

# M200 WHOLE GENOME SEQUENCE OF *TREPONEMA PARALISLEPORIDARUM* ECOVAR LEPUS, STRAIN V3603-13

Petra Pospíšilová<sup>1</sup>, Darina Čojková<sup>2,3</sup>, Pavla Havlíčková<sup>1</sup>, Pavla Fedrová<sup>1</sup>, Lenka Mikalová<sup>1</sup>, David Najt<sup>1</sup>, Nikola Tom<sup>1</sup>, Linda Hisgen<sup>4</sup>, Simone Lueert<sup>4</sup>, Johannes T. Lumeij<sup>5</sup>, Erik O. Ågren<sup>6</sup>, Sascha Knaut<sup>4</sup>, and David Šmajs<sup>1\*</sup>

<sup>1</sup>Department of Biology, Faculty of Medicine, Masaryk University, Kamenice 5/A6, 625 00, Brno, Czech Republic

<sup>2</sup>Veterinary Research Institute, Hudcova 70, 621 00, Brno, Czech Republic

<sup>3</sup>Department of Biomedical Engineering, Brno University of Technology, Technická 12, 616 00 Brno, Czech Republic

<sup>4</sup>Institute of International Animal Health/One Health, Friedrich-Loeffler-Institute, Federal Research Institute for Animal Health, Stödter-10, 17493 Greifswald – Insel Riems

<sup>5</sup>Division of Zoological Medicine, Department of Clinical Sciences of Companion Animals, Faculty of Veterinary Medicine, Utrecht University, P.O. Box 80.154, Utrecht 3508 TD, The Netherlands.

<sup>6</sup>Department of Pathology and Wildlife Diseases, National Veterinary Institute, Uppsala, Sweden

## Background

Lagomorph syphilis is caused by *Treponema paraluisleporidarum* ecovar Cuniculi (*TPeC*) and ecovar Lepus (*TPeL*) infecting rabbits and hares respectively. This pathogen is closely related to the human pathogen *Treponema pallidum*, causing syphilis (*ssp. pallidum*), yaws (*ssp. pertenue*), and bejel (*ssp. endemicum*). In this study we have determined complete genome sequence of *Treponema paraluisleporidarum* ecovar Lepus, isolate V3603-13, using a previously described Pooled Sequence Genome Sequencing (PSGS) technique, and compared it to the previously determined genomic sequence *TPeC* strain Cuniculi A. The agent of hare syphilis has been found similar but distinct from the rabbit syphilis treponeme.

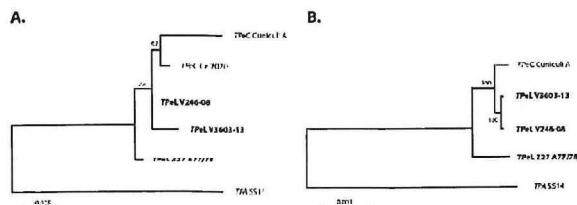
**Table 1.** Sequencing of *TPeL* isolates and the corresponding sequencing parameters

Isolate	Breadth/depth of coverage (no. determined kb/average coverage)				
	Pool 1	Pool 2	Pool 3	Pool 4	Genome
V3603-13	253.56/1,044	243.93/1,414	251.98/905	379.62/762	1,129.05/998
V246-08	249.63/201	246.26/1,249	251.79/1,206	373.59/1,208	1,121.67/992
V247/78	52.53/615	105.03/913	90.38/905	111.37/608	359.3/773

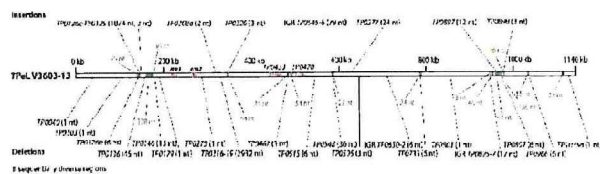
## Results

The *TPeL* V3603-13 genome revealed an overall gene synteny with the *TPeC* Cuniculi A genome and with the human pathogen *T. pallidum*. Compared to the *TPeC* Cuniculi A genome, *TPeL* V3603-13 contained four insertions and 11 deletions larger than three nucleotides (ranging between 6–2,932 nt). In addition, there were 25 indels one to three nucleotides long, altogether spanning 36 nt. Moreover, the nucleotide variants between the *TPeC* Cuniculi A genome and *TPeL* V3603-13 included single nucleotide variants (SNVs, n=293) and double nucleotide differences (n=8, covering 16 nucleotides). Major proteome differences between *TPeL* and *TPeC* were found to be encoded by the *tpr* gene family and by genes encoding outer membrane proteins, which suggests that these components are essential for host adaptation of lagomorph syphilis agent.

**Fig. 3. Molecular phylogeny of *TPeL* and *TPeC* samples.** A. The tree was built on partial *TP0548* sequences (*TPeC* Cz-2020 and *TPeL* SS14) (Jekl *et al.*, 2021; Petrosova *et al.*, 2013). There were a total of 777 positions in the final dataset. B. The tree was built on available genome sequences. There were total of 355133 positions in the final dataset. The trees were inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The bootstrap support is shown next to the branches. The scale shows the number of substitutions per site. As an outgroup, *TPA* SS14 sequence was used. All positions containing gaps and missing data were omitted.



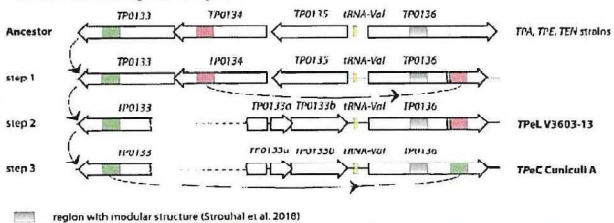
**Fig. 1. An overview of genetic differences between *TPeL* V3603-13 and *TPeC* Cuniculi A genomes.** Insertions and deletions are shown above and below a schematic representation of the *TPeL* isolate V3603-13 chromosome, respectively. Green areas represent sequentially diverse regions and the length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



**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.*, 2008, and by Šmajs *et al.*, 2011, differ, likely as a result of repeat reshuffling during individual laboratory handling of the Cuniculi A strain.



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



## Conclusion

We have determined the first complete genome sequence of *Treponema paraluisleporidarum* ecovar Lepus (*TPeL* V3603-13) and also provided two additional draft genome sequences of two additional *TPeL* strains. The agent of hare syphilis has been found similar but, yet, distinct from the rabbit syphilis treponeme and the major proteome differences comprised *tpr* and outer membrane proteins and/or antigens. The phylogeny revealed that the *TPeC* Cuniculi A genome appears to be evolutionary modern compared to *TPeL*. Based on previous estimations of *TPA* and *TPE* mutation rates, both *TPeL* and *TPeC* appear to be separated by 0.5–3 kiloyears of lagomorph treponeme evolution.



Acknowledgement: The project National Institute of Virology and Bacteriology (Programme EXCELES, ID Project No. LK22NP05103) -Funded by the European Union - Next Generation EU.