Sau-PCR, a Novel Amplification Technique for Genetic Fingerprinting of Microorganisms

Viviana Corich, Alessandro Mattiazzi, Elisa Soldati, Angela Carraro and Alessio Giacomini
Appl. Environ. Microbiol. **2005, 71(10):6401. DOI: 10.1128/AEM.71.10.6401-6406.2005.**

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Oct. 2005, p. 6401–6406 Vol. 71, No. 10
0099-2240/05/$08.00+0 doi:10.1128/AEM.71.10.6401–6406.2005

Sau-PCR, a Novel Amplification Technique for Genetic Fingerprinting of Microorganisms†

Viviana Corich, Alessandro Mattiazzi, Elisa Soldati, Angela Carraro, and Alessio Giacomini*

Dipartimento di Biotecnologie Agrarie, Università di Padova, Legnaro, Padova, Italy

Received 12 January 2005/Accepted 16 May 2005

The proposed technique is based on the digestion of genomic DNA with the restriction endonuclease Sau3AI and subsequent amplification with primers whose core sequence is based on the Sau3AI recognition site. The method was tested on strains of lactic acid bacteria but could be proposed for virtually any culturable organism from which DNA can be extracted.

The need to characterize and distinguish bacterial strains is crucial in many fields of applied microbiology, in order to assess the biodiversity levels within natural microbial populations, to evaluate the compositions of complex starter cultures, or to follow the fates of strains added to natural environments for technological purposes. In these cases, it is important to unambiguously identify such microorganisms at either species or strain levels. Moreover, the increasing number of commercial strains sold on the market alone or in mixed cultures and covered by commercial rights or patents urges the development or improvement of methods for rapidly and clearly confirming their identities and tracing their presence (16).

The introduction of molecular fingerprinting techniques based on genomic DNA amplification has considerably increased the sensitivities and the speeds of execution of these tasks. In particular, the random amplified polymorphic DNA (RAPD [24]) and amplified-fragment-length polymorphism (AFLP [23]) methods are to date the most widely used approaches, as evidenced by the number of scientific papers (200 and 120 per year, respectively, on average) published in international journals during the last 4 years with the names of the techniques appearing in the article titles (Current Contents search; The Institute for Scientific Information, Philadelphia, PA). Each of these approaches has some advantages and at the same time suffers from some problems (10). In particular, the RAPD technique is rather simple and fast, and the analysis generally can be performed starting from a lysate of a bacterial colony without the need to extensively purify the DNA. However, it is well known that RAPD profiles can be sensitive to even modest changes in the reaction conditions, which could lead to problems of reproducibility, particularly regarding the minor faint bands, which are not always well conserved among replicates of the same sample (11, 12, 18, 21). Conversely, the AFLP technique normally displays good levels of reproducibility and reliability, apart from some reported problems mostly related to the initial DNA concentration (2, 5) or to the endonuclease or ligase treatment efficiency (13), but it is quite laborious and time consuming, considering its needs for two enzymatic reactions and for large polyacrylamide gels to reach a good level of band separation.

The molecular fingerprinting technique proposed in this work is based on a genomic DNA enzymatic digestion and subsequent amplification that recall in part the RAPD and AFLP methods and basically retains the random amplification concept from the former and the enzymatic digestion of genomic DNA from the latter. As does the AFLP method, this technique makes use of primers based on the restriction enzyme recognition sequence, but it does not require addition of linkers, and the products can be resolved on agarose gels, as in the RAPD method.

The name of the technique comes from the restriction endonuclease Sau3AI, used to fragment genomic DNA, which was chosen for some interesting features. Its recognition site, GATC, is 4 bp long with a G+C percentage value of 50 that is capable of producing a very large number of short fragments from the digestion of the total DNA of most microorganisms. This enzyme generates sticky ends with four protruding bases, a very uncommon situation for commercially available 4-bp cutters. This feature, allowing the entire GATC sequence to remain on one side of the cut DNA strand, was essential for the design of the primer, whose core sequence is indeed constituted by these four bases, which are present at both extremities of the digested DNA fragments. In addition, all Sau-PCR primers are endowed with the same 7-nucleotide (nt)-long tail, CCGCCGC, at their 5′ ends, which does not participate in the initial annealing event (the low-stringency phase), which is intended to confer, by the use of only C's and G's, a conveniently high melting temperature for the successive high-stringency amplification cycles. The nucleotide sequence of the tail was chosen so as to avoid the possibility of primer self annealing. At the 3′ end of the GATC sequence, as in the AFLP technique and similar methods, a variable number of nucleotides can be added to increase primer selectivity, so that the higher the number of nucleotides is, the lower the theoretical number of amplifiable Sau3AI fragments is.

The digestion of the DNA by a frequent-cutting restriction endonuclease has the scope of generating a numerous population of short fragments, so that target sites could be more easily accessed by the primers than in a bulk of intact genomic DNA. In fact, uncut DNA is present in solution as chromo-

* Corresponding author. Mailing address: Dipartimento di Biotecnologie Agrarie, Università di Padova, viale dell'Università 16, I-35020 Legnaro, Padova, Italy. Phone: 39 049 8272924. Fax: 39 049 8272929. E-mail: alessio.giacomini@unipd.it.

† Supplemental material for this article may be found at http://aem.asm.org/.

TABLE 1. Primers used in this study and annealing temperatures for the high-stringency step of Sau-PCR amplification

Primer	Length (nt)	Sequence (5′ to 3′)	Annealing temperature (°C)
SAUA	12	CCGCCGCGATCA	44
SAUC	12	CCGCCGCGATCC	46
SAUG	12	CCGCCGCGATCG	46
SAUT	12	CCGCCGCGATCT	44
SAG	13	CCGCCGCGATCAG	48
SCA	13	CCGCCGCGATCCA	48
STG	13	CCGCCGCGATCTG	48
SGAG	14	CCGCCGCGATCGAG	52

somes, or at least as very large molecules, due to partial shearing occurring during the extraction procedures, each carrying many target sequences, while primers are more homogeneously dispersed in the solution volume. It is therefore likely that cut DNA fragments can more efficiently move throughout the liquid and have a greater probability to come in contact with the primers.

The strains used in this work are listed in Table S1 in the supplemental material. All lactobacilli and enterococci were grown in MRS medium (Oxoid, Basingstoke, United Kingdom), while streptococci were cultivated in M17 medium (Oxoid), in both cases for 3 days at 37°C. *Escherichia coli* was grown in Luria-Bertani broth at 37°C overnight.

Genomic DNA from lactic acid bacteria (LAB) was extracted from 14-ml liquid cultures following the procedure described by Querol et al. (15), which was modified by the addition of lysozyme (25 mg/ml; Sigma, St. Louis, MO) and mutanolysin (10 U/ml; Sigma) to the lysis solution (1), and the genomic DNA was resuspended in 50 μl of deionized sterile water.

Suitable dilutions of DNA solutions were utilized for quantification with a Picogreen double-stranded DNA quantitation reagent kit (Molecular Probes, Eugene, Oregon) and a Spectrafluor spectrofluorimeter (Tecan Group Ltd., Maennedorf, Switzerland).

Ten μl of DNA solution (200 ng) was digested at 37°C overnight with 10 units of Sau3AI (Amersham Biosciences AB, Uppsala, Sweden) in a final volume of 20 μl by using the appropriate restriction buffer.

The amplification reaction was performed in a volume of 25 μl into 0.2-ml tubes by using a Bio-Rad iCycler thermal cycler (Bio-Rad, Hercules, California), with the following reagent concentrations: 2 μl of Sau3AI-digested DNA (10 ng/μl), 200 μM deoxynucleoside triphosphates (dNTPs; Amersham), 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, 0.5 U *Taq* polymerase (Amersham; 5U/μl), and 2 μM primer (MWG-Biotech, Ebersberg, Germany; high-purity salt-free purified). The primers used in this work are listed in Table 1.

The main steps of the technique are graphically summarized in Fig. 1, and the amplification protocol was designed according to the following description.

Fill-in of protruding Sau3AI ends. The fill-in of Sau3AI-cut ends was carried out by the *Taq* polymerase. Starting at 25°C for 5 s, the temperature was gradually raised (0.1°C/s) to 60°C to reduce the possible partial denaturation effects of the low-G+C-content DNA of LAB and maintained for 30 s.

Amplification: low-stringency phase. In the first amplification cycle, the primers annealed to the template at the Sau3AI site sequence (GATC) plus the 1- to 3-nt elongation in the 3′ end, leaving the 5′ G and C extensions outside. Denaturation was carried out at 94°C for 60 s. The temperature was brought to 50°C for 15 s and then gradually (0.1°C/s) lowered to 25°C. This touchdown cycle was deemed necessary since the primer melting temperature at this stage was very low (around 20°C), as the GC tail was excluded from the annealing. Ramping was therefore aimed at reducing the possibility of unspecific pairing. The temperature was then raised gradually (0.1°C/s) from 25°C to 50°C to avoid possible primer detachment from the template while still allowing a good thermal environment for the *Taq* activity, and it was maintained at 50°C for 30 s. The whole cycle was repeated twice to allow tail formation on both ends of each Sau3AI fragment (Fig. 1).

Amplification: high-stringency phase. At this stage, the target amplified fragments were constituted of the Sau3AI genomic regions flanked by the GC tails and hence completely matched the primer on both ends. It was therefore possible to proceed to the high-stringency phase, which was composed of 15 s of denaturation at 94°C followed by 60 s of annealing at 44 to 50°C (depending on the primer used; Table 1) and 2 min of extension at 65°C and which was repeated 35 times. A final extension step was performed at 65°C for 5 min, and then the temperature was brought to 4°C.

Amplified samples were run on agarose gels and the bands visualized by ethidium bromide staining. Gel concentrations of 1 to 2% were tested, with the latter generally giving the best levels of band separation. Digital images were acquired with an EDAS290 image capturing system (Kodak, Rochester, NY) and analyzed with the GelCompar II package (Applied Maths, Sint-Martens-Latem, Belgium). The similarity matrix from the band profiles was calculated using the cosine product-moment correlation coefficient (19). The dendrogram was obtained by means of the unweighted pair group method using arithmetic averages (UPGMA) clustering algorithm (17).

The number of fragments theoretically obtainable from a Sau-PCR amplification can be roughly estimated first by dividing the genome size, which is for example of about 2 million bp for many lactobacilli, by the experimentally estimated mean length (400 bp; see below) of the Sau3AI fragment population, which gives a number of about 6,000. Hypothesizing a random distribution of the nucleotides on the DNA and using a primer with a 1-nt extension, the possibility of finding a fragment with the identical 4-plus-1 nt GATCX 5′ sequences (where X is the nucleotide of choice for the primer) on both ends decreases to 1/16, accounting for about 400 bands. This value lowers to about 25 (1/256) and 1.5 (1/4,096) fragments for stretches of 4-plus-2 and 4-plus-3 nt, respectively. From these calculations, it appears that a primer with a 2-nt addition could potentially give a fingerprint with a suitable number of bands, although such an estimation has to be adjusted by taking into account the real genomic G+C percentage content and the actual nucleotide distribution, which can lead to considerably different numbers (22).

The effectiveness of the technique was examined using various species and strains of LAB, a category of microorganisms

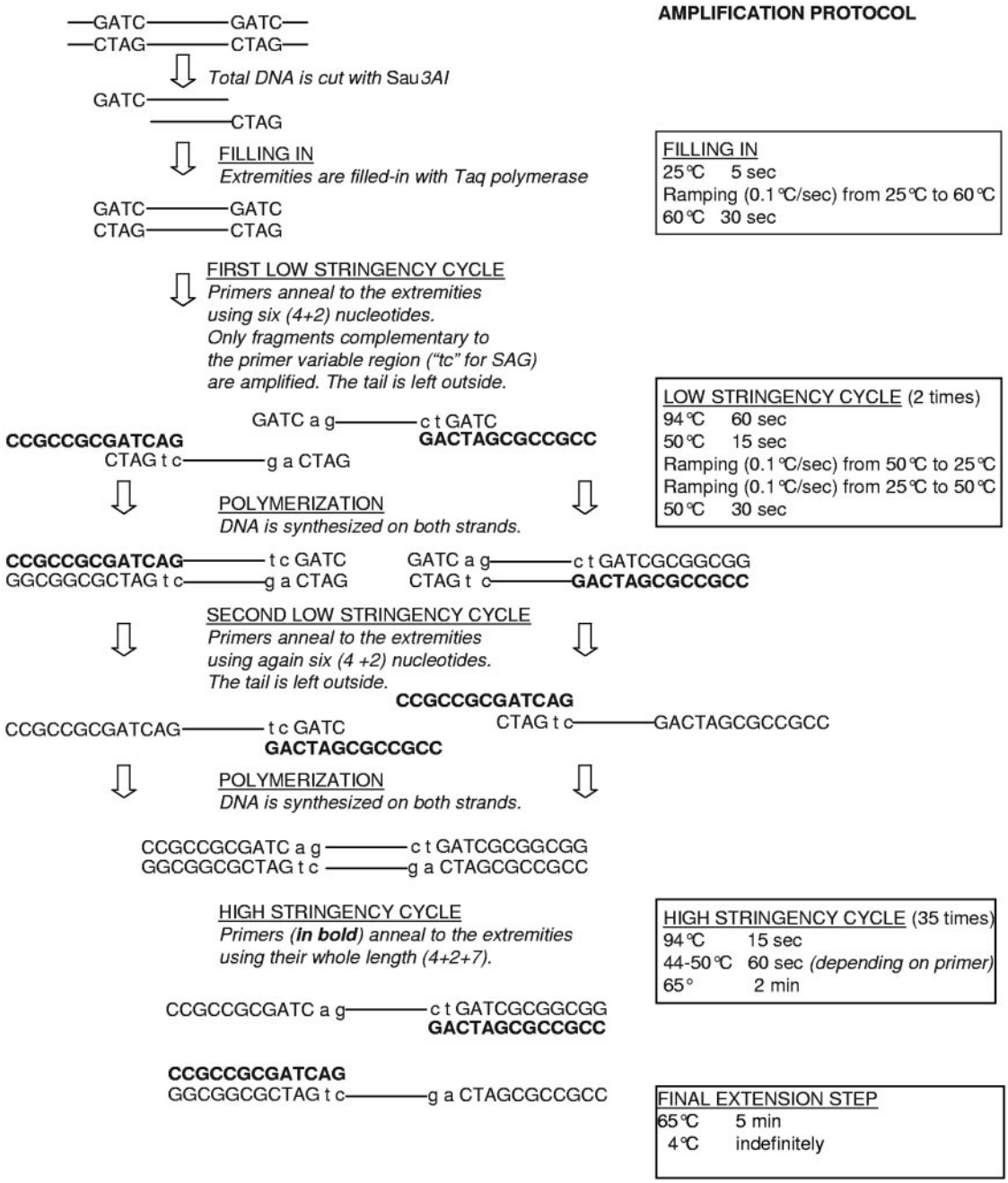

FIG. 1. Scheme of the Sau-PCR technique with the use of primer SAG (drawn in bold). The thermal protocol is reported at the right and can be assembled as a unique PCR program, which takes about 3 hours to complete. Small letters indicate nucleotides that are complementary to the variable part of the primers and that therefore differ depending on the primer used.

that contains numerous species of great technological relevance. As a first approach, four strains from three different LAB species, namely, *Lactobacillus helveticus* DMS20075^T, *L. delbrueckii* subsp. *lactis* DSM20072^T, *L. helveticus* 455, and *L. fermentum* A85, were analyzed. The first two are the type strains of the respective species, while the others were isolated from a natural whey starter culture. The DNA of the four strains, digested with Sau3AI, gave restriction products, with the majority of fragments ranging from about 0.2 to 4.5 kb and with mean values of around 0.3 to 0.4 kb (data not shown). These samples were then subjected to Sau-PCR analysis using primers having one, two, or three 3′ nucleotide extensions (Table 1). Amplification with primer SAUC (see Fig. S1 in the supplemental material), which has a 1-nt extension, generated profiles with limited numbers of fragments (from 5 to 8), but all the bands appeared well defined on the gel. The use of the other three possible primers with 1-nt extensions, namely, SAUA, SAUG, and SAUT, gave similar results (data not

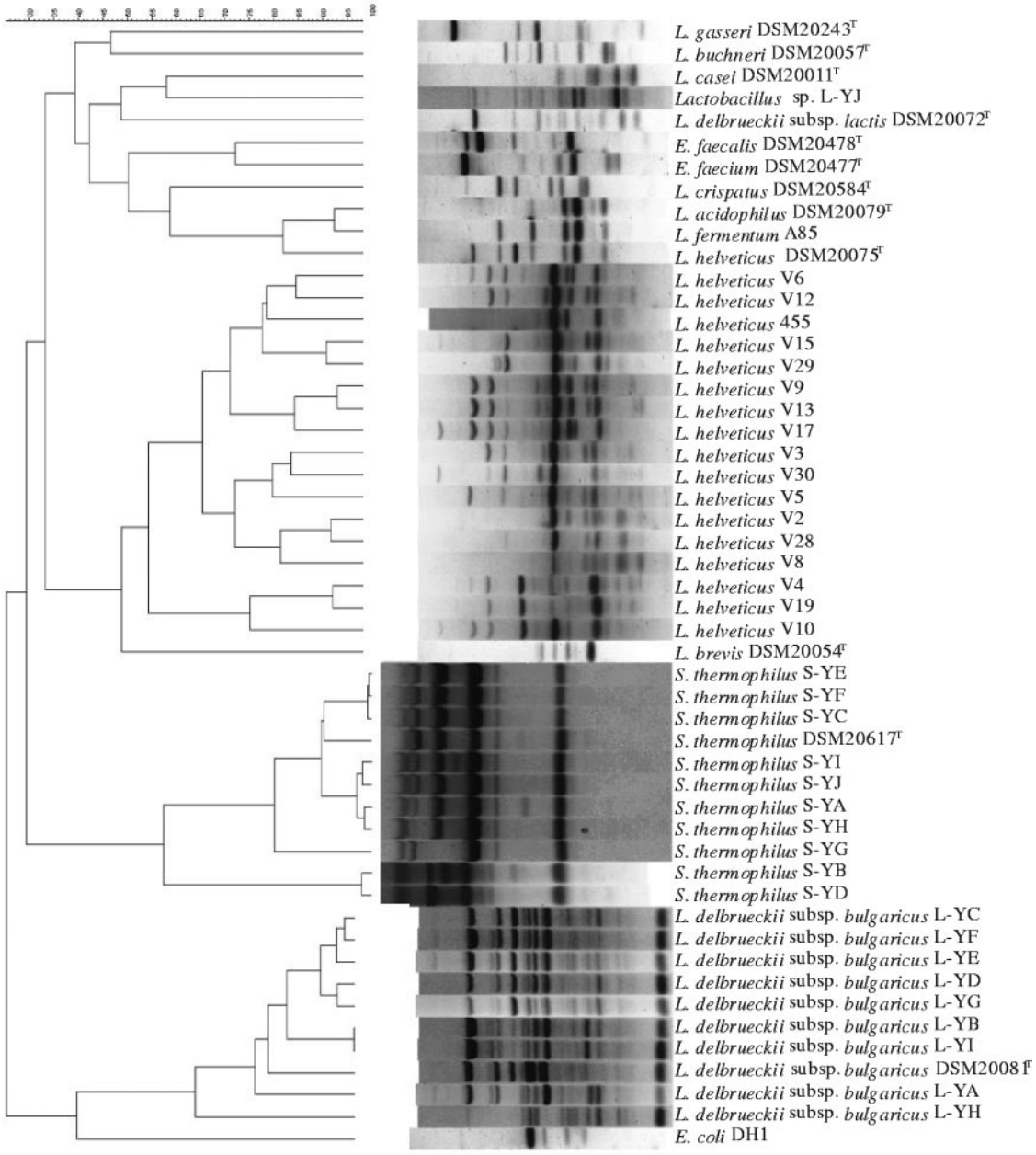

FIG. 2. Dendrogram from Sau-PCR profiles generated by primer SAG showing relationships among 13 LAB species and 36 isolates. *E. coli* DH1 was used as an outgroup. The cosine coefficient was used to calculate the similarity matrix. The upper scale indicates the percentage of similarity. *E. faecalis*, *Enterococcus faecalis*; *E. faecium*, *Enterococcus faecium*.

shown). The number of bands generated with primer STG, which has a 2-nt extension, was higher and ranged from 10 to 15. As in the previous case, the bands appeared intense and well defined. Similar results were produced by primers SAG and SCA (data not shown). When the 3-nt-extended primer SGAG was tested, the profiles showed numbers of bands ranging from 8 to 11, but their intensities were sometimes variable and the replicates were not always identical.

It has to be remarked that DNA fragmentation by endonuclease cleavage plays a key role in Sau-PCR. In fact, the use of uncut genomic DNA, amplified as a control, always produced very few bands completely unrelated to the corresponding Sau-PCR profiles (data not shown). Such poor amplification with nonrestricted DNA clearly suggests that primer annealing is very inefficient unless the Sau3AI termini are exposed.

Since the 2-nt addition to the GATC primer core sequence allowed for the enhanced detection of polymorphisms, as was also theoretically expected, primers SAG and SCA were used on 13 different LAB species, most of which belong to the genus *Lactobacillus*, with the addition of *Escherichia coli* DH1 (8) as an outgroup, to evaluate the levels of detectable polymorphism (see Fig. S2 in the supplemental material). Amplification with

primer SCA produced fragments ranging from about 0.1 to 1.0 kb, and the numbers of bands in the profiles were between 4 and 15. The fingerprints obtained with primer SAG were composed of bands of 0.1 to 1.0 kb as well, and the numbers of fragments of the profiles ranged from 5 to 17.

To verify the capability of Sau-PCR to differentiate strains within the same species, the technique was applied to 17 *L. helveticus* isolates from a natural whey starter culture and to 10 *Streptococcus thermophilus* plus 9 *L. delbrueckii* subsp. *bulgaricus* strains recovered from commercial yogurts (4) by use of primer SAG, which had produced good results for the species tested. The use of the last group of microorganisms also allowed testing of the technique down to the subspecies level. The resulting profiles, along with those from the SAG amplification of the 13 LAB species (see Fig. S2b in the supplemental material), were subjected to computer analysis to calculate similarity relationships and to draw a single dendrogram (Fig. 2) to better visualize and numerically evaluate the overall separating capabilities of the technique. Each group of strains clustered very well, with the respective type strain included, except in the case of *L. helveticus.* This result could be related to the strain's different technological and geographical origin, which was a hard cheese from Switzerland. Such inconsistency has also been observed by others with different fingerprinting techniques (C. Andrighetto, personal communication) and may be resolved by using diverse primers. However, it is also possible that there are important genotypic differences between these strains that have gone undetected without the application of genome fingerprinting.

To numerically define the reproducibility level of the technique, the DNA from three type strains was separately extracted four times and used for as many independent Sau-PCR amplifications with primer SAG. The four profiles of each strain were analyzed with GelCompar II to obtain the similarity values, which were 95.8% for *L. helveticus*, 95% for *S. thermophilu*s, and 97% for *L. delbrueckii* subsp. *bulgaricus*; these values represent the reproducibility threshold values. Moreover, for the fingerprints shown in this work, the presence of very faint or background bands is minimal; this fact positively affects the reproducibility, which is known to be one major drawback of the RAPD technique (20) and is also not completely negligible in the AFLP method (3).

Recent studies on enterobacterial repetitive intergenic consensus (ERIC) PCR, a well-known fingerprinting technique based on the detection of conserved repetitive genomic sequences (9), has revealed that in PCR experiments primers can normally efficiently anneal to nonperfectly complementary regions, even under stringent annealing conditions (6, 7, 14). To inspect this possibility in Sau-PCR, a SAG primer with a different composition of the GC stretch (5′GGCGGCGGAT CAG) was used, and amplification patterns were compared with those generated by a canonical SAG primer. Such evaluation highlighted few differences, which in essence pertained to the appearance or disappearance of one or two bands on fingerprints composed of 7 to 10 bands (data not shown), which suggests a modest involvement of improper annealings.

As a general consideration, the overall time needed to perform a Sau-PCR experiment is equivalent to that needed for the RAPD technique, with the sole addition of the time required for the enzymatic digestion, and it is much shorter than that necessary for an AFLP analysis.

In the present work, only the use of one primer at a time was investigated, but the possibility of coupling different primers having equal or different numbers and/or types of extensions could be explored in the search for the optimal level of fingerprint complexity. In addition, profiles from Sau-PCR amplifications of the same DNA with different primers could be combined and analyzed together as a unique fingerprint.

This technique could be extended to virtually all organisms from which genomic DNA can be extracted, but the amplification protocol might need some adjustments in these cases. With regard to the restriction endonuclease, other enzymes could also be tested, even though those commercially available that have the features described for Sau3AI, namely, MboI (isoschizomer of Sau3AI), FatI (recognition sequence CATG), and TspEI (recognition sequence AATT), are few, as revealed by a search in the REBASE database (http://rebase.neb.com).

In conclusion, we have developed a new technique, named Sau-PCR, which could be considered for DNA fingerprinting-based analyses as a possible alternative to the RAPD technique in cases where reproducibility or polymorphism levels are not satisfactory and as an alternative to the AFLP technique, but with lower costs in terms of time and equipment, when a restriction-plus-amplification approach is preferred.

A preliminary account of this work was presented at the 18th International ICFMH Symposium, "Microbial Adaptation to Changing Environments," held in Lillehammer, Norway, on 18 to 23 August 2002.

Marco Aloisantoni is gratefully acknowledged for technical assistance. We are grateful to Andrea Squartini for helpful comments.

This work was funded in part by a "Progetto di Ateneo" of the University of Padova.

REFERENCES

1. **Andrighetto, C., L. Zampese, and A. Lombardi.** 2001. RAPD-PCR characterization of lactobacilli isolated from artisanal meat plants and traditional fermented sausages of Veneto region (Italy). Lett. Appl. Microbiol. **33:**26–30.
2. **Arens, P., H. Coops, J. Jansen, and B. Vosman.** 1998. Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. Mol. Ecol. **7:**11–18.
3. **Bagley, M. J., S. L. Anderson, and B. May.** 2001. Choice of methodology for assessing genetic impacts of environmental stressors: polymorphism and reproducibility of RAPD and AFLP fingerprints. Ecotoxicology **10:**239–244.
4. **Corich, V., A. Mattiazzi, E. Soldati, A. Carraro, and A. Giacomini.** 2004. Relationships between chemical and microbiological composition of commercial plain yogurts. Ital. J. Food Sci. **2:**221–233.
5. **Gaudeul, M., P. Taberlet, and I. Till-Bottraud.** 2000. Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. Mol. Ecol. **9:**1625–1637.
6. **Gillings, M., and M. Holley.** 1997. Amplification of anonymous DNA fragments using pairs of long primers generates reproducible DNA fingerprints that are sensitive to genetic variation. Electrophoresis **18:**1512–1518.
7. **Gillings, M., and M. Holley.** 1997. Repetitive element PCR fingerprinting (rep-PCR) using enterobacterial repetitive intergenic consensus (ERIC) primers is not necessarily directed at ERIC elements. Lett. Appl. Microbiol. **25:**17–21.
8. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
9. **Hulton, C. J. S., C. G. Higgins, and P. M. Sharp.** 1991. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. Mol. Microbiol. **5:**825–834.
10. **Jones, C. J., K. J. Edwards, S. Castaglione, M. O. Winfield, F. Sala, C. van de Wiel, G. Bredemeijer, B. Vosman, M. Matthes, A. Daly, R. Brettschneider, P. Bettini, M. Buiatti, E. Maestri, A. Malcevschi, N. Marmiroli, R. Aert, G. Volckaert, J. Rueda, R. Linacero, A. Vazquez, and A. Karp.** 1997. Reproducibility of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol. Breeding **3:**381–390.
11. **MacPherson, J., P. Eckstein, G. Scoles, and A. Gajadhar.** 1993. Variability of

the random amplified polymorphic DNA assay among thermal cyclers, and effects of primer and DNA concentration. Mol. Cell. Probes **7:**293–299.
12. **Meunier, J., and P. Grimont.** 1993. Factors affecting reproducibility of random amplified polymorphic DNA fingerprinting. Res. Microbiol. **144:**373–379.
13. **Mueller, U. G., and L. L. Wolfenbarger.** 1999. AFLP genotyping and fingerprinting. Trends Ecol. Evol. **14:**389–394.
14. **Niemann, S., T. Dammann-Kalinowski, A. Nagel, A. Pühler, and W. Selbitschka.** 1999. Genetic basis of enterobacterial repetitive intergenic consensus (ERIC)-PCR fingerprint pattern in *Sinorhizobium meliloti* and identification of *S. meliloti* employing PCR primers derived from an ERIC-PCR fragment. Arch. Microbiol. **172:**22–30.
15. **Querol, A., E. Barrio, and D. Ramon.** 1992. A comparative study of different methods of yeast strain characterization. Syst. Appl. Microbiol. **15:**439–446.
16. **Ramos, M. S., and S. K. Harlander.** 1990. DNA fingerprinting of lactococci and streptococci used in dairy fermentation. Appl. Microbiol. Biotechnol. **34:**368–374.
17. **Romesburg, H. C.** 1984. Cluster analysis for researchers. Lifetime Learning Publications, Belmont, Calif.
18. **Schweder, M., R. Shatters, S. West, and R. Smith.** 1995. Effect of transition interval between melting and annealing temperatures on RAPD analyses. BioTechniques **19:**38–42.
19. **Sneath. P. H. A., and R. R. Sokal.** 1973. Numerical taxonomy. W. H. Freeman and Co., San Francisco, Calif.
20. **Tanaka, J., and F. Taniguchi.** 2002. Emphasized-RAPD (e-RAPD): a simple and efficient technique to make RAPD bands clearer. Breeding Sci. **52:**225–229.
21. **Tyler, K. D., G. Wang, S. D. Tyler, and W. M. Johnson.** 1997. Factors affecting reliability and reproducibility of amplification-based DNA fingerprinting of representative bacterial pathogens. J. Clin. Microbiol. **35:**339–346.
22. **Vincze, T., J. Posfai, and R. J. Roberts.** 2003. NEBcutter: a program to cleave DNA with restriction enzymes. Nucleic Acids Res. **31:**3688–3691.
23. **Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van der Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau.** 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acid Res. **23:**4407–4414.
24. **Welsh, J., and M. McClelland.** 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. **18:**7213–7218.