VERSITA

Cent. Eur. J. Biol. • 3(4) • 2008 • 396-406
DOI: 10.2478/s11535-008-0043-4

Central European Journal of **Biology**

Unusual intragenomic and interspecific variability of *Geobacillus* 16S-23S rRNA internal transcribed spacers

Research Article

Nomeda Kuisiene*, Juozas Raugalas, Donaldas Chitavichius

Department of Plant Physiology and Microbiology, Vilnius University, LT-03101 Vilnius, Lithuania

Received 1 June 2008; Accepted 1 September 2008

Abstract: The aim of this study was to evaluate the inter- and intraspecific as well as intragenomic variability of *Geobacillus* 16S-23S rRNA internal transcribed spacers without tRNA genes and to compare these sequences with sequences bearing tRNA genes. In this study the structural analysis was performed in a unique way because the length and the sequence of the structural blocks were adjusted to fit the structure of 16S-23S rRNA internal transcribed spacers of five different *Geobacillus* species. Our study demonstrated the mosaic-like structure of 16S-23S rRNA internal transcribed spacers in *Geobacillus*. Some characteristics of these spacers of geobacilli were not previously reported for other bacteria: unusually short conserved sequence in the 5' end region, some identical conserved blocks in both 5' and 3' regions of 16S-23S rRNA internal transcribed spacers, the same sequence blocks in both 16S-23S and 23S-5S rRNA intergenic spacers. Our study demonstrated quite uniform arrangement of the sequence blocks in *Geobacillus thermodenitrificans*. This species diverged early in the phylogenetic tree of the genus *Geobacillus*. For the phylogenetically recent species *Geobacillus kaustophilus* and *Geobacillus lituanicus* the low inter- and intraspecific, but high intragenomic variability, as a consequence of recent phylogenetic events, was established.

Keywords: *Geobacillus • 16S-23S internal transcribed spacer • Taxonomy • Conserved region • Thermophilic*

1. Introduction

The 16S-23S rRNA intergenic transcribed spacers (ITS) are the most variable regions of the ribosomal operon. Since the ITS have fewer functional constraints than the adjacent ribosomal genes, which undergo concerted evolution, their sequences can contain traces of ribosomal operon rearrangements and species-specific or even strain-specific traits [1].

Differing levels of variability within bacterial ITS have been described at the intraspecific level, although variability between the ITS of different operons is not always seen as some species possess identical ITS sequences throughout the ribosomal operons. In many organisms, intragenomic variability consists of only a few substitutions over the entire ITS length, and results from the presence or absence of tRNA genes. However, intragenomic variability (mosaic-like structure) has been also detected in the bacterial ITS [2-6]. To date, mosaic-like structures have been detected in the non-thermophilic bacteria *Salmonella enterica*, *Haemophilus parainfluenzae*, *Staphylococcus aureus*, *Vibrio cholerae*, *Vibrio mimicus*, *Vibrio parahaemolyticus*, *Myxobacterium* spp., *Acinetobacter baylyi, Photobacterium damselae* and *Clostridium difficile* [2,3,5,7-9].

Geobacillus ITS sequences with tRNA genes were analysed and used for the design of the genus-specific primers [10]. ITS sequences with tRNA genes have been found to be highly variable upstream of boxA in the *Geobacillus kaustophilus-Geobacillus lituanicus-Geobacillus thermoleovorans-Geobacillus vulcani* species group [11]. To our knowledge, ITS sequences without tRNA genes have never been analysed in the genus *Geobacillus*, and the degree of variability in the latter sequences is not known. The aim of this study was to evaluate the inter- and intraspecific as well as intragenomic variability of *Geobacillus* (thermophilic

* *E-mail: nomeda.kuisiene@gf.vu.lt*

Strain	Accession numbers
G. kaustophilus HTA426	BA000043, rrnA ITS 11974 – 12605 BA000043, rrnB ITS 32344 – 33137 BA000043, rrnC ITS 90966 – 91351 BA000043, rrnD ITS 155856 – 156520 BA000043, rrnE ITS 251446 – 252121 BA000043, rrnF ITS 278100 – 278439 BA000043, rrnG ITS 521371 – 521910 BA000043, rrnH ITS 2190385 – 2190783 BA000043, rrnI ITS 2945348 – 2946082
G. kaustophilus ATCC 8005^T	AF478063
G. lituanicus DSM 15325^T	EF645695, EF645696, EF645697[b], EF645698, EF645699[b], EF645700, EF645701
G. stearothermophilus DSM 13240[a]	Contig303 ITS 2264-1629 Contig343 ITS 26524-27266 Contig394 ITS 61396-62044 Contig378 ITS 52147-52982
G. stearothermophilus ATCC 12980^T	AF478064
G. stearothermophilus 36A	AY144575, AY144576[c], AY144577, AY144578, AY144579[c], AY144580, EF645702[d], EF645703, EF645704[d], EF645705, EF645706[c]
G. stearothermophilus 28	EF645715[e], EF645716[d], EF645717[e], EF645718[f], EF645719, EF645720, EF645721, EF645722, EF645723, EF645724, EF645725
G. stearothermophilus 3	EF645707[d], EF645708, EF645709[f], EF645710, EF645711, EF645712, EF645713, EF645714[d]
G. thermodenitrificans NG80-2	CP000557, rrn001 ITS 12169 – 12743[i] CP000557, rrn002 ITS 30880 – 31452[i] CP000557, rrn003 ITS 88875– 89261[g] CP000557, rrn004 ITS 151881– 152128[h] CP000557, rrn005 ITS 227523– 227911[g] CP000557, rrn006 ITS 260346– 260731[g] CP000557, rrn007 ITS 549083– 549473[g] CP000557, rrn008 ITS 1313774– 1313908 CP000557, rrn009 ITS 2172259– 2172508[h] CP000557, rrn010 ITS 2970839– 2971224[g]

Table 1. Sequences of 16S-23S rRNA internal transcribed spacers used in this study.

G – Geobacillus
a – numbers of contigs are listed according to the data of 02-01-2008 available at http://www.genome.ou.edu/blast/bstearo_blastall.html
b, c, d, e, f, g, h – sequences are identical in the arrangement of the sequence blocks
i – sequences differed in 3' end: two neighbouring stretches of nucleotides (8 nt each) were duplicated in the block U3 in 3' end of rrn001 but not in rrn002.

bacteria) ITS without tRNA genes and to compare the structure of these sequences with those containing tRNA genes. In this study the structural analysis was performed in a unique way because the length and the sequence of the structural blocks were adjusted to fit the ITS structure of five different *Geobacillus* species. Consequently, evidence of rearrangements of ITS during the evolution of *Geobacillus* species have been found in this study.

2. Experimental Procedures

Accession numbers of the ITS sequences used in this work are listed in Table 1. The sequences of *G. lituanicus* DSM 15325^T (DSM – Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH) and *Geobacillus stearothermophilus* wild type strains 3, 28 and 36A were obtained in our previous study [10]. ITS sequences were also extracted from the whole genome sequences of *G. kaustophilus* HTA426 and *Geobacillus thermodenitrificans* NG80-2 as well as from the contigs of the incomplete sequenced genome of *G. stearothermophilus* DSM 13240. The sequences AF478064 (*G. stearothermophilus* ATCC 12980^T, ATCC – American Type Culture Collection) and AF478063 (*G. kaustophilus* ATCC 8005^T) were obtained from GenBank.

Sequences were aligned using MEGA 3.1 program [12]. ITS alignments were adjusted manually. In order to aid alignment, detection and further analyses of the short sequence blocks that were present in only certain sequences, ITS regions from different species were aligned separately. The sequences of *G. stearothermophilus* DSM 13240 and those of *G. stearothermophilus sensu stricto* (strains 3, 28, 36A and ATCC 12980^T) were analysed separately as we have revealed [10] that *G. stearothermophilus* DSM 13240 was misidentified and represents another species of the genus *Geobacillus*, not *G. stearothermophilus*. MEGA 3.1 program was used to identify conserved regions within the different ITS sequences and

among the different *Geobacillus* species and strains. Conserved regions within and between all the examined *Geobacillus* species were defined arbitrarily as five or more consecutive nucleotides either identical or highly similar in sequence. Only block V represents a highly variable stretch of three nucleotides. Additionally, large conserved regions (>20 bp) were tested using BLASTN searches.

The sequence blocks common to at least two *Geobacillus* species were designated as U (universal) blocks. The blocks found only in the ITS sequences of a certain species (unique blocks) were named KP (*G. kaustophilus*), LT (*G. lituanicus*), ST (*G. stearothermophilus sensu stricto*), TD (*G. thermodenitrificans*) and GS (*G. stearothermophilus* DSM 13240).

3. Results

3.1. The universal and unique sequence blocks

All sequences shared the first 5 (U1) and the last 24-28 (U12) bp (Figures 1-5). The block U1 was highly conserved throughout the tested species, but U12 varied in sequence despite the overall similarity of this end of ITS. The interspersed sequences between these conserved ends were composed of a collection of sequence blocks that were not present in all ITS, indicating a mosaic-like structure of geobacilli ITS sequences.

In total, 84 universal and 38 unique sequence blocks were identified in the ITS sequences of *Geobacillus*. Some universal blocks (U1-U6, U8, U11-U16 and U81) as well as unique blocks have been found in all tested species (Figures 1-5). Unique blocks KP were found only in the sequences from different operons of *G. kaustophilus* HTA426 but not in the sequence of *G. kaustophilus* ATCC 8005^T (Figure 1). Unique blocks ST1, ST7, ST8, ST10, ST11 were found in the sequences of *G. stearothermophilus* 3, 28 and 36A. Block ST5 was found in the sequences of *G. stearothermophilus* ATCC 12980^T, 28 and 36A. Blocks ST2 and ST3 were found in the sequences of all examined *G. stearothermophilus* strains. Only three strain-specific ST blocks, ST4, ST6 and ST9, were found in the sequences of the particular strains – *G. stearothermophilus* 3, *G. stearothermophilus* 36A and *G. stearothermophilus* 28, respectively (Figure 2).

The universal blocks varied in length from 5 nt (U1 and U10) to 79 nt (U9). Although thirty-six universal blocks were completely identical in all the species, the majority of blocks showed some degree of variability (data not shown). The most variable in the sequence was block U2, with a consensus sequence of (C)(GAA)(A)GAAAA/CGC/T(G/A)(G/A)(A). The length of this block varied from 7 to 14 nt in different ITS sequences. The degree of variability of block U2 in all species was comparable; most variability was detected between the different sequences in the same locus of the particular strain, consequently, this sequence was interpreted as the distinct block.

tRNA^{Ile} and tRNA^{Ala} genes have been found in some of the ITS sequences during our previous work [10]. The antitermination element boxA was identified in all tested sequences, with sequence EF645721 (*G. stearothermophilus* 28) containing two boxA elements (Figure 2d).

Some universal blocks have been found to be genus-specific. During the BLAST search, no similarity with the sequences of prokaryotes has been detected for the sequence blocks U36, U44, U46, U47, U50, U54, U55, U60, U70 and U75. The sequence blocks U17 and U19 have been found not only in the 16S-23S rRNA intergenic spacer but also in the spacer between 23S rRNA and 5S rRNA genes. Both U17 and U19 were detected in 23S-5S rRNA intergenic spacer in six ribosomal operons (rrn002, rrn003, rrn005, rrn006, rrn007 and rrn008) of *G. thermodenitrificans* NG80-2 as well as in 23S-5S rRNA intergenic spacer of the operons rrnC and rrnD of *G. kaustophilus* HTA426. *G. thermodenitrificans* NG80-2 did not possess these two blocks in ITS, and *G. kaustophilus* HTA426 had U17 and U19 in the ITS of rrnH and not in the ITS of rrnC and rrnD (Figure 1a). *G. lituanicus* DSM 15325^T, *G. stearothermophilus* DSM 13240 and *G. stearothermophilus sensu stricto* also had U17 and U19 in their 5' end of some ITS. These results suggest that U17 and U19 in the ITS are derived from the 23S-5S rRNA intergenic spacer.

3.2. Sequences with tRNA genes *vs.* sequences without them

ITS sequences with tRNA genes were compared to sequences without them in order to reveal the differences in both the composition and the arrangement of the sequence blocks.

The conserved blocks U9 and U16, when present, were found only in the sequences with tRNA genes. Sequences EF645701 (*G. lituanicus* DSM 15325^T), rrnA (*G. kaustophilus* HTA426), the contig 303 of *G. stearothermophilus* DSM 13240, and all the sequences of *G. stearothermophilus sensu stricto*, lacked U9 but contained U16. Both sequences with tRNA genes of *G. thermodenitrificans* NG80-2 contained this block in their 5' end (Figures 1-5). It should be noted that U9 was found only in the sequences of geobacilli

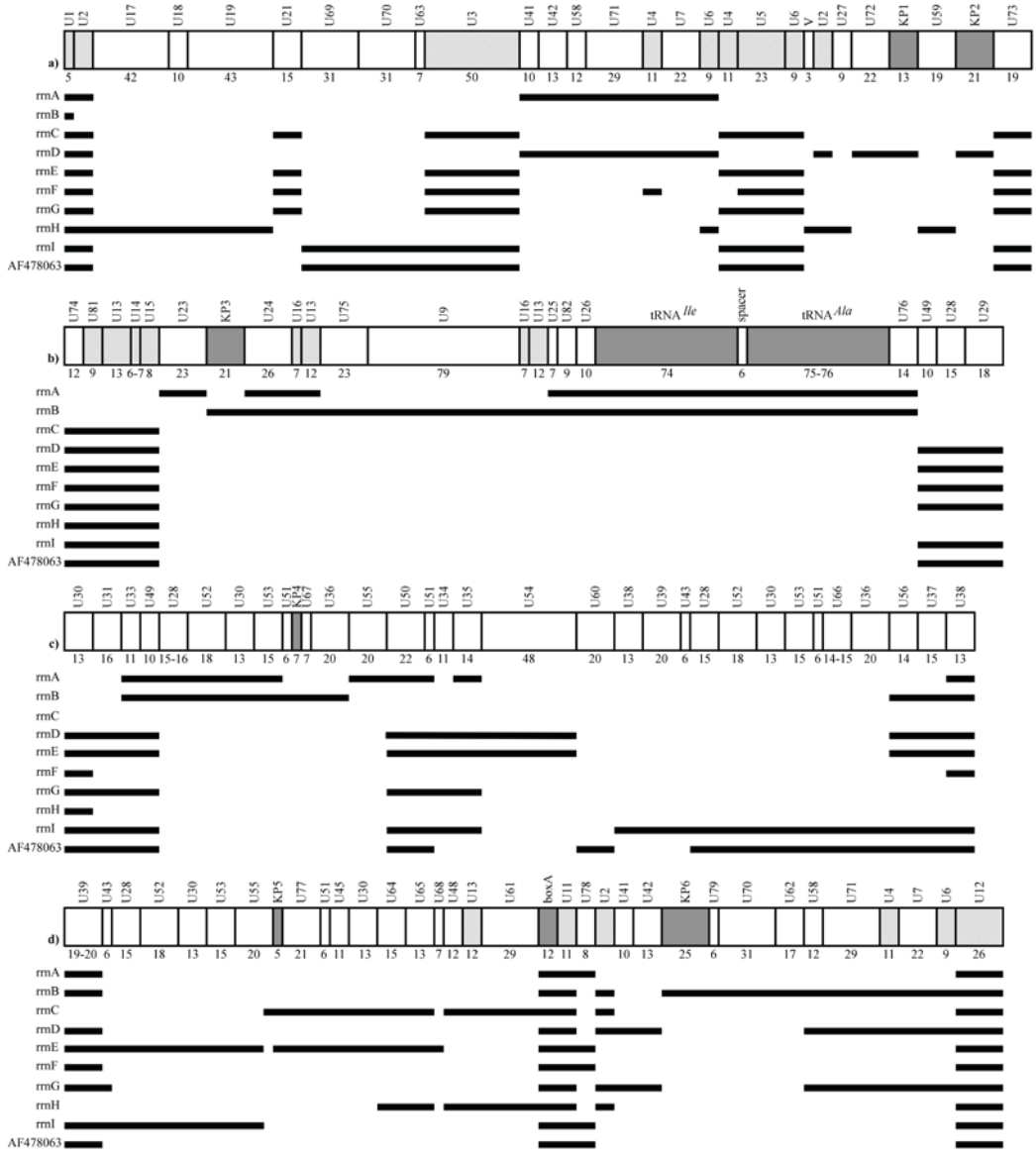


Figure 1. Schematic representation of the mosaic-like structure of 16S-23S internal transcribed spacers of *Geobacillus kaustophilus*. KP blocks, boxA and tRNA genes are shaded black, and the universal blocks, present in all the tested species, are latticed. The number under each block indicates the length of this block in nucleotides in this locus (not sequence length of consensus). The sections a, b, c and d represent the parts of the alignment.

in BLASTN search suggesting this large (79 nt) block is genus-specific.

The cluster U23-U24 was found in the sequences with tRNA genes of all tested species except for *G. thermodenitrificans* NG80-2 and the sequence rrnB of *G. kaustophilus* HTA426. The latter sequence lacked U23, although contained U24 (Figure 1b, 3).

The conserved blocks U81 and U14 were found in nearly all the sequences without tRNA genes but not in the sequences containing tRNA genes. The exceptions were AF478064 of *G. stearothermophilus* ATCC 12980^T as well as AY144578 and AY144580 of *G. stearothermophilus* 36A, which all lacked U14 but contained U81 (Figure 2b,d). In the sequences of all species examined, the cluster of the blocks U16-U13 could be found only in the sequences with tRNA genes (Figure 1b, 2b, 3b, 4a, 5b), while the combination U81-U13-U14-U15 was restricted to the sequences without tRNA genes (Figure 1b, 2b, 3a, 4a, 5a,b).

Some sequences without tRNA genes (rrnH of *G. kaustophilus* HTA426, contig 378 of *G. stearothermophilus* DSM 13240) had the block U18 (Figure 1a, 4a), and some of them (rrnD, rrnE, rrnF, rrnG and rrnI of *G. kaustophilus* HTA426 and

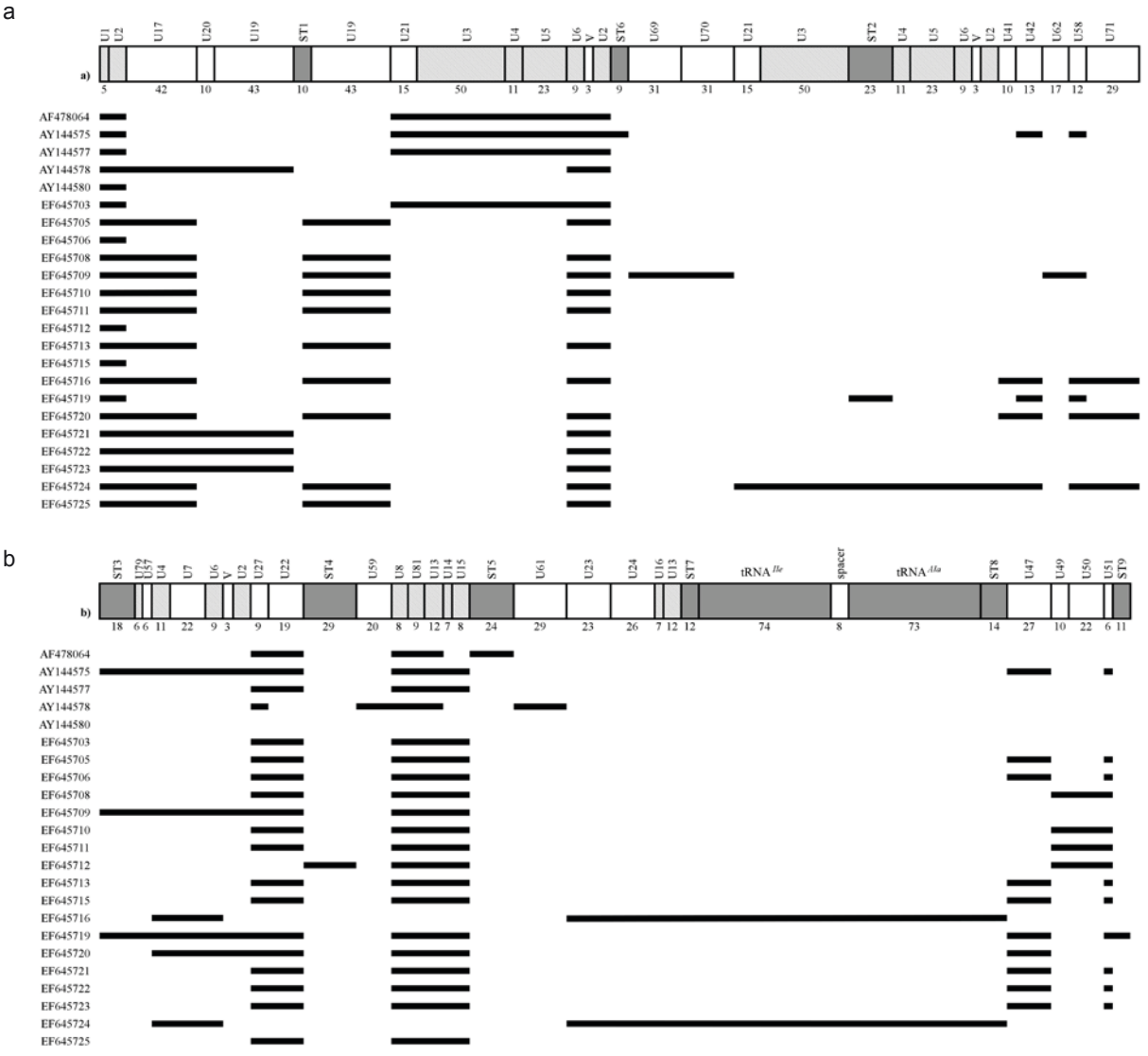
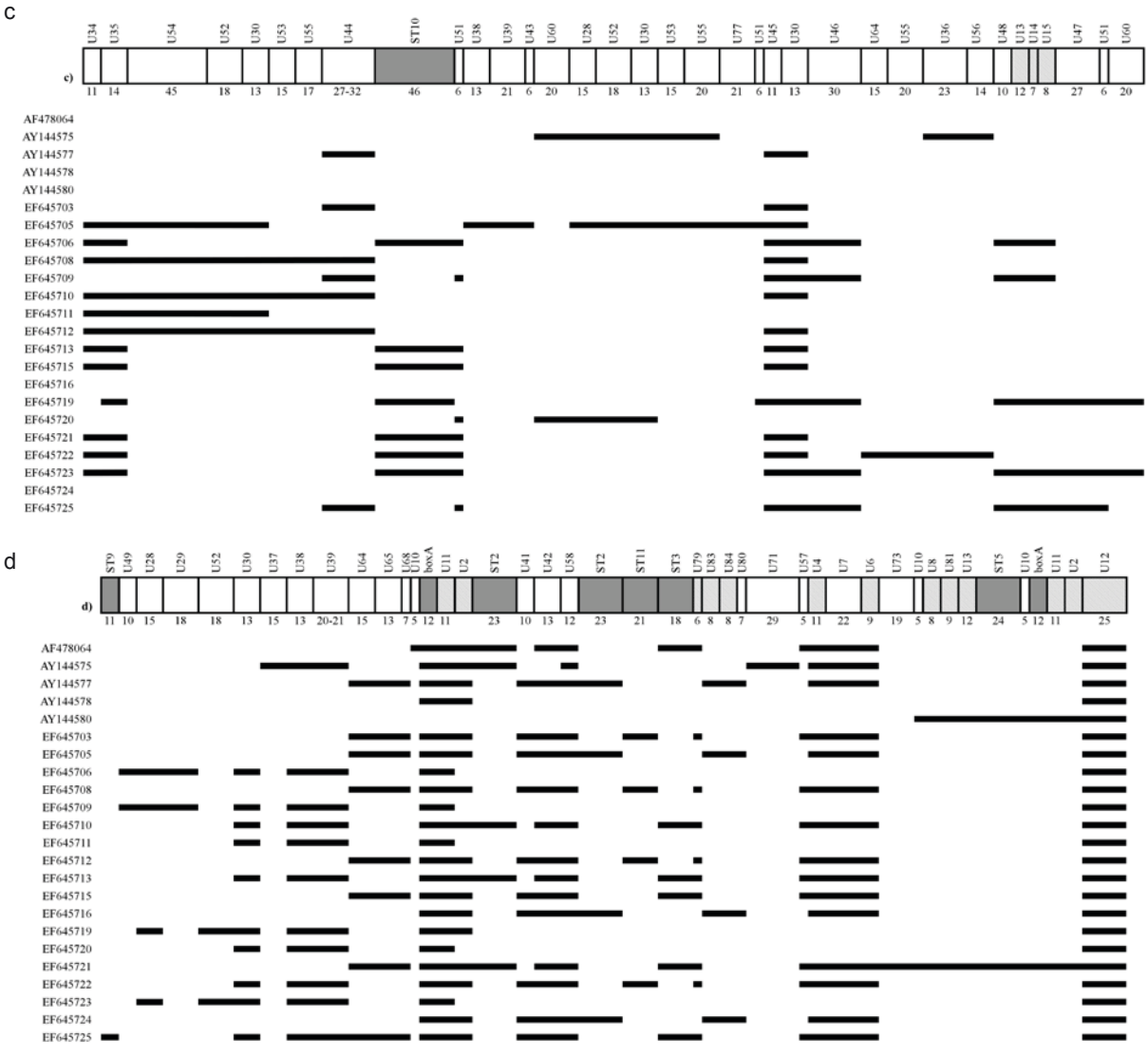


Figure 2. Schematic representation of the mosaic-like structure of 16S-23S internal transcribed spacers of *Geobacillus stearothermophilus sensu stricto*. ST blocks, boxA and tRNA genes are shaded black, and the universal blocks, present in all the tested species, are latticed. The number under each block indicates the length of this block in nucleotides in this locus. The sections a, b, c and d represent the parts of the alignment (figure continued on the next page).

AF478063 of *G. kaustophilus* ATCC 8005^T, EF645706 and EF645709 of *G. stearothermophilus sensu stricto*, contig 394 of *G. stearothermophilus* DSM 13240, EF645695 and EF645696 of *G. lituanicus* DSM 15325^T) possessed the block U29 (Figure 1b, 2d, 4b, 5c). U18 and U29 were found neither in the sequences with tRNA genes nor in any sequence of *G. thermodenitrificans* NG80-2. The sequence block U31 was found in the sequences without tRNA genes rrnD, rrnE, rrnG and rrnI of *G. kaustophilus* HTA426 and AF478063 of *G. kaustophilus* ATCC 8005^T, contig 394 *G. stearothermophilus* DSM 13240, EF645696 of *G. lituanicus* DSM 15325^T (Figure 1c, 4b, 5c). The blocks U47 and U48, when present, were also found only in the sequences without tRNA genes. U47 was present in the following sequences: AY144575, AY144576, AY144579, EF645705, EF645706, EF645713, EF645715, EF645717, EF645719, EF645720, EF645721, EF645722, EF645723 and EF645725 of *G. stearothermophilus sensu stricto*, contigs 343, 378 and 394 of *G. stearothermophilus* DSM 13240, and EF645695 of *G. lituanicus* DSM 15325^T (Figure 2b,c, 4b, 5c). U48 was found in rrnC and rrnH of *G. kaustophilus* HTA426, AY144576, AY144579, EF645706, EF645709, EF645719, EF645723 and EF645725 of *G. stearothermophilus sensu stricto* and contigs 343, 378 and 394 of *G. stearothermophilus* DSM 13240 (Figure 1d, 2c, 4b).

Thus, although the majority of the conserved blocks were found in both the sequences with tRNA genes and



continued Figure 2. Schematic representation of the mosaic-like structure of 16S-23S internal transcribed spacers of *Geobacillus stearothermophilus sensu stricto*. ST blocks, boxA and tRNA genes are shaded black, and the universal blocks, present in all the tested species, are latticed. The number under each block indicates the length of this block in nucleotides in this locus. The sections a, b, c and d represent the parts of the alignment (figure continued from the previous page).

those without these genes, some blocks were identified only in the sequences of the particular type, with or without tRNA genes and irrespective of the species.

3.3. 5' end of ITS *vs.* 3' end of ITS

Some sequence blocks were found in both the 5' and 3' end of ITS. The blocks U2, U3, U4 and U6 were found in the 5' region of all examined species, although not all sequences of these species possessed these blocks. Two copies of U2 were found in the 5'end of sequences rrnD and rrnH of *G. kaustophilus* HTA426, rrn003, rrn005, rrn006, rrn007 and rrn010 of *G. thermodenitrificans* NG80-2, contig 378 of *G. stearothermophilus* DSM 13240, EF645697, EF645698, EF645699 and EF645700 of *G. lituanicus* DSM 15325^T and all the sequences of *G. stearothermophilus sensu stricto* except for AY144576, AY144579, AY144580, EF645706, EF645712 and EF645715 (Figure 1a, 2a,b 3a, 4a, 5a,b). Three copies of U2 were present in the 5' end of sequences AY144575, EF645709, EF645718 and EF645720 of *G. stearothermophilus sensu stricto* (Figure 2a,b). The 3' end regions (downstream of boxA) of the examined ITS also possessed U2, U3, U4 and U6, the only exceptions being U2 in the sequences of *G. thermodenitrificans* NG80-2 and U3 in the sequences of *G. lituanicus* DSM 15325^T, *G. kaustophilus* ATCC 8005^T, *G. kaustophilus* HTA426 and *G. stearothermophilus sensu stricto*. These four blocks

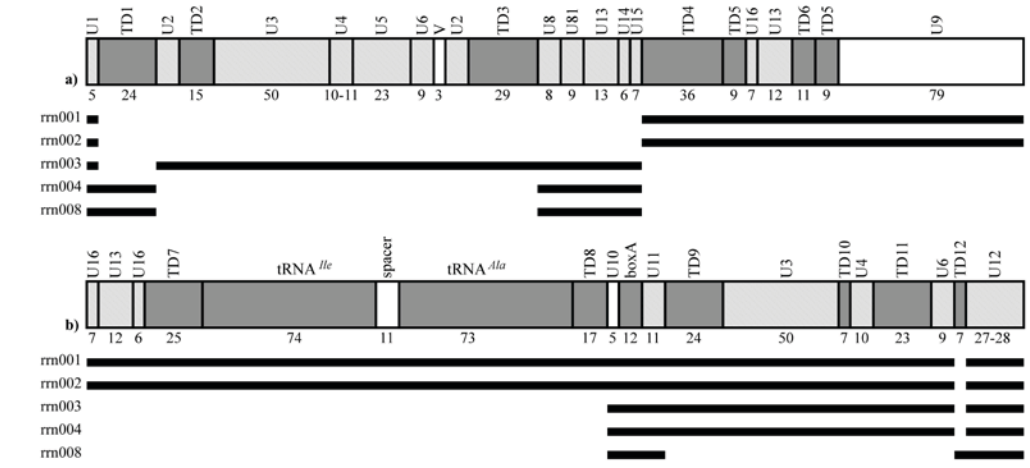


Figure 3. Schematic representation of the mosaic-like structure of 16S-23S internal transcribed spacers of *Geobacillus thermodenitrificans* NG80-2. TD blocks, boxA and tRNA genes are shaded grey, and the universal blocks, present in all the tested species, are latticed. The number under each block indicates the length of this block in nucleotides in this locus. The sections a and b represent the parts of the alignment.

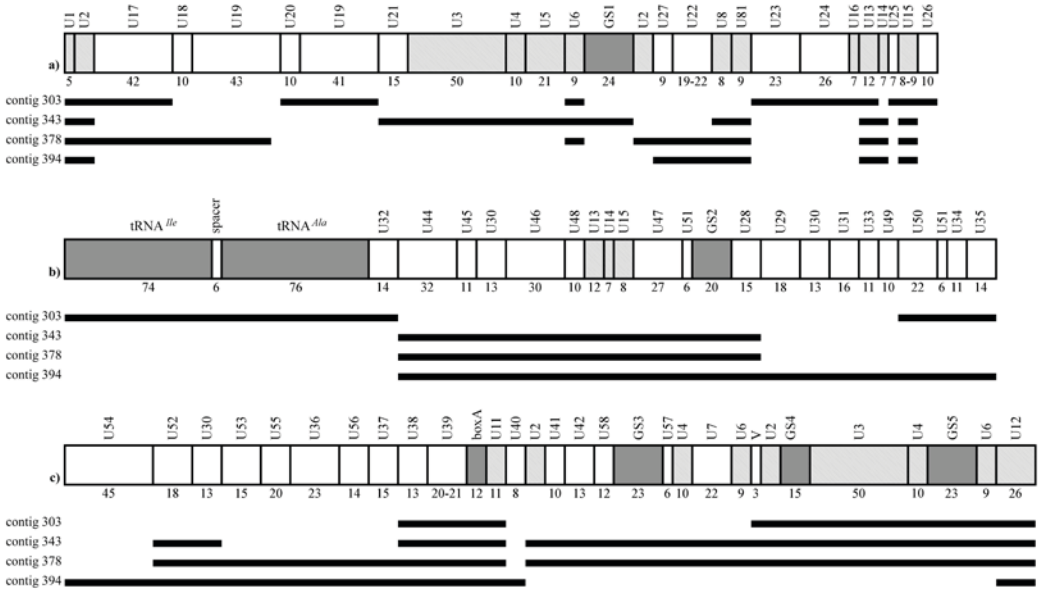


Figure 4. Schematic representation of the mosaic-like structure of 16S-23S internal transcribed spacers of *Geobacillus stearothermophilus* DSM 13240. GS blocks, boxA and tRNA genes are shaded black, and the universal blocks, present in all the tested species, are latticed. The number under each block indicates the length of this block in nucleotides in this locus. The sections a, b and c represent the parts of the alignment.

were not detected in the central part of ITS of the examined species.

U41, U42, U58, U71, U4, U7 and U6 constitute a conserved cluster of the blocks in the 5' end of ITS of rrnA and rrnD of *G. kaustophilus* HTA426, EF645702, EF645704, EF645707, EF645714, EF645716, EF645720 and EF645724 of *G. stearothermophilus sensu stricto,* EF645697, EF645799 and EF645701 of *G. lituanicus* DSM 15325^T (Figure 1a, 2a,b, 5a). The same cluster U41-U42-U58-U71-U4-U7-U6 was identified in the 3' regions of the sequences rrnG and rrnD of *G. kaustophilus* HTA426 as well as EF645701 and EF645700 of *G. lituanicus* DSM 15325^T (Figure 1d, 5d). The sequences rrnD (*G. kaustophilus* HTA426) and EF645701 (*G. lituanicus* DSM 15325^T) had two copies of these concatenated blocks in both the 5' and the 3' end regions of ITS. Sequences of *G. stearothermophilus* DSM 13240 possessed an interrupted cluster in which the pair of blocks GS3-U57 was found to be inserted between U41-U42-U58 and U4-U7-U6 in the 3' end region of the contigs 343 and 378. These two contigs as well as contig 303 also had second

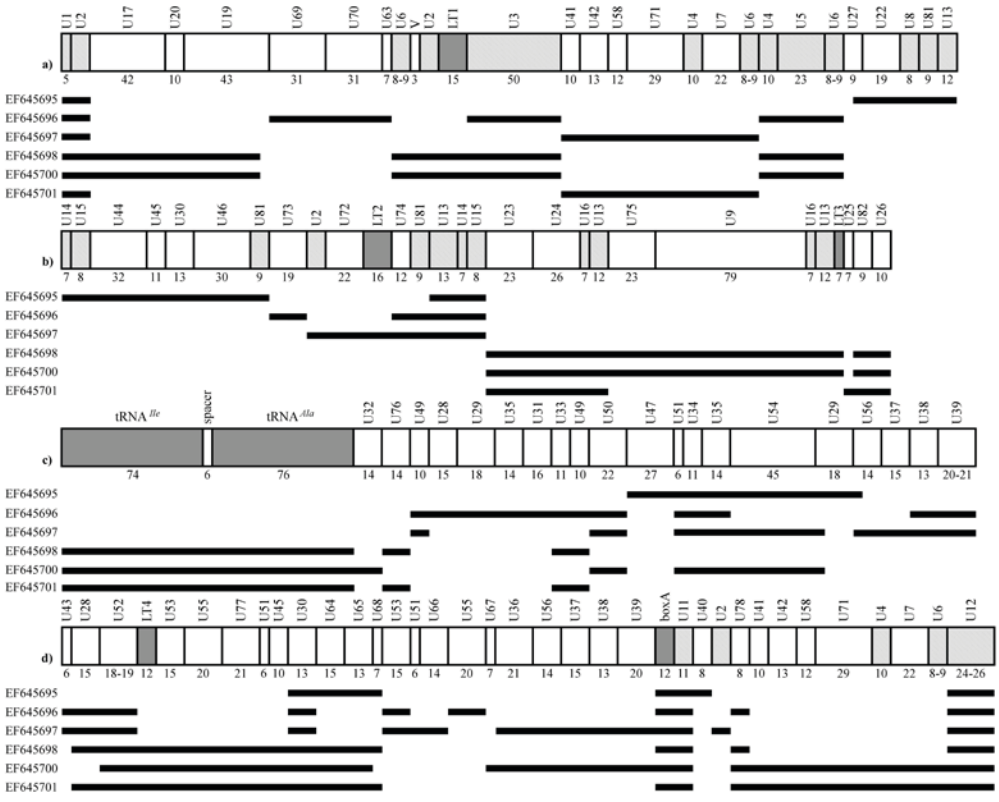


Figure 5. Schematic representation of the mosaic-like structure of 16S-23S internal transcribed spacers of *Geobacillus lituanicus* DSM 15325^T. LT blocks, boxA and tRNA genes are shaded black, and the universal blocks, present in all the tested species, are latticed. The number under each block indicates the length of this block in nucleotides in this locus. The sections a, b, c and d represent the parts of the alignment.

copies of U4 and U6 with GS5 in between them directly upstream of block U12 (Figure 4c). The sequences of *G. stearothermophilus sensu stricto* also contained an interrupted cluster in their 3' end where U41-U42-U58 and U4-U7-U6 were separated by different combinations of ST and universal blocks. Only AY144575 possessed U71 directly upstream of U4-U7-U6 (Figure 2d).

The conserved block U70 was found in the 5' end region of rrnI and AF478063 of *G. kaustophilus* strains HTA426 and ATCC 8005^T, respectively, as well as in the 5' end of EF645696 of *G. lituanicus* DSM 15325^T (Figure 1a, 5a). The same block was found in the 3' region of rrnB of *G. kaustophilus* HTA426 (Figure 1d). The block U70 is concatenated through the block U62 to the cluster U58-U71-U4-U7-U6. The blocks U41 and U42 were not found upstream this cluster. The block U70 was also found in the 5' region of the sequence EF645709 (*G. stearothermophilus* 3). This block is concatenated to U62-U58 as in the 3' end region of rrnB of *G. kaustophilus* HTA426 (Figure 2a). However, we did not find U70 in the 3' end regions of the sequences of *G. stearothermophilus sensu stricto*.

The identical sequence blocks and even the conserved arrangement of the blocks were identified in both the 5' end and the 3' end regions of *Geobacillus* ITS sequences. The presence of these conserved blocks in both ends of ITS does not appear to depend on the presence or absence of the tRNA genes.

3.4. Strain *vs.* strain of the same species

This study evaluated intraspecific variability for two species, *G. kaustophilus* and *G. stearothermophilus* (*sensu stricto*).

The composition and arrangement of the conserved blocks of the sequence AF478063 of *G. kaustophilus* ATCC 8005^T was most similar to rrnI of *G. kaustophilus* HTA426. The 5' end of AF478063 was identical with rrnI, and of all the sequences evaluated, only these two sequences had U69-U70-U63 cluster of the blocks (Figure 1a). The 3' end of AF478063 was identical with that of rrnA, rrnE, rrnF and rrnI of *G. kaustophilus* HTA426 (Figure 1d). The central region of AF478063 was most similar to rrnI despite some differences between these two sequences (Figure 1c,d). For example, AF478063 contained U60 that was not found in the sequence of rrnI. Instead, rrnI possessed

a U34-U35 cluster that was not found in the sequence AF478063. In addition, rrnI had two copies of the clusters U38-U39-U43 and U28-U52-U30-U53, while AF478063 contained only single copies of these two combinations of blocks.

Similar observations were made for *G. stearothermophilus sensu stricto*. For example, the 5' end was identical in the sequences EF645708, EF645710, EF645711, EF645713 (strain 3), EF645725 (strain 28) and EF645705 (strain 36A) (Figure 2a), and the 3' end was identical in the sequences EF645711 (strain 3) and EF645720 (strain 28) (Figure 2d). Furthermore, the arrangement of the blocks was completely identical in the sequences EF645709 (strain 3) and EF645718 (strain 28) as well as in the sequences EF645707, EF645714 (strain 3), EF645716 (strain 28), EF645702, EF645704 (strain 36A).

In conclusion, no significant differences were found in the respective sequences of different strains of the same species, and some sequences of different strains possessed completely identical arrangement of the structural blocks.

3.5. Species *vs.* species

In order to test interspecific variability of ITS sequences in the genus *Geobacillus*, the presence/absence of both universal and unique blocks as well as the arrangement of these blocks were compared in different species of geobacilli.

Some unique blocks have been found to be species-specific. *G. thermodenitrificans* specific blocks TD1, TD3, TD7 (Figure 3) were queried in a BLASTN search and were found to be unique; no sequence with the significant value of similarity was found throughout the prokaryotes. TD8 and TD11 were also found in the sequence EU157953 of ITS of *G. thermodenitrificans* DSM 465^T confirming that these conserved blocks are species-specific. ST11 also appears to be species-specific as it did not show any significant similarity in BLASTN search with the sequences of prokaryotes, although was found in *G. stearothermophilus* 3, 28 and 36A (Figure 2d). Unique conserved blocks KP2 and KP3 (Figure 1a,b) as well as ST4 (Figure 2b) did not match any database entry, suggesting they are specific to *G. kaustophilus* and *G. stearothermophilus sensu stricto,* respectively.

The differences in the arrangement of the blocks were also compared in different species. The conserved block U20 was found in the sequences of *G. stearothermophilus sensu stricto, G. stearothermophilus* DSM 13240 and *G. lituanicus* DSM 15325^T (Figure 2a, 4a, 5a). In the sequences of *G. lituanicus* DSM 15325^T and *G. stearothermophilus* DSM 13240, the block U20 was located only in the sequences with tRNA genes. In contrast, only the sequences without these genes of *G. stearothermophilus sensu stricto* possessed U20. U20 was not found in the sequences of *G. thermodenitrificans* NG80-2, *G. kaustophilus* ATCC 8005^T and *G. kaustophilus* HTA426.

The block U25 was found in sequences of *G. kaustophilus* HTA426, *G. stearothermophilus* DSM 13240 and *G. lituanicus* DSM 15325^T, located directly downstream of U16-U13 in some sequences with tRNA genes of these three species (Figure 1b, 4a, 5b). The other combination U16-U13-U75 was found only in the sequences of *G. kaustophilus* HTA426 (rrnB) and *G. lituanicus* DSM 15325^T (EF645698 and EF645700) (Figure 1b, 5b). Neither the block U25 nor U75 were determined in the sequences of *G. thermodenitrificans* NG80-2 and *G. stearothermophilus sensu stricto*.

Unique blocks ST8 and TD8 were found downstream of tRNA genes in the sequences of *G. stearothermophilus sensu stricto* and *G. thermodenitrificans* NG80-2, respectively (Figure 2b, 3b). In contrast, the sequences of the other three species possessed U76 or U32: rrnA and rrnB (*G. kaustophilus* HTA426) possessed the block U76, contig 303 (*G. stearothermophilus* DSM 13240) – the block U32, and the sequences of *G. lituanicus* DSM 15325T – both U32 (EF645700) and U76 (EF645698 and EF645701) (Figure 1b, 4b, 5c).

The most intriguing results were obtained when the sequences of *G. kaustophilus* HTA426 and *G. ituanicus* DSM 15325^T were compared. The 5' regions of two sequences with tRNA genes, rrnA from *G. kaustophilus* HTA426 and EF645701 from *G. lituanicus* DSM 15325^T, were completely identical in terms of the composition and the arrangement of the blocks (Figure 1a,b, 5a,b). In the 5' end of rrnB of *G. kaustophilus* HTA426 the clusters of the blocks U24-U16-U13 and U25-U82-U26 were interspersed by the cluster U75-U9-U16-U13 (Figure 1b). This cluster was absent in the ITS of rrnA, but the completely identical cluster U75-U9-U16-U13 was found in the sequences EF645698 and EF645700 of *G. lituanicus* DSM 15325^T (Figure 5b). The 3' region of rrnA of *G. kaustophilus* HTA426 was identical with that of EF645698 of *G. lituanicus* DSM 15325^T and some sequences without tRNA genes from both species but differed from the ITS of rrnB. The latter sequence shared one cluster of the blocks with the sequences EF645700 and EF645701 of *G. lituanicus* DSM 15325^T (Figure 1d, 5d). The cluster of the blocks U76-U33-U49-U28-U52 was found downstream of tRNA genes in the sequences rrnA and rrnB of *G. kaustophilus* HTA426 as well

as in the sequences EF645698 and EF645701 of *G. lituanicus* DSM 15325^T (Figure 1b,c, 5c,d). The similarity of the sequences without tRNA genes was also surprising. The 5' and 3' end regions of EF645796 of *G. lituanicus* DSM 15325^T were completely identical with the sequence rrnI of *G. kaustophilus* HTA426 (Figure 1, 5). The only differences were detected in the central part of the sequences (Figure 1c,d, 5d).

4. Discussion

The aim of this study was to evaluate the inter- and intraspecific as well as intragenomic variability of *Geobacillus* 16S-23S rRNA internal transcribed spacers without tRNA genes and to compare the structure of these sequences with those bearing tRNA genes.

Mosaic-like structure has been previously found in the ITS sequences of some mesophilic bacteria [3-6]. Our study also demonstrated the mosaic-like structure in the ITS sequences of thermophilic bacteria of the genus *Geobacillus*. However, some characteristics of ITS of geobacilli differed from those of the other bacteria. The very short sequence block U1 (5 nt) was found in the 5' end region of *Geobacillus* ITS, which was absent in the genomes of other endospore formers and at least some other bacteria. A conserved region of 68 bp in length was common to 5' end of all the sequences of *Bacillus anthracis* and *Bacillus mycoides* [1]. Xu and Côte found that the conserved 5' end of ITS was at least 70 bp long in each strain of *Bacillus* and *Brevibacillus* examined [13]. However, variability in the arrangement of ITS of geobacilli resembled that of mesophilic endospore-formers. Sequences with tRNA genes of the latter bacteria have been shown to differ from the sequences without the genes of tRNA [1,13]. The 5' end of the sequences with and without tRNA genes in thermophilic endospore-forming *Geobacillus* differed markedly. Sequences with tRNA genes differed from the corresponding sequences without tRNA genes in the arrangement of conserved blocks. The 3' end of ITS downstream of boxA was less variable.

Recombination events are thought to be responsible for the mosaic-like structure of ITS [2,6,14]. Our study also suggested the occurrence of recombination events. For example, the sequence EF645721 (*G. stearothermophilus* 28) possessed two boxA elements. The sequence blocks U17 and U19 have been found not only in the 16S-23S rRNA intergenic spacer but also in the spacer between 23S rRNA and 5S rRNA genes. Our results suggest that blocks U17 and U19 in ITS are derived from the 23S-5S rRNA intergenic spacer. Recombination events could be responsible for such transfer. To our knowledge, this is the first report on the presence of the same sequence blocks in both 16S-23S rRNA and 23S-5S rRNA intergenic spacers. Another characteristic, previously not detected in ITS sequences of bacteria, was the presence of the same conserved blocks in both ends of ITS. It is interesting to note that the cluster U41-U42-U58-U71-U4-U7-U6 found in the 5' end of ITS was usually interrupted in the 3' end of ITS. We hypothesize that the emergence of this cluster in the 3' end is also a consequence of the recombination events.

The presence of sequence block U9 was limited only to the ITS sequences. We could not detect this block elsewhere in the completely sequenced genomes of *G. kaustophilus* HTA426 and *G. thermodenitrificans* NG80-2, although both ITS sequences with tRNA genes of *G. thermodenitrificans* NG80-2 contained this block in their 5' end. Both *G. kaustophilus* HTA426 and *G. lituanicus* DSM 15325^T possessed one ITS without U9. We suggest that U9 was horizontally transferred from *G. thermodenitrificans* to the ancestor of *G. kaustophilus-G. lituanicus* lineage. Block U9 is consistently located in close proximity to the tRNA genes. These genes may have been associated with the recombination events responsible for the mosaic-like structure of ITS in bacteria [4,7].

Regarding the phylogenetic analysis, *G. thermodenitrificans* branches separately from the *G. kaustophilus-G. lituanicus-G. stearothermophilus* lineage [10]. This species diverged early in the phylogenetic tree of the genus *Geobacillus*. Sequence analyses indicate that *G. thermodenitrificans* possesses more species-specific unique blocks than other species involved in our study. ITS sequences of this species were the most uniform in the terms of the composition and arrangement of the blocks when compared with other species of geobacilli. Moreover, ITS sequences of *G. thermodenitrificans* were mainly composed of the specific TD blocks as well as of the U blocks universally present in all tested species. The only exceptions were U9 and U10. This finding suggests that other blocks emerged in the ITS sequences during evolution of the other species of *Geobacillus.* The blocks U1-U6, U8, U11-U16 and U81 can be considered as the core sequence blocks of ITS of geobacilli.

The most intriguing results were obtained in inter- and intraspecific analysis of ITS sequences. Intraspecific variability could be evaluated for two species, *G. kaustophilus* and *G. stearothermophilus* (*sensu stricto*), and variability was determined to be low. The similarity between some ITS sequences of *G. lituanicus* DSM 15325^T and *G. kaustophilus* HTA426

(interspecific variability) was surprising and reflected recent phylogenetic events. In contrast, *G. thermodenitrificans* possessed very similar, and in some cases identical, ITS sequences, indicating intragenomic variability of this species was quite low. Gürtler [14] reported higher interspecies than intraspecies variation for ITS of enterococci, and our data were comparable with this for *G. thermodenitrificans*. For *G. kaustophilus* and *G. lituanicus*, however, we found that intragenomic variability of the ITS from some correspondent operons was higher than interspecific variability.

In conclusion, our study demonstrated the mosaic-like structure of ITS in *Geobacillus*, similar to reports for some other bacteria. However, some characteristics of these spacers of geobacilli were not previously reported for other bacteria, including unusually short conserved sequences in the 5' region, some identical conserved blocks in both 5' and 3' regions of 16S-23S rRNA internal transcribed spacers, and the same sequence blocks in both 16S-23S and 23S-5S rRNA intergenic spacers. This study demonstrated low inter- and intraspecific but high intragenomic variability for the phylogenetically recent species of the genus *Geobacillus* as a consequence of recent phylogenetic events.

Acknowledgement

This work was supported by grant of the Lithuanian State Science and Studies Foundation (project number T-05112).

References

[1] Cherif A., Borin S., Rizzi A., Ouzari H., Boudabous A., Daffonchio D., Bacillus anthracis diverges from related clades of the Bacillus cereus group in 16S-23S ribosomal DNA intergenic transcribed spacers containing tRNA genes, Appl. Environ. Microbiol., 2003, 69, 33-40

[2] Carr L.E., Gürtler V., Seviour R.J., Variation of 16S-23S rRNA intergenic spacer regions (ISRs) in Acinetobacter baylyi (strain B2) isolated from activated sludge, Syst. Appl. Microbiol., 2004, 27, 478-491

[3] Osorio C.R., Collins M.D., Romalde J.L., Toranzo A.E., Variation in 16S-23S rRNA intergenic spacer regions in Photobacterium damselae: a mosaic-like structure, Appl. Environ. Microbiol., 2005, 71, 636-645

[4] Privitera A., Rappazzo G., Sangari P., Gianninò V., Licciardello L., Stefani S., Cloning and sequencing of a 16S/23S ribosomal spacer from Haemophilus parainfluenzae reveals an invariant, mosaic-like organization of sequence blocks, FEMS Microbiol. Lett., 1998, 164, 289-294

[5] Sadeghifard N., Gürtler V., Beer M., Seviour R.J., The mosaic nature of intergenic 16S-23S rRNA spacer regions suggests rRNA operon copy number variation in Clostridium difficile strains, Appl. Environ. Microbiol., 2006, 72, 7311-7323

[6] Wenner T., Roth V., Decaris B., Leblond P., Intragenomic and intraspecific polymorphism of the 16S-23S rDNA internally transcribed sequences of Streptomyces ambofaciens, Microbiology, 2002, 148, 633-642

[7] Chun J., Huq A., Colwell R.R., Analysis of 16S-23S rRNA intergenic spacer regions of Vibrio cholerae and Vibrio mimicus, Appl. Environ. Microbiol., 1999, 65, 2202-2208

[8] Gürtler V., Barrie H.D., Typing of Staphylococcus aureus strains by PCR-amplification of variable-length 16S-23S rDNA spacer regions: characterization of spacer sequences, Microbiology, 1995, 141, 1255-1265

[9] Perez-Luz S., Fernandez J., Rodriguez-Valera F., Pascual L., Moreno C., Amo A., et al., Sequence diversity of the internal transcribed spacer (ITS) region of the rRNA operons among different serogroups of Legionella pneumophila isolates, Syst. Appl. Microbiol., 2002, 25, 212-219

[10] Kuisiene N., Raugalas J., Stuknyte M., Chitavichius D., Identification of the genus Geobacillus using genus-specific primers, based on the 16S-23S rRNA internal transcribed spacer, FEMS Microbiol. Lett., 2007, 277, 165-172

[11] Kuisiene N., Raugalas J., Chitavichius D., Comparative sequence analysis of 16S-23S rRNA internal transcribed spacers of the genus Geobacillus, Biologija, 2008, 54, 1-6

[12] Kumar S., Tamura K., Nei M., MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment, Brief. Bioinform., 2004, 5, 150–163

[13] Xu D., Côte J.-C., Phylogenetic relatioships between Bacillus species and related genera inferred from comparison of 3' end 16S rDNA and 5' end 16S-23S ITS nucleotide sequences, Int. J. Syst. Evol. Microbiol., 2003, 53, 695-704

[14] Gürtler V., The role of recombination and mutation in 16S-23S rDNA spacer rearrangements, Gene, 1999, 238, 241-252