



**HAL**  
open science

## Physical and Genetic Map of *Streptococcus thermophilus* A054

Yvonne Roussel, Mireille Pebay, Gérard Guédon, Jean-Marc Simonet, Bernard Decaris

► **To cite this version:**

Yvonne Roussel, Mireille Pebay, Gérard Guédon, Jean-Marc Simonet, Bernard Decaris. Physical and Genetic Map of *Streptococcus thermophilus* A054 . *Journal of Bacteriology*, 1994, 176 (24), pp.7413 - 7422. 10.1128/jb.176.24.7413-7422.1994 . hal-01659598

**HAL Id: hal-01659598**

**<https://hal.univ-lorraine.fr/hal-01659598>**

Submitted on 8 Dec 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Physical and Genetic Map of *Streptococcus thermophilus* A054

YVONNE ROUSSEL, MIREILLE PEBAY, GERARD GUEDON, JEAN-MARC SIMONET,†  
AND BERNARD DECARIS\*

Laboratoire de Génétique et Microbiologie, Unité Associée Institut National de la Recherche Agronomique,  
Université Henri Poincaré Nancy I, 54506 Vandoeuvre-lès-Nancy Cedex, France

Received 23 May 1994/Accepted 28 September 1994

The three restriction endonucleases *Sfi*I, *Bss*HIII, and *Sma*I were found to generate fragments with suitable size distributions for mapping the genome of *Streptococcus thermophilus* A054. A total of 5, 8, and 24 fragments were produced with *Sfi*I, *Bss*HIII, and *Sma*I, respectively. An average genome size of 1,824 kb was determined by summing the total fragment sizes obtained by digestions with these three enzymes. Partial and multiple digestions of genomic DNA in conjunction with Southern hybridization were used to map *Sfi*I, *Bss*HIII, and *Sma*I fragments. All restriction fragments were arranged in a unique circular chromosome. Southern hybridization analysis with specific probes allowed 23 genetic markers to be located on the restriction map. Among them, six *rrn* loci were precisely located. The area of the chromosome containing the ribosomal operons was further detailed by mapping some of the *Apa*I and *Sgr*AI sites. Comparison of macrorestriction patterns from three clones derived from strain A054 revealed two variable regions in the chromosome. One was associated with the tandem *rrnD* and *rrnE* loci, and the other was mapped in the region of the lactose operon.

*Streptococcus thermophilus* is an economically important lactic acid bacterium used in the manufacture of yoghurt and some hard cheeses. It is a gram-positive, facultative anaerobic microorganism that is exclusively isolated from milk and fermented dairy products. Despite its importance in the dairy industry, genetic studies of *S. thermophilus* have been limited until recently. To date, basic methods of gene transfer by transformation, transduction, or conjugative mobilization have been reported (34). Several genes have been cloned and analyzed, including genes involved in glucide metabolism (*lacZ*, *lacS*, *galE*, *galM*, and *ldh* [5, 25, 43, 44, 51]) and DNA recombination (*recA* [15]), ribosomal sequences (4, 23, 38), and the aspartate racemase gene (62). A marked restriction polymorphism has been detected between different strains of this species by analysis of hybridization patterns obtained with a species-specific probe (13) and with probes by genes coding for rRNA (38). Despite the availability of gene transfer systems and cloned genes, the structure and organization of the *S. thermophilus* genome are not well defined.

Pulsed-field gel electrophoresis (PFGE) (9, 52) has proved to be an efficient method for genome size estimation and construction of chromosomal maps, as well as being a useful tool for the characterization of bacterial species. Preliminary results from PFGE have indicated a genome size of 1.75 Mb for strain ST1 of *S. thermophilus* (31) and allowed restriction polymorphism within *S. thermophilus* to be observed (48).

In this study, we used PFGE to estimate the size of the chromosome of *S. thermophilus* A054 and to construct a physical map by using the restriction enzymes *Sma*I, *Sfi*I, *Bss*HIII, *Apa*I, and *Sgr*AI. Hybridization probes were used to locate 23 genetic markers on the chromosomal map. Genetic polymorphism within strain A054 was analyzed by comparing the macrorestriction patterns of three different clones of this

strain. This allowed us to detect and locate two variable regions on the chromosomal map of strain A054.

### MATERIALS AND METHODS

**Bacterial strains and culture methods.** *S. thermophilus* and *Streptococcus salivarius* were shown to be closely related (*S. thermophilus* was formerly named *S. salivarius* subsp. *thermophilus*) but distinct enough to belong to different species (4, 50). Thus, in this paper, the species was named *S. thermophilus*. The industrial strain A054 (35) was selected to establish a reference map for *S. thermophilus*. The NST2210 clone of A054 was used to construct the physical and genetic map. In order to study genetic polymorphism within the A054 strain, various clones obtained by successive subclonings from the NST2210 clone were analyzed. At each subcloning, a culture was grown from one of the isolated colonies and then streaked onto M17 (58) agar medium. Two of the subclones obtained in this manner, NST2216 and NST2221, were analyzed.

**DNA preparation.** For conventional electrophoresis, DNA was prepared according to a modified version of the Marmur procedure (13). For high-molecular-weight DNA preparations, cells were grown in M17 broth supplemented with 20 mM DL-threonine until mid-log phase (optical density at 650 nm of 0.5), when chloramphenicol was added to a final concentration of 170 µg/ml. The culture was then incubated for 1 h before harvesting the cells. High-molecular-weight DNA was then purified in agarose plugs as described previously (29), except that 70 U of mutanolysine (Sigma) per ml was used overnight instead of lysozyme.

**Restriction endonuclease digestion and gel electrophoresis.** The restriction endonucleases were purchased either from Boehringer Mannheim or New England Biolabs. For each digestion reaction, a slice of the agarose plug was rinsed with the appropriate restriction buffer and then incubated with 100 µl of the restriction buffer with 10 to 20 U of enzyme for 4 to 12 h at the required reaction temperature. For partial digestions with *Sfi*I, 10 U of the enzyme was used for 60 min. Multiple digestions were performed sequentially by washing the agarose plugs and changing the buffer to the appropriate one between incubations. Consecutive double digestions were

\* Corresponding author. Mailing address: Laboratoire de Génétique et Microbiologie, Faculté des Sciences, Université de Nancy I, Blvd. des Aiguillettes, B.P. 239, 54506 Vandoeuvre-lès-Nancy Cedex, France. Phone: (33) 83 91 20 96. Fax: (33) 83 91 25 00.

† Present address: Institut de Génétique et Microbiologie. Université de Paris-Sud, Centre National de la Recherche Scientifique Unité de Recherche Associée 1354, 91405 Orsay Cedex, France.

TABLE 1. DNA fragments used for physical and genetic mapping

Probe	Gene (function) description	Source	Reference
<b>Plasmids</b>			
pNST21	<i>purA</i> (adenylosuccinate synthetase)	<i>S. thermophilus</i> NST7	39
pEPK1	<i>lacS</i> (lactose permease)	<i>S. thermophilus</i> A147	44
	<i>galM</i> (aldose 1-epimerase)		
	<i>galE</i> (UDP glucose 4-epimerase)		
pSth1	<i>ldh</i> (L-lactate dehydrogenase)	<i>S. thermophilus</i> A054	5
pTIL45	<i>pepC</i> (aminopeptidase C)	<i>S. thermophilus</i> 302	10
pTG2215	<i>rec</i> -like (DNA recombination)	<i>S. thermophilus</i> A156	54
pVE1044	<i>recA</i> (DNA recombination)	<i>S. thermophilus</i> IL73	15
I41 <sup>a</sup>	23S gene coding for rRNA	<i>S. thermophilus</i> CNRZ368	38
I43 <sup>a</sup>	16S–23S gene coding for rRNA and upstream sequences	<i>S. thermophilus</i> CNRZ368	38
I43.1 <sup>a</sup>	16S gene coding for rRNA and upstream sequences	<i>S. thermophilus</i> CNRZ368	40
I44 <sup>a</sup>	5S gene coding for rRNA and downstream sequences	<i>S. thermophilus</i> CNRZ368	38
pNST42	Unknown (0.6-kb <i>Xho</i> I into pBC KS <sup>+</sup> )	<i>S. thermophilus</i> CNRZ368	Our laboratory
pNST49	Unknown (6-kb <i>Eco</i> RI into pBC KS <sup>+</sup> )	<i>S. thermophilus</i> CNRZ368	Our laboratory
pNST81	Unknown (6-kb <i>Eco</i> RI into pBC KS <sup>+</sup> )	<i>S. thermophilus</i> CNRZ368	Our laboratory
pNST91	Unknown (5-kb <i>Eco</i> RI into pBC KS <sup>+</sup> )	<i>S. thermophilus</i> CNRZ368	Our laboratory
pNST92	Unknown (1.7-kb <i>Eco</i> RI into pBC KS <sup>+</sup> )	<i>S. thermophilus</i> CNRZ368	Our laboratory
pNST93	Unknown (1.1-kb <i>Eco</i> RI into pBC KS <sup>+</sup> )	<i>S. thermophilus</i> CNRZ368	Our laboratory
pKR107	<i>ftf</i> (fructosyltransferase)	<i>S. salivarius</i> ATCC 25975	19
pGS102	<i>gtfJ</i> (glucosyltransferase I)	<i>S. salivarius</i> ATCC 25975	20
pR224	<i>hexA</i> (DNA reparation)	<i>S. pneumoniae</i>	45
<b>Oligonucleotides</b>			
O44	5S gene coding for rRNA, tRNA <sup>Asn</sup> gene (rRNA, tRNA <sup>Asn</sup> )	<i>S. thermophilus</i> CNRZ368	23
TG20	tRNA <sup>Val</sup> gene (tRNA <sup>Val</sup> )	<i>S. thermophilus</i> A054	23, 53
TG1847	<i>lacZ</i> ( $\beta$ -galactosidase)	<i>S. thermophilus</i> A054	54
TG2131	<i>lacS</i> (lactose permease)	<i>S. thermophilus</i> A054	54
<b><math>\lambda</math> clones<sup>b</sup></b>			
NST19	<i>gor</i> (glutathione reductase)	<i>S. thermophilus</i> CNRZ368	Our laboratory
NST20	<i>rmd-rmE</i> locus tandem (genes coding for rRNA, tRNA)	<i>S. thermophilus</i> CNRZ368	Our laboratory
NST102	Unknown	<i>S. thermophilus</i> CNRZ368	Our laboratory
NST201	Unknown	<i>S. thermophilus</i> CNRZ368	Our laboratory
NST202	Unknown	<i>S. thermophilus</i> CNRZ368	Our laboratory
NST203	Unknown	<i>S. thermophilus</i> CNRZ368	Our laboratory

<sup>a</sup> I41, I43, I43.1, and I44 are restriction fragments that were purified from DNA plasmids. A more detailed description of these fragments is given in Fig. 4A.

<sup>b</sup>  $\lambda$  clones were obtained from a genomic library of *S. thermophilus* CNRZ368 in bacteriophage  $\lambda$  GEM11 (41).

performed with individual fragments from the first digest after they had been separated by PFGE. DNA fragments contained in excised gel blocks were then digested with 50 U of enzyme.

PFGE was performed with the contour-clamped homogeneous electric field mode (12) with the Bio-Rad system (model CHEF-DR11). DNA samples were separated in 1 to 1.2% agarose gels in 0.5 $\times$  Tris-borate-EDTA buffer (55) refrigerated at 14°C throughout the experiment. Specific migration conditions are detailed in the figure legends.  $\lambda$  DNA concatemers (Bio-Rad), *Xho*I-digested  $\lambda$  DNA, *Saccharomyces cerevisiae* YNN295 chromosomes (Bio-Rad), and *S. cerevisiae* 288C chromosomes (3) were used as size markers.

To check for the existence of any fragments smaller than 30 kb, all digests were also systematically resolved by conventional gel electrophoreses. For detecting fragments between 15 and 50 kb, 0.3% agarose gels were run for 48 h at 0.7 V/cm in TAE buffer (49). For the detection of fragments in the size range 0.5 to 20 kb, 0.9% agarose gels were run for 14 h at 1.2 V/cm in TAE buffer.

Each restriction fragment was named by the initial letter(s) of the enzyme used to produce it (Sm, *Sma*I; Sf, *Sfi*I; B, *Bss*HII; A, *Apa*I; Sg, *Sgr*AI). Fragments obtained from double digestions were also named with the initial letters of the two enzymes used (BSf, *Bss*HII-*Sfi*I; SfSm, *Sfi*I-*Sma*I; BSm, *Bss*HII-*Sma*I). The fragments from each digest were numbered in order, from the largest to the smallest, except for *Apa*I and

*Sgr*AI, for which the size of the fragment is given. Multiple fragments not resolved were marked a, b, c, etc. Fragments obtained from the partial digestions were named on the basis of the enzyme used to produce them. These were preceded by the letter p (e.g., pSf for *Sfi*I), and fragment sizes were indicated by letters in descending alphabetical order from the largest to the smallest.

**Southern hybridization and genomic DNA probes.** Restriction fragments were transferred onto positively charged nylon membranes (Hybond N+; Amersham) by alkaline transfer (46). The DNA fragments used as probes (random probes and gene probes) are described in Table 1. The fragments were labeled with [ $\alpha$ -<sup>32</sup>P]dCTP with either nick translation or random priming systems (Amersham). High-stringency hybridizations were carried out as previously described (28). For low-stringency hybridizations, prehybridization and hybridization buffers contained 30% formamide and the membranes were washed in 1 $\times$  SSC (0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate at 60°C. Oligonucleotide hybridizations were performed as previously reported (23).

## RESULTS

**Selection of restriction endonucleases and determination of the A054 genome size.** Restriction endonucleases were sought

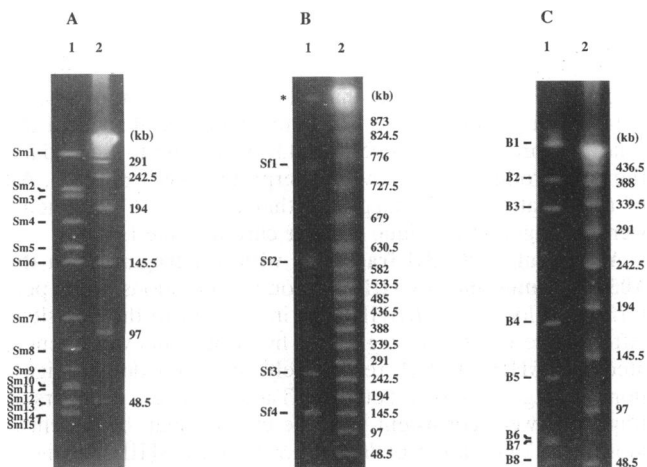


FIG. 1. PFGE separation of *Sma*I, *Sfi*I, and *Bss*HIII digestion fragments of *S. thermophilus* A054 chromosomal DNA. (A) Lane 1, *Sma*I digest; lane 2, lambda DNA concatemers. Migration parameters: 1% agarose, pulse time of 5 to 15 s at 200 V for 27 h. (B) Lane 1, *Sfi*I digest; lane 2, lambda DNA concatemers. Migration parameters: 1.2% agarose, pulse time of 60 s at 200 V for 28 h. (C) Lane 1, *Bss*HIII digest; lane 2, lambda DNA concatemers. Migration parameters: 1.1% agarose, pulse time of 5 to 30 s at 200 V for 28 h. The names of the fragments are indicated on the left of each panel, and the sizes of lambda DNA concatemers are given in kilobases on the right of each panel. The asterisk (B, lane 1) indicates the compression region, which contains partial digestion fragments.

for their ability to digest the chromosome into a small number of large fragments that could be well resolved by PFGE. In view of the low G+C content of *S. thermophilus* (37 to 40%) (17), enzymes that recognize G+C-rich sequences (*Apa*I, *Bss*HIII, *Ecl*XI, *Not*I, *Sfi*I, *Sgr*AI, and *Sma*I) were tested. Only *Sma*I, *Sfi*I, and *Bss*HIII were found to generate restriction fragments with suitable sizes for PFGE (Fig. 1 and data not shown) and therefore were chosen for mapping the genome of *S. thermophilus* A054. With these enzymes, fragments with sizes smaller than 30 kb were not observed in PFGE gels but could be detected by conventional electrophoresis (data not shown). *Sma*I, *Sfi*I, and *Bss*HIII cut the A054 DNA into 24, 5, and 8 fragments, respectively (Table 2). The average of the sum of the total fragment sizes obtained by *Sma*I, *Sfi*I, and *Bss*HIII digestions from the genome of strain A054 was estimated to be 1,824 kb. Among other restriction enzymes tested, *Not*I did not cut the DNA of *S. thermophilus* A054. Endonucleases *Apa*I and *Ecl*XI generated several large fragments, as well as many fragments too small for convenient analysis. *Sgr*AI produced a low number of fragments; however, several were of similar sizes and were not properly separated. Digestions with the enzymes *Apa*I and *Sgr*AI were used to solve ambiguities and to further detail the physical map.

**Alignment of *Sfi*I fragments in the A054 chromosome.** Partial digestions by *Sfi*I were performed to determine the order of the five *Sfi*I fragments (Fig. 2). The production of some partially digested fragments was facilitated by the different susceptibilities of the GGCCN<sub>5</sub>GGCC sites to digestion by *Sfi*I. The largest partial fragment obtained, pSfa (the size of which has been estimated to be larger than 1,700 kb), was likely to correspond to the complete, linearized chromosomal DNA.

TABLE 2. Size of restriction fragments of *S. thermophilus* A054 DNA (clone NST2210)<sup>a</sup>

<i>Sma</i> I		<i>Sfi</i> I		<i>Bss</i> HIII		<i>Bss</i> HIII- <i>Sfi</i> I		<i>Sfi</i> I- <i>Sma</i> I		<i>Bss</i> HIII- <i>Sma</i> I	
Name	Size (kb) <sup>b</sup>	Name	Size (kb)	Name	Size (kb)	Name	Size (kb)	Name	Size (kb)	Name	Size (kb)
Sm1	330 ± 5	Sf1	793 ± 7	B1	642 ± 5	BSf1	349 ± 3	SfSm1	221 ± 2	BSm1	240 ± 4
Sm2	221 ± 2	Sf2	611 ± 4	B2	383 ± 2	BSf2	261 ± 2	SfSm2	212 ± 2	BSm2	182 ± 1
Sm3	212 ± 2	Sf3	261 ± 2	B3	320 ± 2	BSf3	250 ± 3	SfSm3	172 ± 2	BSm3	158 ± 1
Sm4	182 ± 1	Sf4	144 ± 1	B4	180 ± 1	BSf4	197 ± 2	SfSm4	158 ± 1	BSm4	130 ± 1
Sm5	158 ± 1	Sf5	18.1 ± 0.3	B5	125 ± 1	BSf5	180 ± 1	SfSm5	146 ± 1	BSm5	116 ± 1
Sm6	146 ± 1			B6	68 ± 1	BSf6	134 ± 2	SfSm6	144 ± 1	BSm6	107 ± 1
Sm7	107 ± 1			B7	64 ± 1	BSf7	125 ± 1	SfSm7	107 ± 1	BSm7	96 ± 1
Sm8	85 ± 1			B8	48 ± 1	BSf8	124 ± 3	SfSm8	98 ± 1	BSm8	94 ± 1
Sm9	72 ± 1					BSf9	68 ± 1	SfSm9	85 ± 1	BSm9	85 ± 1
Sm10	61 ± 1					BSf10	64 ± 1	SfSm10	72 ± 1	BSm10	72 ± 1
Sm11	58 ± 1					BSf11	48 ± 1	SfSm11	65 ± 1	BSm11	68 ± 1
Sm12	46 ± 1					BSf12	18.1 ± 0.3	SfSm12	61 ± 1	BSm12	64 ± 1
Sm13	46 ± 1					BSf13	12.2 ± 0.3	SfSm13	58 ± 1	BSm13	61 ± 1
Sm14	38 ± 1							SfSm14	46 ± 1	BSm14	48 ± 1
Sm15	36 ± 1							SfSm15	46 ± 1	BSm15	46 ± 1
Sm16	9.2 ± 0.1							SfSm16	36 ± 1	BSm16	46 ± 1
Sm17	4.85 ± 0.03 <sup>c</sup>							SfSm17	30 ± 3	BSm17	40 ± 1
Sm18	3.34 ± 0.01							SfSm18	18.1 ± 0.3	BSm18	38 ± 1
Sm19a-f	0.91 ± 0.01 <sup>d</sup>							SfSm19	9.2 ± 0.1	BSm19	36 ± 1
								SfSm20	8.5 ± 0.1	BSm20	24 ± 1
								SfSm21	8.5 ± 0.1	BSm21	24 ± 1
								SfSm22	4.85 ± 0.03 <sup>c</sup>	BSm22	19 ± 1.7
								SfSm23	3.34 ± 0.01	BSm23	9.2 ± 0.1
								SfSm24a-f	0.91 ± 0.01 <sup>d</sup>	BSm24	5.80 ± 0.07
										BSm25	4.85 ± 0.03 <sup>c</sup>
										BSm26	3.34 ± 0.01
										BSm27a-f	0.91 ± 0.01 <sup>d</sup>

<sup>a</sup> The total sizes of *Sma*I, *Sfi*I, *Bss*HIII, *Bss*HIII-*Sfi*I, *Sfi*I-*Sma*I, and *Bss*HIII-*Sma*I were 1,821, 1,827, 1,830, 1,830, 1,815, and 1,823 kb, respectively.

<sup>b</sup> The size indicated is the average of at least four independent values and is given with the standard deviation from the mean.

<sup>c</sup> Missing bands in NST2216 and NST2222 clones.

<sup>d</sup> Multiple fragments, with the number of fragments per band determined as six for the NST2210 clone (see text).

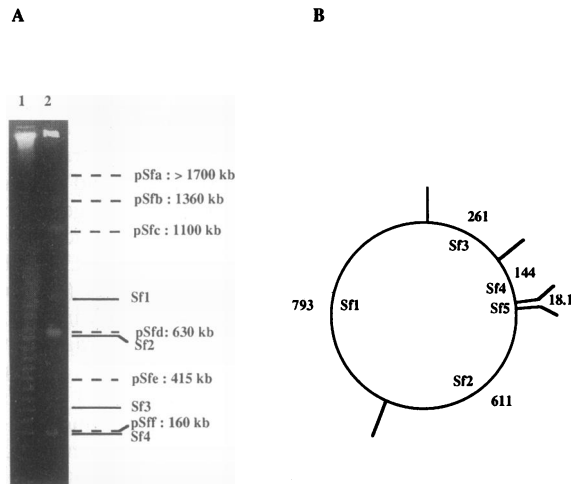


FIG. 2. PFGE separation of *SfiI* partial digestion fragments and deduced *SfiI* restriction map of the *S. thermophilus* A054 chromosome. (A) Pulsed-field agarose gel electrophoresis of *S. thermophilus* A054 chromosomal DNA digested by *SfiI*. Lane 1, lambda DNA concatemers; lane 2, partial *SfiI* digests of A054 *S. thermophilus* DNA. Migration parameters: 1% agarose, pulse time of 50 to 90 s at 200 V for 24 h. Bands that correspond to complete digestion are indicated by solid lines. Partial digestion products are indicated by dotted lines. Sizes are given in kilobases. (B) *SfiI* physical map of *S. thermophilus* A054 chromosome derived by *SfiI* partial digestion analysis. The *SfiI* fragment arrangement (Sf1-Sf2-Sf5-Sf4-Sf3) is given inside the circle, and fragment sizes are given in kilobases outside the circle.

*SfiI* partial digests showed five other partial digestion products: pSfb, pSfc, pSfd, pSfe, and pSff, with sizes of 1,360, 1,100, 630, 415, and 160 kb, respectively. By using the fragment sizes of the partial digestion products, pSfb, pSfd, and pSff were interpreted as Sf1 plus Sf2, Sf2 plus Sf5, and Sf4 plus Sf5, respectively. These data indicated that the four fragments Sf1,

Sf2, Sf5, and Sf4 were arranged in this order. The pSfe product could have resulted from the uncleaved *SfiI* site(s) between the Sf3 and Sf4 fragments or between the Sf3, Sf4, and Sf5 fragments. This indicated that Sf3 and Sf4 fragments were adjacent and that the five fragments could be aligned in the following order: Sf1-Sf2-Sf5-Sf4-Sf3. According to these data, the pSfc product could only be interpreted as Sf1 plus Sf3. All of these data strongly suggested that the five *SfiI* fragments were arranged in a unique circular chromosome (Fig. 2).

**Alignment of *BssHII* fragments with *SfiI* fragments in the A054 chromosome.** *BssHII-SfiI* double digestions were performed to locate the *BssHII* sites in relation to the *SfiI* sites. After single digestions, eight and five fragments were generated by *BssHII* and *SfiI*, respectively, and after double digestion, 13 fragments were detected (Table 2). These numbers of fragments were consistent with the circular map of the chromosome. Comigration of *BssHII*, *SfiI*, and *BssHII-SfiI* digestion products indicated that the three largest *BssHII* fragments (B1, B2, and B3) and three of the *SfiI* fragments (Sf1, Sf2, and Sf4) disappeared after double digestion, indicating that they contained *SfiI* and *BssHII* sites, respectively. Chromosomal DNA digested by *SfiI*, *BssHII*, and *BssHII-SfiI* was hybridized with single-copy probes in order to identify the overlapping fragments. All probes revealed just one fragment on each digestion profile, indicating that these were overlapping fragments on the map. The hybridization results obtained with the single-copy probes on the *SfiI*, *BssHII*, and *BssHII-SfiI* fragments are summarized in Table 3. By using hybridization data and by comparing fragment sizes obtained after double or single digestions, we were able to deduce the origin of each *BssHII-SfiI* fragment and to map all of the *BssHII* fragments in relation to the *SfiI* fragments.

**Alignment of *SmaI* fragments in the A054 chromosome.** Double digestions with *SmaI* and *SfiI* or *BssHII* were performed in order to locate the *SmaI* sites in relation to the *SfiI* or *BssHII* sites. *SfiI-SmaI* and *BssHII-SmaI* double digestions generated 29 and 32 fragments, respectively (Table 2). Comi-

TABLE 3. Hybridization of single-copy DNA probes to restriction fragments

Probes	Hybridizing fragment(s)							
	<i>SfiI</i>	<i>BssHII</i>	<i>BssHII-SfiI</i>	<i>SmaI</i>	<i>BssHII-SmaI</i>	<i>SfiI-SmaI</i>	<i>ApaI</i>	<i>SgrAI</i>
pNST21 <sup>a</sup>	Sf1	B1	BSf3	Sm5	Sm5	Sm5	A-178	Sg-108
pTG2215	Sf1	B1	BSf3	Sm5	Sm5	Sm5	A-178	Sg-28
pNST92	Sf1	B1	BSf3	Sm11	Bsm22	Sm11	A-18	Sg-62
pVE1044	Sf1	B1	BSf3	Sm13	Sm13	Sm13	A-50	Sg-52
pR224	ND <sup>b</sup>	ND	ND	Sm13	Sm13	Sm13	A-50	Sg-52
λNST102 <sup>a</sup>	Sf3	B1	Sf3	Sm8	Sm8	Sm8	ND	ND
pEPK1	Sf5	B2	Sf5	Sm1	Bsm8	Sf5	A-60	Sg-182
TG1847	Sf5	ND	ND	ND	ND	ND	ND	ND
TG2131	Sf5	ND	ND	Sm1	ND	Sf5	A-60	Sg-182
pSth1	Sf2	B2	ND	Sm9	Sm9	Sm9	ND	ND
pNST81	Sf2	B2	BSf1	Sm1	Bsm8	SfSm11	ND	ND
pNST49 <sup>a</sup>	Sf2	B2	BSf1	Sm3	Bsm5	Sm3	ND	ND
λNST202	ND	B2	ND	Sm1/Sm9	Bsm8/Sm9	SfSm11/Sm9	ND	ND
λNST201	Sf2	B3	BSf4	Sm4	Sm4	SfSm3	ND	ND
λNST203	Sf1	B3	BSf8	Sm7	Sm7	Sm7	ND	ND
pTIL45	Sf1	B4	ND	Sm2	Bsm4	Sm2	ND	Sg-221
I44 <sup>a</sup>	Sf1	B4	B4	Sm2	Bsm4	Sm2	A-10	Sg-221
λNST19	Sf1	B5	ND	Sm6	Bsm7	Sm6	A-141	Sg-183
pNST93	Sf1	B5	B5	Sm6	Bsm7	Sm6	ND	ND
pNST91	Sf1	B7	B7	Sm2	B7	Sm2	A-44	Sg-221
pNST42	Sf1	B8	B8	Sm6	B8	Sm6	ND	ND

<sup>a</sup> Probe that revealed several fragments of different intensities. For each of these probes, the most strongly revealed fragment is indicated.

<sup>b</sup> ND, not determined.

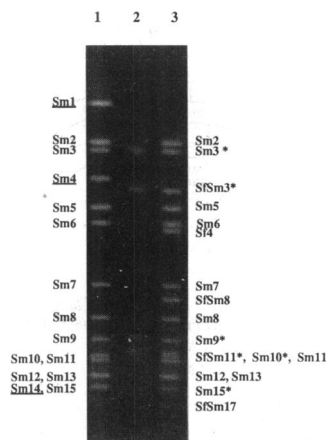


FIG. 3. PFGE separation of *Sma*I, *Sma*I-*Sfi*I, and *Sfi*I-*Sma*I digestion fragments of *S. thermophilus* A054 chromosomal DNA. Lane 1, *Sma*I digest; lane 2, *Sfi*I digested by *Sma*I; lane 3, *Sma*I-*Sfi*I digest. Migration parameters: 1.1% agarose, pulse time of 5 to 15 s at 200 V for 27 h. The names of the fragments are indicated on each side of the panel. The *Sma*I fragments that were digested by *Sfi*I are underlined to the left of the panel. The fragments that were generated after *Sma*I digestion of the *Sfi*I fragment are asterisked on the right of the panel.

gration of *Sma*I and *Sfi*I-*Sma*I digestion products indicated that the Sm1, Sm4, and Sm14 fragments disappeared in the double *Sma*I-*Sfi*I digestion, indicating that they contained at least one *Sfi*I site. In the same way, the Sm1, Sm2, Sm3, Sm6, and Sm11 fragments were found to contain at least one *Bss*HIII site. Moreover, in order to identify the *Sma*I fragments contained within the *Sfi*I or the *Bss*HIII fragments, double consecutive digestions of some isolated *Sfi*I and *Bss*HIII fragments by *Sma*I were performed. Some of the *Sfi*I fragments (Sf1, Sf2, Sf3, and Sf4 and the partially digested fragment pSfc, composed of Sf1 and Sf3) and some of the *Bss*HIII fragments (B1, B2, B3, B4, and B5) were isolated from PFGE gels, digested by *Sma*I, and electrophoresed. An example is given in Fig. 3, where the Sf2 fragment was digested by *Sma*I. The restriction fragments generated by the cleavage of the other *Sfi*I and *Bss*HIII fragments by *Sma*I are indicated in Table 4. In this experiment, the Sf5 fragment too small to be extracted from

TABLE 4. Digestion products of some of the *Sfi*I and *Bss*HIII fragments by *Sma*I (double consecutive digestions) detected by PFGE

First/second enzyme	Restriction fragment	Fragment(s) obtained after digestion by <i>Sma</i> I
<i>Sfi</i> I/ <i>Sma</i> I	pSfc	Sm2, Sm5, Sm6, Sm7, SFSm8, Sm8, Sm11, Sm12, Sm13, Sm14 <sup>a</sup>
	Sf1	Sm2, Sm5, Sm6, Sm7, Sm11, Sm13 <sup>a</sup>
	Sf2	Sm3, SFSm3, Sm9, SFSm11, Sm10, Sm15
	Sf3	SFSm8, Sm8, Sm12, SFSm17
	Sf4	Sf4
<i>Bss</i> HIII/ <i>Sma</i> I	B1	BSm1, Sm5, Sm8, Sm12, Sm13, Sm14 <sup>a</sup>
	B2	BSm5, BSm8, Sm9, Sm10, Sm15
	B3	Sm4, Sm7 <sup>a</sup>
	B4	BSm4, BSm17
	B5	BSm7 <sup>a</sup>

<sup>a</sup> The fragments obtained after the second digestion were separated by PFGE; therefore, fragments with sizes less than 30 kb were not detected, which explains why some of the SFSm and BSm fragments are missing.

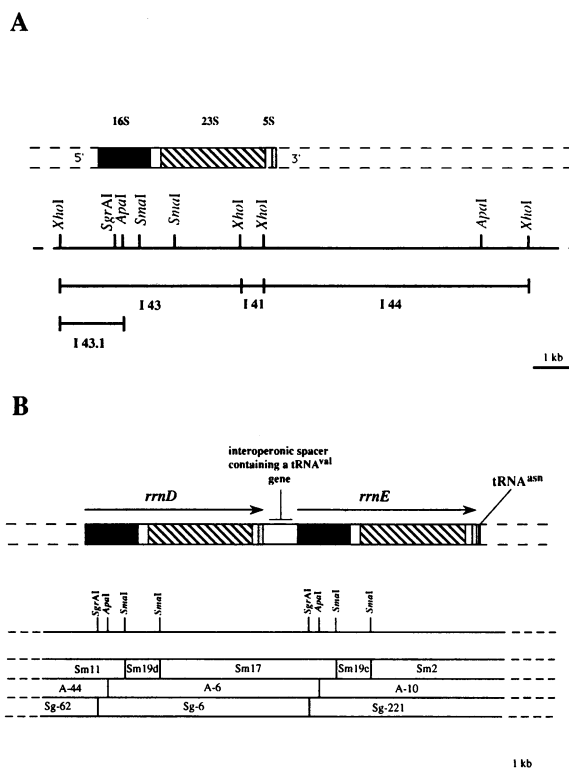


FIG. 4. Organization of *rrn* loci. The genes 16S, 23S, and 5S are shown as boxes. (A) Location of the I43, I43.1, I41, and I44 probes both on the genetic map of an *rrn* locus and on a restriction map with sites used in this study. The I43 probe is a 5-kb *Xho*I fragment containing the 5' region of an *rrn* locus. The I43.1 probe is a 1.7-kb *Apa*I-*Xho*I fragment containing the left part of a 16S gene. The I41 probe is a 0.53-kb *Xho*I fragment including the right part of a 23S gene. The I44 probe is a 7-kb *Xho*I fragment containing the left end of a 23S gene, the whole 5S gene, and downstream sequence. (B) Restriction map and genetic organization of the tandem *rrnD* and *rrnE* operons.

the PFGE gel was isolated from a conventional gel. The redigestion of this fragment by *Sma*I resulted in a fragment of 18 kb, indicating that it was not cleaved by *Sma*I. In order to identify *Sma*I fragments overlapping *Sfi*I and *Bss*HIII fragments, chromosomal DNA digested by *Sma*I, *Sfi*I-*Sma*I, and *Bss*HIII-*Sma*I was hybridized with the probes previously used for *Sfi*I, *Bss*HIII, and *Sfi*I-*Bss*HIII digests. The results are summarized in Table 3. All probes revealed one fragment on each digestion pattern, except for the  $\lambda$ NST202 probe, which contained an *Sma*I site and hybridized with the Sm1 and Sm9 fragments, indicating that they were adjacent on the chromosome. Analysis of double-digestion products and hybridization data obtained from both single and multiple digestions with *Sma*I allowed us to map most of the *Sma*I sites in relation to the *Sfi*I and *Bss*HIII sites.

Southern hybridizations with probes corresponding to the rRNA genes were performed in order to map the restriction fragments that harbor the *rrn* loci. *S. thermophilus* has been previously shown to contain rRNA genes clustered in the order 5'-16S-23S-5S-3' (38). Six copies of *rrn* loci have been detected in the A054 chromosome (23, 40). The location of the ribosomal probes (I43, I43.1, and I41) in relation to the rRNA gene map is illustrated in Fig. 4A. Although no *Sfi*I or *Bss*HIII sites were found in these probes, two *Sma*I sites were detected in the I43 probe. This probe was therefore used to detect all

TABLE 5. Hybridization of multiple-copy DNA probes (rRNA and tRNA genes) to restriction fragments

Probe	Enzyme(s)	Hybridizing fragments
I43	<i>Sma</i> I	Sm2, Sm5, Sm6, Sm11, Sm12, Sm13, Sm14, Sm17, Sm18, Sm19a-f
	<i>Bss</i> HIII	B1, B4, B5
	<i>Sfi</i> I	Sf1, Sf3
	<i>Apa</i> I	A-178, A-50, A-44, A-35, A-24, A-18, A-17, A-13, A-10, A-6
	<i>Sgr</i> AI	Sg-221, Sg-183, Sg-62, Sg-56, Sg-52a, Sg-52b, Sg13, Sg-6
I43.1	<i>Sma</i> I	Sm2, Sm5, Sm11, Sm13, Sm14, Sm17
	<i>Sma</i> I- <i>Bss</i> HIII	Sm5, Sm13, BSm17, Sm14, BSm20, Sm17
	<i>Sma</i> I- <i>Sfi</i> I	Sm2, Sm5, Sm11, Sm13, SfSm17, Sm17
	<i>Apa</i> I	A-178, A-50, A-44, A-24, A-17, A-6
	<i>Sgr</i> AI	Sg-221, Sg-62, Sg-52a, 52b, Sg-13, Sg-6
I41	<i>Sma</i> I	Sm2, Sm6, Sm11, Sm12, Sm17, Sm18
	<i>Sma</i> I- <i>Bss</i> HIII	BSm4, BSm7, Sm12, BSm22, Sm17, Sm18
	<i>Sma</i> I- <i>Sfi</i> I	Sm2, Sm6, Sm11, Sm12, Sm17, Sm18
	<i>Apa</i> I	A-50, A-35, A-18, A-13, A-10, A-6
	<i>Sgr</i> AI	Sg-221, Sg-183, Sg-62, Sg-56, Sg-52b, Sg-6
TG20	<i>Sma</i> I	Sm6, Sm11, Sm17, Sm18
	<i>Sma</i> I- <i>Bss</i> HIII	BSm7, BSm22, Sm17, Sm18
O44	<i>Sma</i> I	Sm2, Sm12
	<i>Sma</i> I- <i>Bss</i> HIII	BSm4, Sm12

restriction fragments containing the *rrn* loci. The 16S probe I43.1 and the 23S probe I41 were used to differentiate between the *Sma*I fragments located upstream of the *rrn* *Sma*I sites and those located downstream. Hybridization results obtained with the three ribosomal probes are summarized in Table 5. If all six *rrn* loci are assumed to have the same organization as the one presented in Fig. 4A, then each *Sma*I fragment revealed by I43.1 could be paired with another *Sma*I fragment revealed by I41, with an Sm19 fragment lying between all of them. For the identification of such pairs, we took into consideration the previously established assignments of *Sma*I fragments toward *Bss*HIII and *Sfi*I or other *Sma*I fragments. For example, the two fragments Sm8 and Sm12 were previously shown to be located between Sm1 and Sm14; however, they could not be ordered. Among these four fragments, only two fragments were revealed by the ribosomal probes: Sm14 by the I43.1 probe and Sm12 by I41 probe. The fragments Sm14 and Sm12 could therefore be paired. Similar analysis was done for the other *Sma*I fragments, and the following pairings were found: Sm11 with Sm17 and Sm2 with Sm6.

The linkage between Sm2 and Sm17 fragments was deduced by analysis of the  $\lambda$ NST20 clone. Restriction and hybridization analyses showed that this clone contained both the Sm17 and the I44 sequences (data not shown). The I44 sequence previously found to harbor the end of the *rrnE* locus was, however, mainly composed of a single-copy sequence (37). When the I44 probe was hybridized with the *Sma*I digest products, six fragments were revealed. The Sm2 fragment was the only one to be strongly revealed, showing that the I44 sequence belonged to the Sm2 fragment. These results demonstrated that Sm17 was adjacent to Sm2, with an Sm19 fragment lying between them.

The arrangement of four *Sma*I fragments (Sm5, Sm13, Sm16, and Sm18) required further analysis by *Apa*I digestions. This enzyme was found to cut the ribosomal I43 probe (Fig.

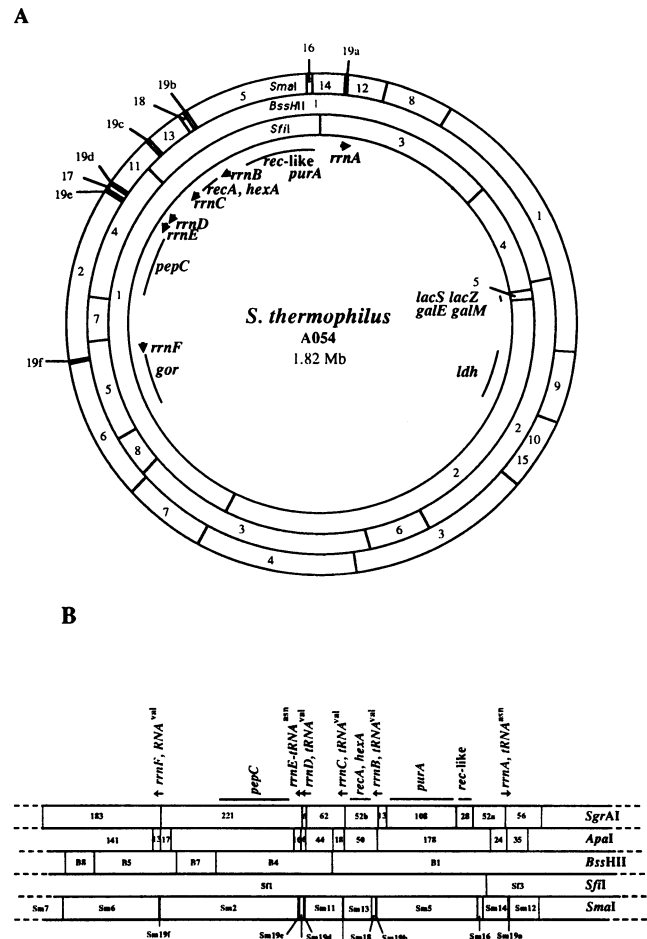


FIG. 5. Physical and genetic map of *S. thermophilus* A054. (A) Map of the A054 chromosome with *Sma*I, *Bss*HIII, and *Sfi*I fragments. The outer circle represents *Sma*I fragments, the middle circle represents *Bss*HIII fragments, and the inner circle represents *Sfi*I fragments. Restriction fragments are numbered in order of their sizes (Table 2). Fragments Sm10 and Sm15 could not be oriented. The inner arcs indicate the areas in which the genes have been located. The location and transcription direction of rRNA genes are indicated by arrows. (B) Map of the A054 chromosome with *Apa*I and *Sgr*AI fragments in the region containing the ribosomal loci. Fragments *Apa*I and *Sgr*AI have been designated according to their sizes.

4A). These four *Sma*I fragments were ordered and oriented by analyzing the hybridization of the single-copy probes pTG2215 and pVE1044 and ribosomal probes I43.1 and I41 on *Apa*I digests. Hybridization of the pVE1044, I43.1, and I41 probes with *Apa*I fragment A-50 showed that it corresponded to the Sm13 and Sm18 fragments. In the same way, hybridization of pTG2215 and I43.1 probes with fragment A-178 demonstrated that Sm5 and Sm18 were adjacent and corresponded to the A-178 fragment. These results showed that these *Sma*I fragments could be aligned in the following order: Sm11, Sm13, Sm18, Sm5, Sm16, and Sm14. The resulting physical map with *Sfi*I, *Bss*HIII, and *Sma*I sites is presented in Fig. 5A.

**Location of *Apa*I and *Sgr*AI sites in the region containing the *rrn* loci in the A054 chromosome.** In the region of the chromosome containing the six *rrn* loci (from fragment Sm6 to Sm12), *Sgr*AI and *Apa*I sites were mapped in relation to the *Bss*HIII, *Sfi*I, and *Sma*I sites. The resulting map is presented in Fig. 5B. First, single-copy probes previously shown to be

located within this region were hybridized on *ApaI* and *SgrAI* digest patterns in order to determine which fragments overlapped the *SmaI*, *BssHII*, and *SfiI* fragments. Results with single-copy probes are reported in Table 3. Second, probes containing ribosomal sequences were hybridized with *ApaI* and *SgrAI* digestion products. Like *ApaI*, *SgrAI* enzyme was also found to have one site in the I43 ribosomal sequence (Fig. 4A). The I43 probe was used to identify all *SgrAI* and *ApaI* fragments bearing ribosomal sequences. The restriction fragments that hybridized with ribosomal probes are given in Table 5. Hybridization of the I41 and I43.1 probes with total DNA digests of *ApaI* and *SgrAI* was then performed to identify the paired fragments. This allowed all *SgrAI* and *ApaI* sites in this region to be mapped (except for some internal *ApaI* fragments within the Sm2 fragment).

**Construction of the genetic map of the A054 chromosome.** Markers used for genetic mapping are described in Table 1. Most of the genes tested (*recA*, *rec*-like, *purA*, *gor*, *ldh*, *lacZ*, *lacS*, *galM*, *galE*, *pepC*, and genes coding for rRNA) were cloned from *S. thermophilus* strains and were mapped in the A054 chromosome by Southern hybridization experiments under high-stringency conditions. The hybridization results are given in Tables 3 and 5. The locations of these cloned sequences on the physical map are shown in Fig. 5A and B. For all of the mapped genes (except for genes coding for rRNA), the genetic markers were associated with specific restriction fragments. For the genes coding for rRNA, the presence of two *SmaI* sites within the *rm* loci allowed their precise locations and transcription direction to be deduced.

The locations of the tRNA<sup>Asn</sup> and tRNA<sup>Val</sup> genes were determined by using specific oligonucleotides. The O44 oligonucleotide, which overlaps the 3' end of the 5S rRNA gene, the 5S-tRNA<sup>Asn</sup> spacer, and two nucleotides of the tRNA<sup>Asn</sup> gene (23), was used as a probe to identify fragments containing the tRNA<sup>Asn</sup> gene associated with *rm* loci. The TG20 oligonucleotide containing the anticodon of a tRNA<sup>Val</sup> gene (23, 53) was then used as a probe to identify all fragments bearing tRNA<sup>Val</sup> genes. Hybridization results obtained with O44 and TG20 on *SmaI* and *SmaI*-*BssHII* digests are summarized in Table 5. Each of the six fragments revealed by the I41 probe was hybridized with either the O44 or the TG20 oligonucleotides. Similar results were obtained by hybridizing the two tRNA probes on *HindIII* digests (23). TG20 hybridized with the four fragments bearing *rmB*, *rmC*, *rmD*, and *rmF* loci, and O44 hybridized with the other two fragments bearing *rmA* and *rmE* loci. These results suggested that a tRNA<sup>Val</sup> gene was located downstream from the 5S rRNA gene of *rmB*, *rmC*, *rmD*, and *rmF* loci and that a tRNA<sup>Asn</sup> gene was located downstream from the 5S rRNA gene of the *rmE* and *rmA* loci. The genetic map of the region containing the tandem *rmD*-*rmE* is described in Fig. 4B.

Genes coming from other related species were also tested by Southern hybridization under low-stringency conditions. Because *S. salivarius* is closely related to *S. thermophilus* (4, 17), probes specific to the *fff* and *gffJ* genes (19, 20) of *S. salivarius* were hybridized on total DNA of A054 strain, but no hybridization signal was detected. The gene involved in DNA repair, *hexA*, cloned from *Streptococcus pneumoniae* (45), was also hybridized with *S. thermophilus* DNA. A hybridization signal was observed with the Sm13 fragment.

**Analysis of genetic polymorphism within the A054 strain.** The NST2210 clone was used as a reference to construct the physical and genetic map of *S. thermophilus*. Two other clones, NST2216 and NST2222, derived from the NST2210 clone by subcloning, were also analyzed in this study. PFGE and conventional electrophoresis with the *SmaI*, *BssHII*, *SfiI*, *ApaI*,

and *SgrAI* enzymes were also performed with the total DNA of clones NST2216 and NST2222. The NST2216 clone exhibited one difference in its restriction patterns compared with those of NST2210. The Sm17, A-6, and Sg-6 fragments were missing in the NST2216 clone. A further difference in the *BssHII* pattern was observed in which the B4 fragment of 181 kb was replaced by a smaller fragment of 175 kb in NST2216. Both the I41 and I43.1 probes hybridized with the Sm17, A-6, Sg-6, and B4 fragments, which were either deficient or modified in the NST2216 clone. These results indicated that a sequence of approximately 6 kb containing the 3' extremity of a *rm* locus and the 5' extremity of another *rm* locus had been deleted in NST2216. This deletion involved the tandem *rmD* and *rmE* loci (Fig. 4B) and resulted in a clone containing only five *rm* loci instead of the original six. This phenomenon has already been observed in *S. thermophilus* CNRZ368 (38).

Analyses of the restriction patterns of the NST2222 clone revealed two differences from those of the NST2210 clone. The first difference corresponded to a deletion of a 6-kb sequence in the tandem *rm* loci, as found in the NST2216 subclone. The second difference was observed in the *SmaI* digestion pattern, where the Sm1 fragment of 330 kb was replaced by a 325-kb fragment, suggesting a size difference of about 5 kb. This variation in size was also seen in the *SgrAI* and *ApaI* digests. In the *SgrAI* digestion pattern, the 182-kb fragment disappeared and a 176-kb fragment appeared in the subclone NST2222. In the *ApaI* digestion, a fragment of 60 kb was replaced by a fragment of 56 kb. The TG2131 probe hybridized with these fragments, indicating that the variable region was located in the region of the *lac-gal* operon.

## DISCUSSION

The chromosome size of *S. thermophilus* A054 has been estimated to be 1.82 Mb. This value is similar to the previous value of 1.75 Mb obtained by Le Bourgeois et al. (31) for the ST1 strain of *S. thermophilus*. The chromosome size of this species is small in comparison with those of some other eubacteria: *Streptomyces ambofaciens* strains at 8 Mb (29), *Anabaena* sp. at 6.4 Mb (2), and *Myxococcus xanthus* at 9.4 Mb (11). These bacteria can undergo complex cycles of morphological differentiation, facilitating their adaptation to poor or fluctuating environments. However, the values obtained with *S. thermophilus* are in the same order of magnitude as the sizes already determined for other lactic acid bacteria: for example, 2.1 to 2.2 Mb for *Streptococcus mutans* (24, 36), *Streptococcus sobrinus* (36), and *S. pneumoniae* (18); 2.0 to 2.7 Mb for *Lactococcus lactis* subsp. *lactis* (30, 31, 57, 59); 1.9 Mb for *Lactobacillus acidophilus* (47); 2.0 Mb for *Lactobacillus gasseri* (47); and 1.8 to 2.1 Mb for *Leuconostoc oenos* (14, 27). These bacteria require numerous factors for their growth, and they generally live in a constant or rich environment. The relatively small size of their chromosome could reflect the absence of functions required for respiration and survival in fluctuating ecosystems.

We have established the first chromosomal map of *S. thermophilus*. The small number of fragments produced by each of the three enzymes used and the relatively small size of the chromosome allowed us to construct a physical map by classical strategies. The data obtained with the partial and double digestions are consistent with the concept of a single and circular chromosome in the A054 strain.

The 23 genes mapped in this study were found to be clustered in two distinct regions. The first region, containing the *rm* loci (fragments B5 to Sm12), included 18 of the 23 markers, and the second region (fragments Sf5 to Sm9)



contained the other five genes. The six *rm* loci have been found clustered in a region of 540 kb, which corresponds to 30% of the chromosome. This *rm* locus grouping has also been observed in other bacteria, such as *Escherichia coli* (1), *Bacillus subtilis* (41), *L. lactis* subsp. *lactis* (30, 59), *Clostridium perfringens* (7), and *Haemophilus influenzae* (32), where the *rm* loci were located in a region representing about one-third to one-half of the chromosome. In the chromosome of A054, the transcription direction of each *rm* locus was also determined: five *rm* loci (*rmB* to *rmF*) were in an anticlockwise orientation, and the sixth (*rmA*) was in a clockwise orientation. The six *rm* loci were all transcribed away from a common region. This homogeneous divergent orientation of *rm* loci relative to a common point is also found in organisms as different as *E. coli* (16), *B. subtilis* (26), *L. lactis* subsp. *lactis* (30, 59), *C. perfringens* (7), and *H. influenzae* (32). In all of these cases, the *rm* loci are not equally distributed around the divergent point; a larger number of *rm* loci are located to one side of the divergent point. If the chromosome is bidirectionally replicated from its origin in the same direction as that of the *rm* locus transcription, as demonstrated in *E. coli* (6) and *B. subtilis* (26), the origin of replication should be located between *rmA* and *rmB*. The location of the *purA* gene in the Sm5 fragment supports this hypothesis. Indeed, in the genetic maps of *B. subtilis* 168 (41) and *E. coli* K-12 (1), the *purA* gene is located fairly close to the origin of replication.

Sequence and hybridization data with O44 and TG20 probes revealed that all six *rm* loci were associated with tRNA genes (reference 23 and this work). Sequence determination of the 5' extremity of the 5S gene and the adjacent downstream region in I44 (*rmE* locus) showed the presence of a single tRNA gene, tRNA<sup>Asn</sup>. This tRNA gene was found to be immediately followed by a transcription terminator, suggesting that this single tRNA gene was associated with the *rmE* locus (23). A similar association has also been described for an rRNA operon in *Acholeplasma laidlawii* (56). Hybridization of the O44 oligonucleotide with two fragments in *Hind*III, *Sma*I, and *Sma*I-*Bss*HII digestions (one corresponding to *rmE* and the other to *rmA*) suggested that another tRNA<sup>Asn</sup> gene is associated with the *rmA* locus. In this last instance, we did not know whether the tRNA<sup>Asn</sup> gene is a unique tRNA gene, as in the *rmE* locus, or whether it is associated with other genes in the tRNA gene cluster, as previously observed in *B. subtilis* (60) and *Lactobacillus bulgaricus* (42). Hybridization of the TG20 oligonucleotide with the four restriction fragments (corresponding to the *rmB*, *rmC*, *rmD*, and *rmF* loci) in digests obtained with high- and low-frequency cutting enzymes indicated that each of the four tRNA<sup>Val</sup> genes was located downstream from a *rm* locus. For one of them, comparison of the P20 sequence (from which TG20 was derived) with the ribosomal sequences indicated that the tRNA<sup>Val</sup> gene was immediately located downstream from the 5S gene (23). For the other three *rm* loci, we did not know whether the tRNA<sup>Val</sup> gene was associated with the *rm* loci or not. In other low-G+C gram-positive bacteria such as *B. subtilis* (21), *A. laidlawii* (56), and *Staphylococcus aureus* (22), the tRNA<sup>Val</sup> gene was shown to be the first tRNA gene of a cluster of numerous tRNA genes immediately located downstream from *rm* loci. The similarity in order of tRNA genes observed in this cluster between these bacteria showed that tRNA genes have evolved from large tRNA gene clusters in the ancestral gram-positive genome (22, 56), and this suggests that clusters of tRNA genes could be found downstream from *rmB*, *rmC*, *rmD*, and *rmF* loci in the genome of *S. thermophilus*.

When the restriction patterns of three different clones of the A054 strain were compared, two variable regions were found

within the chromosome of strain A054. The first region, mapped in the B4 fragment, contained the two closely linked loci *rmD* and *rmE*. A homologous recombination between these *rm* loci probably caused the deletion of one of the loci and the sequence lying between them, resulting in the loss of one *rm* locus. This situation has already been observed in *B. subtilis* (61) and *S. thermophilus* CNRZ368 (39, 40). The detection of a variable sequence in the *rm* gene cluster region is not surprising, because *rm* loci are large repeated elements that can serve as sites of intrachromosomal homologous recombination. Most minor intraspecies variations observed in the chromosome of *L. lactis* subsp. *lactis* (30) and *Clostridium* strains (8) have been found between *rm* loci. The second variable region of the A054 chromosome was located in the region containing the lactose operon. The mutation did not involve *rm* loci, and it probably resulted from a deletional event.

In order to assess the genomic variability at the interspecific level, it was interesting to compare the genetic organization of *S. thermophilus* with that of related species like other streptococci or with another lactic acid bacterium, *L. lactis* subsp. *lactis*. A striking similarity of *rm* loci chromosomal organization has been observed between *L. lactis* subsp. *lactis* strains (30, 59) and *S. thermophilus* A054, suggesting that the ancestor of these two species likely had this organization. These two species possess six *rm* loci that are organized in the same orientation. Moreover, the four internal *rm* loci (*rmA* to *rmD* in *L. lactis* subsp. *lactis* or *rmB* to *rmE* in *S. thermophilus*) are clustered in a region corresponding to less than 10% of the chromosome. Although the two species share a similar distribution of the *rm* loci, the genetic organization of the region containing the *rm* loci seems to be only roughly conserved. These *rm* locus regions were found to contain the genes *pepC*, *recA*, *rec*-like, and *hexA*, but the position and order of these genes were not the same in the two organisms. Likewise, the *ldh* gene was localized in the region opposite the *rm* locus region in the two species, but its chromosomal location was not exactly conserved (33). Physical and genetic maps of other streptococcal strains, *S. pneumoniae* (18) and *S. mutans* (24, 36), have also been established with PFGE. In the *S. pneumoniae* chromosome, six *rm* loci were also found, and three of them have been located on the chromosomal map. However, no similarity was obvious on the basis of the locations of the *rm* and *hexA* genes. In the *S. mutans* GS-5 chromosome, a relative proximity of the *ldh* and *gal* genes was also observed (24).

#### ACKNOWLEDGMENTS

We are grateful to A. Mercenier for supplying *S. thermophilus* A054; P. Slos and A. Mercenier for the *lacS*, *lacZ*, TG20, and *rec*-like probes; N. Bernard and J. Delcour for the *ldh* probe; P. Duwat and A. Gruss for the *recA* probe; P. M. Giffard and N. A. Jacques for the *gtfJ* and *fff* probes; B. Poolman for the *gal* probe; M. Prudhomme and J. P. Claverys for the *hexA* probe; and M. P. Chapot-Chartier and J.-C. Gripon for the *pepC* probe. We thank Lois Silk for help in revising the manuscript.

Y.R. was supported by a fellowship from the Ministère de la Recherche et de la Technologie.

#### REFERENCES

- Bachmann, B. J. 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. 54:130-197.
- Bancroft, I., C. P. Wolk, and E. V. Oren. 1989. Physical and genetic maps of the genome of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. 171:5940-5948.
- Bellis, M., M. Pagès, and G. Roizès. 1987. A simple and rapid method for preparing yeast chromosomes for pulse field gel electrophoresis. Nucleic Acids Res. 15:6749.

4. Bentley, R. W., J. A. Leigh, and M. D. Collins. 1991. Intra-genetic structure of *Streptococcus* based on comparative analysis of small-subunit rRNA sequences. *Int. J. Syst. Bacteriol.* **41**:487-494.
5. Bernard, N., and J. Delcour. Personal communication.
6. Brewer, B. J. 1988. When polymerases collide: replication and the transcription organization of the *E. coli* chromosome. *Cell* **53**:679-686.
7. Canard, B., and S. T. Cole. 1989. Genome organization of the anaerobic pathogen *Clostridium perfringens*. *Proc. Natl. Acad. Sci. USA* **86**:6676-6680.
8. Canard, B., B. Saint-Joanis, and S. T. Cole. 1992. Genomic diversity and organization of virulence genes in the pathogenic anaerobe *Clostridium perfringens*. *Mol. Microbiol.* **6**:1421-1429.
9. Carle, G. F., and M. V. Olson. 1984. Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. *Nucleic Acids Res.* **12**:5647-5664.
10. Chapot-Chartier, M. P., and J.-C. Gripon. Personal communication.
11. Chen, H., I. M. Keseler, and L. J. Shimkets. 1990. Genome size of *Mycococcus xanthus* determined by pulsed-field gel electrophoresis. *J. Bacteriol.* **172**:4206-4213.
12. Chu, G., D. Vollrath, and R. W. Davis. 1986. Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* **234**:1582-1585.
13. Colmin, C., M. Pébay, J. M. Simonet, and B. Decaris. 1991. A species-specific DNA probe obtained from *Streptococcus salivarius* subsp. *thermophilus* detects strain restriction polymorphism. *FEMS Microbiol. Lett.* **81**:123-128.
14. Daniel, P., E. de Waele, and J. N. Hallet. 1993. Optimisation of transverse alternating field electrophoresis for strain identification of *Leuconostoc oenos*. *Appl. Microbiol. Biotechnol.* **38**:638-641.
15. Duwat, P., S. D. Ehrlich, and A. Gruss. 1992. A general method for cloning *recA* genes of gram-positive bacteria by polymerase chain reaction. *J. Bacteriol.* **174**:5171-5175.
16. Ellwood, M., and M. Nomura. 1982. Chromosomal locations of the genes for rRNA in *Escherichia coli* K-12. *J. Bacteriol.* **149**:458-468.
17. Farrow, J. A. E., and M. D. Collins. 1984. DNA base composition, DNA-DNA homology and long-chain fatty acid studies on *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Gen. Microbiol.* **130**:357-362.
18. Gasc, A. M., L. Kauc, P. Barraillé, M. Sicard, and S. Goodgal. 1991. Gene localization, size, and physical map of the chromosome of *Streptococcus pneumoniae*. *J. Bacteriol.* **173**:7361-7367.
19. Giffard, P. M., C. Rathman, E. Kwan, D. W. Kwan, K. L. Bunny, S.-P. Koo, and N. A. Jacques. 1993. The *fff* encoding the cell-bound fructosyltransferase of *Streptococcus salivarius* ATCC25975 is preceded by an insertion sequence and followed by *FUR1* and *clpP* homologues. *J. Gen. Microbiol.* **139**:913-920.
20. Giffard, P. M., C. L. Simpson, C. P. Milward, and N. A. Jacques. 1991. Molecular characterization of a cluster of at least two glucosyltransferase genes in *Streptococcus salivarius* ATCC25975. *J. Gen. Microbiol.* **137**:2577-2593.
21. Green, C. J., and B. S. Vold. 1983. Sequence analysis of a cluster of twenty-one tRNA genes in *Bacillus subtilis*. *Nucleic Acids Res.* **19**:5763-5774.
22. Green, C. J., and B. S. Vold. 1993. *Staphylococcus aureus* has clustered tRNA genes. *J. Bacteriol.* **175**:5091-5096.
23. Guédon, G., M. Pébay, C. Colmin, J. M. Simonet, and B. Decaris. 1992. The 23S-5S spacer of two rRNA loci of *Streptococcus salivarius* subsp. *thermophilus* includes a promoter. *Biochimie* **74**:585-588.
24. Hantman, M. J., S. Sun, P. J. Piggot, and L. Daneo-Moore. 1993. Chromosome organization of *Streptococcus mutans* GS-5. *J. Gen. Microbiol.* **139**:67-77.
25. Herman, R. E., and L. L. McKay. 1986. Cloning and expression of the  $\beta$ -D-galactosidase gene from *Streptococcus thermophilus* in *Escherichia coli*. *Appl. Environ. Microbiol.* **52**:45-50.
26. Jarvis, E. D., R. L. Widom, G. LaFauci, Y. Setoguchi, I. R. Richter, and R. Rudner. 1988. Chromosomal organization of rRNA operons of *Bacillus subtilis*. *Genetics* **120**:625-635.
27. Lamoureux, M., H. Prévots, J. F. Cavin, and C. Diviès. 1993. Recognition of *Leuconostoc oenos* strains by the use of DNA restriction profiles. *Appl. Microbiol. Biotechnol.* **39**:547-552.
28. Larbi, D., C. Colmin, L. Rousselle, B. Decaris, and J. M. Simonet. 1990. Genetic and biological characterization of nine *Streptococcus salivarius* subsp. *thermophilus* bacteriophages. *Lait* **70**:107-116.
29. Leblond, P., F. X. Francou, J. M. Simonet, and B. Decaris. 1990. Pulsed-field electrophoresis analysis of the genome of *Streptomyces ambofaciens* strains. *FEMS Microbiol. Lett.* **72**:79-88.
30. Le Bourgeois, P., M. Lautier, M. Mata, and P. Ritzenthaler. 1992. Physical and genetic map of the chromosome of *Lactococcus lactis* subsp. *lactis* IL1403. *J. Bacteriol.* **174**:6752-6762.
31. Le Bourgeois, P., M. Mata, and P. Ritzenthaler. 1991. Pulsed-field gel electrophoresis as a tool for studying the phylogeny and genetic history of lactococcal strains, p. 140-145. *In* G. M. Dunny, P. P. Cleary, and L. L. McKay (ed.), *Genetics and molecular biology of streptococci, lactococci, and enterococci*. American Society for Microbiology, Washington, D.C.
32. Lee, J. J., H. O. Smith, and R. J. Redfield. 1989. Organization of the *Haemophilus influenzae* Rd genome. *J. Bacteriol.* **171**:3016-3024.
33. Llanos, R. M., A. J. Hillier, and B. E. Davidson. 1992. Cloning, nucleotide sequence, expression, and chromosomal location of *ldh*, the gene encoding L-(+)-lactate dehydrogenase, from *Lactococcus lactis*. *J. Bacteriol.* **174**:6956-6964.
34. Mercenier, A., and Y. Lemoine. 1989. Genetics of *Streptococcus thermophilus*: a review. *J. Dairy Sci.* **72**:3444-3454.
35. Mercenier, A., C. Robert, D. A. Romero, P. Slos, and Y. Lemoine. 1987. Transfection of *Streptococcus thermophilus* spheroplasts, p. 234-237. *In* J. J. Ferretti and R. Curtiss III (ed.), *Streptococcal genetics*. American Society for Microbiology, Washington, D.C.
36. Okahashi, N., C. Sasakawa, N. Okada, M. Yamada, M. Yoshikawa, M. Tokuda, I. Takahashi, and T. Koga. 1990. Construction of a *NotI* restriction map of the *Streptococcus mutans* genome. *J. Gen. Microbiol.* **136**:2217-2223.
37. Pébay, M. 1993. Ph.D. thesis. University of Nancy I, Nancy, France.
38. Pébay, M., C. Colmin, G. Guédon, C. De Gaspéri, B. Decaris, and J. M. Simonet. 1992. Detection of intraspecific DNA polymorphism in *Streptococcus salivarius* subsp. *thermophilus* by a homologous rDNA probe. *Res. Microbiol.* **143**:37-46.
39. Pébay, M., C. Colmin, G. Guédon, J. M. Simonet, and B. Decaris. 1993. Chromosomal genetic instability in *Streptococcus thermophilus*. *Lait* **73**:181-190.
40. Pébay, M., Y. Roussel, J. M. Simonet, and B. Decaris. 1992. High-frequency deletion involving closely spaced rRNA gene sets in *Streptococcus thermophilus*. *FEMS Microbiol. Lett.* **98**:51-56.
41. Piggot, P. J. 1990. Genetic map of *Bacillus subtilis* 168, p. 107-145. *In* K. Drlica and M. Riley (ed.), *The bacterial chromosome*. American Society for Microbiology, Washington, D.C.
42. Pittet, A. C., and H. Hottinger. 1989. Sequence of a hexameric tRNA gene cluster associated with rRNA genes in *Lactobacillus bulgaricus*. *Nucleic Acids Res.* **17**:4873.
43. Poolman, B., T. J. Royer, S. E. Mainzer, and B. F. Schmidt. 1989. Lactose transport system of *Streptococcus thermophilus*: a hybrid protein with homology to the melibiose carrier and enzyme III of phosphoenolpyruvate-dependent phosphotransferase systems. *J. Bacteriol.* **171**:244-253.
44. Poolman, B., T. J. Royer, S. E. Mainzer, and B. F. Schmidt. 1990. Carbohydrate utilization in *Streptococcus thermophilus*: characterization of the genes for aldose 1-epimerase (mutarotase) and UDPglucose 4-epimerase. *J. Bacteriol.* **172**:4037-4047.
45. Prudhomme, M., V. Méjean, B. Martin, and J.-P. Claverys. 1991. Mismatch repair genes of *Streptococcus pneumoniae*: HexA confers a mutator phenotype in *Escherichia coli* by negative complementation. *J. Bacteriol.* **173**:7196-7203.
46. Reed, K. C., and D. A. Mann. 1985. Rapid transfer of DNA from agarose gels to nylon membrane. *Nucleic Acids Res.* **13**:7207-7221.
47. Roussel, Y., C. Colmin, J. M. Simonet, and B. Decaris. 1993. Strain characterization, genome size and plasmid content in the *Lactobacillus acidophilus* group (Hansen & Mocoquot). *J. Appl. Bacteriol.* **74**:549-556.
48. Salzano, G., G. Moschetti, F. Villani, and S. Coppola. 1993. Biotyping of *Streptococcus thermophilus* strains by DNA fingerprinting. *Res. Microbiol.* **144**:381-387.

49. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
50. Schleifer, K. H., M. Ehrmann, U. Krusch, and H. Neve. 1991. Revival of the species *Streptococcus thermophilus* (ex. Orla-jensen, 1919) nom. rev. Syst. Appl. Microbiol. **14**:386–388.
51. Schroeder, C. J., C. Robert, G. Lenzen, L. L. McKay, and A. Mercenier. 1991. Analysis of the *lacZ* sequences from two *Streptococcus thermophilus* strains: comparison with the *Escherichia coli* and *Lactobacillus bulgaricus*  $\beta$ -galactosidase sequences. J. Gen. Microbiol. **137**:369–380.
52. Schwartz, D. C., and C. R. Cantor. 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell **37**:67–75.
53. Slos, P., J.-C. Bourquin, Y. Lemoine, and A. Mercenier. 1991. Isolation and characterization of chromosomal promoters of *Streptococcus salivarius* subsp. *thermophilus*. Appl. Environ. Microbiol. **57**:1333–1339.
54. Slos, P., and A. Mercenier. Personal communication.
55. Smith, C. L., S. R. Klcó, and C. R. Cantor. 1988. Pulsed-field gel electrophoresis and the technology of large DNA molecules, p. 41–72. In K. E. Davies (ed.), Genome analysis: a practical approach. IRL Press, Oxford.
56. Tanaka, R., Y. Andachi, and A. Muto. 1991. Evolution of tRNAs and tRNA genes in *Acholeplasma laidlawii*. Nucleic Acids Res. **19**: 6787–6792.
57. Tanskanen, E. I., D. L. Tulloch, A. J. Hillier, and B. E. Davidson. 1990. Pulsed-field gel electrophoresis of *Sma*I digests of lactococcal genomic DNA, a novel method of strain identification. Appl. Environ. Microbiol. **56**:3105–3111.
58. Terzaghi, B. E., and W. E. Sandine. 1975. Improved medium for lactic streptococci and their bacteriophages. Appl. Microbiol. **29**:807–813.
59. Tulloch, D. L., L. R. Finch, A. J. Hillier, and B. E. Davidson. 1991. Physical map of the chromosome of *Lactococcus lactis* subsp. *lactis* DL11 and localization of six putative rRNA operons. J. Bacteriol. **173**:2768–2775.
60. Vold, B. S. 1985. Structure and organization of genes for transfer ribonucleic acid in *Bacillus subtilis*. Microbiol. Rev. **49**:71–80.
61. Widom, R. L., E. D. Jarvis, G. LaFauci, and R. Rudner. 1988. Instability of rRNA operons in *Bacillus subtilis*. J. Bacteriol. **170**: 605–610.
62. Yohda, M., H. Okada, and H. Kumagai. 1991. Molecular cloning and nucleotide sequencing of the aspartate racemase gene from lactic acid bacteria *Streptococcus thermophilus*. Biochim. Biophys. Acta **1089**:234–240.