G	G	A	G	G	A
TP_0548_29	A	A	A	A	A	A	G	A
TP_0548_31	A	A	G	A	A	A	T	A
TP_0548_32	G	A	G	A	A	A	G	A

B. New alleles of TP0705

Coordinates according to TPA SS14 (CP004011.1)	1498	1828	1844	1948	2111
TPA SS14 and TPA Nichols	C	G	C	C	C
TP_0705_14	T	G	C	C	C
TP_0705_15	C	A	C	C	A
TP_0705_16	C	G	A	A	C

Fig 1. An alignment of the newly identified allelic variants. A. New alleles of TP0548. B. New alleles of TP0705. New alleles are in italics. A complete overview of allelic variants found in this study at the TP0136, TP0548, and TP0705 loci are shown in S1 Fig. No new alleles were found in the TP0136 locus.

https://doi.org/10.1371/journal.pone.0217611.g001

2014/2015 and 13 in 2016/2017), suggesting that the genetic variability of TPA has not been fully described yet.

Macrolide resistance

The prevalence of macrolide resistance causing mutations among TPA isolates in the Czech Republic is shown in Fig 4. While there is an increasing trend in the number of samples containing A2058G mutations (86.7% in 2016/2017), the number of A2059G mutations has decreased over time (3.3% in 2016/2017). This trend was detected in a previous study [4]. In
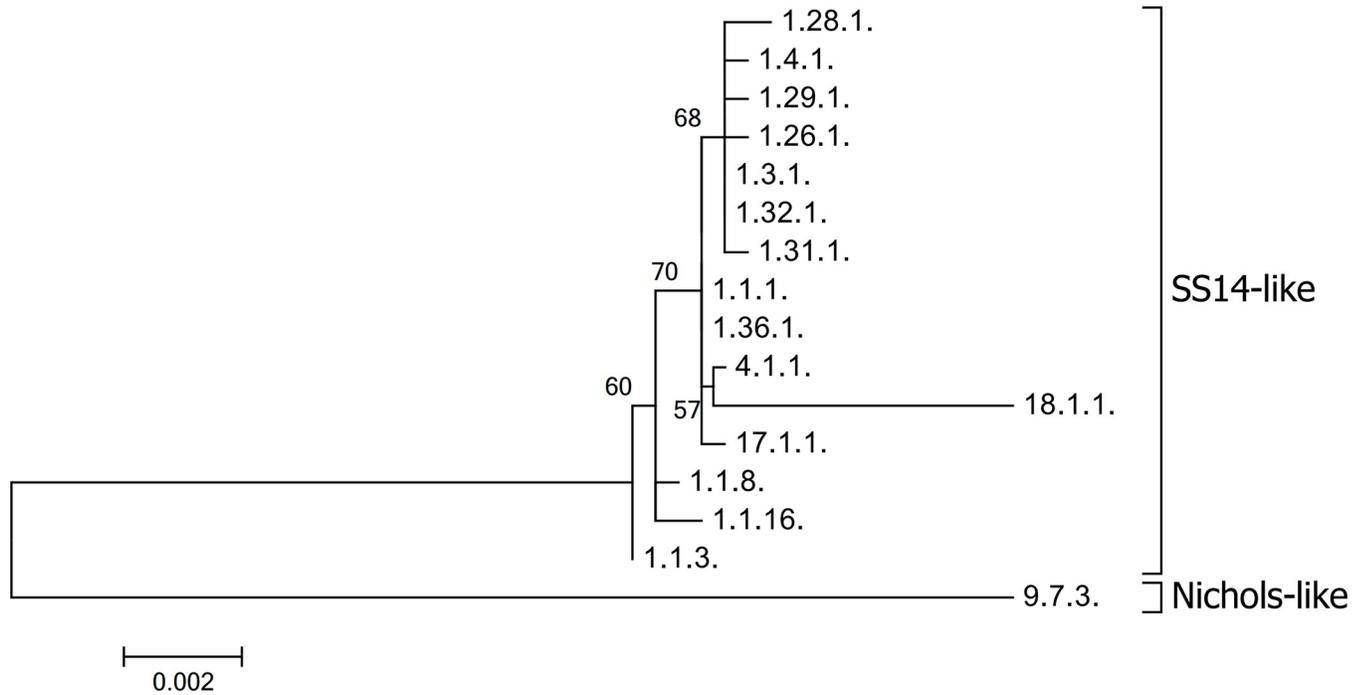
Fig 2. A phylogenetic tree of allelic profiles identified among fully typed samples. The scale shows the number of substitutions per site. Bootstrap values are shown next to branches. The length of concatenated sequences was 2593 bp and contained 95 variable positions. The tree was constructed in MEGA7 [24] using the Maximum Likelihood method [25] with the bootstrap test [26].
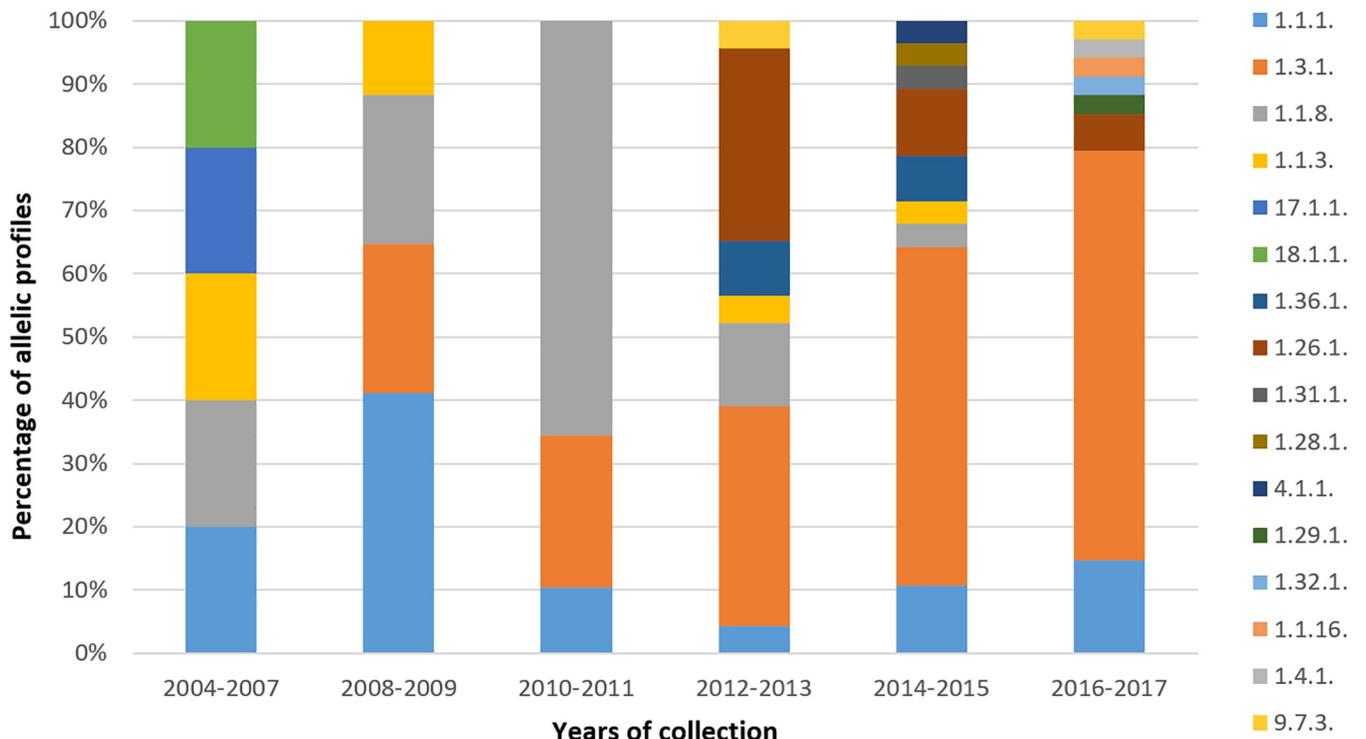
https://doi.org/10.1371/journal.pone.0217611.g002

Fig 3. Distribution of TPA allelic profiles identified in the Czech Republic from 2004 to 2017.

https://doi.org/10.1371/journal.pone.0217611.g003
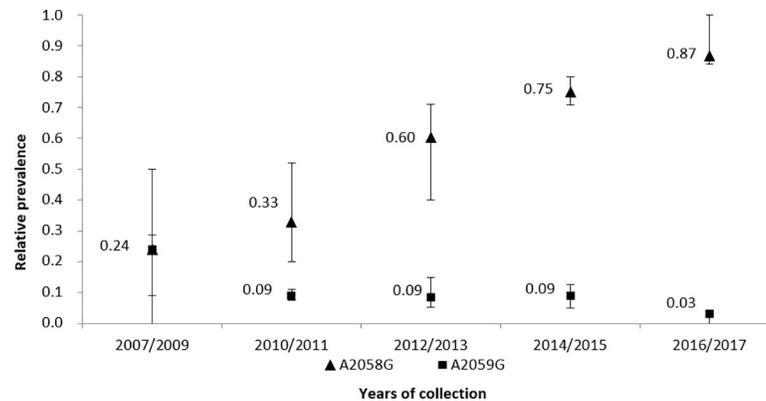
Fig 4. Identified prevalence of macrolide resistance causing mutations in the Czech Republic during the study period. The prevalence over the two- or three-year intervals was calculated as an average; standard errors of the mean are shown. During years 2004 to 2006, none of the mutations were found.

addition, associations of these mutations with different allelic profiles were identified. Allelic profiles 1.26.1. and 1.3.1. were found to be associated with the A2058G mutation ($p = 0.0342$ and $p < 0.0001$, respectively) and allelic profile 1.1.3. was associated with the A2059G mutation ($p < 0.0001$). Moreover, allelic profile 1.1.8. was found to be associated with macrolide susceptibility ($p < 0.0001$).

Associations of allelic profiles with patient characteristics

Besides macrolide resistance or susceptibility, we tested possible associations of allelic profiles with sex, locality, stage of disease and serology. No allelic profile associations with serology (RPR titer) and stage of syphilis were found. Allelic profile 1.1.8. was found more frequently among women than men (p = 0.0143). When compared to Prague, allelic profile 1.26.1. showed a geographical association (p < 0.0001) with the city of Brno.

Discussion

In this study, we examined 675 samples collected from patients suspected of having syphilis between 2004 to 2017 from four clinics, two in Brno and two in Prague. Almost one-half of the samples were PCR-positive and typeable (n = 281, 41.6%). In comparison with other MLST studies [15–17], this study represents largest collection of typed samples. Moreover, a long time period of collecting samples allowed to discover over two dozens of allelic profiles including highly diverse allele in the TP0136 locus.

The majority of the typed samples were swabs from primary or secondary syphilis lesions (82.2%), and the rest were from whole blood or tissue samples (16.7% and 1.1%). While analysis of swab samples resulted in a similar number of fully typed (i.e., samples typed at all three TP0136, TP0548 and TP0705 loci) (n = 128, 55.4%) and partially typed samples (sequenced at least one typing locus) (n = 103, 44.6%); whole blood samples analysis revealed a minority of fully typed samples (n = 6, 12.8%) and a majority of partially typed samples (n = 41, 87.2%). For swab samples, patients that had positive RPR titer were more likely PCR positive compared to RPR-negative patients ($p = 0.0001$) suggesting that the group of patients that contains both RPR-negative and PCR-negative patients is likely to include patients not having syphilis. Altogether, these findings indicate that the swab samples are more suitable for molecular typing of TPA strains and isolates, an observation that was noticed in previous studies [4–5, 27–28].
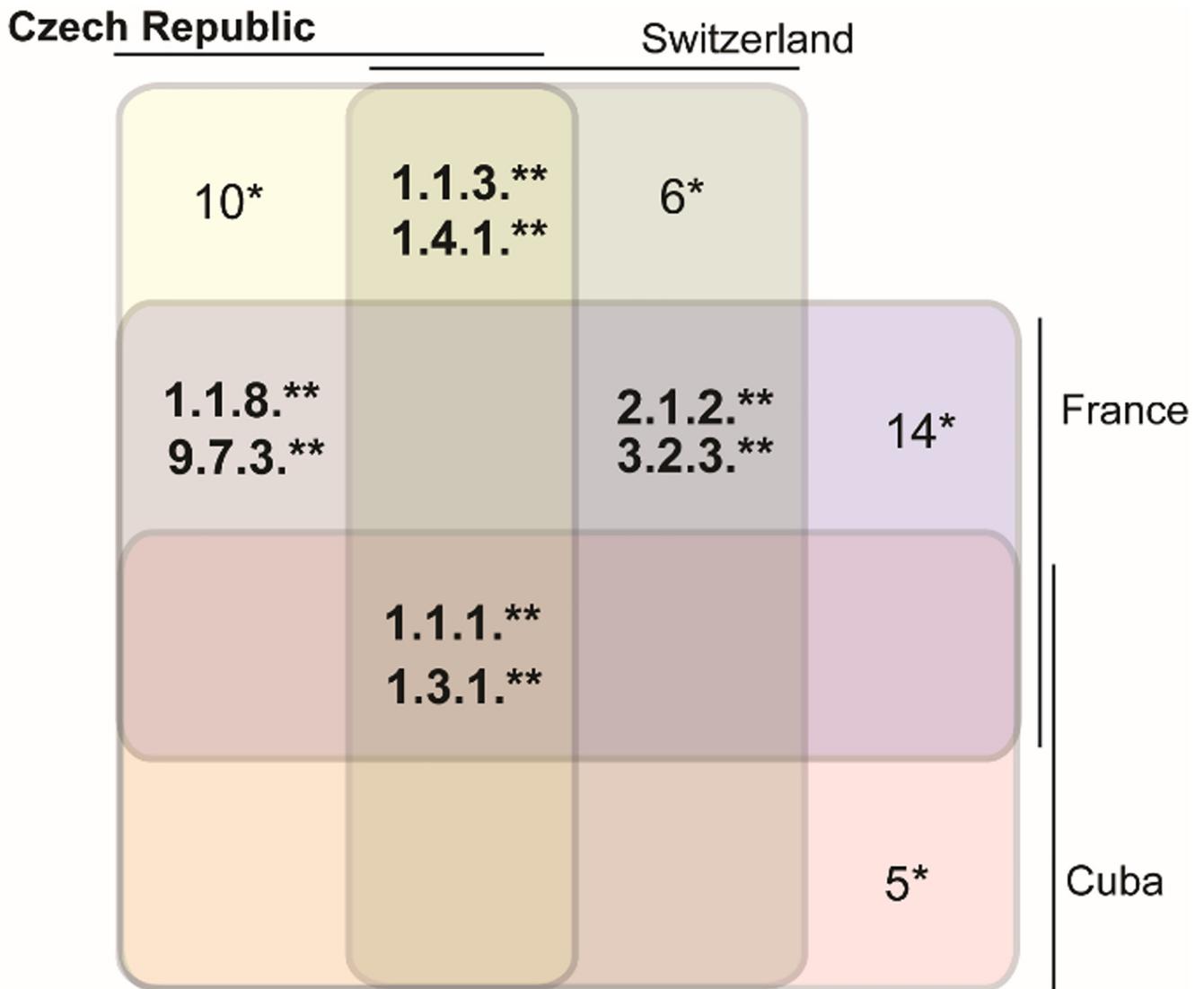
Fig 5. TPA allelic profiles identified in the Czech Republic, France, Switzerland and Cuba [15–18].

https://doi.org/10.1371/journal.pone.0217611.g005

In this study, we have identified 16 different allelic profiles among 136 fully typed TPA-containing samples, in the same collection of samples we would identify 13 different profiles by SBMT. In the Czech Republic, in comparison to samples analyzed in previous studies from other countries [15–17], there were partial overlaps with identified fully determined allelic profiles. Samples collected in Switzerland [18] revealed four allelic profiles also found in this study (i.e., 1.1.1., 1.1.3., 1.3.1. and 1.4.1.), while 20 were different. Similarly, samples collected in France [18] revealed four shared allelic profiles (1.1.1., 1.3.1., 1.1.8., 9.7.3.) with 28 profiles that were different. Cuban samples [18] revealed two shared allelic profiles (1.1.1., 1.3.1.) and 17 different profiles (Fig 5). Only allelic profiles 1.1.1. and 1.1.3. were shared by all four countries

indicating that TPA allelic profile variability is relatively high even in somewhat closely related geographical regions (Fig 5). Comparison of shared allelic profiles among Czech Republic, France, Switzerland and Cuba (Fig 5) revealed that the profiles 1.1.1. and 1.3.1. were detected in every MLST study in most time points whereas the profile 1.4.1. was detected in Switzerland at least two years earlier than in the Czech Republic (see PubMLST database [18]). Allelic profile 9.7.3. shared between Czech Republic and France was detected in similar years and profile 1.1.8. disappeared in 2016 in France and in 2016/2017 in the Czech Republic. The most frequent allelic profiles were thus similarly detected in different countries suggesting supranational spreading of certain syphilis strains. When both partially and fully typed samples were analyzed, a total of eight new alleles were identified in 281 typeable samples. We found five novel allelic variants among fully typed samples and three novel allelic variants among partially typed samples. This finding further shows the relatively high variability of TPA strains circulating in the Czech Republic.

The most common allelic profile in the Czech Republic, 1.3.1. (41.2%), corresponds to the SU2 genotype based on SBMT [19] and to the "g" ECDCT_TP0548 subtype based on ECDCT [13]. This allelic profile is also the most common in Belgium [27], France [16], Switzerland [15], Italy [29], the UK [30], and other countries including Cuba [17], the USA [31], and Australia [32]. The second and third most common allelic profiles found in this study were 1.1.8. (20.6%) and 1.1.1. (14.7%) both having allelic variant TP_0548_1 corresponding to the "f" ECDCT_TP0548 subtype [13], which is the most common in Argentina [33], China [34–39], Taiwan [40], and Russia [41].

In our study, we found differences in the local distribution of allelic profiles between samples from Brno and samples from Prague. In Prague, we found a greater number of allelic profiles, which may be related to Prague's larger population as well as the larger number of people visiting Prague, who could represent potential carriers of new profiles. Despite the lower number of different allelic profiles identified in Brno, one profile (1.26.1.) was exclusively present in the Brno region (found in 12 patients). This finding further extends previous findings that showed differences between individual countries and suggests that sexual networks can differ even within a single country.

In this study, associations of different allelic profiles 1.26.1., 1.3.1., and 1.1.3. with macrolide resistance and 1.1.8. with macrolide susceptibility, were found. Similar associations have been found in other studies [4, 15, 42]. This study confirmed two trends: {i} there are increasing numbers of A2058G mutations and {ii} there are decreasing numbers of A2059G mutations, a trend that was also seen in a previous study [4]. As suggested by Grillová *et al.* (2014) [4], the differences in the occurrence of the A2058G and A2059G mutations could reflect opposite trends in the use of spiramycin and azithromycin in the Czech Republic. While there is a decreasing trend in the use of spiramycin, azithromycin is being prescribed more frequently [4]. While the A2058G mutation does not encode resistance to spiramycin, the A2059G does. In addition to decreased use of spiramycin, the A2059G mutation was predicted to have a higher fitness cost compared to A2058G [42].

A phylogenetic analysis of allelic profiles found in this study clearly differentiated profiles belonging to SS14-like and Nichols-like strains (Fig 2), however, differences within the SS14 strains were supported with low bootstrap values (lower than 70%). With the assumption that the MLST typing system has about 30% of the discriminatory power of whole-genome analyses [15] and the low support for clustering of SS14-like strains, the identified SS14 strains were highly clonal showing low genetic diversity, which is also likely at the whole genome level. In fact, previous whole-genome studies that analyzed SS14-like strains found only limited genetic diversity among SS14-like strains [43–44]. Among the SS14-like allelic profiles identified in this study, allelic profile 18.1.1. (originally found by Flasarová *et al.* 2006) [14] appeared to be

Fig 6. Modular structure of the TP0136 gene in the TPA SS14 strain and in the TPA 18.1.1. clinical isolate. A. A schematic representation of the TP0136 gene in the TPA SS14 strain between coordinates 158434–158678 (CP004011.1). The figure was modified from Strouhal et al. (2018) [45]. The r0 repetitive sequence in the TPA 18.1.1. clinical isolate replaced the first r4 repetitive sequence. B. A list of repetitive (r0, r1, r2, r3, r4, r6) and non-repetitive sequences (r5). The nucleotide differences within r1 and r2 repetitions in the TPA SS14 strain are shown in bold.

https://doi.org/10.1371/journal.pone.0217611.g006

the most divergent (Fig 2), which was predominantly a result of a highly divergent sequence found at the TP0136 locus. In addition, the sequences of allelic profile 18.1.1. in TP0548 and TP0705 were identical to and different from the SS14 sequences, respectively. A detailed analysis of locus TP0136 revealed that the observed genetic diversity was a result of gene reshuffling of modular gene segments (Fig 6). A similar modular structure to that of TP0136 was found among TPE strains [45]. Another similar genetic rearrangement of TP0136 was found in the study of Grillová et al. [17], where four modular regions (r1-r2-r3-r4) were deleted.

The most probable explanation for the sequence differences in the TP0136 gene in clinical isolate 18.1.1 is a gene conversion event copying the r0 sequence instead of the first r4 sequence. This scenario explains why both r0 sequences were identical in the 18.1.1. isolate.

There are already numerous examples of genome rearrangements in treponemal strains and isolates including the *tpr*K gene [46], *tpr*CDIGJK genes [47], the TP0133 gene [45, 48–50], rRNA (*rrn*) operons [51], TP0856 and TP0858, and other genes [45].

Most of the TPA samples analyzed in this study belonged to the SS14-clade [23, 52], and only 2.41% of the 249 determined samples belonged to the Nichols-like clade. The number of identified Nichols-like TPA strains is lower compared to previous studies [15–16]; however, it is not far from the worldwide estimated number of Nichols-like strains (i.e., 5.9%) [50].

In addition to geographical variability found among TPA samples taken from different countries, an analysis of the temporal occurrence of TPA allelic profiles (Fig 3) in the Czech Republic between 2004–2017 revealed an increasing number of identified allelic profiles and also differences in the spectra of identified allelic profiles over the years. While the first finding corresponds to the increasing number of collected samples during recent years, the second observation is consistent with the dynamic character of TPA strains in the infected population. While some allelic profiles remained for the whole study period (e.g., allelic profile 1.1.1.), other allelic profiles showed an increasing prevalence (e.g., 1.3.1.), while others slowly disappear (e.g., 1.1.8.) and some emerged (e.g., allelic profile 1.26.1. which appeared for the first time in 2012/2013 and persisted until 2017). While these findings are at least partly attributable to the random distribution of allelic profiles and to sampling bias, they could also point to possible differences in the fitness and/or pathogenicity of particular allelic profiles. As of now, we are still some distance from a determination of the full genetic diversity of TPA isolates in the global population and also from understanding the role of genetic differences in syphilis epidemiology. Mapping the genetic diversity of TPA strains in the context of additional clinical data will likely help answer at least some of these questions including connection of diverse genetic profiles with several patients´ characteristics.

Supporting information

S1 Fig. Alignment of the different allelic variants in the TP0136, TP0548 and TP0705 loci identified in this study (new alleles are in italics).
(PDF)

S1 Table. Primers used for the nested-PCR.
(PDF)

S2 Table. Allelic profiles identified in partially typed samples.
(PDF)

Acknowledgments

Author Contributions

Formal analysis: Eliška Vrbová.

Funding acquisition: David Šmajs.

Investigation: Eliška Vrbová, Linda Grillová, Lenka Mikalová, Petra Pospíšilová, Eliška Dastychová, Martina Kojanová, Miluše Kreidlová, Daniela Vaňousová, Filip Rob, Přemysl Procházka, Alena Krchňáková, Vladimír Vašků, Vladana Woznicová, Hana Zákoucká.

Methodology: Linda Grillová.

Resources: Radim Strnadel, Alena Krchňáková, Monika Dvořáková Heroldová, Ivana Kuklová, Hana Zákoucká.

Supervision: David Šmajs.

Writing – original draft: Eliška Vrbová.

Writing – review & editing: Eliška Vrbová, Linda Grillová, Petra Pospíšilová, David Šmajs.

References

1. Radolf JD, Deka RK, Anand A, Šmajs D, Norgard MV, Yang XF. *Treponema pallidum*, the syphilis spirochete: making a living as a stealth pathogen. Nat Rev Microbiol. 2016; 14: 744–759. https://doi.org/10.1038/nrmicro.2016.141 PMID: 27721440

2. Newman L, Rowley J, Vander Hoorn S, Wijesooriya NS, Unemo M, Low N, et al. Global Estimates of the Prevalence and Incidence of Four Curable Sexually Transmitted Infections in 2012 Based on Systematic Review and Global Reporting. PLoS One. 2015; 10: e0143304. https://doi.org/10.1371/journal.pone.0143304 PMID: 26646541

3. Peeling RW, Mabey D, Kamb ML, Chen XS, Radolf JD, Benzaken AS. Syphilis. Nat Rev Dis Primers. 2017; 3: 17073. https://doi.org/10.1038/nrdp.2017.73 PMID: 29022569

4. Grillová L, Pětrošová H, Mikalová L, Strnadel R, Dastychová E, Kuklová I, et al. 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 ClinMicrobiol. 2014; 52: 3693–700.

5. Gayet-Ageron A, Lautenschlager S, Ninet B, Perneger TV, Combescure C. Sensitivity, specificity and likelihood ratios of PCR in the diagnosis of syphilis: a systematic review and meta-analysis. Sex Transm Infect. 2013; 89:251–6. https://doi.org/10.1136/sextrans-2012-050622 PMID: 23024223

6. Grange PA, Allix-Beguec C, Chanal J, Benhaddou N, Gerhardt P, Morini JP, et al. Molecular subtyping of *Treponema pallidum* in Paris, France. Sex Transm Dis. 2013; 40: 641–4. https://doi.org/10.1097/OLQ.0000000000000006 PMID: 23859911

7. Grange PA, Mikalová L, Gaudin C, Strouhal M, Janier M, Benhaddou N, et al. *Treponema pallidum* 11qj Subtype May Correspond to a *Treponema pallidum* Subsp. *Endemicum* Strain. Sex Transm Dis. 2016; 43: 517–8. https://doi.org/10.1097/OLQ.0000000000000474 PMID: 27419817

8. Mikalová L, Strouhal M, Grillová L, Šmajs D. The molecular typing data of recently identified subtype 11q/j of *Treponema pallidum* subsp. pallidum suggest imported case of yaws. Sex Transm Dis. 2014; 41: 552–3. https://doi.org/10.1097/OLQ.0000000000000165 PMID: 25118969

9. Noda AA, Grillová L, Lienhard R, Blanco O, Rodríguez I, Šmajs D. Bejel in Cuba: molecular identification of *Treponema pallidum* subsp. *endemicum* in patients diagnosed with venereal syphilis. ClinMicrobiol Infect. 2018; 24: 1210.e1–1210.e5.

10. Kojima Y, Furubayashi K, Kawahata T, Mori H, Komano J. Circulation of Distinct *Treponema pallidum* Strains in Individuals with Heterosexual Orientation and Men Who Have Sex with Men. J ClinMicrobiol. 2019; 57: e01148–18.

11. Pillay A, Liu H, Chen CY, Holloway B, Sturm AW, Steiner B, et al. Molecular subtyping of *Treponema pallidum* subspecies *pallidum*. Sex Transm Dis. 1998; 25: 408–14. PMID: 9773432

12. Katz KA, Pillay A, Ahrens K, Kohn RP, Hermanstyne K, Bernstein KT, et al. Molecular epidemiology of syphilis—San Francisco, 2004–2007. Sex Transm Dis. 2010; 37: 660–3. https://doi.org/10.1097/OLQ.0b013e3181e1a77a PMID: 20601928

13. Marra C, Sahi S, Tantalo L, Godornes C, Reid T, Behets F, et al. Enhanced molecular typing of *Treponema pallidum*: geographical distribution of strain types and association with neurosyphilis. J Infect Dis. 2010; 202: 1380–8. https://doi.org/10.1086/656533 PMID: 20868271

14. Flasarová M, Šmajs D, Matějková P, Woznicová V, Heroldová-Dvoráková M, Votava M. Molecular detection and subtyping of *Treponema pallidum* subsp. *pallidum* in clinical specimens. Epidemiol Mikrobiol Imunol. 2006; 55: 105–11. PMID: 16970074

15. Grillová L, Bawa T, Mikalová L, Gayet-Ageron A, Nieselt K, Strouhal M, et al. Molecular characterization of *Treponema pallidum* subsp. *pallidum* in Switzerland and France with a new multilocus sequence typing scheme. PLoS One. 2018; 13: e0200773. https://doi.org/10.1371/journal.pone.0200773 PMID: 30059541

16. Pospíšilová P, Grange PA, Grillová L, Mikalová L, Martinet P, Janier M, et al. Multi-locus sequence typing of *Treponema pallidum* subsp. *pallidum* present in clinical samples from France: Infecting

treponemes are genetically diverse and belong to 18 allelic profiles. PLoS ONE. 2018; 13: e0201068. https://doi.org/10.1371/journal.pone.0201068 PMID: 30024965

17. Grillová L, Noda AA, Lienhard R, Blanco O, Rodríguez I, Šmajs D. Multilocus sequence typing of *Treponema pallidum* subsp. *pallidum* in Cuba from 2012 to 2017. J Infect Dis. 2018. https://doi.org/10.1093/infdis/jiy604 PMID: 30325448

18. Grillova L, Jolley K, Šmajs D, Picardeau M. A public database for the new MLST scheme for *Treponema pallidum* subsp. *pallidum*: surveillance and epidemiology of the causative agent of syphilis. PeerJ. 2019; 6: e6182. https://doi.org/10.7717/peerj.6182 PMID: 30643682

19. Flasarová M, Pospíšilová P, Mikalová L, Vališová Z, Dastychová E, Strnadel R, et al. 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. 2012; 92: 669–74. https://doi.org/10.2340/00015555-1335 PMID: 22434073

20. Matějková P, Flasarová M, Zákoucká H, Borek M, Kremenová S, Arenberger P, et al. Macrolide treatment failure in a case of secondary syphilis: a novel A2059G mutation in the 23S rRNA gene of *Treponema pallidum* subsp. *pallidum*. J Med Microbiol. 2009; 58: 832–6. https://doi.org/10.1099/jmm.0.007542-0 PMID: 19429763

21. Stamm LV, Bergen HL. A point mutation associated with bacterial macrolide resistance is present in both 23S rRNA genes of an erythromycin-resistant *Treponema pallidum* clinical isolate. Antimicrob Agents Chemother. 2000; 44: 806–7. https://doi.org/10.1128/aac.44.3.806-807.2000 PMID: 10755994

22. Molini BJ, Tantalo LC, Sahi SK, Rodriguez VI, Brandt SL, Fernandez MC, et al. Macrolide Resistance in *Treponema pallidum* Correlates With 23S rDNA Mutations in Recently Isolated Clinical Strains. Sex Transm Dis. 2016; 43: 579–83. https://doi.org/10.1097/OLQ.0000000000000486 PMID: 27513385

23. Nechvátal L, Pětrošová H, Grillová L, Pospíšilová P, Mikalová L, Strnadel R, et al. 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. 2014; 304: 645–53. https://doi.org/10.1016/j.ijmm.2014.04.007 PMID: 24841252

24. Kumar S., Stecher G., and Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol. 2015 (submitted).

25. Tamura K., Nei M., Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl AcadSci USA. 2004; 101: 11030–11035.

26. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution. 1985; 39: 783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x PMID: 28561359

27. Mikalová L, Grillová L, Osbak K, Strouhal M, Kenyon C, Crucitti T, et al. Molecular Typing of Syphilis-Causing Strains Among Human Immunodeficiency Virus-Positive Patients in Antwerp, Belgium. Sex Transm Dis. 2017; 44: 376–9. https://doi.org/10.1097/OLQ.0000000000000600 PMID: 28499290

28. Peng RR, Wang AL, Li J, Tucker JD, Yin YP, Chen XS. Molecular typing of *Treponema pallidum*: a systematic review and meta-analysis. PLoS Negl Trop Dis. 2011; 5: e1273. https://doi.org/10.1371/journal.pntd.0001273 PMID: 22087340

29. Giacani L, Ciccarese G, Puga-Salazar C, Dal Conte I, Colli L, Cusini M, et al. Enhanced Molecular Typing of *Treponema pallidum* subspecies *pallidum* Strains From 4 Italian Hospitals Shows Geographical Differences in Strain Type Heterogeneity, Widespread Resistance to Macrolides, and Lack of Mutations Associated With Doxycycline Resistance. Sex Transm Dis. 2018; 45: 237–242. https://doi.org/10.1097/OLQ.0000000000000741 PMID: 29465698

30. Tipple C, McClure MO, Taylor GP. High prevalence of macrolide resistant *Treponema pallidum* strains in a London centre. Sex Transm Infect. 2011; 87: 486–8. https://doi.org/10.1136/sextrans-2011-050082 PMID: 21917695

31. Grimes M, Sahi SK, Godornes BC, Tantalo LC, Roberts N, Bostick D, et al. Two mutations associated with macrolide resistance in *Treponema pallidum*: increasing prevalence and correlation with molecular strain type in Seattle, Washington. Sex Transm Dis. 2012; 39: 954–8. https://doi.org/10.1097/OLQ.0b013e31826ae7a8 PMID: 23191949

32. Read P, Tagg KA, Jeoffreys N, Guy RJ, Gilbert GL, Donovan B. *Treponema pallidum* Strain Types and Association with Macrolide Resistance in Sydney, Australia: New TP0548 Gene Types Identified. J Clin Microbiol. 2016; 54: 2172–4. https://doi.org/10.1128/JCM.00959-16 PMID: 27194693

33. Gallo Vaulet L, Grillová L, Mikalová L, Casco R, Rodríguez Fermepin M, Pando MA, et al. Molecular typing of *Treponema pallidum* isolates from Buenos Aires, Argentina: Frequent Nichols-like isolates and low levels of macrolide resistance. PLoS One. 2017; 12: e0172905. https://doi.org/10.1371/journal.pone.0172905 PMID: 28235102

34. Zhang RL, Wang QQ, Zhang JP, Yang LJ. Molecular subtyping of *Treponema pallidum* and associated factors of serofast status in early syphilis patients: Identified novel genotype and cytokine marker. PLoS One. 2017; 12: e0175477. https://doi.org/10.1371/journal.pone.0175477 PMID: 28410389

35. Zhu B, Bu J, Li W, Zhang J, Huang G, Cao J, et al. High Resistance to Azithromycin in Clinical Samples from Patients with Sexually Transmitted Diseases in Guangxi Zhuang Autonomous Region, China. PLoS One. 2016; 11: e0159787. https://doi.org/10.1371/journal.pone.0159787 PMID: 27467164

36. Li Z, Wang C, Xiao H, Zhao W, Zheng R, Hou J, et al. Enhanced molecular typing of *Treponema pallidum* identified a new tp0548 Gene type in Shandong, China. APMIS. 2017; 125: 937–9. https://doi.org/10.1111/apm.12724 PMID: 28766763

37. Xiao Y, Liu S, Liu Z, Xie Y, Jiang C, Xu M, et al. Molecular Subtyping and Surveillance of Resistance Genes In *Treponema pallidum* DNA From Patients With Secondary and Latent Syphilis in Hunan, China. Sex Transm Dis. 2016; 43: 310–6. https://doi.org/10.1097/OLQ.0000000000000445 PMID: 27100768

38. Dai T, Li K, Lu H, Gu X, Wang Q, Zhou P. Molecular typing of *Treponema pallidum*: a 5-year surveillance in Shanghai, China. J ClinMicrobiol. 2012; 50: 3674–7.

39. Peng RR, Yin YP, Wei WH, Wang HC, Zhu BY, Liu QZ, et al. Molecular typing of *Treponema pallidum* causing early syphilis in China: a cross-sectional study. Sex Transm Dis. 2012; 39: 42–5. https://doi.org/10.1097/OLQ.0b013e318232697d PMID: 22183845

40. Wu BR, Yang CJ, Tsai MS, Lee KY, Lee NY, Huang WC, et al. Multicentre surveillance of prevalence of the 23S rRNA A2058G and A2059G point mutations and molecular subtypes of *Treponema pallidum* in Taiwan, 2009–2013. ClinMicrobiol Infect. 2014; 20: 802–7.

41. Khairullin R, Vorobyev D, Obukhov A, Kuular UH, Kubanova A, Kubanov A, et al. Syphilis epidemiology in 1994–2013, molecular epidemiological strain typing and determination of macrolide resistance in *Treponema pallidum* in 2013–2014 in Tuva Republic, Russia. APMIS. 2016; 124: 595–602. https://doi.org/10.1111/apm.12541 PMID: 27102715

42. Šmajs D, Paštěková L, Grillová L. Macrolide Resistance in the Syphilis Spirochete, *Treponema pallidum* ssp. *pallidum*: Can We Also Expect Macrolide-Resistant Yaws Strains? Am J Trop Med Hyg. 2015; 93: 678–83. https://doi.org/10.4269/ajtmh.15-0316 PMID: 26217043

43. Arora N, Schuenemann VJ, Jäger G, Peltzer A, Seitz A, Herbig A, et al. Origin of modern syphilis and emergence of a pandemic *Treponema pallidum* cluster. Nat Microbiol. 2016; 2: 16245. https://doi.org/10.1038/nmicrobiol.2016.245 PMID: 27918528

44. Pinto M, Borges V, Antelo M, Pinheiro M, Nunes A, Azevedo J, et al. Genome-scale analysis of the non-cultivable *Treponema pallidum* reveals extensive within-patient genetic variation. Nat Microbiol. 2016; 2: 16190. https://doi.org/10.1038/nmicrobiol.2016.190 PMID: 27748767

45. Strouhal M, Mikalová L, Haviernik J, Knauf S, Bruisten S, Noordhoek GT, et al. Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue* from Indonesia: Modular structure of several treponemal genes. PLoSNegl Trop Dis. 2018; 12: e0006867.

46. Centurion-Lara A, LaFond RE, Hevner K, Godornes C, Molini BJ, Van Voorhis WC, et al. Gene conversion: a mechanism for generation of heterogeneity in the tprK gene of *Treponema pallidum* during infection. Mol Microbiol. 2004; 52: 1579–96. https://doi.org/10.1111/j.1365-2958.2004.04086.x PMID: 15186410

47. Gray RR, Mulligan CJ, Molini BJ, Sun ES, Giacani L, Godornes C, et al. Molecular evolution of the tprC, D, I, K, G, and J genes in the pathogenic genus *Treponema*. MolBiolEvol. 2006; 23: 2220–33.

48. Strouhal M, Šmajs D, Matějková P, Sodergren E, Amin AG, Howell JK, et al. Genome differences between *Treponema pallidum* subsp. *pallidum* strain Nichols and T. paraluiscuniculi strain Cuniculi A. Infect Immun. 2007; 75: 5859–66. https://doi.org/10.1128/IAI.00709-07 PMID: 17893135

49. Godornes C, Giacani L, Barry AE, Mitja O, Lukehart SA. Development of a Multilocus Sequence Typing (MLST) scheme for *Treponema pallidum* subsp. *pertenue*: Application to yaws in Lihir Island, Papua New Guinea. PLoS Negl Trop Dis. 2017; 11: e0006113. https://doi.org/10.1371/journal.pntd.0006113 PMID: 29281641

50. Šmajs D, Strouhal M, Knauf S. Genetics of human and animal uncultivable treponemal pathogens. Infect Genet Evol. 2018; 61: 92–107. https://doi.org/10.1016/j.meegid.2018.03.015 PMID: 29578082

51. Čejková D, Zobaníková M, Pospíšilová P, Strouhal M, Mikalová L, Weinstock GM, et al. Structure of *rrn* operons in pathogenic non-cultivable treponemes: sequence but not genomic position of intergenic spacers correlates with classification of *Treponema pallidum* and *Treponemaparaluiscuniculi* strains. J Med Microbiol. 2013; 62:196–207. https://doi.org/10.1099/jmm.0.050658-0 PMID: 23082031

52. Šmajs D, Mikalová L., Strouhal M., Grillová L. Why there are two genetically distinct syphilis-causing strains? For Immunopathol Dis Therap. 2016; 7: 181–90.