RESEARCH ARTICLE

The hare syphilis agent is related to, but distinct from, the treponeme causing rabbit syphilis

Petra Pospíšilová[1], Darina Čejková[2,3], Pavla Buršíková[1], Pavla Fedrová[1], Lenka Mikalová[1], David Najt[1], Nikola Tom[1], Linda Hisgen[4], Simone Lueert[4], Johannes T. Lumeij[5], Erik O. Ågren[6], Sascha Knauf[4,7], David Šmajs[1]*

1 Department of Biology, Faculty of Medicine, Masaryk University, Brno, Czech Republic, 2 Department of Microbiology and Antimicrobial Resistance, Veterinary Research Institute, Brno, Czech Republic, 3 Department of Biomedical Engineering, Brno University of Technology, Brno, Czech Republic, 4 Institute of International Animal Health/One Health, Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Greifswald, Insel Riems, Germany, 5 Division of Zoological Medicine, Department of Clinical Sciences of Companion Animals, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands, 6 Department of Pathology and Wildlife Diseases, National Veterinary Institute, Uppsala, Sweden, 7 Professorship for International Animal Health/One Health, Faculty of Veterinary Medicine, Justus Liebig University, Giessen, Germany

* dsmajs@med.muni.cz

Abstract

The treponemes infecting lagomorphs include *Treponema paraluisleporidarum* ecovar Cuniculus (*TP*eC) and ecovar Lepus (*TP*eL), infecting rabbits and hares, respectively. In this study, we described the first complete genome sequence of *TP*eL, isolate V3603-13, from an infected mountain hare (*Lepus timidus*) in Sweden. In addition, we determined 99.0% of the genome sequence of isolate V246-08 (also from an infected mountain hare, Sweden) and 31.7% of the genome sequence of isolate Z27 A77/78 (from a European hare, *Lepus europeaus*, The Netherlands). The *TP*eL V3603-13 genome had considerable gene synteny with the *TP*eC Cuniculi A genome and with the human pathogen *T. pallidum*, which causes syphilis (ssp. *pallidum*, *TPA*), yaws (ssp. *pertenue*, *TPE*) and endemic syphilis (ssp. *endemicum*, *TEN*). Compared to the *TP*eC Cuniculi A genome, *TP*eL V3603-13 contained four insertions and 11 deletions longer than three nucleotides (ranging between 6 and 2,932 nts). In addition, there were 25 additional indels, from one to three nucleotides long, altogether spanning 36 nts. The number of single nucleotide variants (SNVs) between *TP*eC Cuniculi A and *TP*eL V3603-13 were represented by 309 nucleotide differences. Major proteome coding differences between *TP*eL and *TP*eC were found in the *tpr* gene family, and (predicted) genes coding for outer membrane proteins, suggesting that these components are essential for host adaptation in lagomorph syphilis. The phylogeny revealed that the *TP*eL sample from the European brown hare was more distantly related to *TP*eC Cuniculi A than V3603-13 and V246-08.

Introduction

In addition to human pathogens such as *Treponema pallidum* ssp. *pallidum* (*TPA*), ssp. *pertenue* (*TPE*), and ssp. *endemicum* (*TEN*), causing syphilis, yaws, and bejel, respectively–the *Treponema* genus also contains other animal pathogens [1, 2]. Lagomorph syphilis is caused by *Treponema paraluisleporidarum* [3], with its ecovar Cuniculus (*TPeC*) and ecovar Lepus (*TPeL*), infecting rabbits and hares, respectively. Current knowledge centers almost exclusively around the *TPeC* strain Cuniculi A, isolated in 1939 in Maryland, USA [4], which is still maintained in the laboratory. *TPeC* strain Cuniculi A is genetically very related to the human pathogen *T. pallidum* (98% genome identity on the DNA level; [2, 5]). This study uses recently introduced name *Treponema paraluisleporidarum* [3], as this is consistent with the literature use in the last decade. However, since the GenBank database uses only validated names, the sequences were submitted as *Treponema paraluiscuniculi*. (*T. paraluiscuniculi* was validated decades ago and would not be eligible for validation according to current rules since it is highly related to *Treponema pallidum* [6]. Like human syphilis, *TPeC* causes sexually transmitted infections and known to be associated with skin ulcers in the anogenital region, the face, and/or the paws of rabbits (*Oryctolagus cuniculus* f. *domestica*) [7]. Sexually transmitted *TPeL* infections in European brown hares (*Lepus europaeus*) and mountain hares (*Lepus timidus*), are in most cases clinically inapparent or associated with orofacial and anogenital proliferative crusty skin lesions occurring at mucocutaneous junctions [7–9].

Based on serological screening studies that analyzed samples from more than 1500 animals, the average seroprevalence of syphilis in wild European lagomorphs ranges from 28% to 66%, depending on the serological test(s) used and the sample's country of origin [10–14]. The prevalence of hare syphilis has no clear geographical gradient, however, it is known that seroprevalence against *TPeL* negatively correlates with altitude of sampling areas [11].

Genetic studies on the causative agent of lagomorph syphilis are limited and almost exclusively based on a single available genome sequence of the *TPeC* strain Cuniculi A, which is composed of circular chromosome of 1,133,390 bp with no plasmids [6]. So far, a single study [3] tested the genetic relatedness between *TPeL* and *TPeC*. The study revealed four nucleotide differences in the 2002 bp-long sequence obtained from partial sequences of the 16S rRNA gene, the DNA region downstream of the 16S rRNA gene, and the sequence within the *TPCCA_0225* gene including one nt difference found within the 16S rRNA genes [3].

Based on infection experiments conducted by Lumeij et al. [3] and on the genetic relatedness between *TPeC* and *TPA* [5, 6], we predicted that hare-infecting strains would be phylogenetically ancestral to rabbit-infecting strains and would be highly related to *TPeC*. In this study, we determined the complete genome sequence of *Treponema paraluisleporidarum* ecovar Lepus, isolate V3603-13. We used the previously described Pooled Sequence Genome Sequencing (PSGS) technique to compare the V3603-13 genome to the once-determined genomic sequence *TPeC* strain Cuniculi A. We found that the treponeme causing hare syphilis was in many ways similar but distinct from the treponeme causing rabbit syphilis.

Materials and methods

Lagomorph samples

We obtained samples from mountain hares (*Lepus timidus*) hunted in Sweden. *TPeL* sample V3603-13 was obtained from a hare sampled close to Enånger in county Gävleborg (latitude 61.32470, longitude 17.02430) in 2013. Sample V246-08 came from a hare sample hunted in 2008 near Orsa in the county of Dalarna (latitude 61.30180, longitude 14.28340). Both samples were taken from hares submitted within the Swedish general (passive) wildlife disease

surveillance program, where hunters or anybody else can report findings of disease or mortality in wildlife. The hares were shot by hunters and when they noted scabs and lesions on the lips and nose, they contacted Swedish Veterinary Agency (SVA). Upon request from SVA, the carcasses were sent to the institute for diagnostic examinations. No further information is available on the location of lesions. The brown hare isolate, Z27 A77/78, was collected in The Netherlands in 2010 and was described in an earlier study [3].

Sampling and extraction of DNA

Samples from the skin lesions (full skin thickness tissue samples) were taken from dead animals and stored in a -80 degree Celsius biobank freezer until thawed and used in this study. DNA was extracted from tissue material using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol with some minor modifications. Briefly, we extracted DNA from tissue samples according to the protocols published by Hisgen et al. [11]. Subsequently, glycogen precipitation was performed to clean and concentrate the DNA as described by Knauf et al. [15]. We measured the DNA yield using a NanoDrop photometer (Thermo Fisher Scientific, Darmstadt, Germany).

Amplification of genomic DNA

The genomic DNA of all three samples was amplified using the multiple displacement amplification approach (REPLI-g kit, QIAGEN, Valencia, CA, USA). Samples were diluted one to 50 times (depending on the positivity of treponemal-specific PCR amplification from undiluted and diluted samples) and then used as a template for PSGS, as previously described [16–18]. The DNA was amplified using PrimeSTAR GXL DNA Polymerase (Takara Bio Inc., Otsu, Japan) with 133 specific primer pairs (published by Mikalová et al. [19], with minor modifications) to obtain overlapping PCR products covering the entire chromosome. PCR products were amplified using touchdown PCR under the following cycling conditions: initial denaturation at 94°C for 1 min; 8 cycles: 98°C for 10 s, 68°C for 15 s (the annealing temperature was gradually reduced at a rate of 1°C per cycle), and 68°C for 6 min; 35 cycles: 98°C for 10 s, 61°C for 15 s, and 68°C for 6 min (43 cycles in total); followed by the final extension at 68°C for 7 min. PCR products were subsequently purified using QIAquick PCR Purification Kits (QIAGEN, Valencia, CA, USA) and mixed in equimolar amounts to produce four distinct pools, which were then used for whole genome sequencing.

Whole genome sequencing and assembly of the genomes

The sequencing library for *TP*eL strain V3603-13 was prepared from 1 ng of RNA-free genomic DNA using a Nextera XT DNA Sample Preparation kit (Illumina, CA, USA). Sequencing was performed using the MiSeq Reagent Kit v3 in the MiSeq system (Illumina, USA) at the Veterinary Research Institute sequencing facility (Brno, Czech Republic). The sequencing reads of individual pools were handled separately and assembled *de novo* with SeqMan NGen v4.1.0 software (DNASTAR, Madison, WI, USA) using the default parameters. Contigs obtained from *TP*eL strains were then aligned to the genome of the *TP*eC strain Cuniculi A (CP002103.1) using Lasergene software (DNASTAR, Madison, WI, USA). Missing parts of genomes were subsequently Sanger sequenced (GATC Biotech, Germany). In addition, MinION sequencing (Oxford Nanopore Technologies (ONT), Oxford, UK) was used to sequence paralogous gene regions, mainly those containing *tpr* genes.

The number of repetitions within the *arp* gene (*TP0433*) was determined by amplification of PCR products with primers 32BrepF1 (5′-CGT TTG GTT TCC CCT TTG TC-3′) and 32BrepR1 (5′-GTG GGA TGG CTG CTT CGT ATG-3′) as described by Harper et al.

(2008); the resulting PCR products were subsequently Sanger sequenced. Repetitive sequences within the *TP0470* gene were amplified and sequenced using primers TPI34F4 (5′-GTC TTG TGC ACA TTA TTC AAG-3′) and TPI34R5 (5′-CTT CGT GCA ACA TCG CTA CG-3′).

Both the V246-08 and Z27 A77/78 genomes were amplified using the PSGS method and then sequenced using the Illumina platform as described for the V3603-13 sample. The quality of the raw reads was checked using FastQC [20]. Raw reads were trimmed in length (a minimal length of 35 nt) and quality (Phred quality score ≥ 20) using Cutadapt [21]. Preprocessed reads were mapped using the BWA MEM algorithm [22] onto the *TP*eL Cuniculi A genome (GenBank Acc. No. CP002103.1). The Cuniculi A reference genome was divided into four parts representing the "pools" described above. The post-processing of the reads' mapping was performed using Samtools [22], Picard (https://broadinstitute.github.io/picard/), GATK [23], and NGSUtils/bamutils [24]. Low-quality mappings were omitted from analyses (i.e., mapping quality; MAPQ < 40) as well as reads mapping to multiple sites (e.g., in repetitive and paralogous regions). A minimal alignment length was set to 35 bps, the maximum number of allowed mismatches was set to 5 (or 5% of the read length), and the maximum soft-clipping was set to 5% of the read length. Consensus sequences were determined using variants detected using Samtools "variant calling" followed by Vcfutils [22] and Seqtk (https://github.com/lh3/seqtk). The breadth and depth of coverages were calculated using GATK software.

Libraries used for MinION (ONT) sequencing were prepared using the 1D Native barcoding genomic DNA protocol with EXP-NBD103 and SQK-LSK108 kits (ONT) and Nanopore MinION Spoton flow cells (FLO-MIN106D, version R9) and sequenced for 48 hours. Basecalling and barcoding were done using Guppy (v4.4.1, ONT) in high-accuracy base-calling mode and q-score 7.

Construction of phylogenetic tree

To determine evolutionary relationships among pathogenic treponemes, corresponding parts of the available treponemal genomes (orthologous sequences) were used and included *TP*eC Cuniculus strain Cuniculi A (CP002103.1), *TP*eC Cz-2020 (MW323408), and *TP*A strain SS14 (CP004011.1). The phylogenetic tree was constructed using the Maximum Likelihood method based on the Tamura-Nei model and MEGA 7 software, which was also used for model testing [25]. Bootstrap values were equal to 1000.

Nucleotide sequence accession number

The complete genome sequence of the *TP*eL V3603-13 isolate was deposited in GenBank under accession number CP097901, Bioproject PRJNA606433. Since *Treponema paraluisleporidarum* ecovar Lepus is not validated, the sequences were submitted as *Treponema paraluiscuniculi* into the GenBank database. This is the reason why sequence CP097901 was submitted as *Treponema paraluiscuniculi* strain L2, which was the original designation of the isolate. The sequencing reads assembled into draft genome sequences of the *TP*eL V246-08 (JAMZQX000000000) and Z27 A77/78 isolates were deposited in GenBank under Bioproject accession numbers PRJNA837879 and PRJNA837886, respectively.

Results

Sequencing of *TP*eL isolates

*TP*eL isolates were sequenced using PSGS, as described earlier [16, 17]. Briefly, DNA was amplified with 133 pairs of specific primers to obtain overlapping PCR products [18]). To facilitate sequencing of paralogous genes containing repetitive sequences, PCR products were

Table 1. Sequencing of *TP*eL isolates and the corresponding sequencing parameters.

	Breadth/depth of coverage (no. determined kb/average coverage)				
Isolate	Pool 1	Pool 2	Pool 3	Pool 4	Genome
V3603-13	253.56/	243.93/	251.98/	379.62/	1129.05/
	1044	1414	905	762	998
V246-08	249.63/	246.26/	251.79/	373.99/	1,121.67/
	201	1249	1206	1208	992
Z27 A77/78	52.53/	105.03/	90.38/	111.37/	359.3/
	615	913	905	608	773

https://doi.org/10.1371/journal.pone.0307196.t001

mixed in equimolar amounts into four distinct pools. The PCR products constituting each pool were labeled with multiplex identifier (MID) adapters and Illumina sequenced. Pools 1 through 4 corresponded to PCR amplicons grouped largely by their location in four genome segments. Sequencing results of *TP*eL isolates, including V3603-13, V246-08, and Z27 A77/78 samples, are shown in Table 1. Sequencing of *TP*eL V3603-13 resulted in 1,247,184, 1,481,259, 979,180, and 1,224,430 mapped deduplicated reads for pools 1, 2, 3, and 4, respectively. The *TP*eL V3603-13 genome was completely determined, leaving no sequencing gaps or ambiguities. Sequencing of *TP*eL V246-08 revealed 210,955, 624,319, 603,376, and 1,062,714, and Z27 A77/78 revealed 529,826, 769,416, 762,979, and 380,436 of mapped reads without duplicates in pools 1, 2, 3, and 4, respectively. Compared to the genome of *TP*eL strain V3603-13, 99.0% and 31.7% of the total genome of V246-08 and Z27 A77/78, respectively, were determined.

The overall genome structure of *TP*eL isolate V3603-13

The complete genome sequence of *TP*eL isolate V3603-13 comprised 1,132,489 nucleotides with no sequencing gaps or ambiguities. The genome sequence also includes variable sites in the *tprK* gene (*TP0897*), where the most common variant (with a frequency greater than 50%) was used for the final whole genome consensus sequence. The *TP*eL V3603-13 genome is 901 nt smaller than the reference genome of *TP*eC strain Cuniculi A (GenBank acc. number CP002103). There is overall gene synteny with the *TP*eC strain Cuniculi A genome and the human pathogen *T. pallidum* (subsp. *TPA*, *TPE*, and *TEN*). The *TP*eL V3603-13 genome showed a different *rrn* spacer pattern compared to the *TP*eC Cuniculi A genome, i.e., tRNA gene for Ala/Ile in V3603-13 compared to tRNA gene for Ile/Ala in the rabbit infecting strain of Cuniculi A. Sequences of the 5S, 16S, and 23S rRNA genes were identical in both operons and were without macrolide resistance coding single nucleotide changes A2058G or A2059G in the 23S rRNA genes [26, 27]. The number of 60 bp-long repetitions within the *arp* gene (*TP0433*) in *TP*eL V3603-13 and *TP*eC strain Cuniculi A genomes differed and was 19 and 21 repetitions, respectively. The number of 24 bp-long repetitions was 19 in the *TP0470* in *TP*eL V3603-13, and 6 in *TP*eC strain Cuniculi A. Table 2 summarizes the overall genome parameters of *TP*eL isolate V3603-13.

Major sequence differences between *TP*eL V3603-13 and *TP*eC Cuniculi A

Despite the overall genome synteny of *TP*eL V3603-13 and *TP*eC strain Cuniculi A, several genetic differences differentiate the genomes, including indels and single nucleotide variants (SNVs). An overview of the genetic differences between the rabbit infecting strain, Cuniculi A, and the V3603-13 genome is shown in Fig 1 and S1 Table. While most of the indels were found within genes, a minority was found in the intergenic regions (IGR). Moreover, most of the indels in the genes did not change the reading frame leading to indels of one to few amino

Table 2. Genomic features of the *TP*eLV3603-13 isolate compared to *TP*eC strain Cuniculi A.

Genome parameter	*TP*eL V3603-13	*TP*eC Cuniculi A
Genome size	1,132,489 bp	1,133,390 bp
G+C content	52.7%	52.8%
No. of predicted genes	1067, including 54 untranslated genes	1070, including 54 untranslated genes
No. of genes encoded on plus/minus DNA strand	577/490	577/493
Average/median gene length	1007/873 bp	1006/873 bp
Intergenic region length	60,749 bp (5.4% of the genome length)	62,494 bp (5.5% of the genome length)
No. of predicted genes encoding proteins similar to proteins of known function	647	650
No. of genes encoding conserved hypothetical proteins	137	139
No. of genes encoding hypothetical proteins	229	227
No. of pseudogenes or gene fragments	55	51
No. of tRNA loci	45	45
No. of rRNA operons	2* (6 genes)	2* (6 genes)
No. of other stable RNAs	3	3

**TP*eL V3603-13 and *TP*eC Cuniculi A genomes showed different *rrn* spacer patterns (Ala/Ile and Ile/Ala).

https://doi.org/10.1371/journal.pone.0307196.t002

acids. Compared to the *TP*eC strain Cuniculi A genome, *TP*eL V3603-13 contains four insertions greater than three nucleotides: a 1,874 nt-long insertion containing a *TP0126c-TP0129* fragment similar to the orthologous *TPA* strain Philadelphia 1 genomic sequence [28], a 79 nt insertion in the IGR between *TP0545-TP0546* similar to the *TPA* strain Philadelphia 1 genomic sequence, a 24 nt insertion in *TP0577* that is similar to the *TEN* strain 11q/j genomic sequence [29], and a 12 nt insertion (with undetected sequence similarity) in *TP0897* (*tprK*) (S1 Table).

Compared to the *TP*eC strain Cuniculi A genome, the *TP*eL V3603-13 genome contained eleven deletions greater than three nucleotides: a 2,932 nt deletion similar (but larger) to the deletion present in *TEN* strains [29] comprising the *tprFG* and *tmpC* (*TP0319*) genes, a six nt-long deletion in *TP0126b*, a 45 nt deletion in *TP0136*, a 13 nt deletion in *TP0146*, a six nt deletion in *TP0515*, a 36 nt deletion in *TP0548*, a six nt deletion in the IGR between *TP0650–TP0652*, a six nt deletion in *TP0733*, a 17 nt deletion in the IGR between *TP0895–TP0897*, a six nt deletion in *TP0897* (*tpr*K), and a six nt deletion in *TP0966* (S1 Table).
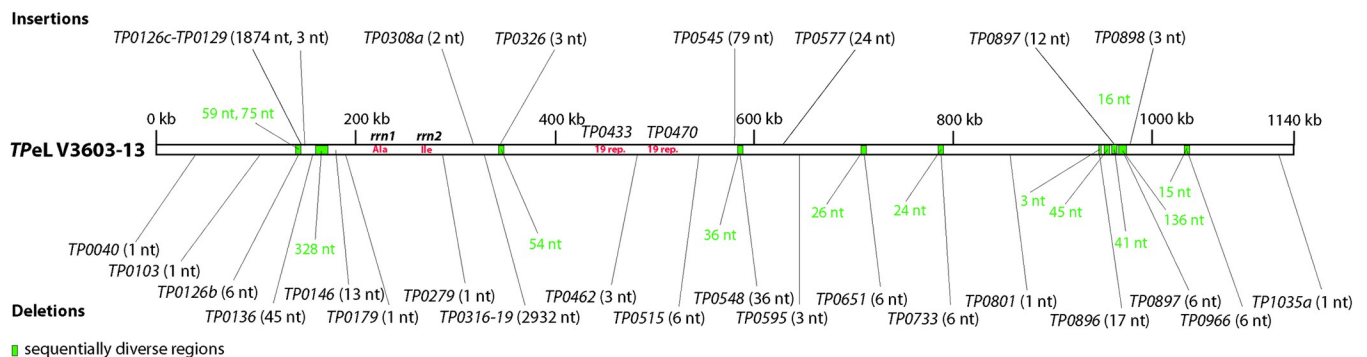

Fig 1. An overview of genome differences between *TP*eL V3603-13 and *TP*eC Cuniculi A. Insertions and deletions are shown above and below the schematic representation of the *TP*eL isolate V3603-13 chromosome, respectively. Green areas represent sequentially diverse regions, and the lengths of these regions are shown in green letters, while indel lengths are shown in black letters.

https://doi.org/10.1371/journal.pone.0307196.g001

In addition, there were 25 additional indels with lengths of 1–3 nucleotides, altogether spanning 36 nt (S1 Table). Three of these changes lead to in-frame insertions or deletions. Other indels located in genes lead to frameshifts, which are shown in S1 Table. A subset of insertion/deletion differences described in S1 Table was also found in *TPeL* isolate V246-08 supporting lack of sequencing artifacts.

Several sequentially diverse chromosomal regions (defined as multiple SNVs) were identified in *TPeL* V3603-13 genome: a 59 nt long region in *TP0126b* showing similarity to *tprK* of *TPA* SS14, a 75 nt long region in *TP0126b* showing similarity to *tprK* of *TPA* X-4, a 328 nt long region in *TP0136* with similarity to *TP0134* of *TPA* Philadelphia 1, a 54 nt long region in position of *TP0326* similar to *bamA* of *TPA* CZ_177zB (*TPA* Nichols-like strain) [30], a 36 nt long region not similar to any known treponemal sequences in position of *TP0548*, a 26 nt long region similar to IGR *TP0650-TP0652* of *TPA* Philadelphia 1 in position of IGR *TP0650-652*, a 24 nt long region not similar to any known treponemal sequences in position of *TP0733*, four regions upstream and in the *tprK* gene (3nt, 45 nt, 41nt, 136 nt) similar to *TPeC* or *TPA* SS14 sequences, and a 15 nt long region similar to *TPE* Fribourg-Blanc in position of *TP0966* (S1 Table).

Moreover, frameshift mutations resulting in two new pseudogenes (*TP0308a*, *TP1035a*) and major sequence changes, including protein elongation or shortening (*TP0040*, *TP0179*, *TP0279*, *TP0471*, *TP0778*, *TP0801*) were described. In addition, another frameshift mutation in existing pseudogenes was found in two pseudogenes (*TP0103*, *TP0146*) (S1 Table). Lastly, the *mglB* gene (*TP0545*) of *TPeL* V3603-13 appears to be fully functional and similar to *TPE* strains (due to the insertion of 79 nts in IGR between *TP0545–TP0546*). Similarly, a nucleotide insertion in V3603-13 *TP0617a* results in a gene version identical to the paralogous *TP0315* gene in the V3603-13 genome. Genes *TP0651* and *TP0896* in V3603-13 are longer but similar to the versions present in *TPE* strains.

As shown in Fig 2, repeat motifs within the *arp* gene (*TP0433*) and the arrangement of the repeat sequences differ between *TPeL* V3603-13 and *TPeC* Cuniculi A genomes.

Single nucleotide sequence differences between the *TPe*L V3603-13 and *TPe*C strain Cuniculi A genomes

In addition to the described differences, 309 additional nucleotide variants are dispersed throughout the genome in *TPeL* V3603-13 compared to the *TPeC* Cuniculi A reference genome. These nucleotide differences were primarily represented by single nucleotide variants
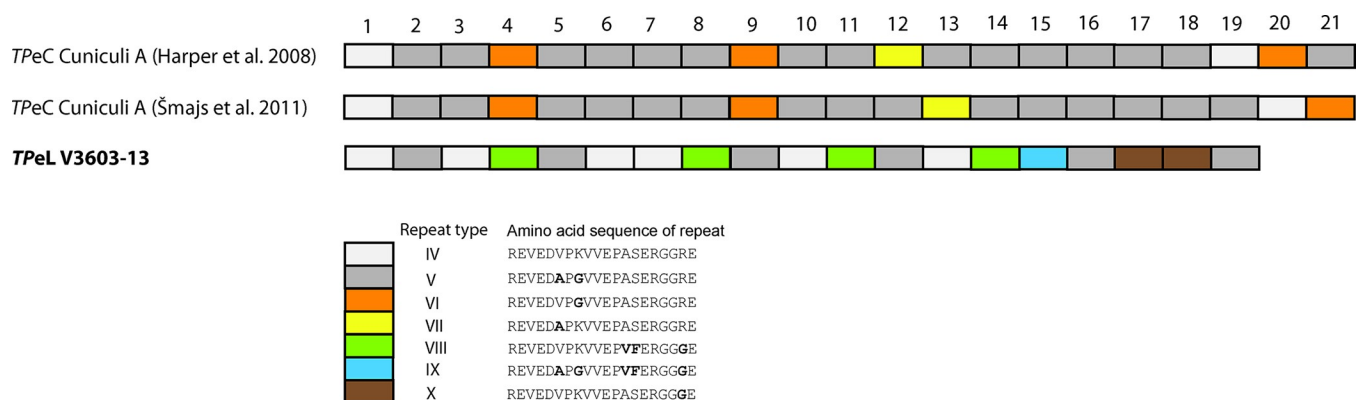

Fig 2. Composition of repeat motif regions observed in the *TPeL* V3603-13 and *TPeC* Cuniculi A genomes. Note that repeat motifs determined by Harper et al. [31], and by Šmajs et al. [6], differ, likely due to repeat reshuffling during individual laboratory handling of the Cuniculi A strain.

https://doi.org/10.1371/journal.pone.0307196.g002

(SNVs) and less frequently by double nucleotide differences (n = 8, covering 16 nucleotides). Analysis of the SNVs between *TP*e*L* samples V3603-13 and V246-08 revealed altogether 35 single nucleotide variants indicating that majority of genetic differences between hare and rabbit syphilis is shared by both V3603-13 and V246-08 samples.

The phylogenetic relationship of lagomorph and human pathogenic *Treponema*

To determine the genetic relatedness of *TP*eL/*TP*eC samples, a tree was built on partial *TP0548* sequences available for two *TP*eC strains/isolates (Cuniculi A; Cz-2020) and three *TP*eL isolates *TP*eL V3603-13, V246-08, and Z27 A77/78 (Fig 3A). Sample Cz-2020 is the only other known member of *TP*eC and only sequences of few loci are available. Therefore, locus TP0548 was selected as the only locus with available sequences in all 5 *T. paraluisleporidarum* samples. *TP*eL isolates clustered together but separate from *TP*eC isolates. To determine the genetic relatedness of both ecovars using a more robust analysis, a tree was built on partial genome sequences available for the complete genome of *TP*eL V3603-13 and *TP*eL isolates V246-08 and *TP*eL Z27 A77/78. Altogether, about one-third of the genome length (355,133 nucleotides) was used to construct the phylogenetic tree (Fig 3B). Both *TP*eL isolates V3603-13 and V246-08, which were isolated from *L. timidus*, clustered closely with the *TP*eC Cuniculi A genome. At the same time, the Z27 A77/78 sample, taken from the *L. europeus* in The Netherlands [3], was more distantly related.

Discussion

In this study, we have determined the first whole genome sequence of the causative agent of hare syphilis. *TP*eL isolate V3603-13 was obtained from an infected Swedish mountain hare. In addition, we generated two partial genome sequences for *TP*eL isolate V246-08, also from an infected mountain hare in Sweden, and Z27 A77/78 that came from a European brown hare in The Netherlands [3]. While the sample quality of V246-08 and Z27 A77/78 (quality refers to the amount of available treponemal DNA) did not allow us to sequence the whole genome, considerable parts of the corresponding genomes were sequenced with the shortest
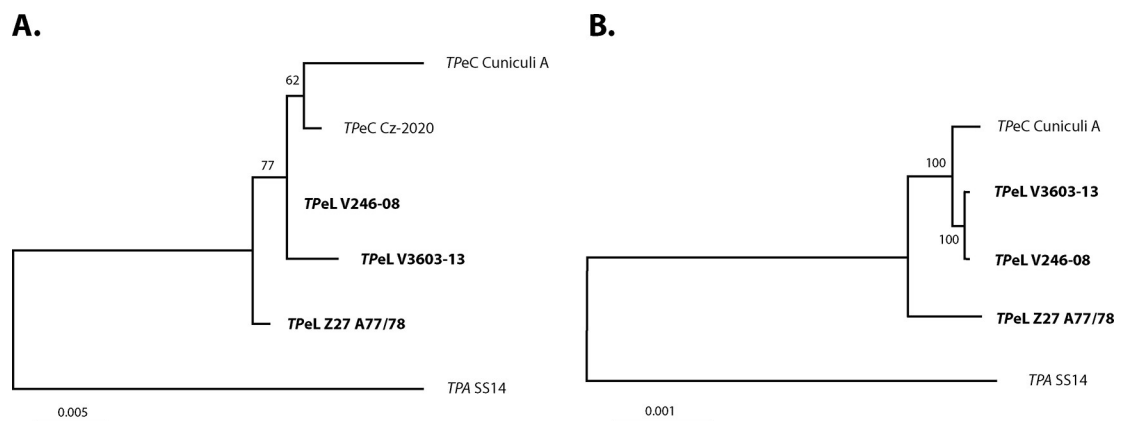

Fig 3. Molecular phylogeny of *TP*eL and *TP*eC samples. A. The tree was built on partial *TP0548* sequences (*TP*eC Cz-2020 and *TPA* SS14) [32, 33]. There was a total of 777 positions in the final dataset. **B.** The tree was built on available genome sequences. There were 355,133 positions in the final dataset. The trees were constructed using the Maximum Likelihood method based on the Tamura-Nei model. Bootstrap support is shown next to the branches. The scale shows the number of substitutions per site. As an outgroup, the *TPA* SS14 sequence was used. All positions containing gaps and missing data were omitted.

https://doi.org/10.1371/journal.pone.0307196.g003

determined sequence from Z27 A77/78 being equal to about a third of the *TP*eC strain Cuniculi A genome, i.e., 31.7%, 355,133 nucleotides.

Until now, only a limited number of studies have been published on the genetics of *TP*eL [3, 34] even though infections caused by *TP*eL appear to be widespread among European hares [11–14, 35]. Since most of the previous and current *TP*eL infections are subclinical and detectable only with serology and/or direct detection of *TP*eL DNA [8], the scarcity of clinical symptoms could explain the lack of scientific data on treponemal infections in hares.

The genome of *TP*eL V3603-13 comprised of 1,132,489 nt and is the smallest genome determined to date of the group of closely related pathogenic treponemes, including human pathogenic *TPA*, *TPE*, and *TEN*. At the whole genome level, genetic similarity between *TP*eL and *TP*eC was > 99.8%, with almost complete gene synteny between the genomes. Interestingly, both the *TP*eL V3603-13 and *TP*eC Cuniculi A genomes showed different *rrn* spacer patterns (Ala/Ile and Ile/Ala, respectively), a feature previously noticed within *TPA* and *TPE* strains as a result of reciprocal intragenomic translocation events [2, 34]. The number of repetitions within the *arp* gene (*TP0433*) in *TP*eL V3603-13 was less than *TP*eC Cuniculi A (19 vs. 21), and the gene was composed of different sequence motifs compared to the *arp* of Cuniculi A (Fig 2) [31]. Among *TPA* isolates, the number and the structure of the *arp* repeats were shown to generally correlate with whole genome phylogeny [36]. Interestingly, the Cuniculi A strain *arp* gene analyzed by Strouhal et al. [5] and Šmajs et al. [6] differed in the sequence of repeat motifs from the same strain analyzed by Harper et al. [31], suggesting that these gene repeat components are prone to frequent reshuffling and accumulation of point mutations. Consistent with this observation, previous studies on the number of *arp* repeat motifs among *TPA* isolates revealed differences in the repeat numbers between whole blood and swab samples taken from the same patient [37], again suggesting the genetic plasticity of this locus. Both the *TP*eC and *TP*eL *arp* genes contained multiple repeat motifs (Cuniculi A, n = 4; V3603-13, n = 5), a feature that has only been described for *TPA* and *TP*eC strains but not for *TPE* (including stains infecting nonhuman primates [38] and *TEN* strains [31].

One major difference between the rabbit infecting *TP*eC Cuniculi A and the hare *TP*eL V3603-13 genome was the presence of an 1874 nt-long insertion in V3603-13. This insertion mainly corresponded to the *TP0126c-TP0129* genes in other *TPA* and *TPE* treponemes, which suggests that the corresponding region was deleted during the evolution of the Cuniculi A strain. In a subpopulation of the *TPA* Nichols strain, a 1.3-kb deletion at this genomic locus (i.e., in *TP0126*) was found in an earlier study [39]; both the *TPA* Nichols and *TP*eC Cuniculi A deletions overlapped by 376 nt. Another previous study revealed that *TP0126*, *TP0126b*, *TP0126c*, *TP0127d*, *TP0128*, and *TP0130* (in addition to other genes) have a modular genetic structure enabling rapid genetic diversification of treponemal strains [40]. At the same time, these regions, due to the presence of repetitive sequences, may also have less genetic stability and be prone to deletion events. Since the Nichols strain has been propagated in rabbits for some time, the deletion of this region may represent one of the adaptations to infecting rabbits (or hares). At least two additional insertions (in *TP0545-546* and *TP0577*) in the V3603-13 genome matched sequences present in other *TPA* and *TPE* strains, suggesting additional deletions in the evolution of the rabbit pathogen, *TP*eC. Based on cross-infection experiments [3], it was predicted that *TP*eL can infect both rabbits and hares while *TP*eC infects only rabbits. Therefore, the deleted regions in *TP*eC may be important relative to hare infections. However, genome manipulation experiments would be needed to prove this and better understand the host-pathogen evolution of treponemal pathogens.

Another striking feature of the *TP*eL V3603-13 genome is a 2932 nt-long deletion in V3603-13 affecting *tprFG* and *tmpC* (*TP0319*) genes (compared to *TP*eC strain Cuniculi A). The genome of *TP*eC Cuniculi A has a similar but smaller deletion in this region (compared to

TPA strain Nichols), suggesting the independent origin of these deletions. Moreover, the surrounding genes, including *TP0309-311*, *TP0313*, *TP0315*, and *TP0318*, contain frameshift mutations in both *TP*eC Cuniculi A and *TP*eL V3603-13 genomes, and it is therefore possible that the entire region is not needed for infection of lagomorphs, including both rabbits and hares. Yet, it is not clear if the *tprFG* region is required for infection of humans since similar but smaller deletions in the *TEN* strains were found in subpopulations of individual *TEN* strains that were experimentally propagated in rabbits [17, 41]. On the other hand, the *TEN* strains isolated directly from Cuban patients suspected of having syphilis [42] contained a deletion at the *tprFG* locus [43] which was similar to one described in the *TEN* reference strains Bosnia A and Iraq B [17, 41]. This suggests that the deletion does not preclude infection in humans nor the emergence of early syphilis-like symptoms [29, 42, 44].

Analysis of the *TP0136* gene in the V3603-13 genome revealed a 328 nt-long region similar to *TP0134* in *TPA* Philadelphia 1 [28] while the *TP0136* gene in Cuniculi A had a sequence similar to *TP0133* (Fig 4). Since the *TP0134* gene (including the donor sequence) was deleted from both the *TP*eC Cuniculi A and *TP*eL V3603-13 genomes, this finding cannot be explained by sequence gene transfer from the *TP0134* locus. An explanation is the presence of a common ancestor, i.e., for the *TP*eC Cuniculi A and *TP*eL V3603-13 strain, with a *TP0134*-like sequence at the *TP0136* locus (Fig 4). Subsequently, a deletion might have occurred. The lineage then became ancestral to the *TP*eL genome while additional gene conversion leading to replacement of the *TP0134*-like sequence by the *TP0133*-like sequence occurred during the evolution of the rabbit infecting *TP*eC genome of Cuniculi A (Fig 4). Although this model suggests *TP*eC being an evolutionarily modern version of *TP*eL, other alternative explanations, including gene recombination events [45] are also possible.

The *TP0136* locus encodes for the fibronectin-binding protein [46] and is known to be a recombinant and positively selected locus in *TPA* treponemes [28, 47]. Moreover, this locus has been predicted to have a modular structure that can lead to relatively frequent sequence changes and indels [40, 48]. However, the part of *TP0136* with the modular structure was not affected during the above-described replacements (Fig 4). All this evidence points to the importance of the *TP0136* locus during human and animal infections.
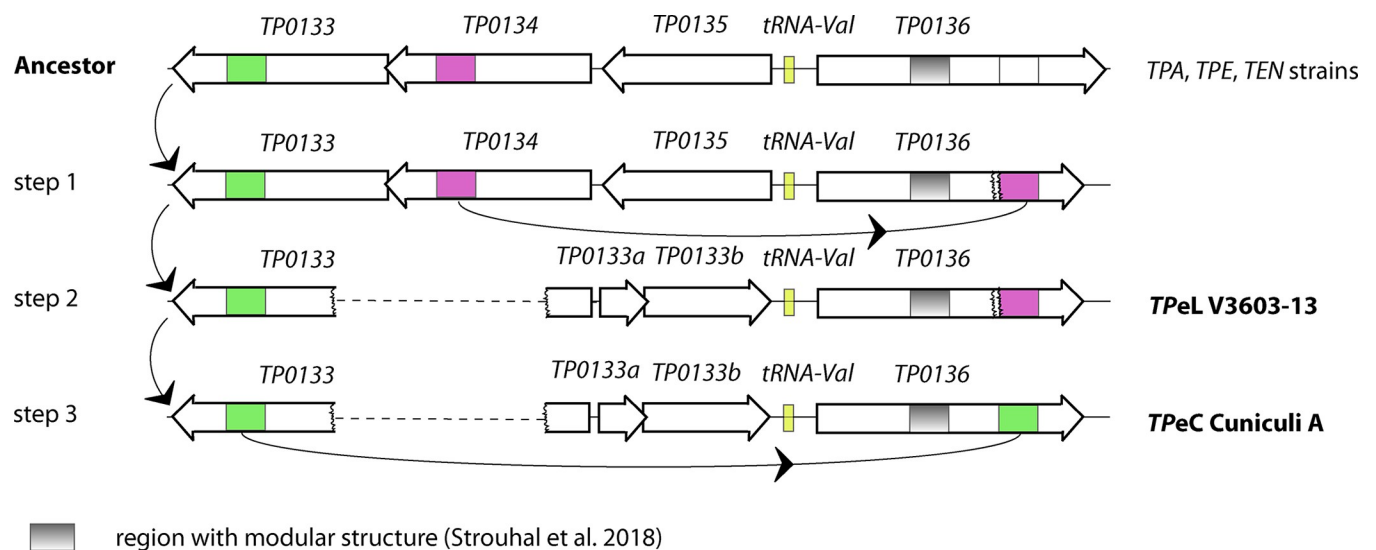

Fig 4. A schematic representation of a possible evolutionary model of the *TP0136* locus in *TP*eL and *TP*eC treponemes. The evolution of this region required several steps, including two gene conversion events and one deletion. The part of *TP0136* showing a modular structure was not affected during these changes.

Besides deletions, insertions, and stretches of nucleotide sequences resulting from gene conversions, we found 309 single nucleotide variants (SNVs) between the *TP*eC Cuniculi A and *Tp*eL V3603-13 genomes. These differences were present in single or two nt sequences (n = 8), suggesting that these events represented accumulated mutations. Although mutation rates in *TP*eL and *TP*eC remain unknown, previous studies estimated the mutation rates of the genetically related treponemes *TPA* and *TPE*. The mutation rate was estimated to be 2.8–4.1×10^{-10} per site per generation [49, 50], which corresponds to a mutation rate of 0.846–1.21×10^{-7} per nucleotide site per year. Other studies, based on temporal analyses with relaxed clock models, estimated slightly higher treponemal mutation rates, i.e., 3.02×10^{-7} [51] and 6.6×10^{-7} [52] per nucleotide site per year, which would correspond to 0.5–3 kiloyears of separate evolution for lagomorph treponemes. We note here that these mutation rate estimates, based on *TPA/E* strains in humans, do not necessarily correspond to lagomorph infecting *TP*eC/L. Moreover, mutation rates could vary over time, especially during the early phases of adaptation to a new host (species), and mutation rates could be even lower in *TPA/TPE* [18, 50].

The mutations causing resistance to macrolide antibiotics (i.e., A2058G or A2059G mutations in the 23S rRNA genes [26, 27, 53]) were not found among the analyzed *TP*eL isolates. *TP*eL strains and *TP*eC strain Cuniculi A remain the only group of pathogenic treponemes where these mutations have not been detected, in contrast to *TPA* [26, 27], *TPE* [54, 55], and *TEN* [56, 57]. The absence of macrolide resistance in *TP*eL and *TP*eC is consistent with the absence of macrolide resistance in *TPE* isolated from wild nonhuman primates [15, 38, 58, 59]. Although speculative and theoretical resistance could quickly emerge in wildlife, too, particularly under selection pressure caused by the mass administration of antimicrobial treatments.

Functionally, the most noticeable differences were found in *tpr* genes (*tprF, -G, -K*), genes that encode for outer membrane proteins/antigens (*TP0136, tmpC, TP0326, TP0433, TP0470, TP0471, TP0515, TP0548, TP0651, TP0733,* and *TP0966*), and transporter/chemotaxis proteins (*TP0040, TP0146, TP0545,* and *TP0577*). Genes coding for metabolic functions (*TP0179, TP0279, mazG,* and *TP0801*) and hypothetical proteins (*TP0126b-129, TP0308a, TP0617a, TP0896,* and *TP1035a*) were relatively infrequent. Out of these functional groups, outer membrane proteins appear to be the most important part of the proteome, i.e., they facilitate immune evasion and adaptation to different hosts.

Conclusions

We determined the first complete genome sequence of *Treponema paraluisleporidarum* ecovar Lepus (*TP*eL V3603-13) isolated from a naturally infected mountain hare in Sweden and provided draft genome sequences of two additional *TP*eL strains. The agent causing hare syphilis was found to be similar but distinct from the rabbit syphilis treponeme; the major predicted proteome differences were associated with the Tpr proteins and outer membrane proteins/antigens. Based on previous estimations of *TPA* and *TPE* mutation rates, both *TP*eL and *TP*eC appear to be separated by 0.5–3 kiloyears of lagomorph treponeme evolution. However, more samples from naturally infected hares and rabbits need to be analyzed to understand better the genetic diversity of lagomorph syphilis and the phylogeny of lagomorph treponemes.

Supporting information

S1 Table. Genetic differences found in the *TP*eL V3603-13 genome when compared to the *TP*eC Cuniculi A genome.
(DOCX)

Acknowledgments

Author Contributions

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(12):744–759. https://doi.org/10.1038/nrmicro.2016.141 PMID: 27721440

2. 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

3. Lumeij JT, Mikalová L, Šmajs D. Is there a difference between hare syphilis and rabbit syphilis? Cross infection experiments between rabbits and hares. Vet Microbiol. 2013; 164(1–2):190–4. https://doi.org/10.1016/j.vetmic.2013.02.001 PMID: 23473645

4. Turner TB, Hollander DH. Biology of the treponematoses based on studies carried out at the International Treponematosis Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. Monogr Ser World Health Organ. 1957;(35):3–266. PMID: 13423342

5. 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(12):5859–66. https://doi.org/10.1128/IAI.00709-07 PMID: 17893135

6. Šmajs D, Zobaníková M, Strouhal M, Čejková D, Dugan-Rocha S, Pospíšilová P et al. Complete genome sequence of Treponema paraluiscuniculi, strain Cuniculi A: The loss of infectivity to humans is associated with genome decay. PloS ONE. 2011; 6:e20415. https://doi.org/10.1371/journal.pone.0020415 PMID: 21655244

7. Smith JL, and Pesetsky BR The current status of *Treponema cuniculi*: review of the literature. Br. J. Vener. Dis. 1967; 43, 117–127. https://doi.org/10.1136/sti.43.2.117 PMID: 5338028

8. Knauf S, Hisgen L, Ågren EO, Barlow AM, Faehndrich M, Voigt U et al. High prevalence and genetic diversity of *Treponema paraluisleporidarum* isolates in European lagomorphs. Microbiol Spectr. 2024; 12(1):e0177423. https://doi.org/10.1128/spectrum.01774-23 PMID: 38095473

9. Lumeij JT, de Koning J, Bosma RB, van der Sluis JJ, Schellekens JF. Treponemal infections in hares in The Netherlands. J Clin Microbiol. 1994 Feb; 32(2):543–6. https://doi.org/10.1128/jcm.32.2.543-546.1994 PMID: 8150971

10. Jakšić BL. Brucelose et syphilis du lievre. Vet. Glas. 1957; 11, 423–425.

11. Hisgen L, Abel L, Hallmaier-Wacker LK, Lueert S, Siebert U, Faehndrich M et al. High syphilis seropositivity in European brown hares (Lepus europaeus), Lower Saxony, Germany. Transbound Emerg Dis. 2020; 67, 2240–2244 https://doi.org/10.1111/tbed.13551 PMID: 32196988

12. Nováková M, Najt D, Mikalová L, Kostková M, Vrbová E, Strouhal M et al. (2019). First report of hare treponematosis seroprevalence of European brown hares (*Lepus europaeus*) in the Czech Republic: seroprevalence negatively correlates with altitude of sampling areas. BMC Vet. Res. 2019; 15, 350. https://doi.org/10.1186/s12917-019-2086-3 PMID: 31627750

13. Posautz A, Leidinger E, Knauer F, Hoffman D, Suchentrunk F, Walzer C, et al. Seroprevalence of *Treponema* sp. in European brown hares (*Lepus europaeus*) in Austria and Germany. Wien. Tierarztl. Monat. 2014; 101, 281–285.

14. Verin R, Pestelli M, Poli A. Treponemal infection in free-ranging European brown hares (Lepus europaeus) in Central Italy: Serology and epidemiology. J Wildl. Dis. 2012; 48, 1079–82. https://doi.org/10.7589/2011-03-069 PMID: 23060513

15. Knauf S, Raphael J, Mitjà O, Lejora IAV, Chuma IS, Batamuzi EK, et al. Isolation of *treponema* DNA from necrophagous flies in a natural ecosystem. EBioMedicine. 2016; 11:85–90. https://doi.org/10.1016/j.ebiom.2016.07.033 PMID: 27488881

16. Zobaníková M, Strouhal M, Mikalová L, Cejková 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. 2013 Apr 18; 7(4):e2172. https://doi.org/10.1371/journal.pntd.0002172 PMID: 23638193; PMCID: PMC3630124.

17. Štaudová B, Strouhal M, Zobaníková M, Čejková D, Fulton LL, Chen L, et al. Whole Genome Sequence of the *Treponema pallidum* subsp. *endemicum* Strain Bosnia A: The Genome Is Related to Yaws Treponemes but Contains Few Loci Similar to Syphilis Treponemes. PLoS Negl Trop Dis. 2014; 8: 3261. https://doi.org/10.1371/journal.pntd.0003261 PMID: 25375929

18. Weinstock G.M., Norris S.J., Sodergren E.J. and Šmajs D. (2000). Identification of Virulence Genes *in Silico*: Infectious Disease Genomics. In Virulence Mechanisms of Bacterial Pathogens (eds Brogden K. A., Roth J.A., Stanton T.B., Bolin C.A., Minion F. Chrisand Wannemuehler M.J. https://doi.org/10.1128/9781555818111.ch17

19. Mikalová L, Strouhal M, Čejková D, Zobaníková M, Pospíšilová P, Norris SJ et al. Genome analysis of *Treponema pallidum* subsp. *pallidum* and subsp. pertenue strains: most of the genetic differences are localized in six regions. PLoS One. 2010; 5:e15713. https://doi.org/10.1371/journal.pone.0015713 PMID: 21209953

20. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: ht*TP*://www.bioinformatics.babraham.ac.uk/projects/fastqc/. 2010.

21. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal. 2011; 17: 10–12. https://doi.org/10.14806/ej.17.1.200

22. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943

23. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20: 1297–1303. https://doi.org/10.1101/gr.107524.110 PMID: 20644199

24. Breese MR, Liu Y. NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. Bioinformatics. 2013; 29: 494–496. https://doi.org/10.1093/bioinformatics/bts731 PMID: 23314324

25. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016; 33(7):1870–4. https://doi.org/10.1093/molbev/msw054 PMID: 27004904

26. 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.

27. Matějková P, Flasarová M, Zákoucká H, Bořek M, Křemenová 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; 58:832–836.

28. Grillová L, Oppelt J, Mikalová L, Nováková M, Giacani L, Niesnerová, et al. Directly sequenced genomes of contemporary strains of syphilis reveal recombination-driven diversity in genes encoding predicted surface-exposed antigens. Front. Microbiol. 2019; 10, 1691. https://doi.org/10.3389/fmicb.2019.01691 PMID: 31417509

29. Mikalová L, Strouhal M., Grange PA, Gaudin C, Janier M, Benhaddou N, et al. Human *Treponema pallidum* 11q/j isolate belongs to subsp. *endemicum* but contains two loci with a sequence in TP0548 and TP0488 similar to subsp. *pertenue* and subsp. *pallidum*, respectively. PLoS Negl. Trop. Dis. 2017; 11: e0005434. https://doi.org/10.1371/journal.pntd.0005434 PMID: 28263990

30. Kumar S, Caimano MJ, Anand A, Dey A, Hawley KL, LeDoyt ME, et al. Sequence Variation of Rare Outer Membrane Protein β-Barrel Domains in Clinical Strains Provides Insights into the Evolution of *Treponema pallidum* subsp. *pallidum*, the Syphilis Spirochete. mBio. 2018; 9(3):e01006–18. https://doi.org/10.1128/mBio.01006-18 PMID: 29895642

31. Harper K.N., Liu H., Ocampo P.S., Steiner B.M., Martin A., Levert K., et al. The sequence of the acidic repeat protein (*arp*) gene differentiates venereal from nonvenereal *Treponema pallidum* subspecies, and the gene has evolved under strong positive selection in the subspecies that causes syphilis. FEMS Immunol. Med. Microbiol. 2008; 53, 322–332. https://doi.org/10.1111/j.1574-695X.2008.00427.x PMID: 18554302

32. Jekl V, Nováková M, Jeklová E, Pospíšilová P, Křenová J, Faldyna M et al. Penicillin treatment failure in rabbit syphilis due to the persistence of *Treponemes* (*Treponema paraluisleporidarum* ecovar Cuniculus) in the focus of infection. Front Vet Sci. 2021; 8:675631. https://doi.org/10.3389/fvets.2021.675631 PMID: 34222401

33. Pětrošová H, Pospíšilová P, Strouhal S, Čejková D, Chen L, Mikalová L, et al. Resequencing of *Treponema pallidum* ssp. *pallidum* strains Nichols and SS14: correction of sequencing errors resulted in increased diversity of syphilis treponeme subclusters. PLoS One. 2013; 8:e74319. https://doi.org/10.1371/journal.pone.0074319 PMID: 24058545

34. 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 *Treponema paraluiscuniculi* strains. Journal of Medical Microbiology. 2013; 62: 196–207. https://doi.org/10.1099/jmm.0.050658–0

35. Horvath I, Kemenes F, Molnar L, Szeky A, Racz I. Experimental syphilis and serological examination for treponematosis in hares. Infect. Immun. 1980; 27, 231–234. https://doi.org/10.1128/iai.27.1.231-234.1980 PMID: 6987170

36. Lieberman NAP, Armstrong TD, Chung B, Pfalmer D, Hennelly CM, Haynes A, et al. High-throughput nanopore sequencing of *Treponema pallidum* tandem repeat genes *arp* and *tp0470* reveals clade-specific patterns and recapitulates global whole genome phylogeny. Front Microbiol. 2022; 13:1007056. https://doi.org/10.3389/fmicb.2022.1007056 PMID: 36204625

37. Mikalová L, Pospíšilová P, Woznicová V, Kuklová I, Zákoucká H, Šmajs D. Comparison of CDC and sequence-based molecular typing of syphilis treponemes: *tpr* and *arp* loci are variable in multiple samples from the same patient. BMC Microbiol. 2013; 13:178. https://doi.org/10.1186/1471-2180-13-178 PMID: 23898829

38. Knauf S, Gogarten JF, Schuenemann VJ, De Nys HM, Düx A, Strouhal M, et al. Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. Emerg. Microbes Infect. 2018; 7:157. https://doi.org/10.1038/s41426-018-0156-4 PMID: 30228266

39. Šmajs D, McKevitt M, Wang L, Howell JK, Norris SJ, Palzkill T, et al. BAC Library of T. *pallidum* DNA in E. coli. Genome Res. 2002; 12: 515–522. https://doi.org/10.1101/gr.207302 PMID: 11875041

40. 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. PLoS Negl Trop Dis. 2018; 12: e0006867. https://doi.org/10.1371/journal.pntd.0006867 PMID: 30303967

41. Mikalová L, Janečková K, Nováková M, Strouhal M, Čejková D, Harper KN, et al. Whole genome sequence of the *Treponema pallidum* subsp. *endemicum* strain Iraq B: A subpopulation of bejel treponemes contains full-length *tpr*F and *tpr*G genes similar to those present in *T. p.* subsp. *pertenue* strains. PLoS One. 2020; 15(4):e0230926. https://doi.org/10.1371/journal.pone.0230926 PMID: 32236138

42. 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. Clinical Microbiology and Infection. 2018; 24: 1210.e1–1210.e5. https://doi.org/10.1016/j.cmi.2018.02.006 PMID: 29454847

43. Vrbová E, Noda AA, Grillová L, Rodríguez I, Forsyth A, Oppelt J et al. Whole genome sequences of *Treponema pallidum* subsp. *endemicum* isolated from Cuban patients: The non-clonal character of isolates suggests a persistent human infection rather than a single outbreak. PLoS Negl Trop Dis. 2022; 16:e0009900. https://doi.org/10.1371/journal.pntd.0009900 PMID: 35687593

44. Kawahata T, Kojima Y, Furubayashi K, Shinohara K, Shimizu T, Komano J et al. Bejel, a Nonvenereal Treponematosis, among Men Who Have Sex with Men, Japan. Emerg Infect Dis. 2019; 25(8):1581–1583. https://doi.org/10.3201/eid2508.181690 PMID: 31310214

45. Pla-Díaz M, Sánchez-Busó L, Giacani L, Šmajs D, Bosshard PP, Bagheri HC, et al. Evolutionary processes in the emergence and recent spread of the syphilis agent, *Treponema pallidum*. Mol Biol Evol. 2022; 39(1):msab318. https://doi.org/10.1093/molbev/msab318 PMID: 34791386

46. Brinkman MB, McGill MA, Pettersson J, Rogers A, Matejková P, Smajs D. A novel *Treponema pallidum* antigen, TP0136, is an outer membrane protein that binds human fibronectin. Infect Immun. 2008; 76 (5):1848–57. https://doi.org/10.1128/IAI.01424-07 PMID: 18332212

47. Maděránková D, Mikalová L, Strouhal M, Vadják Š, Kuklová I, Pospíšilová P, et al. Identification of positively selected genes in human pathogenic treponemes: Syphilis-, yaws-, and bejel-causing strains differ in sets of genes showing adaptive evolution. PLOS Neglected Tropical Diseases. 2019; 13: e0007463. https://doi.org/10.1371/journal.pntd.0007463 PMID: 31216284

48. Zobaníková M, Mikolka P, Čejková D, Pospíšilová P, Chen L, Strouhal M, et al. Complete genome sequence of *Treponema pallidum* strain DAL-1. Stand Genomic Sci. 2012; 7(1):12–21. https://doi.org/10.4056/sigs.2615838 PMID: 23449808

49. Grillová L, Giacani L, Mikalová L, Strouhal M, Strnadel R, Marra C, et al. Sequencing of *Treponema pallidum* subsp. *pallidum* from isolate UZ1974 using Anti-Treponemal Antibodies Enrichment: First complete whole genome sequence obtained directly from human clinical material. PLoS One. 2018; 13: e0202619. https://doi.org/10.1371/journal.pone.0202619 PMID: 30130365

50. Strouhal M, Mikalová L, Havlíčková P, Tenti P, Čejková D, Rychlík I, et al. Complete genome sequences of two strains of *Treponema pallidum* subsp. *pertenue* from Ghana, Africa: Identical genome sequences in samples isolated more than 7 years apart. PLoS Negl Trop Dis. 2017; 11: e0005894. https://doi.org/10.1371/journal.pntd.0005894 PMID: 28886021

51. Lieberman NAP, Lin MJ, Xie H, Shrestha L, Nguyen T, Huang ML, et al. *Treponema pallidum* genome sequencing from six continents reveals variability in vaccine candidate genes and dominance of Nichols clade strains in Madagascar. PLoS Negl Trop Dis. 2021; 15(12):e0010063. https://doi.org/10.1371/journal.pntd.0010063 PMID: 34936652

52. 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

53. 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 (4):678–83. https://doi.org/10.4269/ajtmh.15-0316 PMID: 26217043

54. Mitjà O, Godornes C, Houinei W, Kapa A, Paru R, Abel H, et al. Re-emergence of yaws after single mass azithromycin treatment followed by targeted treatment: a longitudinal study. Lancet. 2018; 391:1599–1607. https://doi.org/10.1016/S0140-6736(18)30204-6 PMID: 29428183

55. John LN, Beiras CG, Houinei W, Medappa M, Sabok M, Kolmau R, et al. Trial of Three Rounds of Mass Azithromycin Administration for Yaws Eradication. N Engl J Med. 2022; 386(1):47–56. https://doi.org/10.1056/NEJMoa2109449 PMID: 34986286

56. 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 Clin Microbiol. 2019; 57(1):e01148–18. https://doi.org/10.1128/JCM.01148-18 PMID: 30381419

57. Shinohara K, Furubayashi K, Kojima Y, Mori H, Komano J, Kawahata T. Clinical perspectives of *Treponema pallidum* subsp. *endemicum* infection in adults, particularly men who have sex with men in the Kansai area, Japan: A case series. J Infect Chemother. 2022; 28(3):444–450. https://doi.org/10.1016/j.jiac.2021.11.012 PMID: 34836779

58. Chuma IS, Roos C, Atickem A, Bohm T, Collins DA, Grillová L, et al. Strain diversity of *Treponema pallidum* subsp. *pertenue* suggests rare interspecies transmission in African nonhuman primates. Sci Rep. 2019; 9, 14243 https://doi.org/10.1038/s41598-019-50779-9 PMID: 31578447

59. Janečková K, Roos C, Fedrová P, Tom N, Čejková D, Lueert S et al. The genomes of the yaws bacterium, Treponema pallidum subsp. pertenue, of nonhuman primate and human origin are not genomically distinct. PLoS Negl Trop Dis. 2023; 17:e0011602. https://doi.org/10.1371/journal.pntd.0011602 PMID: 37703251