

## A Natural Large Chromosomal Inversion in *Lactococcus lactis* Is Mediated by Homologous Recombination between Two Insertion Sequences

MARIE-LINE DAVERAN-MINGOT, NATHALIE CAMPO, PAUL RITZENTHALER,\*  
AND PASCAL LE BOURGEOIS

Laboratoire de Microbiologie et Génétique Moléculaire du CNRS,  
Université Paul Sabatier, 31062 Toulouse, France

Received 6 May 1998/Accepted 13 July 1998

**Comparative analysis of chromosomal macrorestriction polymorphism of the two closely related *Lactococcus lactis* subsp. *cremoris* strains MG1363 and NCDO763 revealed the presence of a large inversion covering half of the genome. To determine what kind of genetic element could be implicated in this rearrangement, the two inversion junctions of MG1363 and NCDO763 chromosomes were cloned and characterized. Nucleotide sequence analysis showed the presence of one copy of the lactococcal IS905 element in each junction. Each copy of this element contained the same nucleotide mutation that inactivates the putative transposase. Comparison of the sequences surrounding the insertion sequence demonstrated that the large inversion arose from a single-step homologous recombination event between the two defective copies of the IS905 element. The large inversion presumably conferred no selective disadvantage on strain NCDO763 because this rearrangement did not alter the *oriC-terC* symmetry of the chromosome and the local genetic environment.**

During the last decade, the increasing number of available chromosome maps has prompted reconsideration of existing theories about the evolution of bacterial genome organization (7, 16). It is now established that bacteria can be classified into two groups based on genomic stability. The first group contains bacteria in which the overall genomic organization is strongly preserved, as for *Escherichia coli* (46), *Salmonella typhimurium* (34), *Clostridium perfringens* (4), *Lactococcus lactis* (30), *Mycoplasma gallisepticum* (56), *Halobacterium salinarum* (23), and *Thermus thermophilus* (54), although chromosome rearrangements do occasionally occur. The second group comprises species in which the gene order is highly rearranged, such as *Salmonella typhi* (37), *Rhodobacter capsulatus* (44), *Helicobacter pylori* (26), *Bacillus cereus* (5), *Leptospira interrogans* (61), and *Pseudomonas aeruginosa* (49).

Until recently, it was generally thought that large ( $\geq 100$ -kb) chromosomal inversions were rarely found in natural conditions. However, there is increasing evidence that this kind of genome rearrangement is not exceptional. In addition to the well-known inversion between *E. coli* K-12 and *S. typhimurium* LT2 (47), comparison of genomic maps at the interspecies level recently revealed other inversions (12, 33, 36, 38, 41, 61). For example, the chromosomes of *S. typhi* Ty2 and *S. typhimurium* LT2 differ by at least three inversions (35). Some large inversions were also found when genome maps comparisons were done at the intraspecies level (Table 1). In one study of 21 *P. aeruginosa* clone C isolates, as many as nine inversions were observed (50).

It has been postulated that ribosomal (*rm*) operons or insertion sequences (IS) are the primary targets involved in inversion formation by homologous recombination. However, experimental data have been obtained only for cultured enteric bacteria. In *E. coli*, the genetic elements implicated in five

spontaneous large inversions were characterized. Recombination between *rmD* and *rmE* operons caused the chromosomal inversion in strain W3110 (24), whereas IS elements were responsible for the remaining inversions (IS3 for strain 1485IN [27], IS5 for strains LN850 and LN1053 [39], and IS5 and IS7 for strain BHB2600 [3]). In *S. typhi*, *rm* loci were implicated in inversions solely on the basis of I-CeuI analysis (37). Only the inversion junctions of strain 1485IN have been characterized by nucleotide sequencing. In gram-positive bacteria, the inversion junctions of only one spontaneous inversion have been characterized at the nucleotide level (25) in *Bacillus subtilis* 166. This inversion is unlikely to have been caused by homologous recombination and probably occurred during the repair process after  $\gamma$ -ray irradiation. The nature of each of the genetic elements involved in the other inversions described in Table 1 is unknown, except for the inversion in a defined lineage of *Neisseria gonorrhoeae* that could probably be mediated by homologous recombination between *pilE1-opaE* and *pilS5-opaD* regions (21).

*L. lactis* is a gram-positive mesophilic bacterium extensively used for health and dairy applications (13). The low-resolution chromosome maps of four independent lactococcal strains have been constructed by pulsed-field gel electrophoresis (PFGE). Two strains, DL11 (58) and IL1403 (30), belong to *L. lactis* subsp. *lactis*, and the other two, MG1363 (31) and FG2 (9), are *L. lactis* subsp. *cremoris* strains. These two subspecies have nucleotide sequences that diverge by 20 to 30% (11), close to the level of divergence between *E. coli* and *S. typhimurium*. Comparison of physical maps of the four strains showed an overall conservation of restriction site locations for the two *L. lactis* subsp. *lactis* strains but not for the *L. lactis* subsp. *cremoris* strain. At the genetic (i.e., gene order) level, different kinds of rearrangements were observed. A large inversion covering almost half of the chromosome (31) was identified by comparing strains of the two different subspecies, and translocation or inversion of four discrete regions had occurred between the two *L. lactis* subsp. *cremoris* strains (9). However, as the genome comparisons were made between genetically

\* Corresponding author. Mailing address: LMGM du CNRS, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France. Phone: (33) 561 33 58 25. Fax: (33) 561 33 58 86. E-mail: ritzenth@ibgc.biotoul.fr.

TABLE 1. Large chromosomal inversions within species

| Organism                                           | No. of inversions | Size of inverted region (%) | Inversion junction(s)      | Reference |
|----------------------------------------------------|-------------------|-----------------------------|----------------------------|-----------|
| Eubacteria (purple)                                |                   |                             |                            |           |
| <i>Bordetella pertussis</i> (6 strains)            | 2                 | 10–25                       | ND                         | 52        |
| <i>Escherichia coli</i> K-12                       |                   |                             |                            |           |
| LN850 and LN1053                                   | 2                 | 40–47                       | IS5 <sup>a</sup>           | 39        |
| W3110                                              | 1                 | 20                          | <i>rrm</i> <sup>a</sup>    | 24        |
| BHB2600                                            | 2                 | 2–13                        | IS1 and IS5 <sup>d</sup>   | 3         |
| 14851N                                             | 1                 | 36                          | IS3 <sup>c</sup>           | 27        |
| <i>Neisseria gonorrhoeae</i> (6 strains)           | 1                 | 30                          | ND                         | 21        |
| <i>Salmonella typhi</i> (127 strains)              | 20                | ND                          | <i>rrm</i> <sup>b</sup>    | 37        |
| <i>Pseudomonas aeruginosa</i> clone C (21 strains) | 9                 | 19–50                       | ND                         | 50        |
| <i>P. stutzeri</i> (20 strains)                    | 4                 | ND                          | ND                         | 22        |
| Eubacteria (gram-positive)                         |                   |                             |                            |           |
| <i>Bacillus subtilis</i>                           |                   |                             |                            |           |
| 166                                                | 1                 | 45                          | Nonhomologous <sup>c</sup> | 25        |
| 60866                                              | 1                 | 40                          | ND                         | 57        |
| <i>Lactococcus lactis</i> (2 strains)              | 1                 | 50                          | ND                         | 31        |
| <i>Mycoplasma hominis</i> (5 strains)              | 1                 | 50                          | ND                         | 28        |
| Eubacteria (spirochetes)                           |                   |                             |                            |           |
| <i>Leptospira interrogans</i> (2 strains)          | 1                 | 40                          | ND                         | 61        |
| Archaeobacteria                                    |                   |                             |                            |           |
| <i>Halobacterium salinarum</i> (3 strains)         | 1                 | 10                          | ND                         | 23        |

<sup>a</sup> Determined by Southern hybridization and genetic studies.

<sup>b</sup> Determined by I-CeuI restriction analysis.

<sup>c</sup> Determined by nucleotide sequence analysis.

<sup>d</sup> Determined by restriction analysis.

unrelated strains, it is not possible to follow the genetic events involved in these rearrangements.

To investigate the lactococcal genome plasticity, the genomes of *L. lactis* subsp. *cremoris* NCDO763 (also called ML3) and MG1363 (19) were compared. Both strains are derivatives of strain NCDO712 (10) and are the lactococcal strains studied most extensively for genetic analysis and molecular biology. Strains NCDO763 and MG1363 differ by only two genome rearrangements: a 30-kb deletion and a large inversion (unpublished data). In this study, we report the identification of

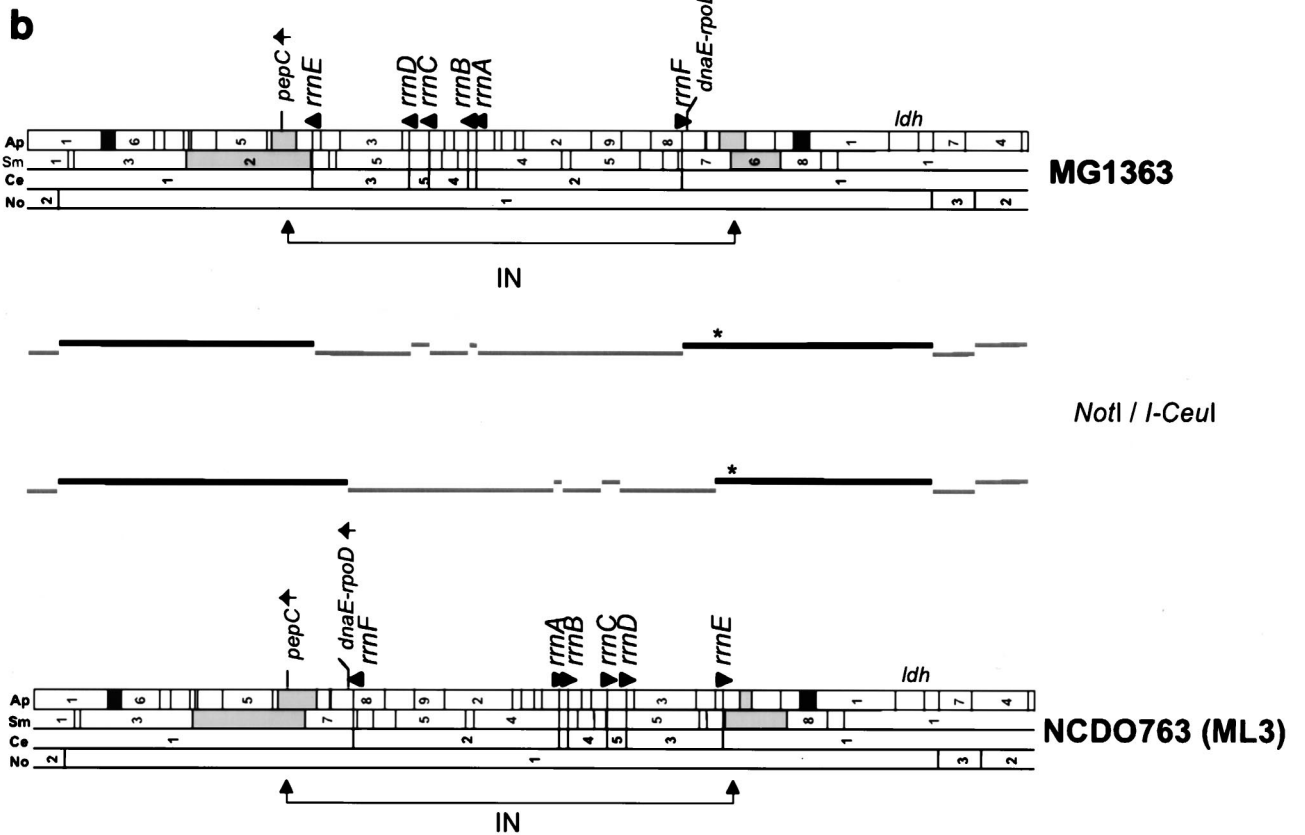
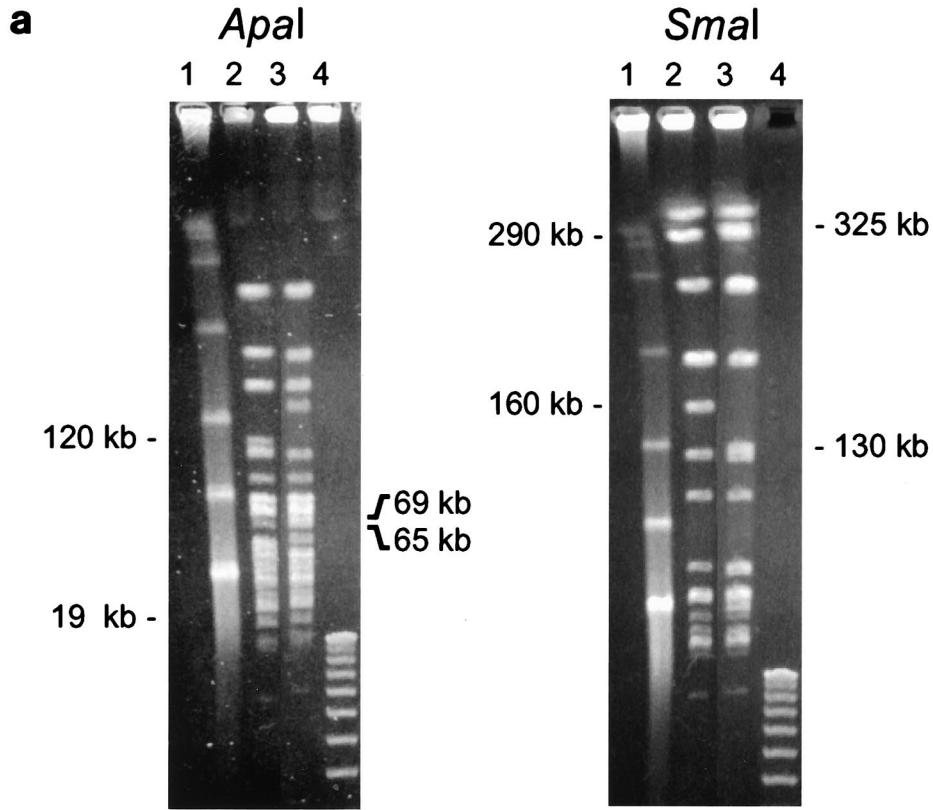
this large chromosomal inversion, cloning of the four inversion junctions (designated MG1363-*inv1* and -*inv2* for strain MG1363 and NCDO763-*inv1* and -*inv2* for strain NCDO763), and nucleotide characterization of the genetic element involved in the rearrangement.

#### MATERIALS AND METHODS

**Bacterial strains and plasmids.** *L. lactis* subsp. *cremoris* NCDO763 (10) was grown in M17 (55), and MG1363 (19) was grown in GM17 (M17 broth containing 0.5% glucose). *E. coli* DH5 $\alpha$  (Life Technologies, Gaithersburg, Md.) was

TABLE 2. Plasmids used

| Name   | Description                                                                   | Reference        |
|--------|-------------------------------------------------------------------------------|------------------|
| pRC1   | Em <sup>r</sup> , 3.5-kb cloning vector constructed from pBluescript KSII (+) | 29               |
| pCL52  | Em <sup>r</sup> , 2.4-kb cloning vector constructed from pBluescript KSII (+) | Unpublished data |
| pCL102 | 5.5-kb <i>Hind</i> III fragment from strain NCDO763 in pRC1                   | This work        |
| pCL132 | 3.2-kb <i>Hind</i> III fragment from strain NCDO763 in pCL52                  | This work        |
| pCL133 | Same as pCL132 but inserted in opposite orientation                           | This work        |
| pCL134 | 9-kb <i>Eco</i> RI fragment from strain NCDO763 in pCL52                      | This work        |
| pCL136 | 1.7-kb <i>Eco</i> RV fragment from pCL134 in pCL52                            | This work        |
| pCL137 | Same as pCL136 but inserted in opposite orientation                           | This work        |
| pCL138 | 1.4-kb PCR product from strain NCDO763 in pCL52                               | This work        |
| pCL139 | Same as pCL138 but inserted in opposite orientation                           | This work        |
| pCL200 | 5.4-kb <i>Sac</i> I fragment from strain MG1363 in pCL52                      | This work        |
| pCL201 | 9-kb <i>Acc</i> I fragment from strain MG1363 in pCL52                        | This work        |
| pCL203 | 1.7-kb <i>Eco</i> RV fragment from pCL200 in pCL52                            | This work        |
| pCL204 | Same as pCL203 but inserted in opposite orientation                           | This work        |
| pCL206 | 1.5-kb <i>Bst</i> XI fragment from pCL201 in pCL52                            | This work        |
| pCL207 | Same as pCL206 but inserted in opposite orientation                           | This work        |
| pCL208 | 1.5-kb <i>Afl</i> III- <i>Bst</i> XI fragment from pCL201 in pCL52            | This work        |
| pCL209 | Same as pCL206 but inserted in opposite orientation                           | This work        |
| pCL210 | 1-kb <i>Eco</i> RV- <i>Sac</i> I fragment from pCL200 in pCL52                | This work        |
| pCL211 | 1-kb <i>Eco</i> RV fragment from pCL200 in pCL52                              | This work        |
| pCL212 | Same as pCL211 but inserted in opposite orientation                           | This work        |



grown in LB broth. Erythromycin was used at concentrations of 150 µg/ml for *E. coli* and 5 µg/ml for *L. lactis*. Plasmids are listed in Table 2.

**DNA manipulation.** Lactococcal chromosome DNA was extracted as follows. Cells were grown overnight in M17 or GM17 broth containing 50 mM DL-threonine; 1 ml of cell suspension was centrifuged for 1 min (13,000 rpm), and the cell pellet was suspended in 0.75 ml of SET buffer (20% sucrose, 20 mM EDTA, 100 mM Tris [pH 8]) containing lysozyme (5 mg/ml). The mixture was incubated at 37°C for 1 h. Then 100 µl of 10% sodium dodecyl sulfate and 20 µl of proteinase K (25 mg/ml) were added, and incubation was continued at 60°C for 15 min. An equal volume of a phenol-trichloroethane-isoamyl alcohol (25/24/1) was added and mixed by vigorous shaking, and the mixture was centrifuged for 10 min (13,000 rpm). The aqueous layer was removed and reextracted with an equal volume of trichloroethane-isoamyl alcohol (24/1). Nucleic acids were precipitated with 2.5 volumes of ethanol after adjustments to 0.3 M NaCl and collected by centrifugation for 10 min (13,000 rpm). The DNA pellet was washed with 1 ml of 70% ethanol, dried for 5 min in a SpeedVac (Savant), and dissolved in 100 µl of ultrapure water (MilliQ; Millipore, Bedford, Mass.) containing RNase A (10 µg/ml). For PFGE analysis, high-molecular-weight chromosomal DNA of *L. lactis* was prepared from agarose-embedded cells as described previously (31). Plasmid DNA from *E. coli* was isolated by using a Qiaprep spin kit (small scale) or Qiagen plasmid midi kit (large scale) as instructed by the manufacturer (Qiagen GmbH, Hilden, Germany). Enzymatic modification of DNA (restriction, ligation, etc.) was performed as instructed by the manufacturers. DNA was introduced in *E. coli* DH5α by electrotransformation in a Gene Pulser (Bio-Rad, Hercules, Calif.) as recommended by the manufacturer.

**PCR amplification and cloning of the NCDO763-inv2 region.** A primer specific for MG1363-inv1 (5'-TTTATTCCTCCAACCTATTA-3') and another specific for MG1363-inv2 (5'-TAGAAATACGGTCTGGTCA-3') were used to PCR amplify (25 cycles of 95°C for 30 s, 48°C for 120 s, and 72°C for 120 s, followed by a final polymerization step at 72°C for 4 min) the NCDO763-inv2 region. The 1.4-kb amplified DNA fragment was purified and cloned into *EcoRV*-digested plasmid pCL52.

**Electrophoresis and Southern hybridization.** Classical 1% agarose gels were made up in 0.1 M TBE (1 M TBE is 1 M Tris base, 1 M boric acid, and 20 mM EDTA). PFGE was performed in a contour-clamped homogeneous electric field system (Pulsaphor Plus; LKB-Pharmacia) in 0.05 M TBE as previously described (32). Standard size markers were phage λ DNA concatemers, obtained as described by Waterbury and Lane (59), and a 1-kb DNA ladder (Life Technologies). Southern hybridization of dried agarose gels was performed as previously described (31). Signals were detected with the BAS1000 bioimaging-analyzer system (Fuji Photo Film Co., Tokyo, Japan) and analyzed with TINA version 2.07c software (Raytest Isotopenβeräte GmbH).

**DNA sequencing and nucleotide analysis.** Automatic double-stranded DNA sequencing was carried out (Génome Express SA, Grenoble, France) for both strands, using universal, reverse, and appropriate synthetic primers. Sequence data were assembled and analyzed with Clone Manager version 4.1 (Scientific & Educational Softwares) and DNAsis version 7.00. DNA and protein sequence similarity searches were carried out with FASTA (45).

**Nucleotide sequence accession numbers.** The GenBank accession numbers for MG1363-inv1 and MG1363-inv2 sequences are AJ223960 and AJ223961, respectively. The accession numbers for NCDO763-inv1 and NCDO763-inv2 sequences are AJ223962 and AJ223963, respectively.

## RESULTS

**Characterization of the large chromosomal inversion in strain NCDO763.** Comparative analysis of *ApaI* and *SmaI* fingerprints of the two genomes (Fig. 1a) revealed that the 69- and 65-kb *ApaI* fragments of MG1363 were replaced by two fragments of 120 and 19 kb and that the 325- and 130-kb *SmaI* fragments of MG1363 were replaced by two fragments of 290 and 160 kb in NCDO763. This restriction polymorphism did not change the overall size of either genome, as the sums of the fragment pairs were the same (134 kb in MG1363 and 139 kb in NCDO763 for *ApaI*; 455 kb and 450 kb for *SmaI*). The simplest explanation for this rearrangement was a chromosomal inversion between the two chromosomes, because a single inversion event does not change the total size of the genome but changes the size of two restriction fragments in each

strain (Fig. 1b). Various hybridization probes were used to match the *ApaI* and *SmaI* fragments of MG1363 with their counterparts in NCDO763. Despite complex hybridization pattern, the 19-kb *ApaI* fragment used as the probe strongly hybridized to the 130-kb *SmaI* and 69-kb *ApaI* fragments of MG1363 and the 160-kb *SmaI* and 19-kb *ApaI* fragments of NCDO763 (Fig. 2b). The 325-kb *SmaI* and 65-kb *ApaI* fragments of MG1363 were associated with the 290-kb *SmaI* and the 120-kb *ApaI* fragments of NCDO763, using the *pepC* gene as a probe (data not shown).

The chromosomal inversion could also be demonstrated by comparison of the restriction patterns generated with *I-CeuI*, an endonuclease that recognizes a consensus sequence in the 23S rRNA gene of bacteria (40), followed by hybridization using the *ldh* gene as the probe. This gene has been accurately mapped to the MG1363 chromosome (31). The *NotI/I-CeuI* double-digested fragments detected were of a size similar to that expected in the case of an inversion (Fig. 1b).

**Cloning the inversion junctions of the NCDO763 and MG1363 chromosomes.** The strategy used for cloning the NCDO763-inv1 region is summarized Fig. 2a. As depicted in this schematic, any restriction fragment located near the left side of an inversion junction (probe 1) should hybridize with the 19-kb *ApaI* and 160-kb *SmaI* fragments of the NCDO763 chromosome and with the 65-kb *ApaI* and the 325-kb *SmaI* fragments of the MG1363 chromosome. A restriction fragment located close to the right side of the junction (Fig. 2a, probe 3) should hybridize with the 19-kb *ApaI* and 160-kb *SmaI* fragments of the NCDO763 chromosome and with the 69-kb *ApaI* and 130-kb *SmaI* fragments of the MG1363 chromosome. However, a fragment carrying the inversion junction (Fig. 2a, probe 2) should give a single hybridization signal with the NCDO763 chromosome (19-kb *ApaI* and 160-kb *SmaI* fragments) and two with the MG1363 chromosome (65- and 69-kb *ApaI* fragments and 130- and 325-kb *SmaI* fragments).

Assuming that the 19-kb *ApaI* fragment of strain NCDO763 as a probe would generate a probe 2-type hybridization pattern, this fragment was used to probe PFGE-separated *SmaI* and *ApaI* fragments of the MG1363 and NCDO763 genomes. The hybridization pattern was more complex than expected (Fig. 2b), suggesting that the 19-kb *ApaI* fragment contained one or several elements repeated along both chromosomes at a high copy number. The unexpected absence of strong hybridization signals with the 325-kb MG1363 and 290-kb NCDO763 *SmaI* fragments could be explained only by a location of the inversion junction near one of the two *ApaI* sites. The above-mentioned strategy was then modified for a two-step cloning procedure. A genomic library of *HindIII*-digested DNA from strain NCDO763 was constructed and probed with the 19-kb *ApaI* fragment. Among the positive clones isolated, two *HindIII* fragments of 3.2 and 5.5 kb (pCL132 and pCL102, respectively [Table 2 and Fig. 3]) generated hybridization patterns consistent with their locations on either side of the inversion endpoint (data not shown). No other *HindIII* fragment containing the inversion endpoint could be cloned. A second NCDO763 genomic library of *EcoRI* fragments was then constructed and probed with pCL132 and pCL102. Both fragments hybridized with a 9-kb *EcoRI* fragment (pCL134). A restriction

FIG. 1. (a) PFGE of chromosomal DNA from strains MG1363 and NCDO763 digested with restriction enzymes *ApaI* and *SmaI*. Lanes: 1, lambda concatemer size standards (each band is a 48.5-kb multimer); 2, NCDO763 DNA; 3, MG1363 DNA; 4, 1-kb DNA ladder. The sizes of fragments correlated with the large chromosomal inversion are indicated beside each gel. Electrophoresis was carried out at 10 V/cm for 13 h at 13°C, with pulse times of 6 s for *ApaI* and 8 s for *SmaI*. (b) Physical maps of MG1363 and NCDO763 chromosomes. Fragments are numbered as in reference 31. Arrows indicate the 5'→3' orientation of the gene. Abbreviations: Ap, *ApaI*; Ce, *I-CeuI*; No, *NotI*; Sm, *SmaI*; IN, inversion region. The locations of fragments generated by a *NotI/I-CeuI* double digestion are shown between the maps. Black lines indicate fragments the size of which has changed; asterisks indicate fragments hybridizing to the *ldh* probe.

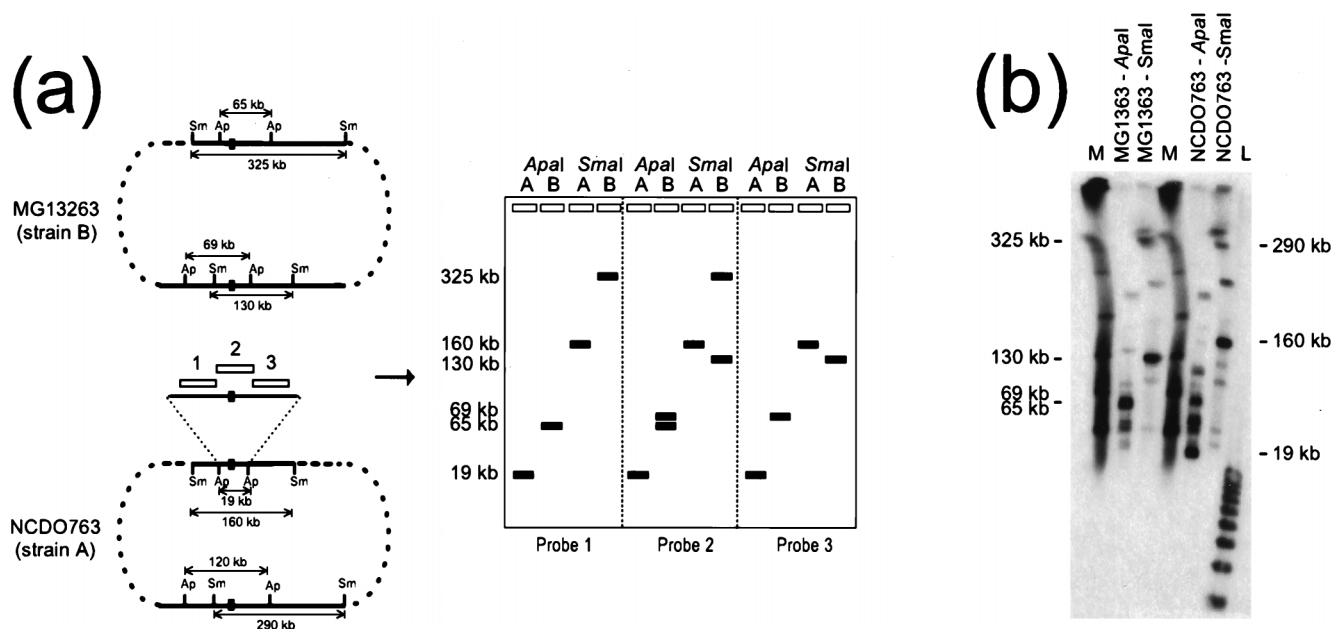


FIG. 2. (a) General strategy for identification of a restriction fragment containing an inversion endpoint. Right, schematic representation of the MG1363 and NCDO763 chromosomes. Open rectangles indicate fragments used as probes. Inversion junctions are indicated by black boxes. Restriction fragments are not drawn to scale. Abbreviations: *Ap*, *Apa*I; *Sma*, *Sma*I. Left, expected hybridization patterns of PFGE-separated *Apa*I and *Sma*I fragments of NCDO763 (A) and MG1363 (B) chromosomes with the various probes. (b) Hybridization of PFGE-separated MG1363 and NCDO763 DNA, using the 19-kb *Apa*I fragment as a probe. A pulse time of 8 s was used. Lane M, lambda concatemer size standards; lane L, 1-kb DNA ladder. Hybridization was performed in the presence of  $^{32}\text{P}$ -labeled lambda and the 1-kb ladder.

map of pCL134 was constructed (Fig. 3, NCDO763-inv1), and hybridization experiments using various subfragments of pCL134 as probes showed that the inversion endpoint was localized into a 1.7-kb *Eco*RV fragment (pCL136) and corresponded to the repeated element (data not shown) responsible for the complex hybridization pattern observed with the 19-kb *Apa*I fragment as probe (Fig. 2b).

The two inversion endpoints of the MG1363 chromosome were cloned with two monocopy regions of pCL134 fragment, located on either side of the repeated element, as the probe. The first endpoint was obtained by isolating a 5.4-kb *Sac*I fragment (pCL200) that hybridized to probe pCL102, and the second corresponded to a 9-kb *Acc*I fragment (pCL201) that hybridized to a subfragment of pCL132. Restriction maps of these two fragments were constructed (Fig. 3, MG1363-inv1 and MG1363-inv2), and hybridization analysis demonstrated that both regions contained the same repeated element as carried by pCL134 fragment. The second inversion endpoint of the NCDO763 chromosome (Fig. 3, NCDO763-inv2) was obtained by PCR amplification using one primer located to the right of the repeated element of the MG1363-inv2 region and the second located to the left of the repeated element of the MG1363-inv1 region. Hybridization analysis using the amplification product (pCL138) as the probe indicated that the same repeated element as found in pCL134, pCL200, and pCL201 was also present in pCL138 (data not shown).

**Sequence analysis of the four inversion endpoints.** Plasmids pCL136, pCL137, pCL138, pCL139, pCL203, pCL204, pCL206, pCL207, pCL208, pCL209, pCL210, pCL211, and pCL212 and parts of pCL132 and pCL133 (Fig. 3 and Table 2) were sequenced to identify the repeated element involved in the large inversion between the two chromosomes. Computer-assisted analysis of the MG1363-inv1 (3,638 bp), MG1363-inv2 (2,942 bp), and NCDO-inv1 (2,896 bp) regions predicted five

open reading frames (ORFs) surrounding a 1,313-bp IS element present in each region (Fig. 4). The nucleotide sequence of the IS element was 99.9% identical to that of the previously characterized lactococcal insertion sequence *IS905a* and 100% identical to that of *IS905b*. The only difference between *IS905a* and *IS905b* is the substitution of a C for a T at position 693 of the published sequence, converting a glutamine codon to a stop codon and truncating about half of the putative transposase (14). A potential transcription terminator with a stem of 20 nucleotides and a loop of 5 nucleotides ( $\Delta G^\circ = -41.2$  kcal/mol) was detected 296 nucleotides beyond ORF 2 of the MG1363-inv1 region. The stem had four T residues at its 3' end, followed by three more T residues that are not part of the stem. This structure is typical of rho-independent transcription termination.

Computer-based comparison of the deduced amino acid sequences of the five ORFs against nonredundant protein sequence databases was performed. The C-terminal part of ORF 2 (53 amino acids) of the MG1363-inv1 region was 45.3% identical to the C-terminal part of the *ceo* gene product from the nisin-sucrose transposon Tn5306 of *L. lactis*. The *ceo* gene encodes an *N*<sup>5</sup>-(carboxyethyl)ornithine synthase (EC 1.5.1.24) involved in the synthesis of *N*-(CA)-amino acids (15). The C-terminal part of ORF 1 (244 amino acids) was 33.7% identical to the *yhaA* gene product of *B. subtilis*. Based on protein domain analysis (8), the ORF 1 and *yhaA* gene products could be related to the aminoacylase (EC 3.5.1.14) enzyme family (51). The 215-amino-acid product of the ORF located just downstream from *IS905* was 58.6% identical to the *pcp* gene product of *Streptococcus pyogenes*. The *pcp* gene encodes a pyrrolidone carboxyl peptidase (PCP; EC 3.4.19.1) which removes N-terminal pyroglutamic acid residues from polypeptides (6). As *E. coli* K-12 strains are deficient in PCP activity (42), it was possible to phenotypically test the pyrrolidonyl peptidase activity of the MG1363-inv1 region by using a simple

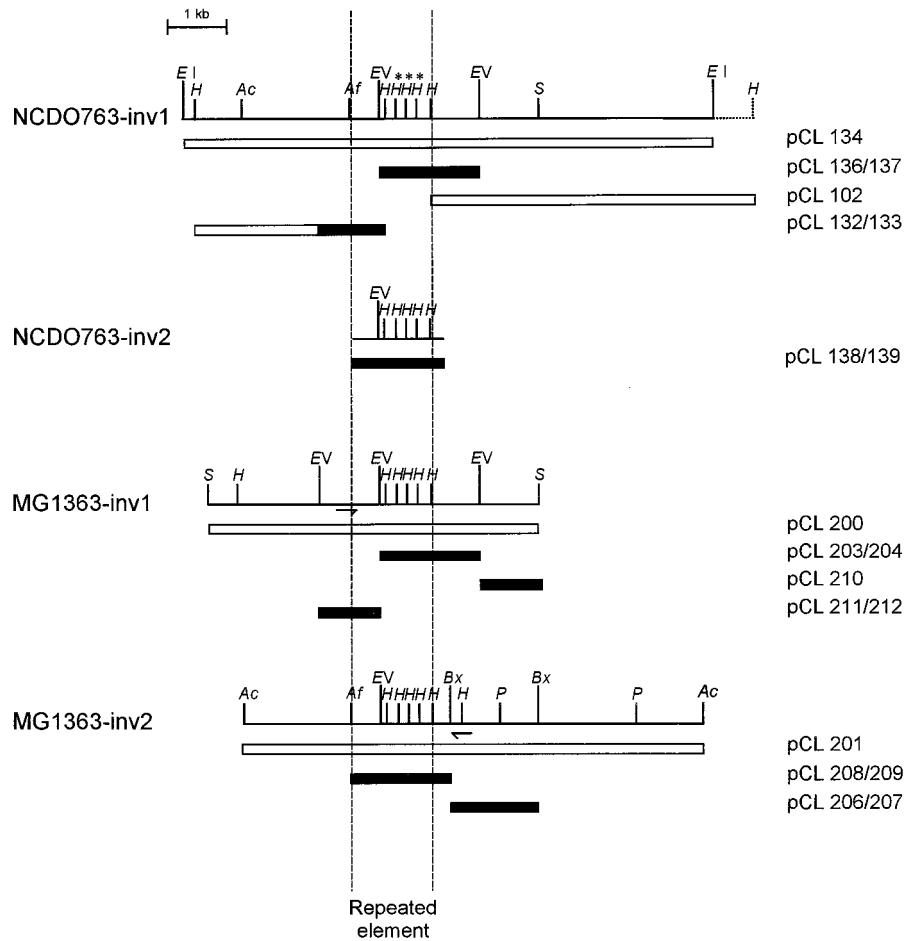


FIG. 3. Restriction maps of the four regions surrounding the inversion junctions. The various restriction fragments cloned are indicated by open rectangles, and the corresponding plasmids are indicated. Black boxes show the DNA regions that were sequenced. Arrows indicate positions of the primers used for PCR amplification of the NCDO763-inv2 region. Abbreviations: EI, *EcoRI*; H, *HindIII*; Ac, *AccI*; Af, *AflIII*; EV, *EcoRV*; S, *SacI*; Bx, *BstXI*; P, *PstI*.

colorimetric assay (43). When pCL102 (containing the entire *pcp* gene) was introduced into *E. coli* DH5 $\alpha$ , the erythromycin-resistant transformants were PCP<sup>+</sup>, indicating that the lactococcal *pcp* gene was functional. The products of ORFs 4 and 5 of the MG1363-inv2 region were 27.7 and 29.2% identical to those of ORFs 4 and 5 of *Lactobacillus casei* bacteriophage A2 (18). The protein encoded by ORF 5 was the putative large subunit of phage terminase, a protein that mediates the specific interaction between the prohead of the virus and its DNA. Sequence similarity comparisons for ORF 5 of the MG1363-inv2 region were based on the in-frame fusion of the two parts of ORF 5, after removal of the nucleotide sequence corresponding to the entire IS905 element, including one of the direct repeats.

The sequences upstream from the IS905 element of MG1363-inv1 and NCDO763-inv2 regions were identical (Fig. 5). Sequence identity was also observed for regions downstream from the IS905 elements of MG1363-inv1 and NCDO763-inv1. No nucleotide divergence other than the reciprocal exchange downstream from the IS elements was detected between MG1363 and NCDO763 chromosomes. Therefore, the large chromosomal inversion occurred by homologous recombination between the two inversely oriented copies of IS905.

## DISCUSSION

This study has established for the first time that a spontaneous chromosomal inversion can be mediated by homologous recombination between two inversely oriented copies of an IS element in gram-positive bacteria. This genetic event led to a large inversion of half of the chromosome, including the six *rm* operons. An inversion involving the same chromosomal regions was detected when the chromosome maps of two lactococcal strains of different subspecies were compared (31). As the putative replication origin (*oriC*) of the lactococcal chromosome was recently located near the *rmA* operon (60), these two inversions did not affect the distance of genes from *oriC*. It is noteworthy that most naturally occurring inversions include either the origin or the terminus of replication (*terC*) and do not disrupt the *oriC-terC* symmetry. It was recently shown that chromosome length variations of natural isolates of *E. coli* are distributed symmetrically with respect to the replication origin and terminus (2). Moreover, inversions restoring the chromosomal balance between *oriC* and *terC* have been found in *S. typhi* and *S. paratyphi* A (35, 36). Altogether, these observations are consistent with there being selective forces maintaining chromosome symmetry (17, 48), with only inversions that do not change the diametrically opposed locations of *oriC* and

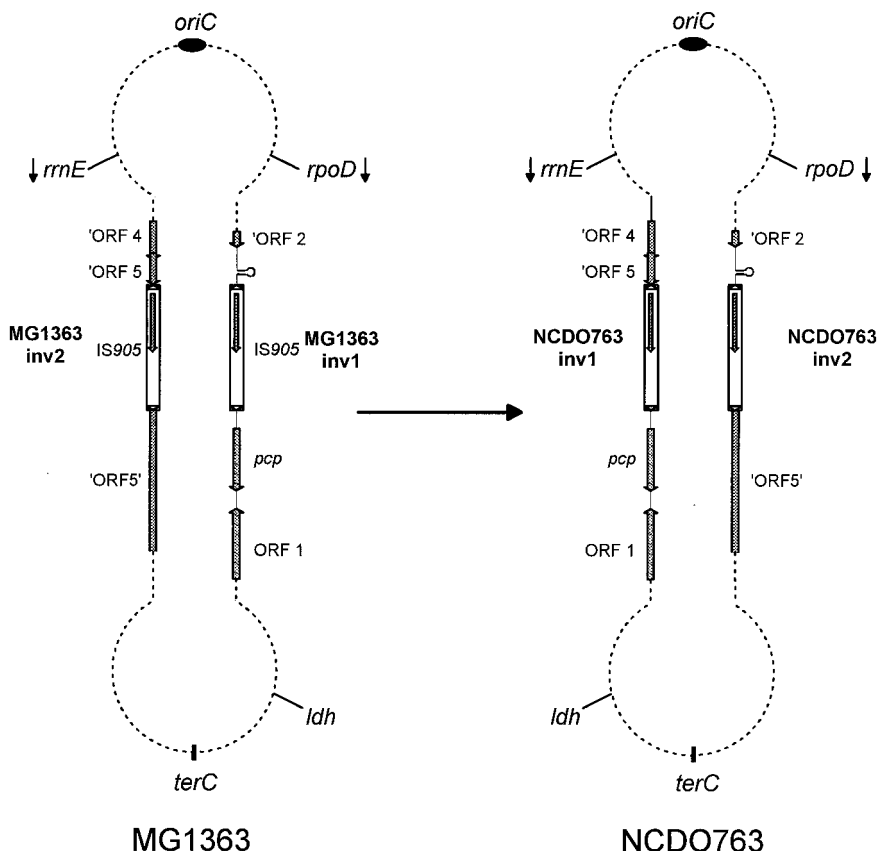


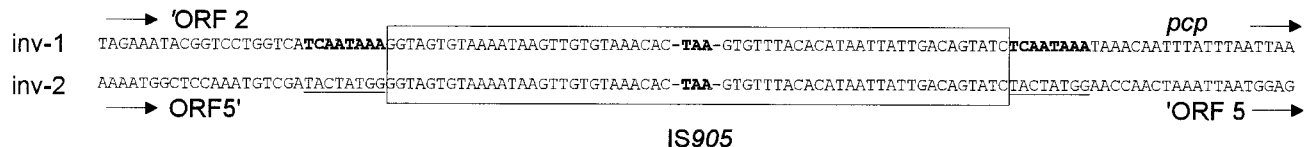
FIG. 4. Structures of the MG1363 and NCDO763 chromosomes. The genetic organization of the regions surrounding the inversion junctions was deduced from sequence analysis. MG1363-inv1, MG1363-inv2, NCDO763-inv1, and NCDO763-inv2 are not drawn to scale of the chromosome. Arrows indicate the direction of gene transcription. The location of *terC* is hypothetical.

*terC*, or that restore this symmetry following other rearrangements, being selectively fixed in natural conditions.

Any genetic event, including transposition, site-specific recombination, double-strand break repair, and homologous re-

combination, may account for the occurrence of large chromosomal inversions in natural conditions. However, inversion junctions have been identified in only a few cases, and so it is not known whether a particular type of genetic event is pref-

MG1363 chromosome



NCDO763 chromosome

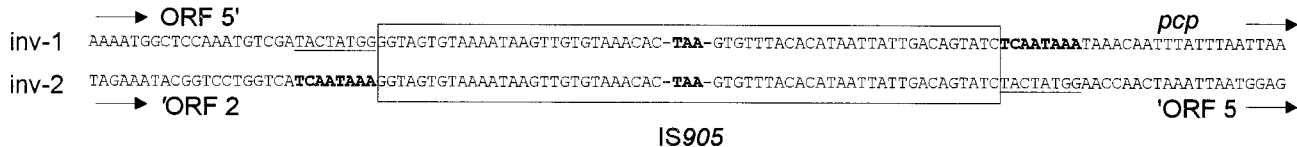


FIG. 5. Nucleotide sequences of part of the IS905 elements and their flanking regions for the two chromosomes. IS905 is boxed, and only the sequences of both 28-bp indirect repeats and the stop codon generated by the C→T mutation are shown. Direct repeats resulting from the IS905 insertion are indicated in bold in the MG1363-inv1 sequence and underlined in the MG1363-inv2 sequence.

entially associated with inversions. Nucleotide sequence analysis of the four chromosomal regions of strains MG1363 and NCDO763 explained the mechanism that generated the inversion in these strains. In contrast to the *E. coli* 1485IN inversion, in which recombination between the two IS3 elements was accompanied by a deletion of more than 50 kb (27) and probably associated with an IS3 transposition event, inversion between strains MG1363 and NCDO763 required only a single homologous recombination event between the two copies of the IS905 element.

Transposition of the IS905 element generates an 8-bp duplication at the target site (14). Sequence analysis of the chromosomal regions at the vicinity of the IS element revealed the presence of these two direct repeats on either side of the IS905 element in strain MG1363 (5'-TCAATAAAA-3' for the MG1363-inv1 region and 5'-TACTATGG-3' for the MG1363-inv2 region). Therefore, strain MG1363 has the parental chromosome structure and strain NCDO763 has the genome structure generated by the inversion. The presence of these different direct repeats unambiguously demonstrates that the two copies of IS905 were inserted into the MG1363 chromosome via independent transposition events, despite the C→T transition that removed half of the transposase. As it is highly improbable that the same mutation arose independently in both IS905 copies after transposition, it is likely that (i) translational readthrough could occur at the stop codon (20), generating a low level of functional transposase, (ii) transposition could generate defective copies of the element at relatively high frequency, or (iii) transposase proteins present into the cell could catalyze the transposition of nonfunctional elements. To our knowledge, there is no precedent for any of these patterns of behavior among prokaryotic IS elements. Thus, there may be a new mode of transposition regulation involved here, in addition to those already described (reviewed in reference 1).

The genetic organization of the MG1363 chromosomal sites where the two IS905 elements are inserted is consistent with the direct-selection hypothesis, which postulates that genomic IS elements can occasionally become fixed in a bacterial population through the beneficial mutations they cause (53). Transposition of the IS905 element should not disturb the expression of the surrounding genes in the MG1363-inv1 region, because this element is integrated just downstream from the putative rho-independent transcription terminator of ORF 2, and the presence of a perfect prokaryotic promoter -35 consensus (5'-TTGACA-3') reading outward 5 bp from the IR-R terminus is likely to create a new hybrid promoter for the *pcp* gene. Transposition of the IS905 element into the MG1363-inv2 region should confer a selective advantage on the strain, because it disrupts a bacteriophage gene that is essential for the lytic cycle.

The large chromosomal inversion described herein should not have phenotypic effects on the fitness of strain NCDO763, in comparison to strain MG1363, because it did not change the symmetry of the chromosome and therefore should not confer a selective disadvantage for replication of the NCDO763 chromosome. In addition, the inversion did not affect the local genetic environment, such as *pcp* gene expression in strain NCDO763. However, sensitive competition experiments are required to confirm that the inversion is neutral for cell fitness.

#### ACKNOWLEDGMENTS

We thank M. Chandler and J. Casadesus for helpful discussions and A. Edelman for reading the manuscript.

This work was supported by grants from the Centre National de la Recherche Scientifique (UPR9007), from the Région Midi-Pyrénées

(RECH 9609795), and from the UE-BIOTECH Programme (CT 96-0498). N. Campo was supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche.

#### REFERENCES

- Arber, W. 1995. The generation of variation in bacterial genomes. *J. Mol. Evol.* **40**:7-12.
- Berghthorson, U., and H. Ochman. 1998. Distribution of chromosome length variation in natural isolates of *Escherichia coli*. *Mol. Biol. Evol.* **15**:6-16.
- Birkenbihl, R. P., and W. Vielmetter. 1989. Cosmid-derived map of *E. coli* strain BHB2600 in comparison to the map of strain W3110. *Nucleic Acids Res.* **17**:5057-5069.
- 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.
- Carlson, C. R., A. Grønstad, and A. B. Kolstø. 1992. Physical maps of the genomes of three *Bacillus cereus* strains. *J. Bacteriol.* **174**:3750-3756.
- Cleuziat, P., A. C. Awadé, and J. Robert-Baudouy. 1992. Molecular characterization of *pcp*, the structural gene encoding the pyrrolidone carboxylate peptidase from *Streptococcus pyogenes*. *Mol. Microbiol.* **6**:2051-2063.
- Cole, S. T., and I. Saint Girons. 1994. Bacterial genomics. *FEMS Microbiol. Rev.* **14**:139-160.
- Corpet, F., J. Gouzy, and D. Kahn. 1998. The ProDom database of protein domain families. *Nucleic Acids Res.* **26**:323-326.
- Davidson, B. E., N. Kordias, M. Dobos, and A. J. Hillier. 1996. Genomic organization of lactic acid bacteria, p. 65-87. In G. Venema, J. H. J. Huis in't Veld, and J. Hugenholtz (ed.), *Lactic acid bacteria: genetics, metabolism and applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Davies, F. L., H. M. Underwood, and M. J. Gasson. 1981. The value of plasmid profile for strain identification in lactic streptococci and the relationship between *Streptococcus lactis* 712, ML3 and C2. *J. Appl. Bacteriol.* **51**:325-337.
- Delorme, C., J. J. Godon, S. D. Ehrlich, and P. Renault. 1994. Mosaic structure of large regions of the *Lactococcus lactis* subsp. *cremoris* chromosome. *Microbiology* **140**:3053-3060.
- Dempsey, J. A. F., A. B. Wallace, and J. G. Cannon. 1995. The physical map of the chromosome of a serogroup A strain of *Neisseria meningitidis* shows complex rearrangements relative to the chromosome of the two mapped strains of the closely related species *N. gonorrhoeae*. *J. Bacteriol.* **177**:6390-6400.
- de Roissart, H., and F. M. Luquet. 1994. Bactéries lactiques: aspects fondamentaux et technologiques. Loriga, Uriage, France.
- Dodd, H. M., N. Horn, and M. J. Gasson. 1994. Characterization of IS905, a new multicopy insertion sequence identified in lactococci. *J. Bacteriol.* **176**:3393-3396.
- Donkersloot, J. A., and J. Thompson. 1995. Cloning, expression, sequence analysis, and site-directed mutagenesis of the Tn5306-encoded N<sup>5</sup>-(carboxyethyl)ornithine synthase from *Lactococcus lactis* K1. *J. Biol. Chem.* **270**:12226-12234.
- Fonstein, M., and R. Haselkorn. 1995. Physical mapping of bacterial genomes. *J. Bacteriol.* **177**:3361-3369.
- François, V., J. Louarn, J. E. Rebollo, and J. M. Louarn. 1990. Replication termination, nondivisible zones, and structure of the *Escherichia coli* chromosome, p. 351-359. In K. Drlica and M. Riley (ed.), *The bacterial chromosome*. American Society for Microbiology, Washington, D.C.
- García, P., J. C. Alonso, and J. E. Suarez. 1997. Molecular analysis of the *cos* region of the *Lactobacillus casei* bacteriophage A2 gene product 3, gp3, specifically binds to its downstream *cos* region. *Mol. Microbiol.* **23**:505-514.
- Gasson, M. J. 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. *J. Bacteriol.* **154**:1-9.
- Gesteland, R. F., and J. F. Atkins. 1996. Recoding: dynamic reprogramming of translation. *Annu. Rev. Biochem.* **65**:741-768.
- Gibbs, C. P., and T. F. Meyer. 1996. Genome plasticity in *Neisseria gonorrhoeae*. *FEMS Microbiol. Lett.* **145**:173-179.
- Ginard, M., J. Lallucat, B. Tümmler, and U. Römling. 1997. Genome organization of *Pseudomonas stutzeri* and resulting taxonomic and evolutionary considerations. *Int. J. Syst. Bacteriol.* **47**:132-143.
- Hackett, N. R., Y. Bobovnikova, and N. Heyrovská. 1994. Conservation of chromosomal arrangement among three strains of the genetically unstable archaeon *Halobacterium salinarum*. *J. Bacteriol.* **176**:7711-7718.
- Hill, C. W., and B. W. Harnish. 1981. Inversions between ribosomal RNA genes of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **78**:7069-7072.
- Jarvis, E. D., S. Chen, and R. Rudner. 1990. Genetic structure and DNA sequences at junctions involved in the rearrangements of *Bacillus subtilis* strains carrying the *tpE26* mutation. *Genetics* **126**:785-797.
- Jiang, Q. J., K. Hiratsuka, and D. E. Taylor. 1996. Variability of gene order in different *Helicobacter pylori* strains contributes to genome diversity. *Mol. Microbiol.* **20**:833-842.
- Komoda, Y., M. Enomoto, and A. Tominaga. 1991. Large inversion in *Escherichia coli* K-12 1485IN between inversely oriented IS3 elements near *lac* and *cdd*. *Genetics* **129**:639-645.



28. Ladefoged, S. A., and G. Christiansen. 1992. Physical and genetic mapping of the genomes of five *Mycoplasma hominis* strains by pulsed-field gel electrophoresis. *J. Bacteriol.* **174**:2199–2207.
29. Le Bourgeois, P., M. Lautier, M. Mata, and P. Ritzenthaler. 1992. New tools for the physical and genetic mapping of *Lactococcus* strains. *Gene* **111**:109–114.
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. Lautier, L. van den Berghe, M. J. Gasson, and P. Ritzenthaler. 1995. Physical and genetic map of the *Lactococcus lactis* subsp. *cremoris* MG1363 chromosome: comparison with that of *Lactococcus lactis* subsp. *lactis* IL1403 reveals a large genome inversion. *J. Bacteriol.* **177**:2840–2850.
32. Le Bourgeois, P., M. Mata, and P. Ritzenthaler. 1989. Genome comparison of *Lactococcus* strains by pulsed-field gel electrophoresis. *FEMS Microbiol. Lett.* **59**:65–70.
33. Liu, S. L., A. Hessel, and K. E. Sanderson. 1993. The *XbaI-Bln I-CeuI* genomic cleavage map of *Salmonella enteritidis* shows an inversion relative to *Salmonella typhimurium* LT2. *Mol. Microbiol.* **10**:655–664.
34. Liu, S. L., and K. E. Sanderson. 1995. *I-CeuI* reveals conservation of the genome of independent strains of *Salmonella typhimurium*. *J. Bacteriol.* **177**:3355–3357.
35. Liu, S. L., and K. E. Sanderson. 1995. Rearrangements in the genome of the bacterium *Salmonella typhi*. *Proc. Natl. Acad. Sci. USA* **92**:1018–1022.
36. Liu, S. L., and K. E. Sanderson. 1995. The chromosome of *Salmonella paratyphi* A is inverted by recombination between *rmH* and *rmG*. *J. Bacteriol.* **177**:6585–6592.
37. Liu, S. L., and K. E. Sanderson. 1996. Highly plastic chromosomal organization in *Salmonella typhi*. *Proc. Natl. Acad. Sci. USA* **93**:10303–10308.
38. Lopez-Garcia, P., A. St Jean, R. Amils, and R. L. Charlebois. 1995. Genomic stability in the archaeae *Haloferax volcanii* and *Haloferax mediterranei*. *J. Bacteriol.* **177**:1405–1408.
39. Louarn, J. M., J. P. Bouché, F. Legendre, J. Louarn, and J. Patte. 1985. Characterization and properties of very large inversions of the *E. coli* chromosome along the origin-to-terminus axis. *Mol. Gen. Genet.* **201**:467–476.
40. Marshall, P., and C. Lemieux. 1991. Cleavage pattern of the homing endonuclease encoded by the fifth intron in the chloroplast subunit rRNA-encoding gene of *Chlamydomonas eugametos*. *Gene* **104**:1241–1245.
41. Michaux-Charachon, S., G. Bourg, E. Jumas-Bilak, P. Guigue-Talet, A. Allardet-Servent, D. O'Callaghan, and M. Ramuz. 1997. Genome structure and phylogeny in the genus *Brucella*. *J. Bacteriol.* **179**:3244–3249.
42. Mulczyk, M., and A. Szewczuk. 1970. Pyrrolidonyl peptidase in bacteria: a new colorimetric test for differentiation of Enterobacteriaceae. *J. Gen. Microbiol.* **61**:9–13.
43. Mulczyk, M., and A. Szewczuk. 1972. Pyrrolidonyl peptidase activity: a simple test for differentiating staphylococci. *J. Gen. Microbiol.* **70**:383–384.
44. Nikolskaya, T., M. Fonstein, and R. Haselkorn. 1995. Alignment of a 1.2 Mb chromosomal region from three strains of *Rhodobacter capsulatus* reveals a significantly mosaic structure. *Proc. Natl. Acad. Sci. USA* **92**:10609–10613.
45. Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**:2444–2448.
46. Perkins, J. D., J. Don Heath, B. R. Sharma, and G. M. Weinstock. 1993. *XbaI* and *BlnI* genomic cleavage maps of *Escherichia coli* K-12 strain MG1655 and comparative analysis of other strains. *J. Mol. Biol.* **232**:419–445.
47. Riley, M., and K. E. Sanderson. 1990. Comparative genetics of *Escherichia coli* and *Salmonella typhimurium*, p. 85–95. In K. Drlica and M. Riley (ed.), *The bacterial chromosome*. American Society for Microbiology, Washington, D.C.
48. Roth, J. R., N. Benson, T. Galitsky, K. Haack, J. G. Lawrence, and L. Miesel. 1996. Rearrangements of the bacterial chromosome: formation and applications, p. 2256–2276. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Resnikoff, M. Riley, M. Schaechter, and H. E. Umberger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. American Society for Microbiology, Washington, D.C.
49. Römling, U., J. Greipel, and B. Tümmler. 1995. Gradient of genomic diversity in the *Pseudomonas aeruginosa* chromosome. *Mol. Microbiol.* **17**:323–332.
50. Römling, U., K. D. Schmidt, and B. Tümmler. 1997. Large genome rearrangements discovered by the detailed analysis of 21 *Pseudomonas aeruginosa* clone C isolates found in environment and disease habitats. *J. Mol. Biol.* **271**:386–404.
51. Sakanyan, V., L. Desmarez, C. Legrain, D. Charlier, I. Mett, A. Kochikyan, A. Savchenko, A. Boyen, P. Falmagne, A. Pirard, and N. Glansdorff. 1993. Gene cloning, sequence analysis, purification, and characterization of a thermostable aminoacylase from *Bacillus stearothermophilus*. *Appl. Environ. Microbiol.* **59**:3878–3888.
52. Stibitz, S., and M. S. Yang. 1997. Genomic fluidity of *Bordetella pertussis* assessed by a new method for chromosomal mapping. *J. Bacteriol.* **179**:5820–5826.
53. Syvanen, M. 1997. Insertion sequences and their evolutionary role, p. 213–220. In F. J. de Bruijn, J. R. Lupski, and G. M. Weinstock (ed.), *Bacterial genomes: physical structure and analysis*. Chapman & Hall, New York, N.Y.
54. Tabata, K., and T. Hoshino. 1996. Mapping of 61 genes on the refined physical map of the chromosome of *Thermus thermophilus* HB27 and comparison of genome organization with that of *T. thermophilus* HB8. *Microbiology* **142**:401–410.
55. Terzaghi, B. E., and W. E. Sandine. 1975. Improved medium for lactic streptococci and their bacteriophages. *Appl. Microbiol.* **29**:807–813.
56. Tigges, E., and F. C. Minion. 1994. Physical map of *Mycoplasma gallisepticum*. *J. Bacteriol.* **176**:4157–4159.
57. Toda, T., and M. Itaya. 1995. *I-CeuI* recognition sites in the *rm* operons of the *Bacillus subtilis* 168 chromosome: inherent landmark for genome analysis. *Microbiology* **141**:1937–1945.
58. 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.
59. Waterbury, P. G., and M. J. Lane. 1987. Generation of lambda concatemers for use as pulsed-field electrophoresis size markers. *Nucleic Acids Res.* **15**:3930.
60. Wind, A., and E. B. Hansen. 1996. Cloning and characterisation of the *dnaA* region in *Lactococcus lactis*, abstr. 276, p. 65. In *Molecular mechanisms in DNA replication and recombination 1996*. Keystone Symposia, Silverthorne, Colo.
61. Zuerner, R. L., J. L. Herrmann, and I. Saint Girons. 1993. Comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity. *J. Bacteriol.* **175**:5445–5451.