

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 paraluisuniculi* strains

Darina Čejková,^{1,2†} Marie Zobaníková,^{1†} Petra Pospíšilová,¹ Michal Strouhal,^{1,2} Lenka Mikalová,¹ George M. Weinstock² and David Šmajš¹

Correspondence

David Šmajš
dsmajs@med.muni.cz

¹Department of Biology, Faculty of Medicine, Masaryk University, Kamenice 5, Building A6, 62500 Brno, Czech Republic

²The Genome Institute, Washington University in St Louis, 4444 Forest Park Avenue, St Louis, MO 63108, USA

This study examined the sequences of the two rRNA (*rrn*) operons of pathogenic non-cultivable treponemes, comprising 11 strains of *T. pallidum* ssp. *pallidum* (TPA), five strains of *T. pallidum* ssp. *pertenue* (TPE), two strains of *T. pallidum* ssp. *endemicum* (TEN), a simian Fribourg-Blanc strain and a rabbit *T. paraluisuniculi* (TPc) strain. PCR was used to determine the type of 16S–23S ribosomal intergenic spacers in the *rrn* operons from 30 clinical samples belonging to five different genotypes. When compared with the TPA strains, TPc Cuniculi A strain had a 17 bp deletion, and the TPE, TEN and Fribourg-Blanc isolates had a deletion of 33 bp. Other than these deletions, only 17 heterogeneous sites were found within the entire region (excluding the 16S–23S intergenic spacer region encoding tRNA-Ile or tRNA-Ala). The pattern of nucleotide changes in the *rrn* operons corresponded to the classification of treponemal strains, whilst two different *rrn* spacer patterns (Ile/Ala and Ala/Ile) appeared to be distributed randomly across species/subspecies classification, time and geographical source of the treponemal strains. It is suggested that the random distribution of tRNA genes is caused by reciprocal translocation between repetitive sequences mediated by a *recBCD*-like system.

Received 2 October 2012
Accepted 14 October 2012

INTRODUCTION

rRNA genes are co-localized in rRNA (*rrn*) operons. The typical bacterial *rrn* operon consists of 16S–23S–5S rRNA genes. In addition, *rrn* operons may contain tRNA genes and regulatory regions. The *rrn* operons are highly transcribed in bacteria (Condon *et al.*, 1992), especially during the exponential phase of growth and in fast-growing bacteria. It is generally believed that bacteria with

a short generation time have multiple *rrn* operons in the genome. Multiple copies of 16S and 23S rRNA genes in an organism are almost identical (Pei *et al.*, 2009, 2010), suggesting homogenization of rRNA genes through homologous recombination (Liao, 2000). The 16S and 23S rRNA genes are widely used in bacterial phylogenetic studies, but the 5S rRNA genes are too short to be useful for this purpose.

In addition to the rRNA genes, the *rrn* operons contain intergenic spacer regions (ISRs). The ISRs are not involved in ribosomal function, so they are not under functional constraints, resulting in higher ISR microheterogeneity among bacterial species and strains (de Vries *et al.*, 2006; Gürtler, 1999). The 16S–23S ISRs vary in length, tRNA composition and intragenomic nucleotide diversity (Stewart & Cavanaugh, 2007), and have been used for bacterial identification, molecular typing (Indra *et al.*, 2010; Sadeghifard *et al.*, 2006) and evolutionary studies (Antón *et al.*, 1998).

†These authors contributed equally to this work.

Abbreviations: CDC, Centers for Disease Control and Prevention; ISR, intergenic spacer region; SNP, single-nucleotide polymorphism; TEN, *T. pallidum* ssp. *endemicum*; TPA, *T. pallidum* ssp. *pallidum*; TPc, *T. paraluisuniculi*; TPE, *T. pallidum* ssp. *pertenue*; WGA, whole-genome amplification; XL-PCR, extra-large PCR.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this study are JX120527–JX120565.

Three supplementary tables are available with the online version of this paper.

In this study, we used the variation present in the *rrn* operons to assess evolutionary relationships among several pathogenic non-cultivable treponemes. The organisms studied comprised *Treponema pallidum* and *Treponema paraluis-cuniculi* species and an unclassified simian isolate (Fribourg-Blanc). The species of *T. pallidum* comprised *T. pallidum* ssp. *pallidum* (TPA), *T. pallidum* ssp. *pertenue* (TPE) and *T. pallidum* ssp. *endemicum* (TEN), the aetiological agents of syphilis, yaws and endemic syphilis, respectively. *T. paraluis-cuniculi* (TPc), the aetiological agent of rabbit syphilis, and the simian Fribourg-Blanc isolate are closely related to the *T. pallidum* spp. (Šmajš *et al.*, 2011a).

Two *rrn* operons have been observed in pathogenic treponemes (Fukunaga *et al.*, 1992) composed of 16S–23S–5S rRNA genes. The 16S–23S ISRs of the TPA Nichols strain (Fraser *et al.*, 1998) contain tRNA-Ile (tRNA-Ile-1; TP_t12) and tRNA-Ala (tRNA-Ala-3; TP_t15) genes within the *rrn1* and *rrn2* operons, respectively. The same spacer pattern (Ile/Ala) has been observed in other complete treponemal genomes (Giacani *et al.*, 2010; Matějková *et al.*, 2008; Šmajš *et al.*, 2011b). In contrast, the TPE CDC-2 and TPE Gauthier strain genomes (Čejková *et al.*, 2012) show an Ala/Ile spacer pattern, where the TP_t12 and TP_t15 orthologues are located within the *rrn2* and *rrn1* operons, respectively.

Stamm *et al.* (2002) used the sequences of 16S–23S ISRs for molecular typing of dermatitis-associated treponemes in cattle. These treponemes are divided into three phylotypes, which cluster within the group of human saprophytic treponemes (*Treponema denticola*, *Treponema phagedenis* and *Treponema vincentii*). Centurion-Lara *et al.* (1996) examined the TPA Nichols and TPE Gauthier strains and found no difference in the 16S–23S ISRs.

Closely related spirochaetes in the genus *Borrelia* contain two distinct *rrn* operon patterns. Whereas Lyme disease agent (*Borrelia burgdorferi sensu lato*) harbours a unique operon composed of 16S–23S–5S–23S–5S rRNA genes, agents of relapsing fever carry an operon consisting of 16S–23S–5S rRNA genes (Fraser *et al.*, 1997; Schwartz *et al.*, 1992). Two typing systems have been developed using the 16S–23S ISR, which includes both the tRNA-Ala and tRNA-Ile genes (Bunikis *et al.*, 2004; Liveris *et al.*, 1996). The typing systems have been applied to differentiate species within *B. burgdorferi sensu lato* in North America (Bunikis *et al.*, 2004), to study populations of tick- and bird-borne *Borrelia garinii* in Eurasia (Comstedt *et al.*, 2009) and to study the association between the *B. burgdorferi sensu stricto* genotype and dissemination of infection (Hanincová *et al.*, 2008; Wormser *et al.*, 2008).

In this study, we compared the sequences of both *rrn* operons among pathogenic treponemes, comprising 11 strains of TPA, five strains of TPE, two strains of TEN, a simian Fribourg-Blanc isolate and a rabbit TPc strain. We also studied 16S–23S ISRs in 30 clinical samples positive for *T. pallidum* DNA.

METHODS

Strains used in this study. The *rrn* operon sequences were examined in 20 strains of the genus *Treponema* (Table 1), comprising a baboon isolate (unclassified *T. pallidum* strain Fribourg-Blanc), a rabbit syphilis strain (TPc) and 18 human strains (TPA, TPE and TEN). Thirty clinical samples (named 2K, 4K, 6K, 15K, 24K, 27K, 34K, 40K, 44K, 47K, 49K, 51K, 52K, 53K, 63K, 73K, 91K, 6000, 9888, 14048, 14207, 16142, RL86Z, RL89BZ, RL95B, RL102B, RL104B, RL110B, RL111B and RL116A) were tested for the presence of 16S–23S ISR sequences encoding either tRNA-Ile or tRNA-Ala, with positive detection of treponemal DNA in all samples. More detailed data on these samples were published recently (Flasarová *et al.*, 2012).

Isolation of treponemal DNA. TPA Nichols and SS14, TPE Samoa D and CDC-2, and TPc Cuniculi A chromosomal DNA was prepared as described previously by Fraser *et al.* (1998) by extracting DNA from experimentally infected rabbits. Treponemes were purified by Hypaque gradient centrifugation (Baseman *et al.*, 1974). Because a high input of DNA was required for the sequencing approach, whole-genome amplification (WGA) (REPLI-g Midi kit; Qiagen) was performed for TPA Nichols DNA according to the manufacturer's instructions. In addition, non-WGA DNAs from TPA Nichols and SS14, TPE Samoa D and CDC-2, and TPc Cuniculi A were used. The Philadelphia 1, Philadelphia 2, DAL-1, Mexico A, Bal 73-1, Grady, MN-3, Madras and Haiti B (TPA), CDC-1, CDC-2, Gauthier and Samoa F (TPE), Bosnia A and Iraq B (TEN), and Fribourg-Blanc (a simian *T. pallidum*) strains were obtained as rabbit testicular tissues containing treponemal cells. After brief centrifugation of the samples at 100 g for 5 min, the DNA enriched for bacterial cells was amplified using the REPLI-g Midi kit.

PCR amplification. The primer pairs RNA1F (5'-GTGTGTG-AGTCTGGCAGGAA-3') and RNA1R (5'-TTATTGCTGTGCGC-ATCTTC-3'), and RNA2F (5'-ACAAGTGAGCGAAGCGTTTT-3') and RNA2R (5'-CCAAGAGACTACCCGTCTG-3'), were used for amplification of the *rrn* operons from treponemal strains. These primer pairs produced extra-large PCR (XL-PCR) products of 5.85 and 5.92 kb, respectively. To obtain these XL-PCR amplicons, a GeneAmp XL PCR kit (Roche Molecular Systems) was used as described by Strouhal *et al.* (2007). XL-PCR products were purified using a QIAquick PCR Purification kit (Qiagen) or ExoSAP-IT kit (GE Healthcare) according to the manufacturer's instructions.

DNA sequencing. DNA sequencing of the XL-PCR products was carried out with a BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems) using a primer-walking approach. Additional internal oligonucleotide sequencing primers (see Table S1, available in JMM Online) were designed using Primer3 software (Rozen & Skaletsky, 2000). The LASERGENE program package (DNASTAR) was used to assemble the consensus sequences.

Phylogenetic analyses. In addition to the *rrn* operons investigated in the 20 strains (Table 1), the *rrn* operons of TPA Chicago (GenBank accession no. CP001752; Giacani *et al.*, 2010) was included in the evolutionary analysis. Concatenated sequences of *rrn1* and *rrn2* operons (Table S2) were used for the construction of evolutionary trees using the neighbour-joining method (Saitou & Nei, 1987) in MEGA4 software (Tamura *et al.*, 2007). The bootstrap consensus trees were determined from 1000 bootstrap resamplings. Branches with <50% bootstrap support were collapsed.

Detection of recombination. To identify genomic rearrangements, *rrn* operons were analysed using the Recombination Detection Program package (version RDP3; Martin *et al.*, 2010). Four methods, including RDP, GENECONV (Sawyer, 1989), MaxChi (Smith, 1992) and

Table 1. *Treponema* strains used in this study

–, Not known.

Strain	<i>Treponema</i> (sub)species	Place of isolation	Date of isolation	Reference	Source of material*
Bal 73-1	TPA	Baltimore, USA	1968	Hardy <i>et al.</i> (1970)	David L. Cox, CDC, Atlanta, GA, USA
Bosnia A	TEN	Bosnia	1950	Turner & Hollander (1957)	Sylvia M. Bruisten, PHL, Amsterdam, NL
CDC-1	TPE	Dersuso, Ghana	1980	Liska <i>et al.</i> (1982)	David L. Cox, CDC, Atlanta, GA, USA
CDC-2	TPE	Akorabo, Ghana	1980	Liska <i>et al.</i> (1982)	Steven J. Norris, UT, Houston, TX, USA
Cuniculi A	<i>paraluiscuniculi</i>	–	pre-1957	Turner & Hollander (1957)	Steven J. Norris, UT, Houston, TX, USA
DAL-1	TPA	Dallas, USA	1991	Wendel <i>et al.</i> (1991)	David L. Cox, CDC, Atlanta, GA, USA
Fribourg-Blanc	Simian isolate	Guinea	1966	Fribourg-Blanc & Mollaret (1969)	David L. Cox, CDC, Atlanta, GA, USA
Gauthier	TPE	Brazzaville, Congo	1960	Gastinel <i>et al.</i> (1963)	Steven J. Norris, UT, Houston, TX, USA
Grady	TPA	Atlanta, USA	1980s	–	David L. Cox, CDC, Atlanta, GA, USA
Haiti B	TPA	Haiti	1951	Turner & Hollander (1957)	David L. Cox, CDC, Atlanta, GA, USA
Iraq B	TEN	Iraq	1951	Turner & Hollander (1957)	Kristin N. Harper, Emory University, Atlanta, GA, USA
Madras	TPA	Madras, India	1954	Laboratory notebook of Rob George CDC	David L. Cox, CDC, Atlanta, GA, USA
Mexico A	TPA	Mexico City, Mexico	1953	Turner & Hollander (1957)	David L. Cox, CDC, Atlanta, GA, USA
MN-3	TPA	Minnesota, USA	–	–	David L. Cox, CDC, Atlanta, GA, USA
Nichols	TPA	Washington, DC, USA	1912	Nichols & Hough (1913)	Steven J. Norris, UT, Houston, TX, USA
Philadelphia 1	TPA	Philadelphia, USA	1988	–	David L. Cox, CDC, Atlanta, GA, USA
Philadelphia 2	TPA	Philadelphia, USA	–	–	David L. Cox, CDC, Atlanta, GA, USA
Samoa D	TPE	Samoa	1953	Turner & Hollander (1957)	Steven J. Norris, UT, Houston, TX, USA
Samoa F	TPE	Samoa	1953	Turner & Hollander (1957)	Steven J. Norris, UT, Houston, TX, USA
SS14	TPA	Atlanta, USA	1977	Stamm <i>et al.</i> (1983)	Steven J. Norris, UT, Houston, TX, USA

*CDC, Centers for Disease Control and Prevention; PHL, Public Health Laboratory; UT, University of Texas.

CHIMAERA (Posada & Crandall, 2001), implemented in the RDP3 package, were applied using default settings.

Analysis of clinical specimens. Skin and mucosal swabs were placed in a tube containing 1.5 ml sterile water and agitated for 5 min at room temperature. The swab was withdrawn and the supernatant was used for DNA isolation. Swab supernatant (0.2–0.4 ml) and whole blood (0.2–0.8 ml) were used for DNA isolation using a QIAamp DNA Mini kit (Qiagen) according to the manufacturer's Blood and Body Fluid Spin Protocol. To detect the presence of treponemal DNA in swab and whole-blood samples, a diagnostic PCR

assay amplifying five different *Treponema*-specific genes including *po1A* (TP0105 locus), *tmpC* (TP0319), TP0136, TP0548 and the 23S rRNA gene was performed. Amplification and subsequent sequencing of TP0136, TP0548 and the 23S rRNA gene have been used, although not for diagnostic purposes, for molecular typing of treponemal strains (Flasarová *et al.*, 2006, 2012; Liu *et al.*, 2001; Matějčková *et al.*, 2009; Woznicová *et al.*, 2007).

The composition of 16S–23S ISR sequences in the *rrn1* and *rrn2* operons, encoding either tRNA-Ile or tRNA-Ala, was determined by another nested PCR. In the first step, each clinical isolate was tested in four parallel reactions with the following primer pairs

(Fig. 1 and Table S3): RNA1Fb and RNA1-tRNA-Ile (first reaction), RNA1Fb and RNA2-tRNA-Ala (second reaction), RNA2Fc and RNA1-tRNA-Ile (third reaction) and RNA2Fc and RNA2-tRNA-Ala (fourth reaction). Using these primer sets, the PCR products revealed the position (*rrn1* or *rrn2*) and composition (tRNA-Ile or tRNA-Ala) of the amplified *rrn* operon. In the second step of the nested PCR, the PCR product of the *rrn1* (from the first and second reactions) region was amplified using TP0225-6aF and TP0225-6bR primers, whilst the PCR product of the *rrn2* (from the third and fourth reactions) region was amplified with RNA2Fa and TP0225-6bR. The second step was not specific for the Ile/Ala or Ala/Ile *rrn* spacer pattern but improved the sensitivity of detection of the PCR product from the first step. Each PCR contained 0.4 μ l 10 mM dNTP mix, 2 μ l 10 \times ThermoPol Reaction buffer (New England BioLabs), 0.1 μ l each primer (100 pmol μ l⁻¹), 0.1 μ l *Taq* DNA polymerase (5000 U ml⁻¹; New England BioLabs), 1 μ l test sample and 16.3 μ l PCR-grade water, giving 20 μ l in total. PCR amplification was performed using a GeneAmp 9800 thermocycler (Applied Biosystems) with the following cycling conditions: 94 °C for 5 min; 40 cycles of 94 °C for 60 s, 72 °C for 20 s and 72 °C for 150 s; and a final extension at 72 °C for 10 min. The second step of the nested PCR

used the same conditions but a lower annealing temperature of 67 °C.

RESULTS

Amplification and sequencing of the *rrn* operons

Two *rrn* operons (16S–23S–5S) have been described in pathogenic *Treponema* genomes with the 16S–23S ISR comprising genes encoding tRNA-Ala or tRNA-Ile (Fraser *et al.*, 1998; Fukunaga *et al.*, 1992; Giacani *et al.*, 2010; Šmajš *et al.*, 2011b). Using XL-PCR, we amplified the *rrn* operons in 20 treponemal strains (Tables 1 and S2) comprising 11 strains of TPA, five strains of TPE, an unclassified simian isolate, two strains of TEN and a rabbit TPc isolate. XL-PCR products were obtained for all 40 investigated regions. However, the assembled sequence of the *rrn2* operon of Iraq B (TEN) was repeatedly ambiguous at several positions, probably due to low DNA quality, so

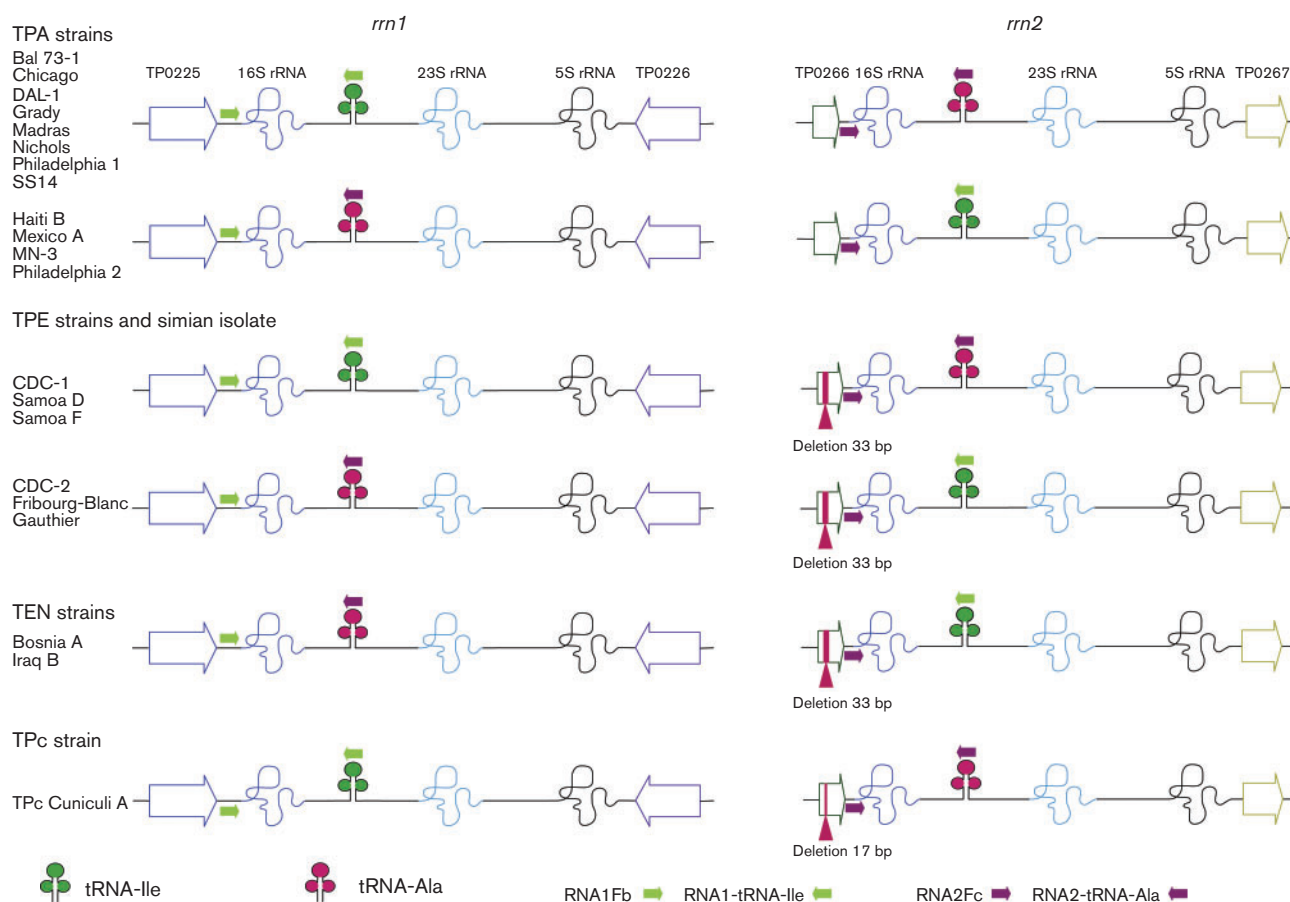


Fig. 1. A schematic representation of the treponemal *rrn* operons consisting of 16S–23S–5S rRNA genes, intergenic regions and a 212 bp DNA sequence upstream of 16S rRNA gene. The positions of the 33 and 17 bp deletions in the non-TPA strains are shown. Please note that both spacer patterns of the 16S–23S ISR encoding either tRNA-Ile or tRNA-Ala were present among TPA and TPE strains in the *rrn1* and *rrn2* regions. Green symbols represent tRNA-Ile, whilst red symbols represent tRNA-Ala. Small coloured arrows in green and purple represent primers used in the clinical samples examined.

Table 2. DNA sequence polymorphisms found among 20 pathogenic *Treponema* strains at the *rrn* operons

SNPs are indicated by underlining, whereas translocation of tRNA is shown in bold. IGR, Intergenic region.

Species	Strain (operon)	Treponemal homologous sequences of rRNA operons, and position downstream (D), upstream (U) or within the RNA gene																		
		IGR (212 bp)			16S rRNA (1537 bp)				IGR (117 or 116 bp)*	tRNA (74 bp)*	IGR (111 or 122 bp)*	23S rRNA (2951 bp)							IGR (50 bp)	5S rRNA (110 bp)
		171–167 U	96 U	93 U	647	1134	1375	1441	71 D	21 U	458	763	766	1092	1359	1546	2104	47 D	81	
TPA	Bal 73-1 (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	Bal 73-1 (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	Chicago (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	Chicago (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	DAL-1 (<i>rrn1</i>)	<u>GGGG</u>	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	DAL-1 (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	Grady (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	Grady (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	Haiti B (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	Haiti B (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	Madras (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	Madras (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	Mexico A (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	Mexico A (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	MN-3 (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	MN-3 (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	Nichols (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	Nichols (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
	Philadelphia 1 (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C
	Philadelphia 1 (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C
Philadelphia 2 (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	A	C	C	
Philadelphia 2 (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	A	C	C	
SS14 (<i>rrn1</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ile	G	G	G	<u>G</u>	G	A	A	<u>G</u>	C	C	
SS14 (<i>rrn2</i>)	GGGGG	A	A	G	G	G	C	G	tRNA-Ala	G	G	G	<u>G</u>	G	A	A	<u>G</u>	C	C	
TPE	CDC-1 (<i>rrn1</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ile	G	G	G	A	G	A	A	A	C	C
	CDC-1 (<i>rrn2</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ala	G	G	G	A	G	A	A	A	C	C
	CDC-2 (<i>rrn1</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ala	G	G	G	A	G	A	A	A	C	C
	CDC-2 (<i>rrn2</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ile	G	G	G	A	G	A	A	A	C	C
	Gauthier (<i>rrn1</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ala	G	G	G	A	G	A	A	A	C	C

Table 2. cont.

Species	Strain (operon)	Treponemal homologous sequences of rRNA operons, and position downstream (D), upstream (U) or within the RNA gene																			
		IGR (212 bp)			16S rRNA (1537 bp)				IGR (117 or 116 bp)*	tRNA (74 bp)*	IGR (111 or 122 bp)*	23S rRNA (2951 bp)							IGR (50 bp)	5S rRNA (110 bp)	
		171–167 U	96 U	93 U	647	1134	1375	1441				71 D	21 U	458	763	766	1092	1359			1546
	Gauthier (<i>rrn2</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ile	G	G	G	A	G	A	A	A	C	C	
	Samoa D (<i>rrn1</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ile	G	G	G	A	G	A	A	A	C	C	
	Samoa D (<i>rrn2</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ala	G	G	G	A	G	A	A	A	C	C	
	Samoa F (<i>rrn1</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ile	G	G	G	A	G	A	A	A	C	C	
	Samoa F (<i>rrn2</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ala	G	G	G	A	G	A	A	A	C	C	
Simian isolate	Fribourg-Blanc (<i>rrn1</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ala	G	<u>A</u>	G	A	G	A	A	A	C	C	
	Fribourg-Blanc (<i>rrn2</i>)	GGGGG	A	<u>G</u>	G	G	G	C	G	tRNA-Ile	G	<u>A</u>	G	A	G	A	A	A	C	C	
TEN	Bosnia A (<i>rrn1</i>)	GGGGG	A	A	G	G	<u>A</u>	C	G	tRNA-Ala	G	G	G	A	G	A	A	A	C	C	
	Bosnia A (<i>rrn2</i>)	GGGGG	A	A	G	G	<u>A</u>	C	G	tRNA-Ile	G	G	G	A	G	A	A	A	C	C	
	Iraq B (<i>rrn1</i>)	GGGGG	A	A	G	G	<u>A</u>	C	G	tRNA-Ala	G	G	G	A	G	A	A	A	C	C	
	Iraq B (<i>rrn2</i>)	GGGGG	A	A	G	G	<u>A</u>	C	G	tRNA-Ile	G	G	G	A	G	A	A	A	C	C	
TPc	Cuniculi A (<i>rrn1</i>)	GGGGG	<u>G</u>	A	<u>A</u>	<u>A</u>	G	<u>T</u>	<u>A</u>	tRNA-Ile	<u>A</u>	G	<u>A</u>	A	<u>A</u>	<u>G</u>	<u>G</u>	<u>G</u>	A	<u>T</u>	<u>A</u>
	Cuniculi A (<i>rrn2</i>)	GGGGG	<u>G</u>	A	<u>A</u>	<u>A</u>	G	<u>T</u>	<u>A</u>	tRNA-Ala	<u>A</u>	G	<u>A</u>	A	<u>A</u>	<u>G</u>	<u>G</u>	<u>G</u>	A	<u>T</u>	<u>A</u>

*The size of sequence between the 16S and 23S rRNA genes (both excluded) varied based on the presence of the tRNA-Ile (117+74+111, in total 302 bp) or tRNA-Ala (116+74+122, in total 312 bp) gene.

the Iraq B sequences were excluded from the construction of phylogenetic trees.

Sequence analysis of *rrn* operons

In the individual TPA genomes, the amplified *rrn1* and *rrn2* regions were identical for 5141 bp (Tables 2 and S2, Fig. 1) including the DNA regions 212 bp upstream of the 16S rRNA, the 16S rRNA (1537 bp), 23S rRNA (2951 bp), 5S rRNA (110 bp) and 23S–5S ISR (50 bp), and a region of 54 bp downstream of the 5S rRNA. Additional identical sequences were located within the 16S–23S ISR downstream of the 16S rRNA (120 bp) and upstream of the 23S rRNA (118 bp) genes (Fig. 2, Table 2). Alternative sequences within the 16S–23S ISR, encoding tRNA-Ile or tRNA-Ala, comprised an additional 64 or 74 bp, respectively (Fig. 2). To extend the comparative analysis over all available data, the TPA Chicago sequences of the *rrn* operons (GenBank accession no. CP001752; Giacani *et al.*, 2010) were added to the sequences of the 20 strains used in this study.

When compared with the TPA strains, a deletion of 33 bp was found in homologous regions of the *rrn2* region in the TPE, TEN and simian strains (Fig. 1), whilst the TPc strain contained a 17 bp deletion at the same position (Fig. 1). These deletions resulted in shortening (33 bp deletion) or truncation (17 bp deletion) of TP0266 orthologues. Among all investigated strains, in addition to the observed deletions, we found only 17 heterogeneous sites within the entire region, excluding the 16S–23S ISR encoding tRNA-Ile or tRNA-Ala. Sixteen sites were single nucleotide changes and one was a single base-pair deletion (Table 2). The *rrn1* operon of the reference TPA Nichols genome (GenBank accession no. AE000520.1; Fraser *et al.*, 1998) showed a deletion within the 16S rRNA gene (data not

shown), whereas all other strains, including the Nichols strain examined in our study, did not. This deletion may represent a sequencing error present in the reference Nichols genome, as dozens of such sequencing errors have already been confirmed (Giacani *et al.*, 2012; Matějková *et al.*, 2008). In contrast, a 1 bp deletion in the TPA DAL-1 genome, upstream of the 16S rRNA gene in the *rrn1* operon, was repeatedly confirmed by Sanger sequencing. The identified nucleotide change at position 2104 of the 23S rRNA gene (differentiating the SS14 strains from other investigated strains) corresponded to the mutation causing macrolide resistance in treponemal strains (Stamm & Bergen, 2000).

All TPA strains differed from the other pathogenic treponemes by a nucleotide change at position 766 of the 23S rRNA gene. The TPE strains and the simian isolate Fribourg-Blanc could be distinguished from the other pathogenic treponemes by a single-nucleotide polymorphism (SNP) localized 93 bp upstream of the 16S rRNA genes. The TPE strains could be differentiated from the simian isolate by a nucleotide sequence change in the 23S rRNA gene (nt 458). The TEN showed a nucleotide change in the 16S rRNA gene, and TPc showed 12 nt changes in the investigated *rrn* sequences (Table 2).

Reciprocal translocation of tRNA genes

In contrast to the phylogenetically conserved SNP distribution in the repetitive sequences of the *rrn* operons, the genes coding for tRNA did not show the same evolutionary pattern (Table 2, Fig. 1). In this study, we observed two 16S–23S ribosomal ISR patterns. The spacer pattern Ile/Ala included the tRNA-Ile gene within the *rrn1* region and the tRNA-Ala gene within the *rrn2* region. The Ile/Ala pattern was observed in the following strains: TPA Nichols, Bal

```

ISR 16S-23S-rrn1      CCTTTCTAAGAGAAAGGGTATGGGCATGGCATGGTGCCGTGTTCCGCCGTGGCCGGAAGC 60
ISR 16S-23S-rrn2      CCTTTCTAAGAGAAAGGGTATGGGCATGGCATGGTGCCGTGTTCCGCCGTGGCCGGAAGC 60
*****

ISR 16S-23S-rrn1      CACACGGTAGGTTTTCTGCTCCTGCACGGCAGTCTCTCCCTTCCCTTTGAAAAGGGG 120
ISR 16S-23S-rrn2      CACACGGTAGGTTTTCTGCTCCTGCACGGCAGTCTCTCCCTTCCCTTTGAAAAGGGG 120
*****

ISR 16S-23S-rrn1      CTTGTAGCTCAGTTGGTTAGAGC-ACTTCTCTGATAAGGAAGGGGTCATTAGTTCAA --- 176
ISR 16S-23S-rrn2      GT-ATAGCTCAACTGGCTAGAGCGACGGCTTTG-CAGGTCGTAGGTCAGGGGTTCGAGTC 178
* ***** ** * * * * * * * * * * * * * * * * * * * * * * * * * * * *

ISR 16S-23S-rrn1      CTCTAATC-----AAGCCCACTATTATTCTTTATGTCCCTTTGTTTTGTTTATGGGG 228
ISR 16S-23S-rrn2      CCCTTATCTCCAGGGAAGCCCACTATTATTCTTTATGTCCCTTTGTTTTGTTTATGGGG 238
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

ISR 16S-23S-rrn1      TAAGGAGTAGGTGGTAGGTGATTTTTGAGAGTATTAGGGTGGGGTGTGAAGTTGAGAAGG 288
ISR 16S-23S-rrn2      TAAGGAGTAGGTGGTAGGTGATTTTTGAGAGTATTAGGGTGGGGTGTGAAGTTGAGAAGG 298
*****

ISR 16S-23S-rrn1      GATGGATAATATGG 302
ISR 16S-23S-rrn2      GATGGATAATATGG 312

```

Fig. 2. Alignment of 16S–23S ISRs in TPA Nichols *rrn* operons. The gene encoding tRNA-Ile (TP_t12) is shown in red, whilst the gene encoding tRNA-Ala-3 (TP_t15) is in blue.

73-1, Grady, SS14, Chicago, DAL-1, Philadelphia 1 and Madras; TPE Samoa D, CDC-1 and Samoa F; and TPc Cuniculi A. The reverse ISR pattern Ala/Ile consisted of the tRNA-Ala gene within the *rrn1* region and the tRNA-Ile gene within the *rrn2* region. The Ala/Ile pattern was found in TPA Mexico A, MN-3, Philadelphia 2 and Haiti B strains; in TPE Gauthier and CDC-2; in the unclassified treponeme Fribourg-Blanc; and in TEN Iraq B and Bosnia A genomes.

The concatenated *rrn* operons, excluding the tRNA genes and their vicinity, clustered according to the species/subspecies classification (Fig. 3a). The TEN Iraq B strain was omitted from the analysis because we were unable to obtain an unambiguous *rrn2* operon sequence. Nevertheless, the *rrn1* operon was identical to another TEN strain, Bosnia A. In contrast, the trees showing concatenated *rrn* operons including tRNA genes (Fig. 3b) branched according to the composition of tRNA in the individual *rrn* operons, and then according to the species/subspecies classification. This phenomenon can be explained by recombination events that have occurred between *rrn* operons.

To predict recombination hot-spot sites within the *rrn* operons, four methods from the RDP3 program were applied. All four methods predicted four recombination sites (Table 3), two sites in each *rrn* operon. The predicted sites corresponded to the same positions within the 16S (nt 783) and 23S (nt 324) rRNA genes in both *rrn* operons.

Structure of the *rrn* operons in clinical isolates containing *T. pallidum* DNA

The composition of the 16S–23S ribosomal ISR (Ile/Ala or Ala/Ile spacer pattern in the *rrn* operons) was tested in 30 recently isolated clinical samples (Flasarová *et al.*, 2012). The results are summarized in Table 4. Only the Ile/Ala pattern was identified in all clinical samples tested, despite the fact that the clinical samples belonged to five different genotypes (Table 4), as revealed by CDC and sequencing-based typing (Flasarová *et al.* 2012; Pillay *et al.* 1998). Nevertheless, all clinical strain genotypes were similar to the SS14 strain genotype.

DISCUSSION

In this study, we examined the *rrn* operons in 20 pathogenic treponemal strains and 30 clinical isolates. All investigated strains contained two copies of the *rrn* operons. Two *rrn* operons with the same composition have also been described in other human and animal treponemes except for *T. Vincentii* containing only one *rrn* operon (Fraser *et al.*, 1998; Matějková *et al.*, 2008; Seshadri *et al.*, 2004; Stamm *et al.*, 2002).

Our results confirmed that there is little diversity within rRNA genes and ISRs. However, our data showed that the *rrn* operon structure displayed blocks of conserved and

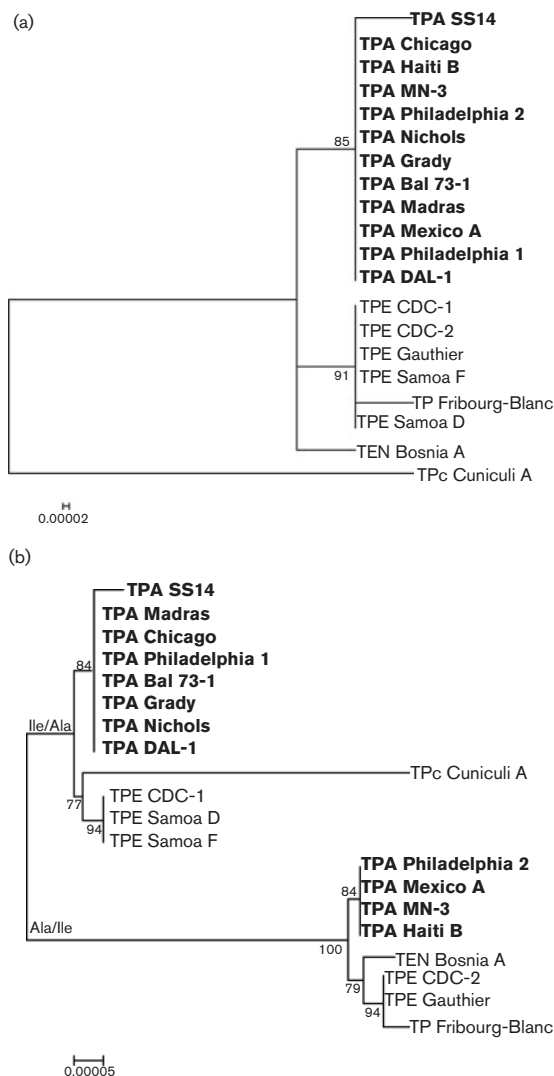


Fig. 3. (a) An unrooted tree constructed from the concatenated sequences of the *rrn* operons excluding the heterologous tRNA genes. The *rrn* operons clustered according to the species/subspecies classification of treponemes. Bar, 0.00002 nt substitutions per site. (b) An unrooted tree constructed from the sequences of the *rrn* operons including heterologous tRNA genes. The *rrn* operons clustered according to the ISR pattern. Bar, 0.00005 nt substitutions per site. Bootstrap values based on 1000 replications are shown next to the branches. The TPA strains causing syphilis are shown in bold.

polymorphic sites. The TPA DAL-1 strain showed a 1 bp deletion upstream of the 16S rRNA gene in the *rrn1* operon. It is known that TPA DAL-1 grows more rapidly in rabbits than other pathogenic strains (Wendel *et al.*, 1991), and it is possible that the different promoter DNA conformation may affect expression of the *rrn1* operon.

Gürtler & Stanisich (1996) used the 16S–23S ribosomal ISR for classification of bacteria. 16S–23S ISRs have been used in several studies (de Vries *et al.*, 2006; Lan & Reeves, 1998;

Table 3. Predicted recombination hot-spot sites using the RDP3 program

Predicted recombination hot-spot site*		Prediction algorithm used in the RDP3 program (P value)			
Start	End	RDP	GENECONV	MaxChi	CHIMAERA
231656	233036	1.33E-58	2.20E-55	8.06E-13	7.80E-13
280058	281448	7.22E-20	1.12E-20	5.91E-04	1.33E-03

*Whole-genome TPE Samoa D coordinates are shown (Čejková *et al.*, 2012; GenBank accession no. CP002374).

Lebuhn *et al.*, 2006), including for treponemal (Centurion-Lara *et al.*, 1996; Stamm *et al.*, 2002) and borrelian samples (Bunikis *et al.*, 2004; Comstedt *et al.*, 2009). Centurion-Lara *et al.* (1996) examined the TPA Nichols and TPE Gauthier strains, and no difference was found. However, they did not examine the genomic positions of individual 16S–23S ISRs. Interestingly, the 16S–23S ISR typing of

Borrelia burgdorferi sensu stricto is in accordance with *ospC* gene typing (Hanincová *et al.*, 2008; Wormser *et al.*, 2008). The *ospC* gene, encoding a protein involved in the initiation of infection in warm-blooded animals, is located on plasmid DNA, whilst the *rrn* operon is on chromosomal DNA. Moreover, different 16S–23S ISR genotypes are associated with different degrees of invasivity (Wormser *et al.*, 2008).

Table 4. Composition of the 16S–23S ribosomal ISR in clinical samples containing TPA DNA

Clinical sample	16S–23S ISR (<i>rrn1/rrn2</i>)	Genotype*	Subtype†
2K	Ile/Ala	SU2R8	14d
4K	Ile/Ala	SSS	14d
6K	Ile/Ala	SSS	14d
15K	Ile/Ala	SSS	14d
24K	Ile/Ala	SSS	14d
27K	Ile/Ala	SSS	14d
34K	Ile/Ala	SSS	14d
40K	Ile/Ala	SSR8	14d
44K	Ile/Ala	SSS	14d
47K	Ile/Ala	SSS	12d
49K	Ile/Ala	SSR8	14d
51K	Ile/Ala	SSS	14d
52K	Ile/Ala	SSS	14d
53K	Ile/Ala	SSS	14d
63K	Ile/Ala	SSR9	15d
73K	Ile/Ala	SU2R8	14d
91K	Ile/Ala	SSS	14d
6000	Ile/Ala	SU2R8	14d
9888	Ile/Ala	SSS	14d
14048	Ile/Ala	SSS	14d
14207	Ile/Ala	SSS	14d
16142	Ile/Ala	SU2R8	14d
RL86Z	Ile/Ala	SU2R8	14d
RL89BZ	Ile/Ala	SSS	14d
RL95B	Ile/Ala	SSS	14d
RL102B	Ile/Ala	SSS	14d
RL104B	Ile/Ala	SU2R8	14d
RL110B	Ile/Ala	SSS	14d
RL111B	Ile/Ala	SU2R8	14d
RL116A	Ile/Ala	XXR8	14e

*Identified according to the method of Flasarová *et al.* (2012).

†Subtype according to the method of Pillay *et al.* (1998).

Despite the low heterogeneity in the *rrn* operons, two different ISR patterns were observed in the pathogenic treponemal samples. Whereas detection of specific nucleotide changes may be of interest in identification of treponemal diseases, the detection of tRNA genes in the 16S–23S ribosomal ISR appears to be of limited use in typing of clinical samples. All clinical samples showed the Ile/Ala spacer pattern in *rrn* operons, so the tRNA-Ile and tRNA-Ala genes are not useful for molecular typing of clinical strains, at least for treponemes present in the population of the Czech Republic.

Due to the conserved machinery of protein synthesis, rRNA genes are expected to be under strong purifying selection and are exposed to the intragenomic homogenization process via gene conversion (Liao, 2000; Nei & Rooney, 2005). Several studies (Acinas *et al.*, 2004; Pei *et al.*, 2009, 2010) have shown that homogenization of multiple rRNA genes is common among bacteria. In addition, Harvey & Hill (1990) successfully constructed several *Escherichia coli* strains with recombined inverted *rrn* operons; however, the recombinants tended to recover the original configuration. The *rrn* operons of treponemal strains are direct repeats: the tRNA-Ala gene is replaced by tRNA-Ile (and vice versa), and the recombination is a common event with no correlation to the otherwise-determined phylogenetic relationship among tested treponemes. It has been postulated that recombination between direct repeats leads to the duplication or deletion of a repeat (Petes & Hill, 1988; Petit, 2005). Whereas tRNA-Ile (TP_t12) is a unique gene in sequenced treponemal genomes, there are three predicted tRNA-Ala genes (TP_t15, TP_t41 and TP_t45; Fraser *et al.*, 1998). As both tRNA-Ile (TP_t12, GenBank accession no. AE000520.1) and tRNA-Ala (TP_t15, AE000520.1) genes need to be maintained in the genomes of pathogenic treponemes, reciprocal translocation, rather than gene conversion, appears to be the mechanism for the observed *rrn* heterogeneity among tested strains. Such a process would

require double cross-overs in both *rrn* operons, and therefore is much less common than insertion/deletion or gene-conversion events (Harvey & Hill, 1990; Hashimoto *et al.*, 2003). Predicted recombination hot-spot sites were located in the 16S and 23S rRNA genes, genes with two identical copies within every strain examined in our study.

During replication of direct-repeat regions, DNA polymerase might lead to strand slippage, thus collapsing a replication fork formation, and recombination enzymes are involved in the DNA repair machinery (Darling *et al.*, 2008; Santoyo & Romero, 2005). Although only the *recF* recombination pathway was predicted in the TPA Nichols genome (Fraser *et al.*, 1998), the *recF* pathway suggests the gene-conversion mechanism (Kobayashi, 1992; Takahashi *et al.*, 1992). Therefore, the reciprocal recombination in pathogenic treponemes may be accompanied by crossing-over, a repair mechanism implemented by the *recBCD* pathway in *E. coli* (Kobayashi, 1992). Recently, *recBCD* orthologues (*addA* and *addB*) were predicted for several investigated treponemal genomes (Čejková *et al.*, 2012; Giacani *et al.*, 2012; Šmajš *et al.*, 2011b), composed of TP0898 and fused TP0899–TP0900 orthologues. However, it would be extremely difficult to prove experimentally the *recBCD*-mediated crossing-over mechanism in *T. pallidum*.

In summary, two different *rrn* spacer patterns (Ile/Ala and Ala/Ile) seem to be distributed randomly across the time and place of original isolation of treponemal strains (e.g. Philadelphia 1 vs Philadelphia 2, CDC-1 vs CDC-2) and the laboratory that provided the treponemal material (Tables 1 and 2). This random distribution of tRNA genes is probably caused by reciprocal translocation between repetitive sequences mediated by a *recBCD*-like system.

ACKNOWLEDGEMENTS

The authors thank Dr David Cox for providing the Bal 73-1, CDC-1, DAL-1, Fribourg-Blanc, Grady, Haiti B, Madras, Mexico A, MN-3, Philadelphia 1 and Philadelphia 2 strains, Dr Steven Norris for the CDC-2, Gauthier, Nichols, Samoa D, Samoa F and SS14 strains, Dr Kristin Harper for the Iraq B strain, Dr Sylvia Bruisten for the Bosnia A strain, and Dr Elizabeth A. Lobos for help with manuscript preparation. This work was supported by grants from the Grant Agency of the Czech Republic (P302/12/0574) and the Ministry of Health of the Czech Republic (NT11159-5/2010) to D. S.

REFERENCES

- Acinas, S. G., Marcelino, L. A., Klepac-Ceraj, V. & Polz, M. F. (2004). Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J Bacteriol* **186**, 2629–2635.
- Antón, A. I., Martínez-Murcia, A. J. & Rodríguez-Valera, F. (1998). Sequence diversity in the 16S–23S intergenic spacer region (ISR) of the rRNA operons in representatives of the *Escherichia coli* ECOR collection. *J Mol Evol* **47**, 62–72.
- Baseman, J. B., Nichols, J. C., Rump, J. W. & Hayes, N. S. (1974). Purification of *Treponema pallidum* from infected rabbit tissue: resolution into two treponemal populations. *Infect Immun* **10**, 1062–1067.
- Bunikis, J., Garpmo, U., Tsao, J., Berglund, J., Fish, D. & Barbour, A. G. (2004). Sequence typing reveals extensive strain diversity of the Lyme borreliosis agents *Borrelia burgdorferi* in North America and *Borrelia afzelii* in Europe. *Microbiology* **150**, 1741–1755.
- Čejková, D., Zobaniková, M., Chen, L., Pospíšilová, P., Strouhal, M., Qin, X., Mikalová, L., Norris, S. J., Muzny, D. M. & other authors (2012). Whole genome sequences of three *Treponema pallidum* ssp. *pertenue* strains: yaws and syphilis treponemes differ in less than 0.2% of the genome sequence. *PLoS Negl Trop Dis* **6**, e1471.
- Centurion-Lara, A., Castro, C., van Voorhis, W. C. & Lukehart, S. A. (1996). Two 16S–23S ribosomal DNA intergenic regions in different *Treponema pallidum* subspecies contain tRNA genes. *FEMS Microbiol Lett* **143**, 235–240.
- Comstedt, P., Asokliene, L., Eliasson, I., Olsen, B., Wallensten, A., Bunikis, J. & Bergström, S. (2009). Complex population structure of Lyme borreliosis group spirochete *Borrelia garinii* in subarctic Eurasia. *PLoS ONE* **4**, e5841.
- Condon, C., Philips, J., Fu, Z. Y., Squires, C. & Squires, C. L. (1992). Comparison of the expression of the seven ribosomal RNA operons in *Escherichia coli*. *EMBO J* **11**, 4175–4185.
- Darling, A. E., Miklós, I. & Ragan, M. A. (2008). Dynamics of genome rearrangement in bacterial populations. *PLoS Genet* **4**, e1000128.
- de Vries, M. C., Siezen, R. J., Wijman, J. G., Zhao, Y., Kleerebezem, M., de Vos, W. M. & Vaughan, E. E. (2006). Comparative and functional analysis of the rRNA-operons and their tRNA gene complement in different lactic acid bacteria. *Syst Appl Microbiol* **29**, 358–367.
- Fiasarová, M., Šmajš, D., Matějková, P., Woznicová, V., Heroldová-Dvoráková, M. & Votava, M. (2006). [Molecular detection and subtyping of *Treponema pallidum* subsp. *pallidum* in clinical specimens]. *Epidemiol Mikrobiol Imunol* **55**, 105–111 (in Czech).
- Fiasarová, M., Pospíšilová, P., Mikalová, L., Vališová, Z., Dastychová, E., Strnadel, R., Kuklová, I., Woznicová, V., Zákoucká, H. & Šmajš, D. (2012). 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* **92**, 669–674.
- Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G., Clayton, R., Lathigra, R., White, O., Ketchum, K. A., Dodson, R. & other authors (1997). Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* **390**, 580–586.
- Fraser, C. M., Norris, S. J., Weinstock, G. M., White, O., Sutton, G. G., Dodson, R., Gwinn, M., Hickey, E. K., Clayton, R. & other authors (1998). Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* **281**, 375–388.
- Fribourg-Blanc, A. & Mollaret, H. H. (1969). Natural treponematosis of the African primate. *Primates Med* **3**, 113–121.
- Fukunaga, M., Okuzako, N., Mifuchi, I., Arimitsu, Y. & Seki, M. (1992). Organization of the ribosomal RNA genes in *Treponema phagedenis* and *Treponema pallidum*. *Microbiol Immunol* **36**, 161–167.
- Gastinel, P., Vaisman, A., Hamelin, A. & Dunoyer, F. (1963). [Study of a recently isolated strain of *Treponema pertenue*]. *Prophyl Sanit Morale* **35**, 182–188 (in French).
- Giacani, L., Jeffrey, B. M., Molini, B. J., Le, H. T., Lukehart, S. A., Centurion-Lara, A. & Rockey, D. D. (2010). Complete genome sequence and annotation of the *Treponema pallidum* subsp. *pallidum* Chicago strain. *J Bacteriol* **192**, 2645–2646.
- Giacani, L., Chattopadhyay, S., Centurion-Lara, A., Jeffrey, B. M., Le, H. T., Molini, B. J., Lukehart, S. A., Sokurenko, E. V. & Rockey, D. D. (2012). Footprint of positive selection in *Treponema pallidum* subsp. *pallidum* genome sequences suggests adaptive microevolution of the syphilis pathogen. *PLoS Negl Trop Dis* **6**, e1698.

- Gürtler, V. (1999). The role of recombination and mutation in 16S–23S rDNA spacer rearrangements. *Gene* **238**, 241–252.
- Gürtler, V. & Stanisich, V. A. (1996). New approaches to typing and identification of bacteria using the 16S–23S rDNA spacer region. *Microbiology* **142**, 3–16.
- Hanincová, K., Liveris, D., Sandigursky, S., Wormser, G. P. & Schwartz, I. (2008). *Borrelia burgdorferi* sensu stricto is clonal in patients with early Lyme borreliosis. *Appl Environ Microbiol* **74**, 5008–5014.
- Hardy, J. B., Hardy, P. H., Oppenheimer, E. H., Ryan, S. J., Jr & Sheff, R. N. (1970). Failure of penicillin in a newborn with congenital syphilis. *JAMA* **212**, 1345–1349.
- Harvey, S. & Hill, C. W. (1990). Exchange of spacer regions between rRNA operons in *Escherichia coli*. *Genetics* **125**, 683–690.
- Hashimoto, J. G., Stevenson, B. S. & Schmidt, T. M. (2003). Rates and consequences of recombination between rRNA operons. *J Bacteriol* **185**, 966–972.
- Indra, A., Blaschitz, M., Kernbichler, S., Reischl, U., Wewalka, G. & Allerberger, F. (2010). Mechanisms behind variation in the *Clostridium difficile* 16S–23S rRNA intergenic spacer region. *J Med Microbiol* **59**, 1317–1323.
- Kobayashi, I. (1992). Mechanisms for gene conversion and homologous recombination: the double-strand break repair model and the successive half crossing-over model. *Adv Biophys* **28**, 81–133.
- Lan, R. T. & Reeves, P. R. (1998). Recombination between rRNA operons created most of the ribotype variation observed in the seventh pandemic clone of *Vibrio cholerae*. *Microbiology* **144**, 1213–1221.
- Lebuhn, M., Bathe, S., Achouak, W., Hartmann, A., Heulin, T. & Schloter, M. (2006). Comparative sequence analysis of the internal transcribed spacer 1 of *Ochrobactrum* species. *Syst Appl Microbiol* **29**, 265–275.
- Liao, D. (2000). Gene conversion drives within genic sequences: concerted evolution of ribosomal RNA genes in bacteria and archaea. *J Mol Evol* **51**, 305–317.
- Liska, S. L., Perine, P. L., Hunter, E. F., Crawford, J. A. & Feeley, J. C. (1982). Isolation and transportation of *Treponema pertenuis* in golden hamsters. *Curr Microbiol* **7**, 41–43.
- Liu, H., Rodes, B., Chen, C.-Y. & Steiner, B. (2001). New tests for syphilis: rational design of a PCR method for detection of *Treponema pallidum* in clinical specimens using unique regions of the DNA polymerase I gene. *J Clin Microbiol* **39**, 1941–1946.
- Liveris, D., Wormser, G. P., Nowakowski, J., Nadelman, R., Bittker, S., Cooper, D., Varde, S., Moy, F. H., Forseter, G. & other authors (1996). Molecular typing of *Borrelia burgdorferi* from Lyme disease patients by PCR-restriction fragment length polymorphism analysis. *J Clin Microbiol* **34**, 1306–1309.
- Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D. & Lefevre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **26**, 2462–2463.
- Matějčková, P., Strouhal, M., Šmajš, D., Norris, S. J., Palzkill, T., Petrosino, J. F., Sodergren, E., Norton, J. E., Singh, J. & other authors (2008). Complete genome sequence of *Treponema pallidum* ssp. *pallidum* strain SS14 determined with oligonucleotide arrays. *BMC Microbiol* **8**, 76.
- Matějčková, P., Flasarová, M., Zákoucká, H., Bůrek, M., Kremenová, S., Arenberger, P., Woznicová, V., Weinstock, G. M. & Šmajš, D. (2009). 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* **58**, 832–836.
- Nei, M. & Rooney, A. P. (2005). Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* **39**, 121–152.
- Nichols, H. J. & Hough, W. H. (1913). Demonstration of *Spirochaeta pallida* in the cerebrospinal fluid: from a patient with nervous relapse following the use of salvarsan. *JAMA* **60**, 108–110.
- Pei, A., Nossa, C. W., Chokshi, P., Blaser, M. J., Yang, L., Rosmarin, D. M. & Pei, Z. (2009). Diversity of 23S rRNA genes within individual prokaryotic genomes. *PLoS ONE* **4**, e5437.
- Pei, A. Y., Oberdorf, W. E., Nossa, C. W., Agarwal, A., Chokshi, P., Gerz, E. A., Jin, Z., Lee, P., Yang, L. & other authors (2010). Diversity of 16S rRNA genes within individual prokaryotic genomes. *Appl Environ Microbiol* **76**, 3886–3897.
- Petes, T. D. & Hill, C. W. (1988). Recombination between repeated genes in microorganisms. *Annu Rev Genet* **22**, 147–168.
- Petit, M.-A. (2005). Mechanisms of homologous recombination in bacteria. In *The Dynamic Bacterial Genome*, pp. 3–32. Edited by P. Mullany. New York: Cambridge University Press.
- Pillay, A., Liu, H., Chen, C.-Y., Holloway, B., Sturm, A. W., Steiner, B. & Morse, S. A. (1998). Molecular subtyping of *Treponema pallidum* subspecies *pallidum*. *Sex Transm Dis* **25**, 408–414.
- Posada, D. & Crandall, K. A. (2001). Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci U S A* **98**, 13757–13762.
- Rozen, S. & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365–386.
- Sadeghifard, N., Gürtler, V., Beer, M. & Seviour, R. J. (2006). The mosaic nature of intergenic 16S–23S rRNA spacer regions suggests rRNA operon copy number variation in *Clostridium difficile* strains. *Appl Environ Microbiol* **72**, 7311–7323.
- Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Santoyo, G. & Romero, D. (2005). Gene conversion and concerted evolution in bacterial genomes. *FEMS Microbiol Rev* **29**, 169–183.
- Sawyer, S. (1989). Statistical tests for detecting gene conversion. *Mol Biol Evol* **6**, 526–538.
- Schwartz, J. J., Gazumyan, A. & Schwartz, I. (1992). rRNA gene organization in the Lyme disease spirochete, *Borrelia burgdorferi*. *J Bacteriol* **174**, 3757–3765.
- Seshadri, R., Myers, G. S., Tettelin, H., Eisen, J. A., Heidelberg, J. F., Dodson, R. J., Davidsen, T. M., DeBoy, R. T., Fouts, D. E. & other authors (2004). Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. *Proc Natl Acad Sci U S A* **101**, 5646–5651.
- Šmajš, D., Mikalová, L., Čejková, D., Strouhal, M., Zbaníková, M., Pospíšilová, P., Norris, S. J. & Weinstock, G. M. (2011a). Whole genome analyses of treponemes: new targets for strain- and subspecies-specific molecular diagnostics. In *Syphilis – Recognition, Description and Diagnosis*, pp. 19–34. Edited by N. S. Sato. Rijeka, Croatia: InTech.
- Šmajš, D., Zbaníková, M., Strouhal, M., Čejková, D., Dugan-Rocha, S., Pospíšilová, P., Norris, S. J., Albert, T., Qin, X. & other authors (2011b). Complete genome sequence of *Treponema paraluis-cuniculi*, strain Cuniculi A: the loss of infectivity to humans is associated with genome decay. *PLoS ONE* **6**, e20415.
- Smith, J. M. (1992). Analyzing the mosaic structure of genes. *J Mol Evol* **34**, 126–129.
- Stamm, L. V. & Bergen, H. L. (2000). 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* **44**, 806–807.
- Stamm, L. V., Kerner, T. C., Jr, Bankaitis, V. A. & Bassford, P. J., Jr (1983). Identification and preliminary characterization of *Treponema*

- pallidum* protein antigens expressed in *Escherichia coli*. *Infect Immun* **41**, 709–721.
- Stamm, L. V., Bergen, H. L. & Walker, R. L. (2002)**. Molecular typing of papillomatous digital dermatitis-associated *Treponema* isolates based on analysis of 16S–23S ribosomal DNA intergenic spacer regions. *J Clin Microbiol* **40**, 3463–3469.
- Stewart, F. J. & Cavanaugh, C. M. (2007)**. Intragenomic variation and evolution of the internal transcribed spacer of the rRNA operon in bacteria. *J Mol Evol* **65**, 44–67.
- Strouhal, M., Šmajš, D., Matějková, P., Sodergren, E., Amin, A. G., Howell, J. K., Norris, S. J. & Weinstock, G. M. (2007)**. Genome differences between *Treponema pallidum* subsp. *pallidum* strain Nichols and *T. paraluisuniculi* strain Cuniculi A. *Infect Immun* **75**, 5859–5866.
- Takahashi, N. K., Yamamoto, K., Kitamura, Y., Luo, S.-Q., Yoshikura, H. & Kobayashi, I. (1992)**. Nonconservative recombination in *Escherichia coli*. *Proc Natl Acad Sci U S A* **89**, 5912–5916.
- Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007)**. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596–1599.
- Turner, T. B. & Hollander, D. H. (1957)**. Biology of the treponematoses based on studies carried out at the International Treponematoses Laboratory Center of the Johns Hopkins University under the auspices of the World Health Organization. *Monogr Ser World Health Organ* **35**, 3–266.
- Wendel, G. D., Jr, Sánchez, P. J., Peters, M. T., Harstad, T. W., Potter, L. L. & Norgard, M. V. (1991)**. Identification of *Treponema pallidum* in amniotic fluid and fetal blood from pregnancies complicated by congenital syphilis. *Obstet Gynecol* **78**, 890–895.
- Wormser, G. P., Brisson, D., Liveris, D., Hanincová, K., Sandigursky, S., Nowakowski, J., Nadelman, R. B., Ludin, S. & Schwartz, I. (2008)**. *Borrelia burgdorferi* genotype predicts the capacity for hematogenous dissemination during early Lyme disease. *J Infect Dis* **198**, 1358–1364.
- Woznicová, V., Šmajš, D., Wechsler, D., Matějková, P. & Flasarová, M. (2007)**. Detection of *Treponema pallidum* subsp. *pallidum* from skin lesions, serum, and cerebrospinal fluid in an infant with congenital syphilis after clindamycin treatment of the mother during pregnancy. *J Clin Microbiol* **45**, 659–661.