Microbiology, Vol. 70, No. 1, 2001, pp. 59–63. Translated from Mikrobiologiya, Vol. 70, No. 1, 2001, pp. 70–76.

EXPERIMENTAL ARTICLES

DNA Relatedness, Divergence, and Sibling Species of the Lactic Acid Bacterium *Streptococcus thermophilus*

A. M. Lysenko*, S. G. Botina, V. I. Ganina***, and V. V. Sukhodolets****

**Institute of Microbiology, Russian Academy of Sciences, pr. 60-letiya Oktyabrya 7, k. 2, Moscow, 117811 Russia*
***State Research Institute of Genetics and Selection of Industrial Microorganisms, Pervyi Dorozhnyi proezd 1, Moscow, 113545 Russia*
****Moscow State University of Applied Biotechnology, Moscow, Russia*

Received January 11, 2000; in final form, May 29, 2000

Abstract—Previously, five distinct groups with 80–90% intragroup DNA homology values were revealed among 19 lactic acid–producing bacterial strains. The study of 39 new strains of thermophilic streptococci in the present work allowed us to reveal the sixth DNA homology group. The nine strains of this group are close, at 55–70% DNA homology levels, to the type strain *Streptococcus thermophilus* ATCC 19258. Group VI showed a low level of DNA–DNA reassociation (20–30%) with the DNA homology groups I, II, III, and V. The intergroup DNA–DNA reassociation values determined from DNA renaturation rates varied from 20 to 50%. These data were interpreted as indicative of the existence of at least four sibling species among the thermophilic streptococci studied.

Key words: DNA reassociation in solution, DNA hybridization level, thermophilic streptococci, sibling species.

Thermophilic streptococci are widely used as starters in the production of yogurt and some cheeses. Unlike other lactic acid bacteria, including lactococci, thermophilic streptococci are able to grow at elevated temperatures (44–52°C). In 1984, Farrow and Collins reclassified the species *Streptococcus thermophilus* into the subspecies *S. salivarius* [1]. This reclassification, however, was later disproved [2, 3]. *S. thermophilus* strains are very diverse in their biochemical properties and nucleotide sequences (see [4] and relevant references in this publication).

Earlier, we reported on the DNA reassociation diversity of 19 *S. thermophilus* strains used as starters in the Russian dairy industry [5, 6]. Analysis showed that these strains, which were isolated in different regions of Russia, can be divided into five DNA homology groups (I through V) with intra- and intergroup DNA reassociation values of 80–90 and 20–60%, respectively [6]. When DNA–DNA hybridization values are determined from DNA renaturation rates, intraspecies homology values usually vary from 60 to 100% [7]. Therefore, the DNA relatedness of strains within 20–50% should be considered as an indication that these strains belong to different species.

The aim of the present work was to study the relatedness of 39 *S. thermophilus* strains based on DNA–DNA reassociation values determined from DNA renaturation rates.

MATERIALS AND METHODS

Thirty-nine thermophilic streptococcal strains used as starters in the Russian dairy industry were analyzed. Of these, 12 strains (CK1019 through CK1030) were from the collection of V.I. Ganina (in Table 1, the former designations of these and some of the other strains studied are given in parentheses). The remaining sixteen strains, hereinafter referred to as Collection 2 (Table 1, nos. 14 through 29), were isolated independently from starter cultures at the All-Russia Research Institute of the Dairy Industry in Moscow. Ten strains, referred to as Collection 3 (Table 1, nos. 30 through 39), were isolated from yogurt, cheese, cheese made from sheep's milk, and other fermented dairy products in Krasnodar (CK1002–CK1004), Baku (CK1013, CK1014), Uglich (CK1017), Mordovia (CK1005, CK1010, and CK1016), and France (CK1012). Strain IL1704 was a generous gift from S. Kulakauskas (France). The type strain *S. thermophilus* ATCC 19258 was obtained from the American Type Culture Collection.

In DNA–DNA hybridization experiments, we also used the DNA samples isolated from five reference strains representing the five DNA homology groups revealed earlier, namely, strain B3371 from group I, strain 108 from group II, strain 5 from group III, strain 722 from group IV, and strain 6kb from group V. The origin of these strains is described elsewhere [5].

The strains were grown in M21 medium [8] at 37°C. The biomass was collected by centrifugation and

0026-2617/01/7001-0059$25.00

Table 1. DNA–DNA hybridization values of *S. thermophilus* strains with respect to the reference strains of different DNA homology groups

No.	Strain	G+C content, mol %	DNA homology group**	DNA homology (%) with respect to	
				reference strain*	ATCC 19258
1	ATCC 19258	39.6	T	–	100
2	CK1019 (CT138)	38.2	VI	100	70
3	CK1020 (CT9)	38.7	VI	98	67
4	CK1021 (CT14)	39.0	VI	84	65
5	CK1024 (CT132)	38.6	VI	79	55
6	CK1025 (TP20)	39.3	VI	96	54
7	CK1026 (T48)	39.4	VI	99	65
8	CK1027 (CT95)	39.9	VI	93	ND
9	CK1028 (T24)	39.8	VI	98	57
10	CK1022 (CT161)	38.4	V	78	41
11	CK1023 (T13)	39.6	V	81	45
12	CK1029 (CT154)	39.8	I	84	48
13	CK1030 (T10/7)	40.0	I	95	53
14	B1364	39.2	V	92	54
15	B1363	39.4	V	99	53
16	CK1001	39.6	V	84	ND
17	CK1035 (HKM)	38.5	V	93	50
18	CK1036 (410)	38.0	V	95	50
19	CK1037 (635)	38.4	V	100	53
20	CK1038 (229)	37.8	V	100	49
21	B6402	38.8	V	98	51
22	B6410	38.6	V	80	ND
23	B6415	38.5	V'	71	41
24	B6403	39.3	III	83	47
25	B6448	39.4	I	98	57
26	B6453	39.5	I	92	ND
27	B6454	38.6	I	95	ND
28	B6469	38.4	I	97	ND
29	B6472	39.2	I	95	ND
30	CK1002	38.3	I	77	51
31	CK1010	39.1	I	72	ND
32	CK1014	39.5	II	100	38
33	CK1003	38.6	II	90	36
34	CK1004	38.2	II	91	ND
35	CK1013	39.1	II	99	ND
36	CK1017	39.8	II	95	ND
37	CK1012	39.4	VI	95	53
38	CK1005	39.2	T(?)	100	61
39	CK1016	38.8	T(?)	68	ND
40	IL1704	38.7	T	–	89

*Reference strains were as follows: CK1019 (for strains nos. 2, 3, 5–9, and 37), CK1020 (for strain no. 4), strain 6kb (for strains nos. 10, 11, and 14), B3371 (for strains nos. 12, 13, 25, and 30), CK1002 (for strain no. 31), B1364 (for strains nos. 15–23), strain 5 (for strain no. 24), B6448 (for strains nos. 26–29), CK1014 (for strains nos. 32–36), and CK1005 (for strains nos. 38 and 39). "ND" stands for "not determined". T is the homology group of the type strain ATCC 19258.

Table 2. DNA–DNA hybridization values of *S. thermophilus* strains from different DNA homology groups

No.	Strain	DNA homology group	DNA homology (%) with respect to the reference strains				
			B3371 (group I)	108 (group II)	5 (group III)	722 (group IV)	6kb (group V)
1	ATCC 19258	T	54	39	43	47	43
2	CK1019	VI	33	21	29	50	20
3	CK1020	VI	30	21	31	ND	25
4	CK1022	V	63	23	38	54	78
5	CK1023	V	58	19	35	48	81
6	CK1029	I	84	31	ND	35	ND
7	CK1030	I	95	30	33	31	ND
8	B1364	V	54	33	45	51	92
9	B1363	V	50	30	47	50	90
10	CK1036	V	50	31	46	53	95
11	CK1038	V	48	28	41	48	89
12	B6415	V	46	18	33	14	58
13	B6403	III	32	26	83	50	45
14	B6448	I	98	25	34	31	51
15	CK1002	I	77	32	27	29	32
16	CK1003	II	35	83	23	39	23
17	CK1014	II	31	81	25	36	21
18	CK1005	T(?)	39	22	30	41	23
19	CK1012	VI	31	18	35	47	19
20	IL1704	T	62	31	48	50	41
21	B3371	I	100				
22	108	II	34	100			
23	5	III	23	25	100		
24	722	IV	30	32	51	100	
25	6kb	V	35	25	61	50	100

Note: "ND" stands for "not determined".

washed twice with 0.15 M NaCl containing 0.1 M Na_2-EDTA. DNA was isolated by the method of Marmur [9] with the following modifications: (a) after the first cycle of deproteinization, crude DNA preparations were kept at 4°C for 10–14 days to raise the degree of deproteinization [5]; (b) after treatment with RNase A, DNA preparations were treated with pronase; and (c) final DNA preparations were cleared by centrifugation [10].

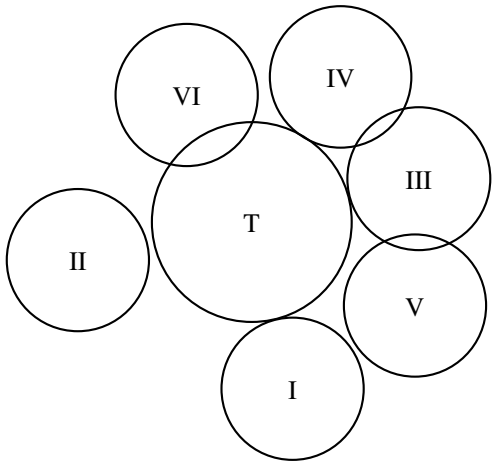
DNA–DNA hybridization measurements were carried out by the method of De Ley *et al.* as described previously [6, 10]. Standard deviation in these measurements did not exceed 3–5%.

The G+C content of DNA was calculated by the formula: G+C (mol %) = $2.08T_m - 106.4$, where T_m is the melting temperature of DNA determined from its melting profile in 0.1 SSC (0.15 M NaCl + 0.015 M sodium citrate, pH 7.0) [11].

RESULTS

The G+C content of the DNA of all 39 thermophilic streptococcal strains studied varied from 37.8 to 40.0 mol % (Table 1), which is a typical range of G+C values for representatives of the species *S. thermophilus* [1].

A comparison of the DNA–DNA hybridization values of strains CK1019–CK1030 showed that eight (CK1019–CK1021 and CK1024–CK1028) belong to one DNA homology group with DNA reassociation values of 80–90% (data not shown). At the same time, the intergroup DNA–DNA hybridization values were considerably lower (20–30%). This allowed a particular species to be reliably assigned to a certain DNA homology group based on its DNA–DNA reassociation value with the respective reference strain (the data in Table 1 represent the maximum DNA–DNA reassociation values observed). The new group of eight aforementioned strains (Group VI) is sufficiently distant



Schematic representation of the DNA homology groups of *S. thermophilus* strains. For explanation, see text.

from the five DNA homology groups that were established earlier. The members of this group are very close to the type strain ATCC 19258 and show DNA–DNA reassociation values with this strain at a level of 65–70% (Table 1). Group VI was found to be closest to Group I, whose members have 50–60% DNA homology with the type strain ATCC 19258 [6]. Two other strains from Ganina's collection, CK1029 and CK1030, were attributed to the homology Group I, while strains CK1022 and CK1023 were placed in Group V (Table 1).

The 16 strains of Collection 2 were found to fall into the DNA homology Group I (five strains), Group III (one strain), and Group V (nine strains) (Table 1). The sixteenth strain of this collection, B6415, was the closest to the members of Group V. However, the insufficiently high level of DNA homology (60–70%) of this strain with the other strains of this group allowed us to classify it into an arbitrary Group V', which is presumably another new DNA homology group of lactic acid–producing thermophilic streptococci. It should be noted that strain B6415 showed abnormally low DNA homology values (18 and 14%, respectively) with the reference strains of Groups II and IV.

Among the 10 strains of Collection 3, five were found to belong to Group II (Table 1, nos. 32 through 36). Two strains, CK 1002 and CK1010, were assigned to Group I, and one strain, CK1012, was assigned to the new DNA homology Group VI. The last two strains of this collection, CK1005 and CK1016, which were isolated in Mordovia, were provisionally Placed in the DNA homology group of the type strain, Group T, since they showed very low DNA homology (30–40%) with the reference strains of all six homology groups (Tables 1 and 2). The strain IL1704, isolated in France, had a high DNA homology (89%) with the Group T strain (Table 1).

Table 2 presents the results of determination of the DNA homology of 20 of the 39 strains under study with respect to the reference strains of the five DNA homology groups established earlier [5]. Data for the other strains under consideration are incomplete and, for this reason, are not presented. It can be seen that Groups III, IV, and V are close, have high DNA homology values between each other (from 50 to 60%) and, as a rule, have low DNA homology values (35%) with the reference strains of Groups I and II. The exceptions are some strains of Group V, which showed DNA reassociation values with members of Group I, ranging from about 50% (the reference strain B3371) to 63% (strain CK1022) (see Table 2 and [6]). DNA homology between Groups I and II is also very low (20–30%).

Such low DNA homology values may indicate that the respective groups of strains belong to different taxa at the species level, since, as noted above, intraspecies DNA–DNA hybridization values determined from DNA renaturation rates are usually not lower than 60% [7]. Therefore, the new DNA homology Group VI, which is characterized by relatively high DNA relatedness (55–70%) to the type strain ATCC 19258, can be considered a subspecies within the species *S. thermophilus* sensu stricto. As for Groups I and II, and the group cluster III–V, they most likely represent taxa analogous to sibling species (for details, see the *Discussion* section).

It should be noted that the DNA–DNA reassociation values of the all strains under study with the type strain ATCC 19258 are not lower than 35–40% (see Tables 1 and 2 and publication [6]). This suggests that the DNA homology group T is ancestral with respect to Groups I–VI.

DISCUSSION

Based on the DNA–DNA hybridization values derived, 39 strains of thermophilic streptococci were found to belong to the previously established DNA homology Groups I (nine strains), II (five strains), III (one strain), and V (twelve strains). Furthermore, nine strains were attributed to the new DNA homology Group VI, and three strains (two of which provisionally) were attributed to the DNA homology group of the type strain *S. thermophilus* ATCC 19258. Taking into account that the new Group VI shows a relatively high level of DNA homology (55–70%) with the type strain ATCC 19258, this group can be considered a separate taxon within the species *S. thermophilus.* At the same time, the DNA homology values of all other groups with the type strain are rather low (20–50%), suggesting that they have substantially diverged from the type strain and cannot be considered as belonging to the same species.

The results presented in this paper confirm the distribution pattern of thermophilic streptococci over the DNA homology groups revealed in our earlier studies [5, 6]. This distribution reflects the divergent evolution

of thermophilic streptococci. The situation is schematically given in the figure, where various DNA homology groups are represented by circles, whose overlap measures the degree to which the groups are related to or diverge from each other. The overlap is minimal with respect to the DNA homology group of the type strain (this group is represented by the central circle marked by the uppercase letter T), and this corresponds to the low level of DNA homology (40–50%) between the type strain and the majority of the other strains studied.

An elementary unit of the vertical evolution of eukaryotic and prokaryotic organisms is thought to be the species level (see, for instance, [12]). Species integrity is maintained by gene recombination, which implies a sufficiently high level of DNA homology within species. The DNA–DNA hybridization values between particular strains ranging between 20 and 40% may indicate that these strains belong to different species. The data accumulated so far may indicate that Groups I and II and cluster III–V are sibling species of *S. thermophilus* sensu stricto.

Reportedly, hybridization between the sibling species of the genus *Saccharomyces* gives nonviable progeny [13]. Such hybridization experiments with the thermophilic streptococci are impossible. However, the low level of DNA–DNA hybridization between the strains studied suggests that regular genetic interactions between them are absent and therefore, they belong to genetically isolated populations that are analogous to the sibling species of yeasts and other organisms. The aforementioned absence of a viable progeny from the hybridized twin *Saccharomyces cerevisiae* species (*S. paradoxus* and *S. bayanus*) can be accounted for by the low level of DNA homology between them: the levels of DNA homology between *S. cerevisiae* and *S. paradoxus*, between *S. cerevisiae* and *S. bayanus*, and between *S. paradoxus* and *S. bayanus* were found to be 50, 20, and 30%, respectively [13, 14].

Thus, the 39 new strains of lactic acid–producing thermophilic streptococci used in Russia as starters in the dairy industry fall into four DNA homology groups, some of which are analogous in the degree of DNA–DNA hybridization to sibling *S. cerevisiae* species. Based on DNA–DNA reassociation values, we infer that the DNA homology Groups T, I, and II and the cluster III–V represent sibling species, provisionally designated as *S. thermophilus* sensu stricto, *S. thermophilus* I, *S. thermophilus* II, and *S. thermophilus* III, respectively.

ACKNOWLEDGMENTS

This work was supported, in part, by the Russian Foundation for Basic Research project no. 00-04-48147.

REFERENCES

1. Farrow, J.A.E. and Collins, M.D., DNA Base Composition, DNA–DNA Homology and Long-Chain Fatty Acid Studies on *Streptococcus thermophilus* and *Streptococcus salivarius*, *J. Gen. Microbiol.*, 1984, vol. 130, pp. 357–362.

2. Schleifer, K.H. and Kilpper-Balz, R., Molecular and Chemotaxonomic Approaches to the Classification of Streptococci, Enterococci and Lactococci: A Review, *Syst. Appl. Microbiol.*, 1987, vol. 10, pp. 1–19.

3. Whiley, R.A. and Hardie, J.M., *Streptococcus vestibularis* sp. nov. from the Human Oral Cavity, *J. Syst. Bacteriol.*, 1988, vol. 38, pp. 335–339.

4. O'Sullivan, T.F. and Fitzgerald, G.F., Comparison of *Streptococcus thermophilus* Strains by Pulse Field Gel Electrophoresis of Genomic DNA, *FEMS Microbiol. Lett.*, 1998, vol. 168, pp. 213–219.

5. Averina, O.V., Lysenko, A.M., Ermakova, L.M., Ogai, D.K., and Sukhodolets, V.V., Comparative Study of DNA Homology in Thermo- and Mesophilic Lactic Acid–producing Streptococci of Different Origin, *Mikrobiologiya*, 1998, vol. 67, no. 6, pp. 792–798.

6. Lysenko, A.M., Karpushina, S.G., and Sukhodolets, V.V., DNA Homology Divergence of *Streptococcus thermophilus* Strains, *Mikrobiologiya*, 1999, vol. 68, no. 4, pp. 514–518.

7. Johnson, J.L., Nucleic Acids in Bacterial Classification, *Bergey's Manual of Systematic Bacteriology*, Krieg, N.R. and Holt, J.G., Eds., London: Williams & Wilkins, 1984, pp. 8–11.

8. Nechaeva, A.A. and Sukhodolets, V.V., Genetic Study of Industrial *Lactococcus lactis* Strains: Search for Transmissible Plasmids Responsible for Lactose Fermentation, *Genetika*, 1996, vol. 32, no. 2, pp. 218–227.

9. Marmur, J., A Procedure for the Isolation of DNA from Microorganisms, *J. Mol. Biol.*, 1961, vol. 3, pp. 208–218.

10. Lysenko, A.M., Gal'chenko, V.F., and Chernykh, N.A., Taxonomic Study of Obligately Methanotrophic Bacteria by the DNA–DNA Hybridization Method, *Mikrobiologiya*, 1988, vol. 57, no. 5, pp. 816–822.

11. Owen, R.J., Hill, R.L., and Lapage, S.P., Determination of DNA Base Composition from Melting Profiles in Dilute Buffers, *Biopolymers*, 1969, vol. 7, pp. 503–516.

12. Sukhodolets, V.V., The Mechanisms of Vertical Evolution, *Usp. Sovrem. Biol.*, 1997, vol. 117, no. 5, pp. 517–533.

13. Naumov, G.I., Genetic Basis for Classification and Identification of the Ascomycetous Yeasts, *Studies Mycol.*, 1987, vol. 30, pp. 469–475.

14. Vauhgan-Martini, A., *Saccharomyces paradoxus* comb. nov., a Newly Separated Species of the *Saccharomyces* sensu stricto Complex Based upon DNA/DNA Homologies, *Syst. Appl. Microbiol.*, 1989, vol. 12, pp. 179–182.