ELSEVIER

FEMS Microbiology Letters 248 (2005) 101–109

FEMS MICROBIOLOGY Letters

www.fems-microbiology.org

Minisatellite polymorphism as a tool to distinguish closely related *Lactococcus lactis* strains

Pascal Quénée [a], Elodie Lepage [a], Woojin Scott Kim [b], Gilles Vergnaud [c,d], Alexandra Gruss [a,*]

[a] *Recherches Laitières et Génétique Appliquée – URLGA, Institut National de la Recherche Agronomique, Domaine de Vilvert, 78352 Jouy en Josas Cedex, France*

[b] *School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW 2052, Australia*

[c] *Laboratoire GPMS, Institut de Génétique et Microbiologie, Bat 400, Universite Paris-Sud, 91405 Orsay Cedex, France*

[d] *Division d'Analyses Microbiologiques, Centre d'Etudes du Bouchet BP3, 91710 Vert le Petit, France*

Received 5 January 2005; received in revised form 11 March 2005; accepted 16 May 2005

First published online 31 May 2005

Edited by A. Klier

Abstract

Genome plasticity is considered as a means for bacteria to adapt to their environment. Plasticity in tandem repeat sequences on bacterial genomes has been recently exploited to trace the epidemiology of pathogens. Here, we examine the utility of minisatellite (i.e., a repeat unit of six nucleotides or more) typing in non-pathogenic food bacteria of the species *Lactococcus lactis*. Thirty-four minisatellites identified on the sequenced *L. lactis* ssp. *lactis* strain IL1403 genome were first analyzed in 10 closely related ssp. *lactis* strains, as determined by randomly amplified polymorphic DNA (RAPD). The selected tandem repeats varied in length, percent identity between repeats, and locations. We showed that: (i) the greatest polymorphism was in *orfs* encoding exported proteins or in intergenic regions; (ii) two thirds of minisatellites were little- or non-variable, despite as much as 90% identity between tandem repeats; and (iii) dendrograms based on either RAPD or minisatellite analyses were similar. Seven minisatellites identified in this study are potentially useful for lactococcal typing. We then asked whether tandem repeats in *L. lactis* were stable upon very long-term (up to two years) storage. Despite large rearrangements previously reported in derivative strains, just one of 10 minisatellites tested underwent an alteration, suggesting that tandem repeat rearrangements probably occur during active DNA replication. We conclude that multiple locus minisatellite analysis can be a valuable tool to follow lactococcal strain diversity.
© 2005 Published by Elsevier B.V. on behalf of the Federation of European Microbiological Societies.

Keywords: Biodiversity; Polymorphism; DNA stability; RAPD; MLVA

1. Introduction

Genomic direct tandem repeat sequences (microsatellites, 1–5 nucleotides; minisatellites, six or more nucleotides) [1] are common, and have long been used in eukaryotes for typing and detection of genetic diseases [2]. More recently, analyses of tandem DNA repeats has been applied to phylogenetic and epidemiological typing of prokaryotes, mainly pathogens [3,4], including *Bacillus anthracis* [5], *Haemophilus influenzae* [6], *Staphylococcus aureus* [7] and *Mycobacterium tuberculosis* species [8]. Those studies have proven valuable in tracing the origins of disease-causing organisms and the sources of outbreaks.

* Corresponding author. Tel.: +33 1 34 65 21 68; fax: +33 1 34 65 20 65.

E-mail addresses: alexandra.gruss@jouy.inra.fr (A. Gruss).

0378-1097/$22.00 © 2005 Published by Elsevier B.V. on behalf of the Federation of European Microbiological Societies.
doi:10.1016/j.femsle.2005.05.026

Tandem repeat polymorphism may also affect protein functions or gene expression (see [9], for review). Minisatellites occurring in *orfs* may encode functionally relevant peptide repeats, such as in antigenic or phase variation, e.g., the *Streptococcus agalactiae* alpha C protein [10] or the *H. influenzae* hemoglobin binding proteins [11]. In *S. agalactiae*, the surface protein alpha3 C undergoes deletions within a 246 nucleotide repeat unit when passaged through mice co-injected with anti-alpha C antibody, suggesting that stress conditions could create an environment that favors emergence of deletions [10]. Polymorphism in the short (microsatellite) repeats has also been observed within a collection of non-pathogenic *Escherichia coli* strains [12], suggesting that polymorphism occurs for reasons other than antigenic escape, and most likely has a more general role in adaptation. Indeed, the rate of microsatellite polymorphism was stimulated by addition of a stress-inducing agent, hydrogen peroxide [13]. Microsatellites have been shown to affect specificity of an oxidoreductase, such that a reversible expansion results in a change of enzymatic activity [14]. Also, their presence in intergenic regions can have a regulatory role, affecting gene expression [15]. Thus, tandem repeat instability may facilitate rapid adaptation to different environmental conditions, and in pathogens, mediate antigenic escape.

Our goal was to examine the feasibility of using multiple locus minisatellite typing (referred to as MLVA for Multiple Locus Variable number of tandem repeat Analysis) on mainly alimentary, rather than pathogenic bacteria for molecular typing of isolates of the species *Lactococcus lactis*. *L. lactis* strains were initially classified phenotypically into two main subspecies, *lactis* and *cremoris*. More recently, they were analyzed using taxonomic tools; these include DNA–DNA hybridization [16], gene polymorphism [17,18], and hybridization with rRNA-specific oligonucleotide probes [19,20] or with cloned genes [21]. In the case of more closely related strains, pulse field gel electrophoresis was used to detect events such as large inversions [22]. The most comprehensive study of *L. lactis* diversity was performed using randomly amplified polymorphic DNA (RAPD) [23], where some 280 strains were analyzed. RAPD results of some of these strains classified Lactococci into three broad groups, referred to as G1, G2, and G3: G1, the largest group, contained subsp. *lactis* strains; G2 corresponded to subsp. *cremoris*, and G3 to subsp. *lactis* strains, although patterns of the G3 group clearly differed from G1.

We performed minisatellite studies on 10 *L. lactis* subsp. *lactis* strains that were previously typed by RAPD analysis [23] and belong to the G1 group. Our results show that these more closely related lactococci can be differentiated using minisatellite analysis but that the genetic diversity within the species makes difficult the design of primers able to amplify all the strains, as was previously observed, for instance, in *Legionella pneumophila* [24]. Interestingly, the most polymorphic repeats were all located in *orfs* encoding exported proteins, or in intergenic regions. Seven minisatellites were identified as being useful for further studies of polymorphism. We also followed minisatellite stability within a single strain that was maintained under starvation conditions over a two-year period. With one exception, minisatellites were stable under long-term conservation conditions.

2. Materials and methods

2.1. Strains and media

We selected 10 strains (Table 1), including IL1403, that are closely related by RAPD typing, and were previously classified into subsp. *lactis* group G1 (see Tailliez et al. [23]). Phenotypic tests were previously used to

Table 1
Strains used in this study[a]

Strain name	Identification	Origin
IL 1403	*L. lactis* subsp. *lactis diacetylactis*	Poligny, France
IL 585	*L. lactis* subsp. *lactis diacetylactis*	Commercial starter
IL 938	*L. lactis* subsp. *lactis diacetylactis*	Commercial starter
IL 583	*L. lactis* subsp. *lactis diacetylactis*	Commercial starter
IL 933	*L. lactis* subsp. *lactis*	Raw milk, France
IL 953	*L. lactis* subsp. *lactis*	Commercial starter
IL 1321	*L. lactis* subsp. *lactis*	Milk, Mexico
IL 738	*L. lactis* subsp. *lactis*	Cream, France
IL 801 (CNRZ 142T[b])	*L. lactis* subsp. *lactis*	Orla-Jensen, Denmark
IL 2961	*L. lactis* subsp. *lactis*	Radish, China
LL41-1 and derivatives	*L. lactis* subsp. *lactis*	[25]

[a] Strains other than LL41-1 and derivatives were kindly given by M.C. Chopin, this institute. They are referred to under the following names in the publication of Tailliez et al.: IL938 = A81; IL933 = A61; IL953 = A108; IL738 = A310; IL801 = NCDO604; IL2961 = NCDO2091.
[b] *L. lactis* subsp. *lactis* type strain.

identify the biovar *diacetylactis* strains [23]. Strains were isolated on M17 solid medium, propagated in M17 broth containing 1% glucose at 30 °C, and stored at −20 or −80 °C in M17 broth containing 15% glycerol. To follow minisatellite stability in long-term surviving strains, we used strain LL41-1 (CRC Culture Collection), three derivatives isolated after one-year storage, and one derivative isolated after two-year storage [25].

2.2. DNA preparation

A single colony of the strain stocks first isolated on plates was used to inoculate cultures. Total cellular DNA was extracted from late exponential phase cultures inoculated with a single colony as previously described [26]. Total DNA was quantified by agarose gel (0.8%) electrophoresis.

2.3. Bacterial genome DNA sequence, minisatellite finder

The *L. lactis* subsp. *lactis* strain IL1403 genome sequence ([27]; http://www.ncbi.nlm.nih.gov/) was used to identify minisatellites (http://minisatellites.u-psud.fr/). The following selection criteria were used: unit length between 6 and 285 bp, copy number of at least 2, and variable percent identity, i.e., between 48% and 100% identity between repeats. Primers used for minisatellite identification are listed in Table 2, and were determined using the Primer3 program (http://frodo.wi.mit.edu).

2.4. Minisatellite PCR amplification and electrophoresis

PCR amplifications were carried out in 15 μl containing 1–10 ng of DNA, 1.5 mM $MgCl_2$, 0.3 μM of each flanking primer, 200 μM of each dNTP (Boehringer Mannheim, France), 1.25 unit of *Taq* DNA polymerase (Qbiogen, France) in 10 mM Tris–HCl, pH 9. A Perkin–Elmer 2400 thermal cycler was used to PCR-amplify minisatellite sequences, using the following program: 96 °C for 5 min, followed by 30 cycles comprising 96 °C for 30 s, then 30 s annealing at appropriate temperature, as determined by T_m values (between 54 and 64 °C), and 72 °C (between 30 s and 2 min, according to the PCR product size). PCR reaction products (5 μl) were deposited in agarose gels (1% or 2% Seakem GTG, or 3% Metaphor agarose; Tebu, France), prepared in Tris-borate–EDTA buffer. A 123 bp ladder (Gibco-BRL, France) was used to estimate fragment sizes greater than 150–200 bp in the 1% and 2% gels, and a 10 bp DNA ladder (Invitrogen, USA) was used for fragments smaller than 150–200 bp. To confirm that amplification was indeed specific, eight PCR products were chosen at random and sequenced; in all cases, the sequence was that expected for minisatellite amplification (data not shown).

Table 2
Primers used for minisatellite amplifications (5′–3′)

ms#	Sequences
ms02	L: ACTAGACAGCAACGGTCAAGTTGTT R: TGCATCAGTTTTGATCCCCTTAATA
ms04	L: AGAAGTGTTAAAAGCAAGTCAAGCA R: TTTTTCTCTTCGAGTTCTGCCTTAG
ms06	L: GATAGTATCATTCACCGCATGGAC R: TCTTGGATAATAAAACACCCGCTAA
ms08	L: TAATGGTGAAGGAAAAGGCTTCAA R: AAACGACGGTCATCACGTTTTC
ms11	L: TGAAGAAGTTCCTGCTGAATAAAAA R: TTTACTGACAGAGCTGCAATTCAT
ms12	L: GAAGGCAAAGAAGTTGTGTATTTGA R: GGACGATTGCGTATTTTATGAATTT
ms15	L: GACGAAGAAGAAGAAAAACCAGCTA R: GCTGATGCCTCACTTTTTATGTGTT
ms17	L: TCAAAGATGTCAGTGAATTAGAAGAGC R: CATTGTAAATGAATCTCCTGCTCTCA
ms18	L: AATGCAGCTTCAGGAAACACTAAAG R: GGTTTTTATATGTCTTTTTCGCGAGT
ms19	L: CAGAAAAAACCTAAAGCTAGTGAACCA R: AGCAACATTTTCTGGTTTTACTTCA
ms21	L: TGGTGCTTCAACAGTTAATCCAAAT R: CTCATTGCGCTTAATGAACTTGAAA
ms22	L: GAATTTGGTGACGAAGCCTTTGAT R: AGATTAATTTCCGCCGCTTCAATTT
ms24	L: AAGTTTTGGCTTCAATGGATATTGT R: GATTTCATTACGATAATCGACTCGT
ms27	L: TAAGAATCAACCAGAAAAACGCTTG R: ATTATTTTGATTTTGGCCTGCATT
ms28	L: ATTCACATCATGCGTCTTAATGAAA R: AGTTCTGTCAATTCCATCGTGTTTT
ms30	L: AAATGACAGAAGCAATCCTGAAAAA R: ATCACATGCTGATCCCAAAGTTTAT
ms32	L: AACGGAATCATTTTGTTTTGTTGT R: AGTTGCGGTAATTTCTTTGCTGAC
ms37	L: ACAGTCTCTCCTTGCTCTTTAGCTG R: GGCTGACGATGAATATAAAGGTGTT
ms40	L: CAGAATTATTTGGCGTCGGTGTA R: ACAGGTGGCGGAAATACAACTAAC
ms42	L: ATGACATATCGGGCATTCCAGAAC R: TTCATTAAGCAATTCAATCAGTCCA
ms46	L: TCAAACATTTTAACATGAGTCCCTTT R: ATATTGTGTGAATGGTCATGGAAAA
ms47	L: GGTGCTTGAGTTGTTGTAGAGGAAT R: TGATTTTAGGAATGACGACTTTTGG
ms49	L: TTTTATCACTTTCATAAGATTTGTCAAC R: GGTATAGATAAAGTGTGTTCCGACAAA
ms52	L: GCCATATGAAAAGGCACCTAAAA R: ATTTTGATAAAGCTGTAATTCATCTAGTTT
ms53	L: CACTCATTTTACCAGCGTTATTGAA R: GGAGTGATGATGTATGGACAACAAA

(*continued on next page*)

Table 2 (*continued*)

ms#	Sequences
ms54	L: CCTCCTCCAAAGAAACTTGAGAAT R: CTGCTTATGACCAATATGGTGAAG
msA1	L: CGCAGTTAACACAGTCAGCCCTAAA R: AAGCCCGAATTCTCAGTCTTTTGAC
msA2	L: ATGCTTCAAGTAGGACAAGTATTGC R: ATGCTGTCACCATTTTGAATTGTAT
msA3	L: GATTCATCGGTTGAGAAGAATGAAA R: TTCTTCAATACTCGTGGAATTGACC
msA4	L: TTTGTAATTTAAAAGACAGGATTTGATT R: CCTCAGTCTTAATTTTGAACTGTTCTT
msA5	L: CCATTCTTTCTTTTCCATATAGGTCA R: ACAGTGACTATTTAGCTGTCAAAACG
msA6	L: TTTCCAAATTAGGGCTACTACTAAGG R: AATAATTCATCAACTCATTCTTCAAA
msA7	L: TTTGAGCTTTTGTGCTTGCTTG R: CAAGTGCGCAGTCTGCAAAG
msA8	L: TCTGAATCTGAACTATCATTTGAACCA R: CACCATCAAATCCGACAGAAAGTAA

2.5. *MLVA data analysis*

For each minisatellite, polymorphism was observed as a change in band migration on agarose gels compared to migration of the IL1403 control. The number of IL1403 amplicons were as expected from the genome sequence; figures presented were rounded off to the nearest whole numbers, as by convention. Minisatellite amplicons from the other strains displayed sizes corresponding to the gain or loss of one or more repeat units, with sometimes clear evidence for intermediate sizes, which were coded as ##.5, or as ##.3 and ##.7, if two distinct intermediate sizes were distinguished (for ms40). Lack of amplification (confirmed by at least two attempts) was coded as 0, and was given the equivalent weight in the analyses as an amplified segment. The resulting data were analyzed as a character dataset using the Bionumerics program (Applied-Maths, Belgium). Similarity values were based on the categorical coefficient and clustering of patterns were determined using the unweighted pair group method with arithmetic averages (UPGMA) [28].

3. Results and discussion

3.1. *Minisatellite analysis of 10* L. lactis *subsp.* lactis *strains*

Thirty-four minisatellites were selected on the IL1403 genome sequence (criteria and programs used are described in Section 2). Two criteria were applied to select minisatellites: a set of 26 minisatellites were chosen for a minimal unit length of 9 bp, and a copy number of at least 5; the percent identity between repeats was between 48% and 96%. Eight additional minisatellites were chosen for homologies between repeats of over 90%, without other restrictions. Note that the inter- or intragenic locations were not pre-screened in the selection of minisatellites under study.

Our laboratory previously classified 280 *L. lactis* strains present in our strain collections by RAPD analyses [23]. These strains vary widely in their origins, both in terms of geographical location and sample (isolated from milk, cream, cheese, industrial starter cultures, or plants). From this characterized collection, we initially selected 10 *L. lactis* strains representative of the three groups classified by Tailliez et al. [23]. In those preliminary experiments, we noted that most of the chosen minisatellites did not give an amplification product from six strains outside of the 'G1' group. We therefore restricted the study to 10 more closely related *L. lactis* subsp. *lactis* strains from the 'G1' group (Table 1) to examine heterogeneity of minisatellite profiles.

The minisatellite profiles obtained with the lactococcal strains could be classified into three groups (exemplified in Fig. 1, and not shown). Twelve minisatellites (35%) were highly polymorphic (polymorphic group; Fig. 1(a)). Twenty minisatellites (~60%) were invariable, or contained just one polymorphic sequence out of 10 (invariable group). Two minisatellites presumably correspond to poorly conserved, or 'host-specific' regions (host-specific group; Fig. 1(b)); in this group, just one, or two strains generated amplified fragments. Seven minisatellites, selected as generating at least four different sized amplicons (i.e., with a diversity index of greater than 0.75; see Fig. 2 and [29]), and as amplifying DNA of nine or all strains (indicated by * in Fig. 1), are potentially useful for future *L. lactis* strain typing.

3.2. *Characteristics of minisatellites*

I. Polymorphic minisatellites are often in orfs encoding putative exported proteins. The selected repeats were examined for whether they were localized in extragenic or intragenic regions. For the 19 out of 34 minisatellites located in *orfs*, the potential functions and locations of the gene products were determined (using http://www.ncbi.nlm.nih.gov/BLAST/ and 'DAS' Transmembrane Prediction server http://www.sbc.su.se/~miklos/DAS/; Table 3). Overall, we noted that most intragenic minisatellites are present in *orfs* encoding putatively exported proteins. Significantly, among the polymorphic minisatellites, all intragenic repeats are in export protein *orfs*. This observation reflects previous reports in pathogens, where minisatellite sequence variation is associated with 'antigen escape' by rapid surface protein modification (referenced in Section 1). As lactococci are mainly non-pathogenic, we suggest that the potential surface

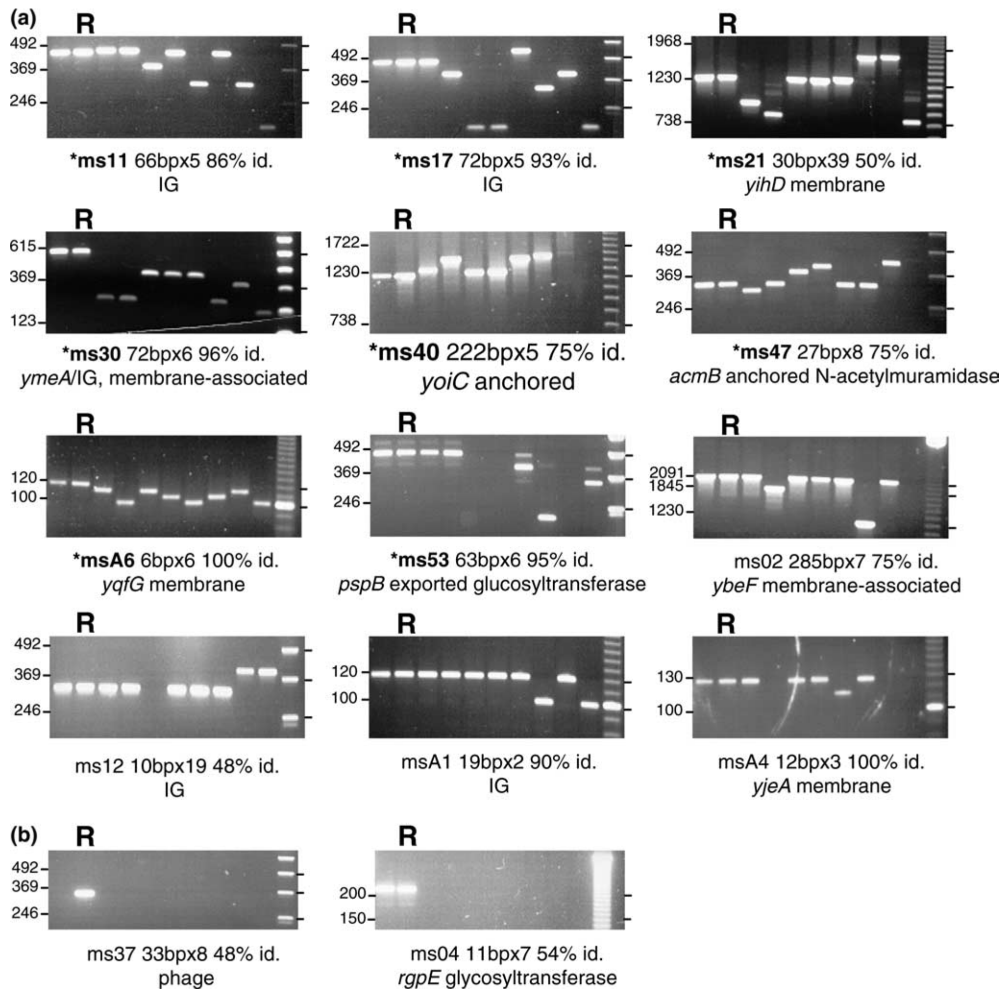

Fig. 1. Examples of minisatellite polymorphism in *L. lactis*. (a) Polymorphic minisatellites. From 34 tested minisatellites, the 12 shown above were the most variable. (b) Strain-specific minisatellites. Note that ms37 putatively corresponds to a phage *orf* and is specific to IL1403. Below each tested minisatellite, the following information is given: the ms number (as per Table 3); length of the repeat and the estimated number of times it is present in the reference strain (marked 'R'); the percent identity between repeat sequences; the gene name of *orf* containing the repeat, or 'IG' for intergenic sequence; if relevant, the putative function of the *orf*, and its hypothetical location. Strains tested, from left to right, are IL585, IL1403 (R, sequenced reference strain), IL938, IL583, IL933, IL953, IL1321, IL738, type strain CNRZ142T, and IL2961. DNA marker sizes (in bp) are indicated for each gel.

variability conferred by minisatellite variation may nonetheless be associated with environmental adaptation. Notably, amplification or deletion in these minisatellites would maintain the coding frame, as the unit length of repeats were multiples of three (Table 3).

II. Invariable minisatellites. Twenty minisatellites were invariable in the strains tested, or showed very low polymorphism. Nearly all the invariable minisatellites were located in *orfs*; about half were embedded in *orfs* encoding exported proteins, and several others were present in stress or repair-related *orfs*. The invariability of these minisatellites may reflect a role of amino acid repeats in protein function rather than in allowing gene plasticity in stress conditions.

III. Strain-specific minisatellites. One of the two 'strain-specific' minisatellites (Fig. 1(b)) is embedded in an *orf* encoding a putative phage function. It is thus not surprising that it was present in just one of the 10 tested strains. The other minisatellite was present in a conserved *orf* (encoding glycosyl transferase); the inability to amplify this minisatellite may be due to the choice of primers and/or genetic variability in this region.

IV. Minisatellite variability is greater in longer tandem repeats. In general, the larger repetitions (21–285 bp) were among the more variable minisatellites, even when percent identity between repeats was low. In contrast, the short tandem repeats (<21) were mostly invariable, despite identity of up to 90% (Table 3 and Fig. 2).

3.3. Comparison of dendrograms derived from MLVA versus RAPD analyses

Dendrograms to evaluate strain relatedness were generated using the ensemble of minisatellite results on the

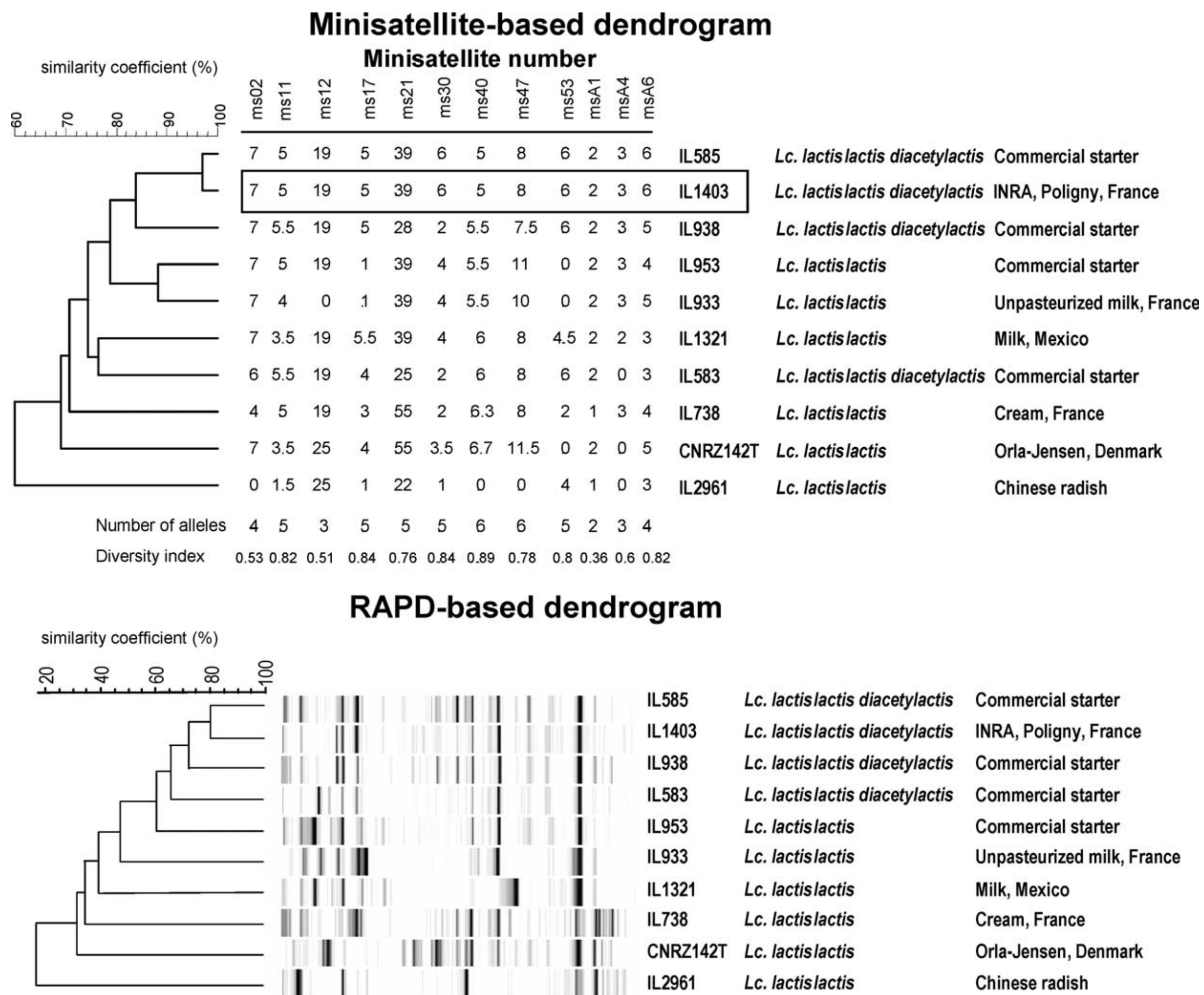

Fig. 2. Dendrograms of 10 *L. lactis* strains based on MLVA and RAPD analyses. Minisatellites were classified using the Bionumerics program (Applied-Maths, Belgium), and as in 'MLVA data analysis' in Section 2. The dendrogram was generated using results of all minisatellites. The estimated numbers of repeats are presented for the 12 polymorphic minisatellites (see Table 3); '0' indicates that no amplification was obtained, and a fraction indicates the amplicon likely contains a partial repeat segment. Bands obtained with IL1403 served as standards (in rectangle). Strains IL585 and IL1403 were discriminated using minisatellite ms37. Below the MLVA dendrogram for the 10 tested strains, are given the number of alleles present per minisatellite, and the diversity index (i.e., probability that any two consecutively sampled strains would be of a different type) [29]. The RAPD dendrogram was obtained by comparison (Program Gel Compare, Applied-Maths, Belgium) and clustering (UPGMA method) of merged RAPD patterns (reported in [23]); data was extracted from earlier comprehensive studies of 280 strains, reported in [23].

10 tested strains, and compared to those obtained by RAPD (Fig. 2). The UPGMA methods used to generate dendrograms produce a hierarchy of profiles, going from higher to lower similarity coefficients. Except for moderate differences (e.g., the position of IL583 differs in MLVA and RAPD dendrograms), the two dendrograms globally give similar classifications of the 10 strains. The MLVA dendrogram does not group the *diacetylactis* isolates. However, (i) the *diacetylactis* designation is associated with the presence of a plasmid encoding citrate permease [30] and (ii) in a more extensive RAPD typing survey, *diacetylactis* and other *lactis* isolates are mixed [23]. Minisatellite typing may, however, give additional data; using RAPD, strains having more than 80% similarity are considered as indistinguishable (due to limits in reproducibility of the technique; [31]). Thus, while IL585 and IL1403 are considered as the same strain by RAPD, they were distinguished using ms37 (within a phage gene), indicating that they differ at the very least by their state of lysogeny (Fig. 1(b)). As a more reproducible method, minisatellite-based typing may provide more accurate results, and thus more objective means to distinguish strains, compared to RAPD.

3.4. *Minisatellite analysis of derivatives of a single* L. lactis *subsp.* lactis *species*

The long-term survival capacity of an *L. lactis* subsp. *lactis* strain, LL41-1, was previously examined [25].

Table 3
Minisatellite characteristics[a]

Gene (ms#[b])	Unit length	Copy no.	Repeat identity (%)	In predicted Orf	Hypothetical localization[c]	Repeat position[d]
Polymorphic						
ybeF (ms2)	285	7	75	Hypothetical	Membrane	152433–154399
yoiC (*ms40)	222	5	75	Hypothetical	Anchored	1484740–1485849
pspB (ms53)	63	6	95	Glucosyltransferase	Secreted	2305857–2306205
yihD (*ms21)	30	39	50	Hypothetical	Membrane	875349–876529
acmB (*ms47)	27	8	75	*N*-acetylmuramidase	Anchored	1977060–1977272
yjeA (msA4)	12	3	100	Hypothetical	Membrane	939926–939959
yqfG (*msA6)	6	6	100	Hypothetical	Membrane	1659843–1659875
ymeA/IG[e] (*ms30)	72	6	96	Hypothetical	Membrane	1242385–1242**826**
IG (*ms17)	72	5	93			757516–757893
IG (*ms11)	66	5	86			552732–553041
IG (msA1)	19	2	90			160209–160247
IG (ms12)	10	19	48			581757–581958
Invariable						
yqfG (msA8)	282	2	88	Hypothetical	Membrane	1659921–1660627
usp45 (msA7)	21	2	90	Hypothetical	Secreted	2314380–2314429
ylbB (ms27)	21	5	66	Transporter permease	Membrane	1112180–1112275
ypjG (msA5)	18	2	90	ABC transporter ATP bind. protein	Membrane	1595875–1595911
potA (ms28)	18	6	54	ABC transporter ATP bind. protein	Membrane	1176730–1176827
ychF (ms06)	16	5	57	ABC transporter permease protein	Membrane	277700–277771
acmD (msA2)	15	2	93	*N*-acetylmuramidase	Cell wall	528335–528364
ponA (msA3)	15	2	93	Cell wall crosslink	Anchored	530811–530841
recN (ms22)	13	5	58	DNA repair protein RecN	Cytoplasmic	883600–883664
ffh (ms42)	12	5	69	Signal recognition particle Ffh	Cytoplasmic	1657304–1657362
infB (ms19)	12	5	66	Transl. Init. factor IF-2	Cytoplasmic	776186–776247
thiE (ms32)	12	5	60	Thiamine-phosphate pyrophosphorylase	Cytoplasmic	1293**606**–1293662
uvrA (ms46)	12	5	59	Exinuclease ABC subunit A	Cytoplasmic	1887989–1888046
vacB1 (ms24)	12	5	56	Ribonuclease	Cytoplasmic	967721–967777
rheA (ms08)	12	11	55	ATP-dependent RNA helicase	Cytoplasmic	354435–354557
dnaJ (ms54)	9	9	62	DnaJ protein/heat shock	Cytoplasmic	2308067–2308149
rpoE (ms15)	9	22	51	RNA polymerase	Cytoplasmic	623908–624105
lmrP/IG (ms52)	9	8	59	Integral membrane protein LmrP	End *lmrP*/start IG	2242265–2242**343**
IG/*pdc* (ms49)	13	6	56	IG + start Pdc		2010511–20105**94**
IG (ms18)	17	9	55			763633–763790
Host-specific						
pi314 (ms37)	33	8	48	Bacteriophage tail lysin	Membrane	1427347–1427610
rgpE (ms04)	11	7	54	Glycosyltransferase	Secreted	207639–207721

[a] All gene assignments, minisatellite unit and repeat lengths, and genome positions correspond to IL1403 sequence annotation (http://www.ncbi.nlm.nih.gov/genomes/framik.cgi?db=Genome&gi=171; [27]).
[b] The minisatellite number given in parentheses (msXX) is as used in figures; *, minisatellites showing the greatest polymorphism.
[c] Location assignments are based on BLAST and hydrophobicity analyses (using http://www.ncbi.nlm.nih.gov/BLAST/ and 'DAS' Transmembrane Prediction server http://www.sbc.su.se/~miklos/DAS/).
[d] Positions correspond to IL1403 sequence (http://www.ncbi.nlm.nih.gov/genomes/framik.cgi?db=Genome&gi=171; [27]); positions in bold indicate that minisatellites span the beginning or end of the indicated gene and the adjacent intergenic region.
[e] IG, intergenic region. Minisatellite covers both *orf* and IG region, where *orf*/IG, or IG/*orf* is indicated.

Derivatives of this strain could be isolated after a one- or two-year period in starvation conditions if the initial medium was low in glucose. However, cell morphologies were greatly altered, and growth of isolates was very slow. The one- and two-year isolates underwent both chromosomal and plasmid rearrangements, as detected by pulse-field gel electrophoresis [25]. As minisatellites expectedly correspond to unstable regions of the DNA genome, we considered that they might be particularly unstable and susceptible to amplification or deletion under the extreme stress conditions imposed by long-term starvation. We therefore looked for minisatellite rearrangements in three surviving variants isolated after one year, and one variant isolated after two years, by comparing 10 profiles with those of the LL41-1 parent (ms02, ms11, ms12, ms17, ms21, ms30, ms40, ms47, ms53, and ms54; data not shown). We were initially surprised that nearly no minisatellite variation was seen in the isolates having undergone extreme starvation. Just one of the 10 minisatellites tested (ms21) showed variation in the two-year old surviving isolate (an estimated 57 copies in the parent, compared to 35 copies in the

two-year isolate). The gene affected, *yihD*, encodes a 1063 amino acid, serine-rich putative surface protein. Possibly, the loss of repeats in *yihD* could have a physiological role in improving strain survival.

Why are minisatellites so stable under extreme long-term starvation conditions? We suggest that the lack of, or very slow DNA replication under these conditions reduces the frequency of replication slippage, a proposed means of generating tandem repeat rearrangements [1]. In contrast, large rearrangements or deletions might be replication-independent and would occur in non-replicative conditions. Further studies in this area may reveal whether minisatellite amplifications and deletions in lactococci occur primarily during periods of active growth and DNA replication.

4. Conclusions

These results show that relatedness of lactococcal isolates can be estimated according to the polymorphism of genomic tandem repeat sequences. Seven minisatellite markers, which generated a band for most or all tested strains but were nevertheless highly polymorphic, could constitute useful markers for future MLVA testing. An *L. lactis* subsp. *cremoris* genome sequence (two genome sequencing projects are in progress; SK11 and MG1363) will be valuable in generating minisatellite probes with a larger spectrum. A very small number of lactococci have been isolated from animal sources [32]. MLVA analyses as presented here could be a useful means of differentiating the food strains from potentially harmful isolates of this species. Our results suggest that tandem repeat rearrangements occur infrequently during stationary phase conservation, which is of particular interest for strain identification purposes.

Acknowledgements

We are grateful to Patrick Tailliez and Jean-Claude Ogier for valuable discussions in the course of this study. This work was supported in part by a grant from the 'Programme Recherche Fondamentale en Microbiologie: Maladies Infectieuses et Parasitaires'.

References

[1] Bennett, P. (2000). Mol. Pathol. 53, 177–183.
[2] Yauk, C.L. (2004) Advances in the application of germline tandem repeat instability for in situ monitoring. Mutat. Res. 566, 169–182.
[3] van Belkum, A. (1999) Short sequence repeats in microbial pathogenesis and evolution. Cell Mol. Life Sci. 56, 729–734.
[4] Pourcel, C., Andre-Mazeaud, F., Neubauer, H., Ramisse, F. and Vergnaud, G. (2004) Tandem repeats analysis for the high resolution phylogenetic analysis of *Yersinia pestis*. BMC Microbiol. 4, 22.
[5] Denoeud, F. and Vergnaud, G. (2004) Identification of polymorphic tandem repeats by direct comparison of genome sequence from different bacterial strains: a web-based resource. BMC Bioinformatics 5, 4.
[6] van Belkum, A., Scherer, S., van Leeuwen, W., Willemse, D., van Alphen, L. and Verbrugh, H. (1997) Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. Infect. Immun. 65, 5017–5027.
[7] Tohda, S., Maruyama, M. and Nara, N. (1997) Molecular typing of methicillin-resistant *Staphylococcus aureus* by polymerase chain reaction: distribution of the typed strains in hospitals. Intern. Med. 36, 694–699.
[8] Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B. and Locht, C. (2000) Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol. Microbiol. 36, 762–771.
[9] Bayliss, C.D., Dixon, K.M. and Moxon, E.R. (2004) Simple sequence repeats (microsatellites): mutational mechanisms and contributions to bacterial pathogenesis: a meeting review. FEMS Immunol. Med. Microbiol. 40, 11–19.
[10] Puopolo, K.M., Hollingshead, S.K., Carey, V.J. and Madoff, L.C. (2001) Tandem repeat deletion in the alpha C protein of group B streptococcus is *recA* independent. Infect. Immun. 69, 5037–5045.
[11] Dawid, S., Barenkamp, S.J. and St Geme 3rd, J.W. (1999) Variation in expression of the *Haemophilus influenzae* HMW adhesins: a prokaryotic system reminiscent of eukaryotes. Proc. Natl. Acad. Sci. USA 96, 1077–1082.
[12] Metzgar, D., Thomas, E., Davis, C., Field, D. and Wills, C. (2001) The microsatellites of *Escherichia coli*: rapidly evolving repetitive DNAs in a non-pathogenic prokaryote. Mol. Microbiol. 39, 183–190.
[13] Jackson, A.L. and Loeb, L.A. (2000) Microsatellite instability induced by hydrogen peroxide in *Escherichia coli*. Mutat. Res. 447, 187–198.
[14] Ritz, D., Lim, J., Reynolds, C.M., Poole, L.B. and Beckwith, J. (2001) Conversion of a peroxiredoxin into a disulfide reductase by a triplet repeat expansion. Science 294, 158–160.
[15] Murphy, G.L., Connell, T.D., Barritt, D.S., Koomey, M. and Cannon, J.G. (1989) Phase variation of gonococcal protein II: regulation of gene expression by slipped-strand mispairing of a repetitive DNA sequence. Cell 56, 539–547.
[16] Jarvis, A. and Jarvis, B.D.W. (1981) Deoxyribonucleic acid homology among lactic streptococci. Appl. Environ. Microbiol. 41, 77–83.
[17] Rodrigues, U.M., Aguirre, M., Facklam, R.R. and Collins, M.D. (1991) Specific and intraspecific molecular typing of lactococci based on polymorphism of DNA encoding rRNA. J. Appl. Bacteriol. 71, 509–516.
[18] Urbach, E., Daniels, B., Salama, M.S., Sandine, W.E. and Giovannoni, S.J. (1997) The ldh phylogeny for environmental isolates of *Lactococcus lactis* is consistent with rRNA genotypes but not with phenotypes. Appl. Environ. Microbiol. 63, 694–702.
[19] Beimfohr, C., Krause, H., Amann, R., Ludwig, W. and Schleifer, K.H. (1993) In situ identification of lactococci, enterococci and streptococci. Syst. Appl. Microbiol. 16, 450–456.
[20] Salama, M., Sandine, W. and Giovannoni, S. (1991) Development and application of oligonucleotide probes for identification of *Lactococcus lactis* subsp. *cremoris*. Appl. Environ. Microbiol. 57, 1313–1318.
[21] Godon, J., Delorme, C., Ehrlich, S.D. and Renault, P. (1992) Divergence of genomic sequences between *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris*. Appl. Environ. Microbiol. 58, 4045–4047.
[22] Le Bourgeois, P., Daveran-Mingot, M.L. and Ritzenthaler, P. (2000) Genome plasticity among related ++Lactococcus strains:

identification of genetic events associated with macrorestriction polymorphisms. J. Bacteriol. 182, 2481–2491.

[23] Tailliez, P., Tremblay, J., Ehrlich, S.D. and Chopin, A. (1998) Molecular diversity and relationship within *Lactococcus lactis*, as revealed by randomly amplified polymorphic DNA (RAPD). Syst. Appl. Microbiol. 21, 530–538.

[24] Pourcel, C., Vidgop, Y., Ramisse, F., Vergnaud, G. and Tram, C. (2003) Characterization of a tandem repeat polymorphism in *Legionella pneumophila* and its use for genotyping. J. Clin. Microbiol. 41, 1819–1826.

[25] Kim, W.S., Park, J.H., Ren, J., Su, P. and Dunn, N.W. (2001) Survival response and rearrangement of plasmid DNA of *Lactococcus lactis* during long-term starvation. Appl. Environ. Microbiol. 67, 4594–4602.

[26] De los Reyes-Gavilan, C.G., Limsowtin, G.K.Y., Tailliez, P., Séchaud, L. and Accolas, J.-P. (1992) A *Lactobacillus helveticus*-specific DNA probe detects restriction fragment length polymorphisms in this species. Appl. Environ. Microbiol. 58, 3429–3432.

[27] Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarme, K., Weissenbach, J., Ehrlich, S.D. and Sorokin, A. (2001) The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. Genome Res. 11, 731–753.

[28] Sokal, R.R. and Michener, C.D. (1958) A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull. 38, 1409–1438.

[29] Hunter, P.R. and Gaston, M.A. (1988) Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J. Clin. Microbiol. 26, 2465–2466.

[30] Kempler, G. and McKay, L. (1979) Characterization of plasmid deoxyribonucleic acid in *Streptococcus lactis* subsp. *diacetylactis*: evidence for plasmid-linked citrate utilization. Appl. Environ. Microbiol. 37, 316–323.

[31] Tailliez, P., Quénée, P. and Chopin, A. (1996) Estimation de la diversité parmi les souches de la collection CNRZ: application de la RAPD á un groupe de lactobacilles. Le Lait 76, 147–158.

[32] Elliott, J.A., Collins, M.D., Pigott, N.E. and Facklam, R.R. (1991) Differentiation of *Lactococcus lactis* and *Lactococcus garvieae* from humans by comparison of whole-cell protein patterns. J. Clin. Microbiol. 29, 2731–2734.