## 16S-23S rRNA intergenic spacer region sequence variation in *Streptococcus thermophilus* and related dairy streptococci and development of a multiplex ITS-SSCP analysis for their identification

Diego Mora,<sup>1</sup> Giovanni Ricci,<sup>1</sup> Simone Guglielmetti,<sup>1</sup> Daniele Daffonchio<sup>2</sup> and Maria Grazia Fortina<sup>1</sup>

Dipartimento di Scienze e Tecnologie Alimentari e Microbiologiche, sezione Microbiologia Industriale<sup>1</sup> and sezione Microbiologia Agraria, Alimentare e Ecologica<sup>2</sup>, Università degli Studi di Milano, Milano, Italy

The 16S–23S rRNA internal transcribed spacer (ITS) region of several *Streptococcus thermophilus* strains and some related dairy streptococci, *S. macedonicus, S. salivarius* and *S. bovis*, was analysed by sequence analysis. All the *Streptococcus* species were easily discriminated on the basis of sequence variations principally located upstream and downstream of the region encompassing the double-stranded processing sites and the tRNA<sup>Ala</sup> gene. Comparison between tRNA<sup>Ala</sup> gene sequences highlighted a high level of sequence conservation among the *Streptococcus* species investigated despite their belonging to separated phylogenetic clusters, i.e. the *S. salivarius* and *S. bovis* rRNA groups. A low but significant degree of variability was detected among the *S. thermophilus* strains, allowing the identification of four different ITS sequences. Similarity analysis of the ITS sequences showed that the *Streptococcus* species were clustered in two main branches, one containing *S. macedonicus* and *S. bovis* strains, and one containing *S. thermophilus* and *S. salivarius* strains. With the aim of developing a rapid tool for the identification of the dairy streptococci species a multiplex ITS-SSCP analysis of two discrete regions within the ITS locus was carried out.

Diego Mora diego.mora@unimi.it

Correspondence

Received1August 2002Revised3December 2002Accepted9December 2002

#### INTRODUCTION

Thermophilic dairy streptococci are routinely identified as *Streptococcus thermophilus* or as '*S. thermophilus*-like' microorganisms. The latter show a similar or a slightly atypical carbohydrate fermentation pattern, compared with those of *S. thermophilus*, but are a different genotype. The taxonomic position of these anomalous strains has been clarified starting from the description of the species *Streptococcus macedonicus* and *Streptococcus waius* (Flint *et al.*, 1999; Tsakalidou *et al.*, 1998). These species are phylogenetically located in the *Streptococcus bovis* rRNA group while *S. thermophilus* is closely related to *Streptococcus salivarius*.

Within the bacterial chromosome, the ribosomal locus has often been investigated with taxonomic and/or phylogenetic purpose; the internal transcribed spacer (ITS) region located

Abbreviations: ITS, internal transcribed spacer; MIS, multiplex ITS-SSCP (analysis); SSCP, single-strand conformational polymorphism.

The GenBank accession numbers for the sequences determined in this work are given in the text.

between the 16S and the 23S rRNA genes provides an excellent tool for a finer identification at the species/strain level, while ribosomal genes may lose resolution at the lower taxa level (Rodriguez-Valera & Garcia-Martinez, 2000). ITS regions show highly conserved sequences in the areas encompassing the tRNA genes and relatively stable regions located at the ends of the spacer, where secondary structures are formed by pairing with stretches upstream of the 16S rRNA gene and downstream of the 23S rRNA gene. ITS sequence analysis has been applied successfully for the identification and differentiation of several bacterial species (Ehrenstein *et al.*, 1996; Maes *et al.*, 1997; Tilsala-Timisjarvi & Alatossova, 1997; Vaneechoutte *et al.*, 1992) and a freely available spacer database has recently been developed (Rodriguez-Valera & Garcia-Martinez, 2000).

In this study, several strains of *S. thermophilus*, *S. macedonicus* and *S. waius* isolated from various sources, such as yogurt, whey, natural cheese starter, cheese and biofilms on stainless steel surfaces of dairy manufacturing plants, were analysed by sequencing the 16S–23S rRNA ITS. The genetic variability of the 16S–23S ribosomal ITS locus

of thermophilic dairy streptococci has been investigated with the double aim of obtaining a detailed picture of this biotechnologically important group of lactic acid bacteria and finding an effective method of identifying them. The availability of rapid and simple methods for the identification of these dairy streptococci would facilitate the investigation of the microbial composition of traditional dairy products produced with raw milk without the use of selected starter cultures. In addition, a species/strain-specific multiplex ITS-SSCP analysis (MIS) of the regions upstream and downstream of the tRNA gene was carried out.

#### **METHODS**

**Bacterial strains.** All *Streptococcus* strains used in this work were routinely maintained at 4 °C after growth at 37 °C for 18 h in M17 broth (Difco). For long-term maintenance, stock cultures were stored in 20 % (v/v) glycerol, 80 % (v/v) M17 broth at -80 °C. The strains of lactic acid bacteria used in this work, their origin and some relevant characteristics are shown in Table 1.

**DNA extraction and PCR experiments.** Total bacterial DNA was extracted as previously described (Mora *et al.*, 2000) starting from 100  $\mu$ l of M17 broth culture. A DNA 16S–23S rRNA ITS region amplification was performed in a volume of 100  $\mu$ l containing: 3  $\mu$ l bacterial genomic DNA solution, 10  $\mu$ l 10 × PCR reaction buffer, 200  $\mu$ mol l<sup>-1</sup> of each dNTP, 2 mmol l<sup>-1</sup> MgCl<sub>2</sub>, 0.5  $\mu$ mol l<sup>-1</sup> of

each primer (ITSF 5'-GTCGTAACAAGGTAGCCGTA-3' and ITSR 5'-CAAGGCATCCACCGT-3') and 0.5 U *Taq* DNA polymerase (Amersham-Pharmacia). The temperature profile was the following: 5 cycles consisting of 94 °C for 1 min, 55 °C for 1 min and 72 °C for 2 min; and 30 cycles consisting of 92 °C for 45 s, 60 °C for 45 s and 72 °C for 2 min. A final extension at 72 °C for 7 min was performed. All amplification reactions were performed in a Gene Amp PCR System 2400 (Applera). Following amplification, 2 µl of product was analysed at 5 V cm<sup>-1</sup> by agarose electrophoresis (1.5% agarose gel, 0.2 µg ethidium bromide ml<sup>-1</sup>) or polyacrylamide electrophoresis (6%, acrylamide: bisacrylamide, 29:1, w/w) in TBE buffer (90 mmol l<sup>-1</sup> Tris/borate, 2 mmol l<sup>-1</sup> Na<sub>2</sub>EDTA, pH 8) and imaged under UV light.

The 16S rRNA gene was amplified as previously described (Mora et al., 2000). The amplification of the *sodA* gene was carried out in 100  $\mu$ l containing: 3 µl bacterial genomic DNA solution obtained as described above, 10  $\mu l$  10  $\times\,$  PCR reaction buffer, 200  $\mu mol \; l^{-1}$  of each dNTP, 2 mmol  $l^{-1}$  MgCl<sub>2</sub>, 0.5 µmol  $l^{-1}$  of each primer (sodAF 5'-CGATGCAGAAACAATGACATT-3', sodAR 5'-GGATTGTCTTG-GTTAGCTGT-3', and sodAR1 5'-GGAGTATCTTGGTTAGCAGT-3') and 0.5 U Taq polymerase (Amersham-Pharmacia). The primer set for the sodA gene was designed on the basis of the sequences reported by Poyart et al. (2002); sodAF, sodAR and sodAR1 were used together in the same PCR reaction to allow the amplification of sodA gene both in Streptococcus infantarius and in Streptococcus lutetiensis (previously S. infantarius subsp. coli) (Schlegel et al., 2000). The temperature profile was the following: 94 °C for 2 min and 35 cycles consisting of 94 °C for 45 s, 58 °C for 35 s and 72 °C for 50 s. A final extension at 72 °C for 7 min was performed.

Streptococcus strain Origin and relevant characteristics					
S. thermophilus	DSM 20617 <sup>T</sup>				
S. thermophilus	A1, B1, C1, D1, E1, F1, G1, H1, I1, I2, L1, N1, N2, P1, Q1; strains isolated from yogurt (Mora et al., 2002)				
S. thermophilus*	ST20, isolated from whey starter; ST24, isolated from Toma cheese; ST52, isolated from Provolone cheese; ST29, isolated from whey starter from Provolone cheese (Mora <i>et al.</i> , 2002)				
S. thermophilus†	<i>thermophilus</i> <sup>†</sup> ST32, isolated from natural cheese starter; ST35, isolated from Crescenza cheese; ST69, isolated from whey starter (Mora <i>et al.</i> , 2002)				
Streptococcus sp.	42†, 48† isolated from Grana cheese; A2.3, A2.7, A2.11, A2.9, B1.21, B1.23, B1.24, B1.18 isolated from Toma cheese				
S. macedonicus	LMG 15061, sour mash (Tsakalidou et al., 1998)				
S. macedonicus	LMG 18487, LMG 18488 <sup>T</sup> Greek Kasseri cheese (Tsakalidou <i>et al.</i> , 1998)				
S. macedonicus‡	ASG 38, ASG 62, isolated from Asiago cheese; MON30, isolated from Montasio cheese, Italy				
'S. thermophilus-like'‡	ASG 19, ASG 69, isolated from Asiago cheese, Italy				
<i>'S. thermophilus</i> -like'§	RALF 1, raw milk of Modicana cow race; RACF7, curd of Ragusano cheese; RAC2413, 'pasta filata' of Ragusano cheese; RCF15807, Ragusano cheese after 15 days of ripening				
'S. thermophilus-like'	TA3.1, TA3.4, TA3.5, TA3.7, TA3.9, TB1.4, TB1.9, TB1.12, TB1.15 isolated from Italian Toma cheese				
S. waiusl	NZRCC 21100 <sup>T</sup> , isolated from biofilm on stainless steel; T, TLST3 isolated from casein; 7c, 6/2, isolated from biofilm on stainless steel (Flint <i>et al.</i> , 1999)				
S. bovis	DSM 20480 <sup>T</sup>				
S. salivarius	DSM 20560 <sup>T</sup>				

#### Table 1. Strains tested and their origin

\*Strains kindly provided by Dr Giorgio Giraffa, Istituto Sperimentale Lattiero Caseario, Lodi, Italy.

†Strains kindly provided by Dr Mauro Scarpellini, Dipartimento di Scienze e Tecnologie Alimentari e Microbiologiche, sezione Microbiologia Agraria Ambientale e Ecologica, Università di Milano, Italy.

‡Strains kindly provided by Dr Angiolella Lombardi, Veneto Agricoltura, Thiene, Italy.

\$Strains kindly provided by Professor Sandra Torriani, Università degli Studi di Verona, Italy.

IlStrains kindly provided by Dr Steve Flint, New Zealand Dairy Research Institute, Palmerston North, New Zealand. NZRCC, New Zealand Dairy Research Institute Culture Collection, Palmerston North, New Zealand.

**ITS sequence analysis.** After amplification, ITS products were purified (NucleoSpin Extract, Machery-Nagel) and sequenced using ITSF as a sequencing primer and a model 310 automatic DNA sequencer (Applera) with fluorescent dideoxy chain terminators. ITS sequences were manually aligned with the published sequences of *Streptococcus* species. Phylogenetic reconstructions were done using Jalview software (http://www2.ebi.ac.uk/~michele/jalview/contents.html).

**SSCP analysis.** To develop a species-specific assay, the amplified ITS regions of all *Streptococcus* strains was used in SSCP analysis. Two different SSCP analyses were carried out: (i) an analysis performed on the entire ITS amplified using the primer set ITSF-ITSR, and (ii) an analysis performed on ITS-16Sf and on ITS-23Sf fragments separately amplified using the primer sets ITSF-ITSRcpR and ITSscpF-ITSR (see Fig. 3). BLASTN analysis (http://www.ncbi.nlm.-nih.gov/BLAST) carried out on ITSscpR and ITSscpF sequences showed significant matches only for ITS region of species belonging to the genus *Streptococcus*.

The SSCP analyses were performed on MDE gels prepared as described by the manufacturer (FMC Bioproducts) or on a standard polyacrylamide gel. For all SSCP analyses, 2–4 µl of each PCR product plus 5 µl gel loading solution was denatured at 100 °C for 15 min and immediately cooled on ice. After 1 min, 3 µl of the solution was analysed by MDE or polyacrylamide gel electrophoresis in TBE buffer using a Miniprotean III apparatus (Bio-Rad). The gel was run at 5 V cm<sup>-1</sup> in TBE buffer, stained in a solution containing 0·5 µg ethidium bromide ml<sup>-1</sup> and photographed in UV light. The reproducibility of the method was evaluated by using amplified fragments obtained from three independent amplifications and repeating the electrophoresis on two independent gels.

### RESULTS AND DISCUSSION

#### ITS amplification and sequence analysis

DNA was extracted from a total of 23 S. thermophilus strains, six S. macedonicus strains, five S. waius strains, and the type strains of S. bovis and S. salivarius, and used for the amplification of the 16S-23S rRNA ITS region. Length polymorphism analysis of amplified ITS regions has been used as an efficient tool for species differentiation by several authors (Mora et al., 1997; Tyrrell et al., 1997; Daffonchio et al., 2000) but it was not applicable for the differentiation of several species of dairy streptococci including those analysed in this study. A single ITS amplification product of about 350 bp was detected in S. thermophilus and S. salivarius strains using a standard agarose gel electrophoresis as previously reported by Tilsala-Timisjärvi & Alatossava (1997) and by Moschetti et al. (1998). Furthermore, (data not shown) a 350 bp amplified fragment was also detected in S. bovis, S. macedonicus and S. waius, increasing the number of Streptococcus species showing identical ITS profiles and often isolated from cheeses produced with raw milk and natural starter (Andrighetto et al., 2002; Tsakalidou et al., 1998). Identical results were obtained for all 'S. thermophiluslike' and Streptococcus sp. strains listed in Table 1.

The ITS sequence analysis carried out on *S. thermophilus* and related species showed that the length of their ITS between the 16S and the 23S rRNA genes ranged between 273 and 274 bp. All *Streptococcus* species were easily discriminated

on the basis of sequence variations principally located upstream and downstream of the region encompassing the double-stranded processing sites and the tRNA<sup>Ala</sup> gene. Comparison between tRNA<sup>Ala</sup> gene sequences highlighted a high level of sequence conservation among Streptococcus species despite their belonging to different phylogenetic clusters, i.e. the S. salivarius and S. bovis rRNA groups. With regard to S. thermophilus strains, ITS sequence comparison highlighted single insertions/deletions and substitutions that allowed the identification of four different ITS sequences (Fig. 1), designated ITS-St-I, ITS-St-II, ITS-St-III and ITS-St-IV. The sequence type ITS-St-I of 274 bp showed 100% similarity with the published ITS sequence of S. thermophilus ATCC 19987 (U32965) and was characteristic of the strains DSM 20617<sup>T</sup>, ST29, ST32, ST69, E1, Q1, G1, N2, N1, D1 and C1; strains A1, H1, P1, I1, L1, ST20, ST24 and ST35 showed the sequence type ITS-St-II of 273 bp due to an adenine deletion in position 282 (Fig. 1). A single substitution in position 284, an insertion in position 286 and a deletion in position 282 allowed the identification of the sequence ITS-St-III for the strain ST52 while the sequence type ITS-St-IV of strains B1 and I2 was characterized by two substitutions, at positions 21 and 275. Sequence ITS-St-IV showed the highest similarity with the ITS sequence of S. salivarius DSM  $20560^{T}$  (Fig. 2). Specifically, the substitution in position 264 (A instead of G) determined the lack of one of the two HaeIII restriction sites that were reported to be species-specific for S. thermophilus strains by Moschetti et al. (1998). These authors also reported the presence of an anomalous S. thermophilus strain showing a HaeIII ITS restriction profile identical to that obtained for S. salivarius but characterized by typical S. thermophilus phenotypic traits. As substantiated by our sequence data, HaeIII restriction analysis highlights an ITS sequence polymorphism in S. thermophilus but fails to differentiate this species from the closely related S. salivarius.

A discrete level of ITS sequence variation was also observed between *S. salivarius* DSM 20560<sup>T</sup> and the ITS sequence of *S. salivarius* ATCC 13419 obtained from the EMBL/ GenBank database.

While S. thermophilus strains showed ITS sequence polymorphism that allowed the identification of four different ITS alleles, an extremely high level of sequence conservation (100% sequence similarity) was detected among the S. macedonicus and S. waius strains. This high level of sequence similarity was unexpected considering the different geographical origin of the strains, i.e. New Zealand and Greece. Furthermore, ITS sequencing did not allow differentiation between S. macedonicus and S. waius strains according to our previous observation based on genotypic and phenotypic characterization which provided evidence that S. macedonicus and S. waius should be considered synonyms (Manachini et al., 2002). ITS sequence variation was also detected within the S. bovis species, confirming the genetic heterogeneity of this taxon (Farrow et al., 1984; Poyart et al., 2002).

		1		tRNA <sup>Ala</sup>	120		
C .	thermophilus DSM 20617 <sup>T</sup> (ITS-St-I		AGT TTOTTATTY	ATTTTGAGAGGTCTTGTGGGGCCTTAGCTCAGCT	CARAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
s ·	thermophilus Al (ITS-St-II)	CTAAGGAAAAA-CGGAATGTACTTC	AGT TTOTTATT	AGTTTTTGAGAGGTCTTGTGGGGGCCTTAGCTCAGCTC	CGAGAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
S.	thermophilus ST52 (ITS-St-III)	CTAAGGAAAAA-CGGAATGTACTTC	AGT TTCTTATTT	AGTTTTGAGAGGTCTTGTGGGGGCCTTAGCTCAGCT	KGAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
S.	thermophilus I2 (ITS-St-IV)	CTAAGGAAAAA-CGGAATGTGCTTC	AGT TTCTTATTT	AGTTTTGAGAGGTCTTGTGGGGCCTTAGCTCAGCTC	KGAGAGCCCCCCCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
s. :	salivarius DSM 20560 <sup>T</sup>	CTAAGGAAAAA - CGGAATGTACTTC	AGT TTCTTATTT	AGTTTTGAGAGGTCTTGTGGGGGCCTTAGCTCAGCTC	GGAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
s	salivarius ATCC 13419	CTAAGGAAAA CGGAATGTACTTO	AGT TTCTTATTT	AGTTTTGAGAGGTCTTGTGGGGGCCTTAGCTCAGCTC	365AGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
<i>s</i> .	bovis DSM 20480 <sup>T</sup>	CTAAGGATAAA - CGGAA - GCACGTT	TTGGA GTATTGTTT	AGTTTTGAGAGGTCTTGTGGGGCCTTAGCTCAGCTC	GAGAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
S. 1	macedonicus LMG 18488 <sup>T</sup>	CTAAGGAAAAAACGGAA-GCACGTI	TGGGAAGTCTTGTTT	AGTTTTGAGAGGTCTTGTGGGGCCTTAGCTCAGCTC	GGAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
<i>s</i> . i	bovis ATCC 27960	AAGGATAAA-CGGAA-GCACGT	TTGGGTATTGTTT	AGT1TTGAGAGGTCTTGTGGGGGCCTTAGCTCAGCTC	3GGAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCC		
Str	eptococcus sp.42	CTAAGGATAAA-GGGAA-GCACGTTTGGGTATTGTTTAGTTTTGAGAGGTCTTGTGGGGCCTTAGCTCAGCTGGGAGAGCGCCTGCTTTGCACGCAGGAGAGTCAGCGGTTCGATCCC					
				ITSscpR			
	12	1 tRNA			240		
<i>s</i> .	thermophilus DSM 20617 <sup>T</sup> (ITS-St-I	) GCTAGGCTCCATTGAATCGAAAGAT	TTCAAGT ATTGTCCI	ATTGAAAATTGAATATCTATATCAAATTCCATATGI	-AAGTAATT-ACATATAGATAGTAACAAGAAAATAAACCGAAA-		
<i>s</i> .	thermophilus A1 (ITS-St-II)	GCTAGGCTCCATTGAATCGAAAGAT	TCAAGT ATTGTCC	ATTGAAAATTGAATATCTATATCAAATTCCATATGI	- AAGTAATT-ACATATAGATAGTAACAAGAAAATAAACCGAAA-		
s.	thermophilus ST52 (ITS-St-III)	GCTAGGCTCCATTGAATCGAAAGAT	TTCAAGT ATTGTCC	ATTGAAAATTGAATATCTATATCAAATTCCATATG1	:-AAGTAATT-ACATATAGATAGTAACAAGAAAATAAACCGAAA-		
s.	thermophilus I2 (ITS-St-IV)	GCTAGGCTCCATTGAATCGAAAGAT	TCAAGT ATTGTCC	ATTGAAAATTGAATATCTATATCAAATTCCATATG7	-AAGTAATT-ACATATAGATAGTAACAAGAAAATAAACCGAAA-		
<i>s</i> .	salivarius DSM 20560 <sup>T</sup>	GCTAGGCTCCATTGAATCGAAAGAT	TCAAGT ATTGTCC	ATTGAAAATTGAATATCTATATCAAATTCCATATGI	-AATTAAAT-ACATATAGATAGTAACAAGAAAATAAACCGAAA-		
s.	salivarius ATCC 13419	GCTAGGCTCCATTGAATCGAAAGAT	TCAAGT ATTGTCC	ATTGAAAATTGAATATCTATATCAAATTCCATATGI	-AATTACTT-ACATATAGATAGTAACAAGAAAATAAACCGAAA-		
<i>s</i>	bovis DSM 20480 <sup>T</sup>	GCTAGGCTCCATTGAATCGAAAGAT	TCAAA GATTGTCC	ATTGAAAATTGAATATCTATATCAAATTCCACGATI	CAAGAAATTGAATTGTAGATAGTAACAAGAAA-TAAACCGAAA-		
<i>S</i> . 1	macedonicus LMG 18488 <sup>T</sup>	GCTAGGCTCCATTGAATCGAAAGAT	TCAAGTTGATTGTCC	ATTGAAAATTGAATATCTATATCAAATTCCACAATI	AGAAATAATTGTAGAAAGTAACAAGAAA-TAAACCGAAA-		
<i>s</i> . 1	bovis ATCC 27960	GCTAGGCTCCATTGAATCGAAAGAT	TCAAA GATTGTCCI	<b>TTGAAAATTGAATATCTATATCAAATTCCACGAT</b>	CAAGAAATTGAATTGTAGATAGTAACAAGAAA-TAAACCGAAAG		
Str	eptococcus sp.42	GCTAGGCTCCATTGAATCGAAAGAT	TCAAGT ATTGTCC	ATTGAAAATTGAATATCTATATCAAATTCCATATGI	-AAGTAATT-ACATATAGATAGTAACAAGAAAATAAACCGAAA-		
		***	*****	******************************	* * * * **** ************		
	24	11 ITSscpF		286			
<i>s</i> .	thermophilus DSM 20617 <sup>T</sup> (ITS-St-I	) CGCTGTGAATATTTAATGAGTTA	GGTCGAAAGGCCAAA	ATAA-			
s.	thermophilus Al (ITS-St-II)	CGCTGIGAATATTTAATGAGTT	GGTCGAAAGGCCAAA	A-TAA-			
s.	thermophilus ST52 (ITS-St-III)	CGCTGTGAATATTTAATGAGTT#	GGTCGAAAGGCCAAA	A-TTAA			
s.	thermophilus I2 (ITS-St-IV)	CGCTGTGAATATTTAATGAGTTA	GGTCGAAAGACCAAA	ATAA~			
<i>s</i>	salivarius DSM 20560 <sup>T</sup>	CGCTGTGAATATTTAATGAGTT AGGTCGAAAGACCAAAA - TAA -					
<i>s</i> .	salivarius ATCC 13419	CGCTGTGAATATTTAATGAGTTA	AGGTCGAAAGACCAAA	AATAA -			
<i>s</i> . 1	bovis DSM 20480 <sup>T</sup>	CGCTGTGATTTANTGAGITTANGGTCAACAGA-CCAAAATAA-					
S. 1	macedonicus LMG 18488 <sup>T</sup>	CGCTGTGATTTAATGAGTTTAAGGTCAGAAGA-CCAAAATAA-					
<i>s</i> . 1	bovis ATCC 27960	CGCTGTGA TTTAATGAGTTTAA	GGTCAACAGAACCAA	AATAA -			
Str	eptococcus sp.42	CGCTGTGAATATTTAATGAGTTA	AGGTCGAAAGGCCAAA/	AATAA -			
	· · · · · · · · · · · · · · · · · · ·						

**Fig. 1.** 16S–23S rRNA ITS sequence alignment of *Streptococcus* species. The tRNA<sup>Ala</sup> gene is delimited by arrows. Target sequences of primers ITSscpF and ITSscpR are underlined. The ITS sequence types detected in *S. thermophilus* strains are indicated in parentheses after the strain names. The ITS sequence accession numbers are the following: *S. thermophilus* sequence type ITS-*St*-I (U32965), *S. thermophilus* sequence type ITS-*St*-II (AJ439456), *S. thermophilus* sequence type ITS-*St*-II (AJ439456), *S. thermophilus* sequence type ITS-*St*-II (AJ439456), *S. thermophilus* sequence type ITS-*St*-IV (AJ439457), *S. salivarius* DSM 20560<sup>T</sup> (AJ439458), *S. salivarius* ATCC 13419 (X83760), *S. bovis* DSM 20480<sup>T</sup> (AJ439459), *S. bovis* ATCC 27960 (U39766), *Streptococcus* sp. 42 (AJ439460), *S. macedonicus/waius* (AF088899).

The phylogenetic analysis based on ITS sequences (Fig. 2) showed that the dairy *Streptococcus* species investigated clustered into two main branches, one grouping *S. macedonicus* and *S. bovis*, and one grouping *S. thermophilus* and *S. salivarius* strains. Further investigation should be carried out to verify if the phylogenetic diversity of *S. thermophilus* strains may reflect some relevant phenotypic traits of this species.

#### **ITS-SSCP** analysis

On the basis of the ITS sequence polymorphism detected, SSCP analyses on the entire ITS, and on a discrete region within the ITS locus, were carried out with the aim of developing a species-specific assay. The SSCP analysis of the entire amplified ITS failed to differentiate streptococcal species (data not shown) due to the high molecular mass of the amplified fragment (about 350 bp) and the limited level of sequence divergence present among the species analysed. Two new sets of primers were then designed for the amplification of the regions upstream and downstream of the highly conserved tRNA<sup>Ala</sup> gene. The new experimental design focused on the amplification of the most variable regions within the ITS locus by the exclusion of the conserved tDNA (Fig. 3). The amplified fragments obtained from the region upstream and downstream of the tRNA<sup>Ala</sup> gene, designated ITS-16Sf and ITS-23Sf, showed dimensions of about 110 bp and 190 bp respectively, according to the sequence data. The SSCP analysis of ITS-16Sf and ITS-23Sf showed better discriminatory power than the use of the entire ITS. Comparison of the SSCP patterns of the ITS-16Sf and ITS-23Sf fragments showed the latter region to be the most effective in addressing the identification at species level, allowing the differentiation of most of the Streptococcus species analysed. Interestingly, the ITS-23Sf fragment generated three single-strand electrophoretic signals, suggesting sequence variation, not detected in the sequence analysis, in the ITS region downstream of the



Fig. 2. Neighbour-joining tree based on the 16S-23S rRNA ITS sequences, showing relationships among *Streptococcus* species. The ITS sequence types detected in *S. thermophilus* strains are indicated in parentheses after the strain names.

tRNA<sup>Ala</sup> gene present in the different copies of the ribosomal operon of the bacterial chromosome. *Streptococcus* sp. 42 and 48 showed an ITS-23Sf SSCP pattern identical to that obtained for all *S. thermophilus* strains except B1 and I2. The SSCP of the ITS-16Sf fragment allowed the identification of *S. macedonicus* strains but failed to differentiate between *S. salivarius* and *S. thermophilus*. Furthermore, an identical ITS-16Sf SSCP profile was shown by *S. bovis* DSM 20480<sup>T</sup> and *Streptococcus* sp. 42 and 48 (data not shown).

An efficient *Streptococcus* species identification was achieved by MIS analysis of ITS-16Sf and ITS-23Sf fragments as shown in Fig. 4. All the species analysed showed unique SSCP profiles and a strain-specific SSCP profile was detected for *S. thermophilus* I2 which easily discriminated it from the other *S. thermophilus* strains by the single-strand conformation polymorphism generated from the ITS-23Sf fragment. The species-specific signature profiles were identified in the SSCP profile of the ITS-16Sf and/or ITS-23Sf fragments depending on the species. *Streptococcus thermophilus* strains were differentiated from *S. salivarius* by the SSCP profile generated from the ITS-23Sf fragment and from *Streptococcus* sp. 42 and 48 by the SSCP profile generated from the ITS-16Sf fragment, while *S. macedonicus* strains were discriminated from *S. bovis* and *S. thermophilus* by the SSCP profile of both the amplified fragments (Fig. 4). In conclusion, the MIS approach, designed with the aim of reducing the influence of the highly conserved tDNA<sup>Ala</sup>, is effective for the identification of dairy and non-dairy streptococcal species.



Fig. 3. Schematic representation of the 16S-23S rRNA ITS of dairy streptococci and position of the primers used in the MIS approach.



**Fig. 4.** MIS profiles of amplified ITS-16Sf and ITS-23Sf fragments of dairy streptococci. The molecular species present in the SSCP analysis are indicated by arrows. ds, double strand of ITS-16Sf or ITS-23Sf amplified fragments; ss, single strand of ITS-16Sf or ITS-23Sf amplified fragments. M, DNA molecular mass marker 50 bp ladder (Amersham Biosciences).

# Application of MIS analysis for the identification of taxonomically undefined *Streptococcus* strains

With the aim of verifying the efficacy of MIS analysis as an identification tool for dairy streptococci, all '*S. thermophilus*-like' and *Streptococcus* sp. strains listed in Table 1 isolated from Italian traditional cheese were tested. The results obtained by MIS analysis were confirmed by direct sequencing of the entire ITS and when necessary by sequence analysis of the 16S rRNA and *sodA* genes.

Among 25 Streptococcus strains, 8 were identified as S. thermophilus (A2.3, A2.7, A2.11, A2.9, B1.21, B1.23, B1.24, B1.18) while 15 were identified as S. macedonicus (ASG19, ASG69, RALF 1, RACF7, RAC2413, RCF15807, TA3.1, TA3.4, TA3.5, TA3.7, TA3.9, TB1.4, TB1.9, TB1.12, TB1.15), underlining the wide presence of this species in dairy products. The taxonomic position of Streptococcus sp. 42 and 48, which showed a MIS profile different from the other strains tested, was investigated by sequencing the 16S-23S rRNA ITS, the 16S rRNA gene and a discrete region from the manganese-dependent superoxide dismutase gene (sodA). Interestingly, Streptococcus sp. strains 42 and 48 isolated from cheese showed an ITS sequence with 100% similarity to S. thermophilus strains in the region downstream of the tRNA<sup>Ala</sup> gene and 99 % similarity with *S. bovis* DSM  $20480^{T}$  in the region upstream of the tRNA<sup>Ala</sup> gene (Fig. 1). The phylogenetic analysis based on ITS sequences showed that strains 42 and 48 clustered in the S. bovis/S. macedonicus branch (Fig. 2). Further analysis was carried out by sequencing the first 600 bp of the 16S rRNA gene of strain 42 (AJ439568). The sequence obtained was compared with those present in the EMBL database using the BLASTN service and the results obtained showed 100% identity with the small ribosomal gene of S. lutetiensis (previously *S. infantarius* subsp. *coli*) HDP90246<sup>T</sup> (AF429763) (Poyart et al., 2002) and S. bovis ATCC 27960 (AB002481), and 99% identity with S. infantarius HDP 90056<sup>T</sup> (AF429762). Because of the high sequence similarity existing between the 16S rRNA genes of S. lutetiensis, S. infantarius and S. bovis, the identification of strains 42 and 48 was subsequently addressed by sequencing a discrete region of the *sodA* gene as recently suggested by Poyart *et al.* (2002). Specifically, 268 bp of the coding region of the amplified sodA gene (AJ439567) was compared with the sequences present in the EMBL database; the results obtained showed 99% identity with the sodA gene sequence of S. infantarius CIP 103233<sup>T</sup> (AJ297184), while lower levels of identity (88%, 86%) were obtained with the homologous gene of S. lutetiensis CIP56.23 (AJ297212) and S. bovis ATCC 33317 (Z95896) respectively. According to the sodA sequence analysis, strains 42 and 48 were therefore classified as S. infantarius. This species has been isolated from infant faeces, human clinical samples and dairy products (Schegel et al., 2000; Poyart et al., 2002) - as for strains 42 and 48, which were isolated from Italian Grana cheese. In this context, the reason for the presence of S. infantarius in dairy products should be investigated with the aim of clarifying if this species has a technological role in milk transformation or if it is the result of occasional contamination.

#### ACKNOWLEDGEMENTS

This work was supported by a grant of the Ministry of the University and Technological and Scientific Research (FIRST 2001). We thank Dr Giorgio Giraffa, Dr Mauro Scarpellini, Dr Angiolella Lombardi, Professor Sandra Torriani and Dr Steve Flint for providing streptococcal strains.

#### REFERENCES

Andrighetto, C., Borney, F., Barmaz, A., Stefanon, B. & Lombardi, A. (2002). Genetic diversity of *Streptococcus thermophilus* strains isolated from Italian traditional cheeses. *Int Dairy J* **12**, 141–144.

**Daffonchio**, **D.**, **Cherif**, **A. & Borin**, **S. (2000)**. Homoduplex and heteroduplex polymorphisms of the amplified ribosomal 16S-23S internal transcribed spacers describe genetic relationships in the "*Bacillus cereus* group". *Appl Environ Microbiol* **66**, 5460–5468.

Ehrenstein, B., Bernards, A. T., Dijkshoorn, L., Gerner-Smidt, P., Towner, K. J., Bouvet, P. J., Daschner, F. D. & Grundmann, H. (1996). *Acinetobacter* species identification by using tRNA spacer fingerprinting. *J Clin Microbiol* **34**, 2414–2420.

Farrow, J. A. E., Kruze, J., Phillips, B. A., Bramley, A. J. & Collins, M. D. (1984). Taxonomic studies on *Streptococcus bovis* and *Streptococcus equinus*: description of *Streptococcus alactolyticus* sp. nov. and *Streptococcus saccharolyticus* sp. nov. *Syst Appl Microbiol* 5, 467–482.

Flint, S. H., Lawrence, J., Ward, H. & Brooks, J. D. (1999). *Streptococcus waius* sp. nov., a thermophilic streptococcus from a biofilm. *Int J Syst Bacteriol* **49**, 759–767.

Maes, N., De Gheldre, Y., De Ryck, R., Vaneechoutte, M., Meugnier, H., Etienne, J. & Struelens, M. J. (1997). Rapid and accurate identification of *Staphylococcus* species by tRNA intergenic spacer length polymorphism analysis. *J Clin Microbiol* 35, 2477–2481.

Manachini, P. L., Flint, S. H., Ward, L. J. H., Kelly, W., Fortina, M. G., Parini, C. & Mora, D. (2002). Comparison between *Streptococcus macedonicus* and *S. waius* strains and reclassification of *Streptococcus waius* (Flint *et al.*, 1999) as *Streptococcus macedonicus* (Tsakalidu *et al.*, 1998). *Int J Syst Evol Microbiol* 52, 945–951.

Mora, D., Fortina, M. G., Nicastro, G., Parini, C. & Manachini, P. L. (1997). Genotypic characterization of thermophilic bacilli: a study on

a new soil isolates and several reference strains. *Res Microbiol* 149, 711–722.

Mora, D., Fortina, M. G., Parini, C., Ricci, G., Gatti, M., Giraffa, F. & Manachini, P. L. (2002). Genetic diversity and technological properties of *Streptococcus thermophilus* strains isolated from dairy products. *J Appl Microbiol* **93**, 278–287.

Moschetti, G., Blaiotta, G., Aponte, I., Catzeddu, P., Villani, F. & Coppola, S. (1998). Random amplified polymorphic DNA and amplified ribosomal DNA spacer polymorphism – powerful methods to differentiate *Streptococcus thermophilus* strains. *J Appl Microbiol* **85**, 25–36.

**Poyart, C., Quesne, G. & Trieu-Cout, P. (2002).** Taxonomic dissection of the *Streptococcus bovis* group by sequencing the manganese-dependent superoxide dismutase gene (*sodA*) sequences: reclassification of *Streptococcus infantarius* subsp. *coli*' as *Streptococcus lutetiensis* sp. nov. and of *Streptococcus bovis* biotype II.2 as *Streptococcus pasteurianus* sp. nov. *Int J Syst Evol Microbiol* **52**, 1247–1255.

Rodriguez-Valera, F. & Garcia-Martinez, J. (2000). Spacers online. ASM News 66, 712–713.

Schlegel, L., Grimont, F., Collins, M. D., Régnault, B., Grimont, P. A. D. & Bouvet, A. (2000). *Streptococcus infantarius* sp. nov., *Streptococcus infantarius* subsp. *infantarius* subsp. nov. and *Streptococcus infantarius* subsp. *coli* subsp. nov., isolated from humans and food. *Int J Syst Evol Microbiol* **50**, 1425–1434.

Tilsala-Timisjärvi, A. & Alatossava, T. (1997). Development of oligonucleotide primers from the 16S-23S rRNA intergenic sequences for identifying different dairy and probiotic lactic acid bacteria by PCR. *Int J Food Microbiol* **35**, 49–56.

Tsakalidou, E., Zoidou, E., Pot, B., Wassill, L., Ludwig, W., Devries, L. A., Kalantzopoulos, G., Schleifer, K. H. & Kersters, K. (1998). Identification of streptococci from Greek Kasseri cheese and description of *Streptococcus macedonicus* sp. nov. *Int J Syst Bacteriol* **48**, 519–527.

Tyrrell, G. J., Bethune, R. N., Willey, B. & Low, D. (1997). Species identification of enterococci via intergenic ribosomal PCR. *J Clin Microbiol* 35, 1054–1060.

Vaneechoutte, M., Rossau, R., De Vos, P. & 7 other authors (1992). Rapid identification of bacteria of the Comamonadaceae with amplified ribosomal DNA-restriction analysis (ARDRA). *FEMS Microbiol Lett* **93**, 227–234.