Microbiology (2010), **156**, 420–430 DOI 10.1099/mic.0.028308-0

Genome diversity in the genera *Fructobacillus*, *Leuconostoc* and *Weissella* determined by physical and genetic mapping

Ivo M. Chelo,† Líbia Zé-Zé‡ and Rogério Tenreiro

Centro de Biodiversidade, Genómica Integrativa e Funcional (BioFIG), Faculdade de Ciências da Universidade de Lisboa, Edificio ICAT, Campus da FCUL, Campo Grande, 1749-016 Lisboa, Portugal

Correspondence
Ivo M. Chelo
imchelo@igc.gulbenkian.pt

Received 13 February 2009
Revised 19 October 2009
Accepted 2 November 2009

Pulsed-field gel electrophoresis analysis of chromosomal single and double restriction profiles of 17 strains belonging to three genera of '*Leuconostocaceae*' was done, resulting in physical and genetic maps for three *Fructobacillus*, six *Leuconostoc* and four *Weissella* strains. *Asc*I, I-*Ceu*I, *Not*I and *Sfi*I restriction enzymes were used together with Southern hybridization of selected probes to provide an assessment of genomic organization in different species. Estimated genome sizes varied from 1408 kb to 1547 kb in *Fructobacillus*, from 1644 kb to 2133 kb in *Leuconostoc* and from 1371 kb to 2197 kb in *Weissella*. Other genomic characteristics of interest were analysed, such as *oriC* and *terC* localization and *rrn* operon organization. The latter seems markedly different in *Weissella*, in both number and disposition in the chromosome. Comparisons of intra- and intergeneric features are discussed in the light of chromosome rearrangements and genomic evolution.

INTRODUCTION

The genera *Fructobacillus*, *Leuconostoc* and *Weissella* are composed of obligate heterofermentative bacterial species that, with species from the genus *Oenococcus*, constitute a single line of descent, the Leuconostoc group (Yang & Woese, 1989; Martinez-Murcia & Collins, 1990). Like other lactic acid bacteria (LAB) they are Gram-positive, non-spore-forming, inhabit nutrient-rich environments such as milk, meat, vegetable products and fermented drinks (Kandler & Weiss, 1986) and have lactic acid as their main end product. The phylogenetic structure of these genera has been defined based on the analysis of 16S rRNA gene sequences (Martinez-Murcia & Collins, 1990; Collins *et al.*, 1993; Endo & Okada, 2008) and in studies involving different molecular markers (Chelo *et al.*, 2007; De Bruyne *et al.*, 2007; Endo & Okada, 2008). However, some taxonomic issues remain to be resolved, including the 'temporary' polyphyly of *Leuconostoc* as a result of the newly formed genus *Fructobacillus* (Endo & Okada, 2008).

As a supra-generic group, the Leuconostocs are phylogenetically related to *Lactobacillus* and *Pediococcus* (Vandamme *et al.*, 1996; Makarova & Koonin, 2007). Although there is currently high interest in LAB genomics (Klaenhammer *et al.*, 2002, 2005; Makarova *et al.*, 2006; Makarova & Koonin, 2007) only three Leuconostoc genome sequences have been published so far, those of *Oenococcus oeni* PSU-1, *Leuconostoc mesenteroides* subsp. *mesenteroides* ATCC 8293^T and *Leuconostoc citreum* KM20 (Makarova *et al.*, 2006; Kim *et al.*, 2008); four others (*Leuconostoc mesenteroides*, *Leuconostoc gasicomitatum*, *Oenococcus oeni* and *Weissella paramesenteroides* strains) are being sequenced (data from the ENTREZ Genome Project database at NCBI). The *L. mesenteroides* and *W. paramesenteroides* strains are relatively new additions to the ongoing sequencing projects and in the case of *Weissella* the only available information regarding genomic features such as chromosome size or number of *rrn* operons comes from the physical and genetic map of *W. paramesenteroides* DSM 20288^T (Chelo *et al.*, 2004). Perhaps due to this scarcity of information, comparative genomic analyses in the Leuconostoc group have been restricted to comparisons with species from other genera (Makarova *et al.*, 2006; Makarova & Koonin, 2007; Marcobal *et al.*, 2007) or are limited to a single species (Zé-Zé *et al.*, 2000, 2008).

In this study we analysed the chromosomes of 17 strains of the genera *Fructobacillus*, *Leuconostoc* and *Weissella* by restriction with *Asc*I, I-*Ceu*I, *Not*I and *Sfi*I and electrophoretic separation by PFGE. Together with Southern

†Present address: Instituto Gulbenkian de Ciência, Apartado 14, P-2781-901 Oeiras, Portugal.

‡Present address: Centro de Estudos de Vectores e Doenças Infecciosas Dr Francisco Cambournac, Instituto Nacional de Saúde Dr Ricardo Jorge I.P., Av. Liberdade no. 5, 2965-575 Águas de Moura, Portugal.

Abbreviation: LAB, lactic acid bacteria.

Supplementary material is available with the online version of this paper.

hybridization of selected probes we were able to construct 13 new physical and genetic maps on which genomic features of interest were localized. A map generated by *in silico* analysis of the *L. citreum* KM20 chromosome was also included in this study. The comparison of different strains in each genus, representing different species, has enabled us to unveil the main types of macrogenomic evolutionary events that shaped the diversification of this Leuconostoc group.

METHODS

Bacterial strains and growth conditions. All strains studied (see Table 1) were grown in MRS medium, pH 6.5, at 30 °C without shaking (with the exception of *Leuconostoc gelidum* DSM 5578^T, which was grown at 22 °C). For *Fructobacillus* strains, the growth medium was supplemented with D-fructose at 2 %. *Escherichia coli* JM109 and XL-1 Blue MRF′ (Stratagene) were grown at 37 °C in LB broth, supplemented with 100 μg ampicillin ml^{-1} when required.

Isolation and cleavage of chromosomal DNA, and DNA fragment nomenclature. Intact genomic DNA was prepared in agarose plugs and single or double digested with the restriction enzymes *Asc*I, I-*Ceu*I, *Not*I and *Sfi*I as previously described (Tenreiro *et al.*, 1994; Zé-Zé *et al.*, 1998, 2000). Restriction fragments produced by digestion with a single enzyme are indicated by the initial letter of the endonuclease. Nomenclature for fragments resulting from double digestion, co-migrating fragments and hybridization results (see the supplementary material available with the online version of this paper) follow Chelo *et al.* (2004). DNA extraction and purification was done using the method described by Pitcher *et al.* (1989). Some DNA fragments were recovered after gel electrophoresis using the Jet Quick Gel extraction spin kit (Genomed). In the case of *L. citreum* KM20 the relative localization of restriction sites was done with pDRAW32 1.0 (AcaClone).

PFGE. PFGE was carried out in the Gene Navigator system (Pharmacia) with contour-clamped homogeneous electric field (CHEF) as previously described (Zé-Zé *et al.*, 1998). A different run was also used for the separation of fragments ranging from 4 to 50 kb. With a total time of 12 h at 260 V, this run consists of steps of 4 h with a pulse time of 0.8 s, 4 h with a pulse time of 1 s, and 4 h with a pulse time of 1.2 s. Staining was done with ethidium bromide or SYBR Green. The mean size of each fragment was estimated from at least two (fragments larger than 1000 kb) or six (fragments smaller than 1000 kb) runs by linear interpolation with two or more flanking size standards using KODAK 1D 2.0 software. Lambda ladder, Low-Range PFG Markers (New England Biolabs), DNA Size Markers – *Sschizosaccharomyces pombe* chromosomal DNA and 2.5 kb Molecular Ruler (Bio-Rad) – were used as molecular mass standards as well as intact chromosomes of *Saccharomyces cerevisiae* (also from Bio-Rad).

PCR conditions and plasmid construction. Most PCR conditions and plasmids used in this work have been previously described (Chelo *et al.*, 2004). A complete list of primers and amplification conditions can be found in Supplementary Table S4.

Table 1. Strains used and genome characteristics

Strain	No. of fragments generated by restriction				Estimated chromosome size in kb*	Map (no. of enzymes)
	*Asc*I	I-*Ceu*I	*Not*I	*Sfi*I		
Leuconostoc mesenteroides subsp. *mesenteroides* CECT 219^T†	2	4	5	5	2032 (78)	Yes (4)
Leuconostoc mesenteroides subsp. *cremoris* CECT 872^T	6	4	14	NA	1644 (16)	Yes (3)
Leuconostoc mesenteroides subsp. *dextranicum* CECT 912^T	10	4	5	6	1796 (48)	Yes (4)
Leuconostoc citreum CECT 4018	15	4	0	NA	1810 (31)	Yes (2)
Leuconostoc citreum CECT 4025^T	5	5	6	4	1850 (33)	Yes (4)
Leuconostoc gelidum DSM 5578^T	6	4	5	6	1953 (83)	Yes (4)
Leuconostoc pseudomesenteroides CECT 4027^T	12	4	7	3	2133 (34)	Yes (4)
Leuconostoc fallax DSM 20189^T	0	4	0	2	1648 ± 45	No
Fructobacillus ficulneus DSM 13613^T	4	5	15	>35	1547 (5)	Yes (3)
Fructobacillus pseudoficulneus LC51^T	4	4	23	>30	1408 (13)	Yes (3)
Fructobacillus fructosus DSM 20349^T	3	4	12	>30	1419 (11)	Yes (3)
Weissella paramesenteroides DSM 20288^T†	3	8	6	4	2026 (8)	Yes (4)
Weissella hellenica DSM 7378^T	4	8	2	2	1850 (48)	Yes (4)
Weissella confusa DSM 20196^T	13	8	10	13	2197 (16)	Yes (3)
Weissella kandleri CECT 4307^T	8	6	7	4	1371 (10)	Yes (4)
Weissella halotolerans CECT 573^T	>15	NA	>25	>35	NA	No
Weissella viridescens DSM 20410^T	2	7	9	9	1520 (22)	Yes (4)

NA, Non-available. Restriction profiles always presented several fragments resulting from incomplete digestion, or fragments of very low intensity.
*Deviations from the mean values are given by the maximal size deviation (kb) from the observed genome size to the mapped size when a map was obtained (values in parentheses), or in the case of *L. fallax* as the SEM of the different enzymes.
†Although genomic maps of these strains have already been published (Chelo *et al.*, 2004; Makarova *et al.*, 2006) they are included here for comparative purposes.

DNA probes. All probes used in this work (see Table 2 for probes used and some of their characteristics) were labelled with digoxigenin using PCR Dig labelling mix (Roche). Transposase B, Permease (AA) and Transport ABC probes were used as linking clones, since they contain restriction sites of the enzymes used in this study (*Not*I site for Transposase B and *Asc*I sites for the others). Sequences for these probes were obtained from the US DOE Joint Genome Institute (JGI) in an early version of the sequencing project of *L. mesenteroides* subsp. *mesenteroides* ATCC 8293^T [Transposase B corresponds to locus tag LEUM_A02, Permease (AA) to LEUM_1791 and Transport ABC to LEUM_0693].

***In silico* analysis.** The localization of markers in *L. citreum* KM20 was done using the sequenced genome annotation when possible or by BLAST analysis (Altschul *et al.*, 1990). For the comparison of *L. mesenteroides* subsp. *mesenteroides* CECT 219^T and *L. citreum* KM20 strains, TBLASTX alignments were done with a cut-off value of 10^{-4} in the WebACT site (http://www.webact.org/WebACT/home).

ITS assignment. Identification of the different sets of tRNA genes in the ITS regions of specific *rrn* operons of *Weissella* strains was based on explicit I-*Ceu*I fragment PCR amplification. In *Fructobacillus* and *Leuconostoc* strains genomic DNA amplification always gave rise to a single band of the size expected for tRNA^{Ala}, as previously seen for most strains (Endo & Okada, 2008). Detailed explanation of the method can be found in Chelo *et al.* (2004).

RESULTS

Macrorestriction fragments and genome sizes of *Fructobacillus*, *Leuconostoc* and *Weissella* strains

The numbers of restriction fragments generated with *Asc*I, *Not*I, *Sfi*I and the homing endonuclease I-*Ceu*I as well as the estimated genome sizes are presented in Table 1. Genome sizes are averages of estimates with the different enzymes (two to four enzymes) and they match genome-mapped sizes when a physical map was obtained. The fully assembled genome sequence of *L. mesenteroides* subsp. *mesenteroides* ATCC 8293^T has become available, revealing a chromosome size of 2038 kb (Makarova *et al.*, 2006; data available at the NCBI and JGI databases). This strain corresponds to the strain CECT 219^T that was analysed in this study but since this map was built without the prior knowledge of the assembled sequence, we chose to present it for comparative purposes. In this way, the error inherent in our approach is also reflected in this map, although we have a priori knowledge of its suitability.

The analysis of the number of restriction fragments generated in this study reveals diverse results. The number

Table 2. DNA sequences used as probes in this study

Probe	Gene (function/description)*	Approx. fragment size (kb)	G+C content (mol%)	Reference for probes
rrs	Small-subunit rRNA [LEUM_r0019]	1.5	51.4	Chelo *et al.* (2004)
rrl	Large-subunit rRNA	1.1	50.3	Chelo *et al.* (2004)
dnaA	Chromosomal replication initiator protein DnaA [LEUM_0001]	0.7	37.7	Chelo *et al.* (2004)
gyrB	DNA gyrase subunit B [LEUM_0005]	1.8	45.6	Chelo *et al.* (2004)
rpoC	DNA-directed RNA polymerase subunit β' [LEUM_1824]	1.4	42.3	Chelo *et al.* (2004)
dnaK	Chaperone protein DnaK [LEUM_1347]	0.6	51.4	Chelo *et al.* (2004)
recA	RecA protein [LEUM_0584]	0.25	49.6	Chelo *et al.* (2004)
GACA3	Undetermined, no coding	3.5	43.8	Chelo *et al.* (2004)
intC8	3-Oxoacyl reductase; possible RNA methyltransferase†	2.8	34.1	Chelo *et al.* (2004)
tgt, clpB	Queuine tRNA-ribosyltransferase/CLPB ATP-binding protein [LEUM_0376]	6	42.2	Zé-Zé *et al.* (2000)
h*arosynth*	3-Phosphoshikimate 1-carboxyvinyltransferase [LEUM_1164]	0.8	40.9	Chelo *et al.* (2004)
h*ctpsynth*	CTP synthase [LEUM_0519]	0.6	41.7	Chelo *et al.* (2004)
h*red*, h*transp*	Aldo-keto reductase; putative transposase†	1.4	32.8	Chelo *et al.* (2004)
h*polC*	DNA polymerase III catalytic subunit, PolC type [LEUM_0689]	0.4	43.0	Chelo *et al.* (2004)
h*rpoA*	DNA-directed RNA polymerase subunit α [LEUM_0221]	0.4	42.0	Chelo *et al.* (2004)
h5′*nucl*	Putative 5-nucleotidase†	0.7	39.8	Chelo *et al.* (2004)
Permease (AA)	Amino acid transporter [LEUM_1791]	1.4	38.6	This study
Transport ABC	Uncharacterized ABC-type transport system, nucleoside-binding protein [LEUM_0693]	0.65	37.9	This study
αATPase	ATP synthase F1 subcomplex α subunit [LEUM_1871]	1.1	43.4	This study
Transposase B	Transposase B, hypothetical protein [LEUM_A02]	0.6	38.0	This study

*When available, locus tags are given for probes that hybridized to the *L. mesesenteroides* subsp. *mesenteroides* genome.
†Possible function of the DNA sequences was obtained by BLAST searches (BLASTX or BLASTP) as described previously (Chelo *et al.*, 2004).

of *Asc*I, *Not*I and *Sfi*I restriction fragments shows considerable variation, ranging from 2 to more than 15 for *Asc*I, 0 to more than 25 for *Not*I and 2 to more than 35 for *Sfi*I. The high number of restriction sites often prevents their use for mapping purposes. Taking into consideration the different degrees of evolutionary proximity of most of the strains used in this study (Chelo *et al.*, 2007), we can also see that in some cases a high proportion of the variation can be found in closely related strains. This is the case for the number of *Asc*I fragments in *L. citreum* CECT 4018 (15 fragments) and *L. citreum* CECT 4025^T (5 fragments) or *Not*I fragments in *L. mesenteroides* subsp. *cremoris* CECT 872^T (14 fragments) and *L. mesenteroides* subsp. *mesenteroides* (5 fragments). A similar situation is seen for the *Not*I fragments in the *Weissella* strains. These show a much greater difference between *W. confusa* DSM 20196^T (10 fragments) and *W. hellenica* DSM 7378^T (2 fragments) than between *W. confusa* DSM 20196^T and *W. viridescens* DSM 20410^T (9 fragments), which are evolutionarily less related. In contrast to this inconsistency of fragment number and evolutionary relatedness, *Fructobacillus* strains represent a case where the variation is reduced.

The number of I-*Ceu*I-generated fragments, indicative of the number of *rrn* operons in the chromosome, reveals a different situation. In addition to an overall smaller degree of variation in comparison with the other restriction profiles, a distinction between *Weissella* and the other two genera can be readily made. In both *Fructobacillus* and *Leuconostoc*, restriction with I-*Ceu*I seems to consistently give four fragments, with the exception of *L. citreum* CECT 4025^T and *Fructobacillus ficulneus* DSM 13613^T, whereas in *Weissella* this number varies between six and eight.

Chromosome sizes also provide an interesting view of the degree of variation in these genomes. The highest and lowest values can be found in *W. confusa* DSM 20196^T (2197 kb) and *W. kandleri* CECT 4307^T (1371 kb) respectively; *Fructobacillus* and *Leuconostoc* show smaller ranges of variation – 1408 kb to 1547 kb in *Fructobacillus* and 1644 kb to 2133 kb in *Leuconostoc*. This is consistent with a later diversification between the *Fructobacillus* and *Leuconostoc* genera. As for the number of some restriction sites, a large variation in chromosome size is observed even in closely related strains, for example *L. mesenteroides* subsp. *mesenteroides* CECT 219^T and *L. mesenteroides* subsp. *cremoris* CECT 872^T, which differ by about 21 % (considering a mean chromosome size of 1838 kb).

Comparison of physical and genetic maps in *Fructobacillus* and *Leuconostoc*

Given both the historical association of *Fructobacillus* and *Leuconostoc* species and the presence of many common characteristics in these genera, their genomic features are described together. The analysis of double digestions together with Southern hybridization of selected probes (see Table 2 and supplementary material) gives the relative localization of many restriction sites and genes, allowing the construction of the physical and genetic maps in Figs 1, 2 and 3. Connecting lines between successive maps identify the position of the same marker in different chromosomes and indicate whether this marker is localized in a conserved region (blue line) or in a region that might have been subject to an inversion event (red line). If no such assessments can be made the connecting line is grey. This classification was obtained by the following procedure. Considering the linear order of shared markers in two circular chromosomes as ordered vectors $X=x_1...x_n$ and $Y=y_1...y_n$, a marker $x_i=y_j$ is said to be in a conserved region (blue) if $x_{i-1}=y_{j-1}$ or $x_{i+1}=y_{j+1}$. Otherwise, if $x_{i-1}=y_{j+1}$ or $x_{i+1}=y_{j-1}$ we consider that an inversion (red) has occurred. Since this process is applied to each shared marker it is possible that an inversion line is plotted for a single marker (as in Figs 1 and 3). When there are alternative mapping possibilities these rules can also be applied as long as the final result is always the same, as in the comparison between *F. ficulneus* DSM 13613^T and *F. pseudoficulneus* LC51^T in Fig. 1. When several markers are allocated to the same fragment, they do not provide relative information and thus we assume they cannot support or contradict assessments made with the other markers (for instance, in the comparison of *L. gelidum* DSM 5578^T with *F. ficulneus* DSM 13613^T in Fig. 1, *rpoC* is considered to be in a conserved region since *tgt/clpB* provides a valid adjacent marker).

Data including restriction profiles, hybridization results and circular maps required for mapping purposes are given for a representative strain of each genus in the supplementary material (Supplementary Files S1–S3; data for other strains are available from the authors on request). The probes used should provide a good estimate of genomic organization dynamics in the strains studied, as revealed by their consistency with nearby markers obtained from whole genome sequencing (Supplementary Fig. S4).

In Fig. 1 a comparison of linear maps of type strains of *Fructobacillus* spp. and *Leuconostoc* spp. is provided. Most of the distinctive features that can be evaluated at this level seem to be well conserved. The origin of transcription *oriC* (assessed by the position of the *dnaA* and *gyrB* markers) is surrounded by *rrn* operons that are transcribed divergently from it. The distribution of the *rrn* operons is largely circumscribed around *oriC*, being localized in 22.1 % (in *L. gelidum* DSM 5578^T) to 42.4 % (in *F. fructosum* DSM 20439^T) of the chromosome. It is noteworthy that when *rrn* operons occupy a more significant part of the chromosome (in *Fructobacillus* strains), the absolute range (around 600 kb) is similar in all chromosomes. Although *rrn* operons are usually equally distributed on both sides of *oriC* there are a few exceptions such as in *F. pseudoficulneus* LC51^T and *F. fructosus* DSM 20349^T, where they have a 3 : 1 configuration (three on one side of *oriC* and one on the other), or in *L. gelidum* DSM 5578^T, where the asymmetry is the greatest, in a 4 : 0 configuration.

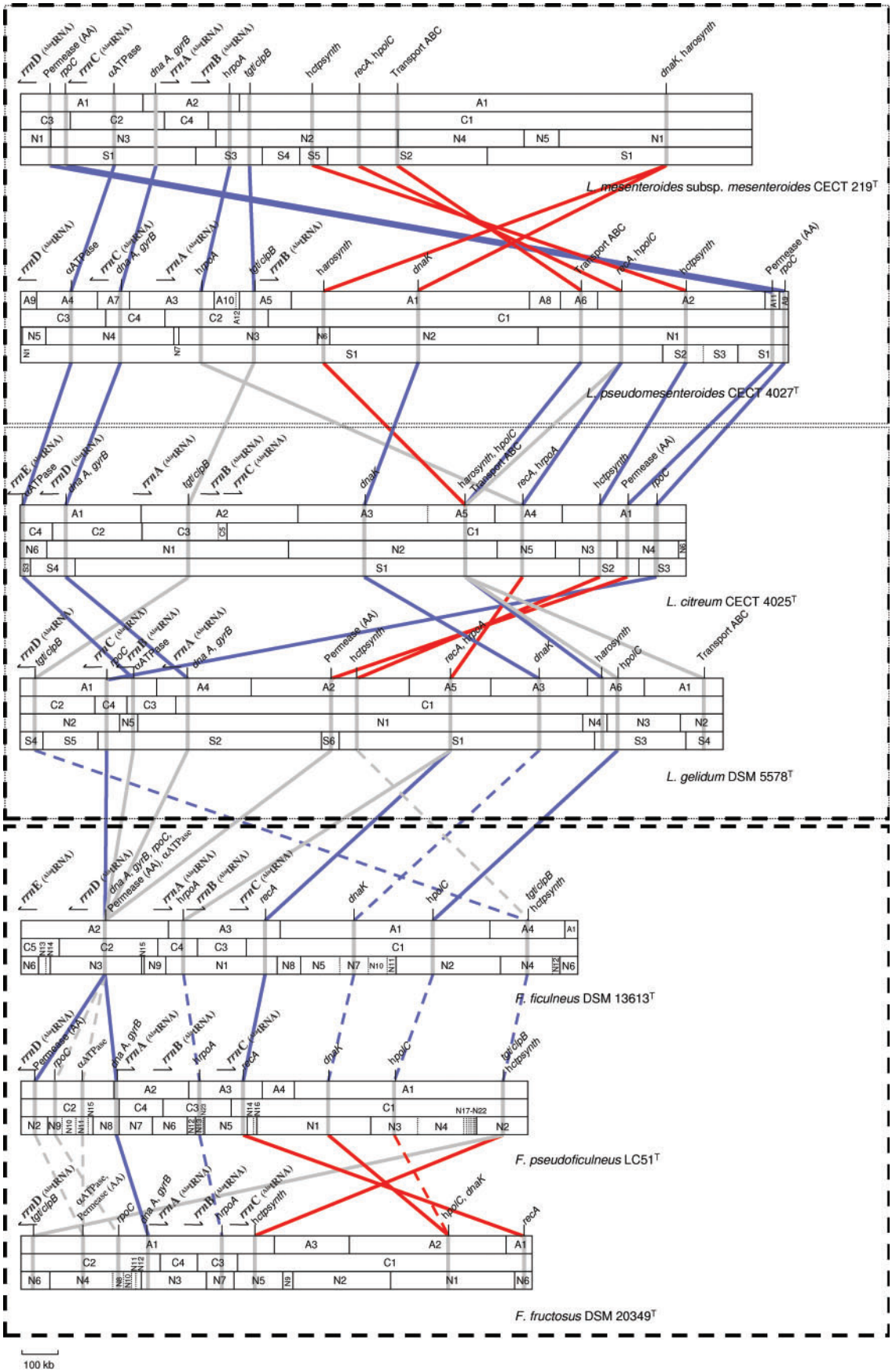
L. mesenteroides subsp. mesenteroides CECT 219T
L. pseudomesenteroides CECT 4027T
L. citreum CECT 4025T
L. gelidum DSM 5578T
F. ficulneus DSM 13613T
F. pseudoficulneus LC51T
F. fructosus DSM 20349T
100 kb

Fig. 1. Physical and genetic maps of type strains of *Leuconostoc* spp. and *Fructobacillus*. Restriction sites for *Asc*I, I-*Ceu*I, *Not*I and *Sfi*I are indicated. The circular genomes (see supplementary material) were linearized from the most distant *rrn* operon regarding *dnaA* that was localized in the counter-clockwise direction. Genetic markers are placed in the median position of the smallest fragment where they hybridized. Connecting lines between maps are colour-coded as follows: markers in conserved genomic regions in blue; markers in genomic regions subjected to inversions in red; and markers in other regions in grey (see text for details). Dashed lines in restriction maps indicate that the relative order of flanking fragments is unknown. Lines connecting markers in such regions are also dashed. The black dashed rectangles around maps limit comparisons within different genera.

The presence and order of genetic markers in these chromosomes also indicates a high degree of maintenance of genomic organization. Of the 13 markers analysed in these strains only Transport ABC and h*arosynth* are not present in all genomes, being absent in the *Fructobacillus* strains. A few markers are consistently found close to *oriC*. That is the case for αATPase, Permease (AA) and also *rpoC*. On the right side of the linear maps, where the terminus of replication *terC* is probably localized (corresponding to a 180° position in the circular maps) it is possible to find the presence of the *dnaK* marker. The presence of many markers in conserved regions is however reduced in comparisons of strains belonging to different clusters. Still, it should be noted that many grey connecting lines are just the result of the impossibility of separating markers in some chromosomes. Probably the most noteworthy feature concerning genomic rearrangements is the presence of large inversions. This is clearly seen in the comparisons of *L. mesenteroides* subsp. *mesenteroides* CECT 219^T with *L. pseudomesenteroides* CECT 4027^T and of *F. pseudoficulneus* LC51^T with *F. fructosus* DSM 20349^T, but it might also have occurred in the evolution of *L. citreum* CECT 4025^T or *L. gelidum* DSM 5578^T. From what is seen in the first two cases these inversions involve a large part of the chromosome, about 50 %, and are symmetrical with respect to *oriC*. Although the analysis of the sequenced genomes of *L. mesenteroides* subsp. *mesenteroides* CECT 219^T and *L. citreum* KM20 (Supplementary Fig. S4) cannot be fully comparable to that presented in Fig. 1 it is nevertheless consistent with the presence of a generally conserved genomic backbone and genomic inversions.

At a more restricted level, the presence of rearrangements can also be inferred even if their exact nature is unknown. This is the case for the linear order of *rpoC*-*rrn* operon-αATPase-*rrn* operon-*dnaA*, *gyrB* that can be seen not only in *L. pseudomesenteroides* CECT 4027^T but also in *L. citreum* CECT 4025^T and *L. gelidum* DSM 5578^T and that is changed in *L. mesenteroides* subsp. *mesenteroides* CECT 219^T.

An evaluation of genomic organization at the intra-specific level can be made from Fig. 2, which shows linear maps of *L. mesenteroides* and *L. citreum* strains; the linear map of the sequenced *L. citreum* KM20 (Kim *et al.*, 2008) is also included in this figure. Both in the comparison of *L. mesenteroides* subsp. *mesenteroides* CECT 219^T with *L. mesenteroides* subsp. *dextranicum* CECT 912^T and in the comparison of *L. citreum* CECT 4025^T and *L. citreum* KM20 the order of markers is highly conserved. It must be noted that the apparent discrepancy in the order of the h*polC*, Transport ABC, h*arosynth* and *dnaK* markers in the two *L. citreum* strains may just be the result of a misplacement of fragments A3 and A5 in *L. citreum* CECT 4025^T (in cases of lack of resolution the fragments are arbitrarily displayed in descending order of size, from left to right in the linear map). In contrast, and even taking into account possible inaccuracies, *L. mesenteroides* subsp. *cremoris* CECT 872^T and *L. citreum* CECT 4018^T show a greater number of differences in the order of markers. These indicate the occurrence of genomic rearrangements, which cannot be described in a simple way. In the case of the *L. mesenteroides* subsp. *cremoris* strain the differences in the order of markers from the other two *L. mesenteroides* strains would require at least two independent transposition events [involving *recA* and Permease (AA)] to be invoked. Nevertheless, differences in the disposition of the *rpoC*, αATPase and *dnaA*/*gyrB* and *rrn* operons imply that genomic rearrangements also occurred that changed the order of these markers and possibly the direction of gene transcription. In this respect, it should be pointed out that the *L. mesenteroides* subsp. *cremoris* strain seems to be more similar to the type strains of other close *Leuconostoc* species, possibly revealing the ancestral order of these markers. The comparison of the three *L. mesenteroides* maps is also interesting as regards genomic size. There are marked differences in the sizes of the chromosomes (Fig. 2), which seem to be circumscribed to the region flanked by the Transport ABC and *dnaK* markers, in contrast to the region surrounding *oriC*, where genomic size is more conserved.

Another significant result is the ubiquitous presence of the Transposase marker in the chromosomes of two *L. mesenteroides* strains and especially *L. citreum* CECT 4018. This marker is a linking clone (has an *Asc*I site) in *L. mesenteroides* subsp. *mesenteroides* CECT 219^T, where it is present in a plasmid (Makarova *et al.*, 2006; this work). The co-occurrence of an increased number of *Asc*I sites and hybridization targets for the Transposase marker in those chromosomes is a strong indication of the maintenance of that *Asc*I restriction site. The propagation of a transposase thus becomes the most likely explanation for the high number of *Asc*I fragments in the chromosomes of *L. mesenteroides* subsp. *dextranicum* CECT 912^T, *L. mesenteroides* subsp. *cremoris* CECT 872^T and *L. citreum* CECT 4018.

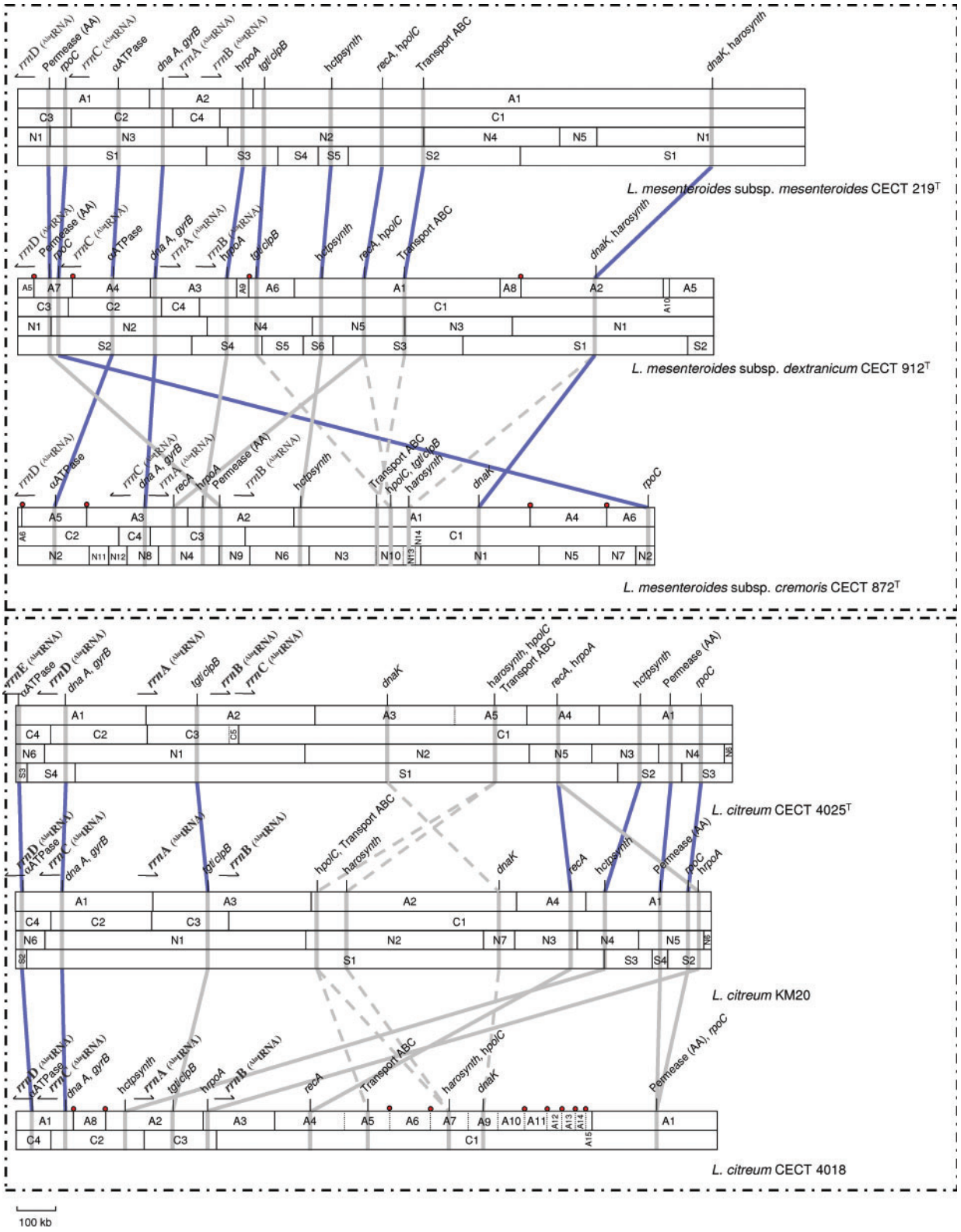

Fig. 2. Intra-specific comparison of physical and genetic maps of *Leuconostoc* strains. Restriction sites for *Asc*I, I-*Ceu*I, *Not*I and *Sfi*I are indicated. The circular genomes (see Supplementary material) were linearized from the most distant *rrn* operon regarding *dnaA* that was localized in the counter-clockwise direction. Genetic markers are placed in the median position of the smallest fragment where they hybridized. Small red circles indicate the location of the Transposase marker. Connecting lines between maps are colour-coded as in Fig. 1. Dashed lines in restriction maps indicate that the relative order of flanking fragments is unknown. Lines connecting markers in such regions are also dashed. The black dashed rectangles around maps limit comparisons within different species.

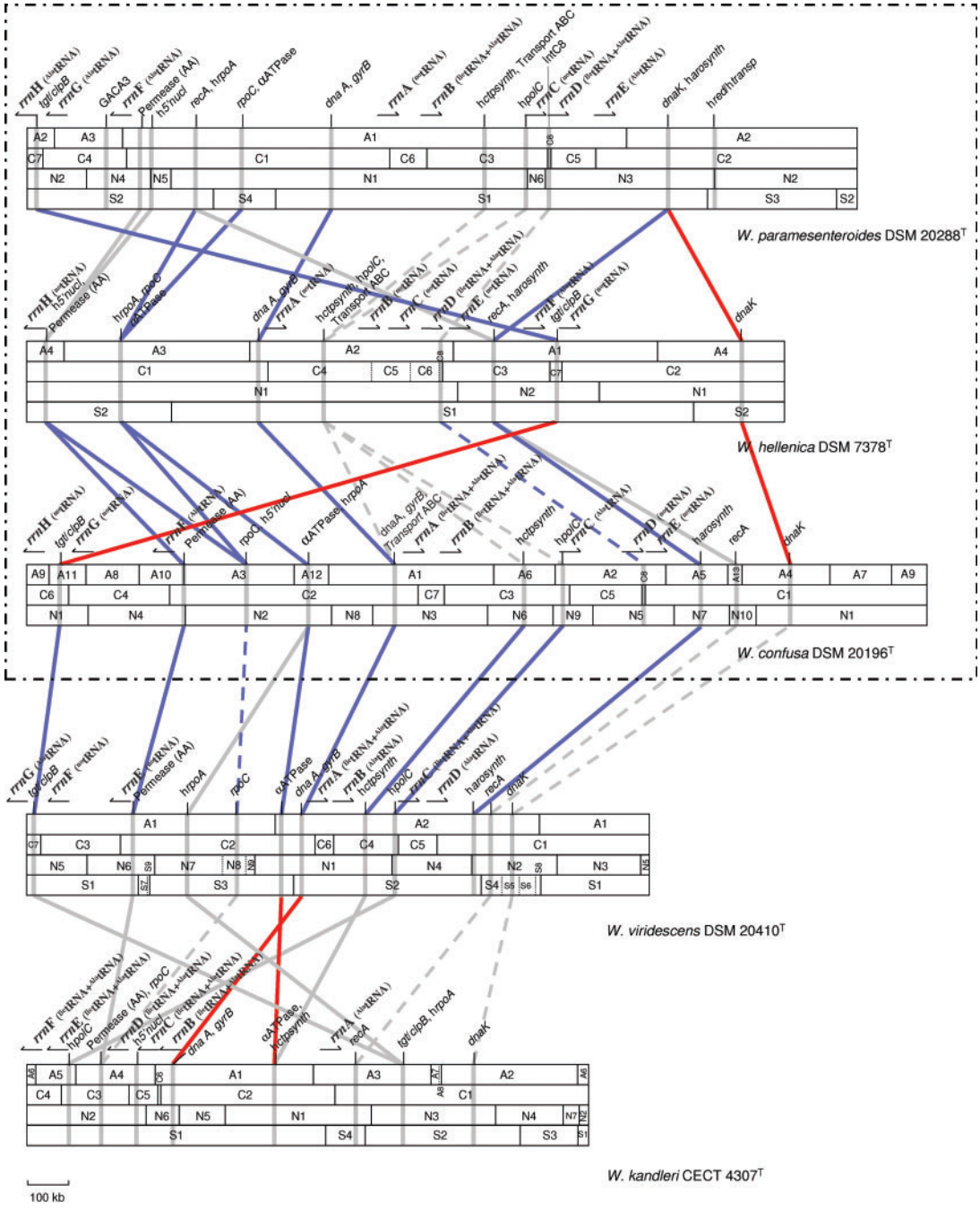

Fig. 3. Physical and genetic maps of type strains of *Weissella* spp. Restriction sites for *Asc*I, I-*Ceu*I, *Not*I and *Sfi*I are indicated. The circular genomes (see supplementary material) were linearized from the most distant *rrn* operon regarding *dnaA* that was localized in the counter-clockwise direction. Genetic markers are placed in the median position of the smallest fragment where they hybridized. Connecting lines between maps are colour-coded as in Fig. 1. Dashed lines in restriction maps indicate that the relative order of flanking fragments is unknown. Lines connecting markers in such regions are also dashed. The black dashed rectangle around maps limits the monophyletic group referred to as the *W. paramesenteroides* group.

Comparison of physical and genetic maps in *Weissella*

When comparing the genomic organization of *Weissella* strains (Fig. 3) with those described for *Fructobacillus* and *Leuconostoc* strains (Figs 1 and 2) it is possible to see that only a few general characteristics are preserved. Once again the ribosomal operons are transcribed divergently from *oriC* and the markers αATPase and *rpoC* are consistently found close to *dnaA*. *dnaK* again appears to be located in

the same region as *terC*. Apart from these conserved features many differences are evident. Concerning *rrn* operon organization in *Weissella*, *rrn* copy number can vary from 6 in *W. kandleri* CECT 4307^T to 8 in strains of the *W. paramesenteroides* cluster. Furthermore, *rrn* operons are more dispersed in the chromosomes of *Weissella*. In *W. kandleri* CECT 4307^T they are present in about 55 % of the chromosome and in *W. confusa* DSM 20196^T this value increases to 69 %. As part of this same organization *oriC* now lies in a large I-*Ceu*I fragment, in contrast to both *Fructobacillus* and *Leuconostoc* maps (Fig. 1), even if it still seems to be located close to a *rrn* operon. In *Wiessella* strains, the *rrn* operons are also unlike *Fructobacillus* or *Leuconostoc rrn* operons, in which they can have different tRNAs coded in their ITS regions. However, this characteristic fails in its possible use to match *rrn* operons in different strains since even clearly homologous *rrn* operons can have different tRNA sets (see *rrn* operons flanking the intC8 marker in Fig. 3), possibly as a result of intra-genomic recombination.

As previously seen, the genomic organization can be more different in more evolutionarily closely related strains than in more distant ones. The *W. paramesenteroides* DSM 20288^T chromosome appears to be more different from the *W. hellenica* DSM 7378^T chromosome than from that of *W. confusa* DSM 20196^T, to which it is less related (Chelo *et al.*, 2007). This is most evident in the order of *rrn* operons. Both the *rrn* disposition around *oriC* (3 : 5) and the relative distance between operons are more similar in the less related strains. In the evolution of the *W. hellenica* DSM 7378^T line it is also possible that an inversion involving *tgt*/*clpB* and *dnaK* markers has occurred. However, looking only at the comparison with the *W. paramesenteroides* strain the possibility of a transposition of the *dnaK* marker region cannot be discarded. When the less related *W. viridescens* DSM 20410^T and *W. kandleri* CECT 4307^T are compared with the strains of the *W. paramesenteroides* cluster (more particularly with *W. confusa* DSM 20196^T) the degree of conservation of gene order in the case of *W. viridescens* DSM 20410^T is remarkable, especially taking into account the differences in chromosome size. The region between the *rrnG* operon in *W. viridescens* DSM 20410^T (corresponding to the *rrnH* operon in *W. confusa* DSM 20196^T) and *rrnA* appears to have been less subject to changes in size than the rest of the chromosome. The chromosome of *W. kandleri* CECT 4307^T, which is the smallest of this study, also reveals a great number of differences, making it difficult to distinguish the presence of any rearrangements or even the correct placement of the map for comparative purposes.

DISCUSSION

The construction of physical and genetic maps, based on PFGE analysis of macrorestriction fragments and hybridization of genetic markers of several *Fructobacillus*, *Leuconostoc* and *Weissella* strains, enabled the comparison of genomic organization at different taxonomic levels in the Leuconostoc group of LAB. One of the first findings is that evolutionary relatedness is generally well reflected at the level of genomic organization in terms of the degree of synteny or maintenance of genomic features such as *rrn* operon organization. Thus, surprisingly, it is possible that even close strains show what seem to be marked differences, as in the case of the number of restriction fragments generated by rare-cutting enzymes. However, the strains used in this study show that large differences in the number of *Asc*I sites and to a lesser extent of *Not*I or *Sfi*I sites may not be significant to the overall chromosome organization and are thus meaningless at the level of genomic comparisons. It was possible to associate the presence of transposable elements with the increase in *Asc*I sites in several *Leuconostoc* strains. Interestingly, the Transposase marker that was used for this purpose was first found in a plasmid in *L. mesenteroides* subsp. *mesenteroides*, suggesting that it may be a plasmid with integration ability.

Another characteristic that was found to be highly variable even in close strains is chromosome size. The highest and lowest values were found in *W. confusa* DSM 20196^T (2197 kb) and *W. kandleri* CECT 4307^T (1371 kb) respectively, which are the most extreme values ever found for Leuconostocs (Lamoureux *et al.*, 1993; Tenreiro *et al.*, 1994) even if they are in the range expected for LAB species (Klaenhammer *et al.*, 2005). The relatively small size of LAB genomes is generally attributed to loss of genes (Makarova *et al.*, 2006) as the result of a continuous adaptation to specific nutrient-rich environments. In fact, in the case of *L. mesenteroides* subsp. *cremoris* CECT 872^T, and to a lesser extent in the case of *W. hellenica* DSM 7378^T, the possible loss of genes that is translated into smaller genome sizes and chromosomal rearrangements must have been accompanied by physiological and probably ecological changes. *L. mesenteroides* subsp. *cremoris* is well known for its markedly reduced metabolic spectrum when compared with other *L. mesenteroides* subspecies with which it has high DNA/DNA homology (Garvie, 1983; Holzapfel & Schilinger, 1991). In the same way, *W. hellenica* also presents a less diverse carbohydrate usage profile than *W. paramesenteroides* or *W. confusa* (Collins *et al.*, 1993). This effect is also seen in *W. kandleri* and *W. viridescens*, whose type strains also have smaller genomes.

The type of genomic rearrangement most relevant in the chromosomes under study is undoubtedly genomic inversion. Particularly relevant in *Fructobacillus* and *Leuconostoc*, inversions like these, which involve a large proportion of the chromosome and are symmetrical with respect to *oriC*, are nevertheless thought to be common. Chromosomal changes of this kind have already been observed in many natural strains of *Lactococcus lactis* (Le Bourgeois *et al.*, 2000), *Lactobacillus* (Klaenhammer *et al.*, 2005) and even in less related genera (Eisen *et al.*, 2000), and also in experimentally evolved strains (Campo *et al.*, 2004). They are considered not to be very deleterious since

both the average distance of each gene to the origin of replication and the transcription direction of most genes are largely maintained. The occurrence of inversions usually requires the presence of repeated sequences in the genome such as those from insertion sequences, prophages or multigenic families. Even if our study was not directed to finding these types of sequences we can at least say that neither *rrn* operons nor the transposable element associated with the Transposase marker seem to have been involved in the detected inversions.

A major part of this work deals with the determination of the organization of ribosomal operons in these genomes and its possible implications. Ribosomal operons can be important elements in the evolution of chromosomal organization. Since they constitute a multigenic family with a great degree of sequence conservation they are often associated with genomic rearrangements such as duplications, deletions and inversions. In fact, although variation in *rrn* operon copy number at the intraspecific level is uncommon (Acinas *et al.*, 2004) it was possible to identify a duplication in *Leuconostoc citreum*. This event could be attributed to the type strain *L. citreum* CECT 4025^T since the two other strains of this species have only four copies, which is the most frequent *rrn* number in *Fructobacillus* and *Leuconostoc*. Ribosomal operons seem also to be involved in, or at least are able to reveal, rearrangements in the three genera studied. A situation that appears to have resulted in the most extreme asymmetry is seen in *L. gelidum* DSM 5578^T, in which all the *rrn* operons are localized to one side of the replication origin.

The analysis of the *rrn* operon organization enables us to distinguish strains belonging to the three groups of Leuconostoc genera: *Fructobacillus* plus *Leuconostoc*, *Oenococcus* and *Weissella*. In *Fructobacillus* and *Leuconostoc* the four *rrn* copies, which may be considered as the ancestral value, are clustered around *oriC* and always have ^{Ala}tRNAs coded in their ITS regions. This type of ITS is also seen in both copies of the *O. oeni* chromosome (Zé-Zé *et al.*, 2000). In this case the *rrn* operons are usually several hundred kilobases apart, with one of them always close to *oriC*. In this respect, the study of the recently described *Oenococcus kitaharae* (Endo & Okada, 2006) would be of great interest in order to confirm the reduced number of *rrn* operons in the genus. *Weissella* is characterized by having the largest and most diverse set of *rrn* operons. The number of *rrn* operons varies between six and eight. This higher bound may be the result of duplications in specific *Weissella* groups, as can be deduced from the presence of the C8 fragment in strains of the *W. paramesenteroides* cluster. The higher number of *rrn* copies together with the existence of different sets of tRNAs in their ITS regions may reveal a closer proximity with other genera such as *Lactococcus*, *Lactobacillus* and *Pediococcus* (de Vries *et al.*, 2005; Klaenhammer *et al.*, 2002). As in *O. oeni* the origin of replication is localized in one large I-*Ceu*I fragment although there is always a group of *rrn* operons nearby.

It is not straightforward to provide an explanation for the relationship of these different sets of *rrn* operons with the diversification of the three genera. If the variation in the number of *rrn* operons reflects adaptations to different environments with different degrees of stability of resources, as was previously shown (Klappenbach *et al.*, 2000) and proposed for *W. paramesenteroides* DSM 20288^T (Chelo *et al.*, 2004), then it is possible that the fast diversifying process proposed for these genera (Chelo *et al.*, 2007) reflects radical changes in the ecological niches that were occupied.

Regarding the current taxonomic classification of this group, and specifically its most recent changes (Endo & Okada, 2008), the genomic characteristics presented here support the recognition of the *Fructobacillus* strains as a monophyletic group. This is reflected in the general conservation of number of restriction sites, genome size and presence/absence of hybridization signals. However, it must be noted that, in what seems to be the most distinctive feature of the remaining Leuconostoc genera (*rrn* number and disposition), *Fructobacillus* strains are very similar to *Leuconostoc* strains.

The physical and genetic maps described here constitute a study at different taxonomic levels that provides an analysis of the evolution of the genomes of different Leuconostoc species, enabling further insights into the evolution and diversification of the species belonging to the genera *Leuconostoc*, *Oenococcus* and *Weissella*.

ACKNOWLEDGEMENTS

The financial support by FCT research grants to I. M. Chelo (SFRH/BD/10675/2002 and SFRH/BDP/27887/2006) and L. Zé-Zé (SFRH/BPD/3653/2000) is gratefully acknowledged.

REFERENCES

Acinas, S. G., Marcelino, L. A., Klepac-Ceraj, V. & Polz, M. F. (2004). Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J Bacteriol* **186**, 2629–2635.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Campo, N., Dias, M. J., Daveran-Mingot, M. L., Ritzenthaler, P. & Le Bourgeois, P. (2004). Chromosomal constraints in Gram-positive bacteria revealed by artificial inversions. *Mol Microbiol* **51**, 511–522.

Chelo, I. M., Zé-Zé, L., Chambel, L. & Tenreiro, R. (2004). Physical and genetic map of the *Weissella paramesenteroides* DSMZ 20288^T chromosome and characterization of different *rrn* operons by ITS analysis. *Microbiology* **150**, 4075–4084.

Chelo, I. M., Zé-Zé, L. & Tenreiro, R. (2007). Congruence of evolutionary relationships inside the *Leuconostoc-Oenococcus-Weissella* clade assessed by phylogenetic analysis of the 16S rRNA gene, *dnaA, gyrB, rpoC* and *dnaK*. *Int J Syst Evol Microbiol* **57**, 276–286.

Collins, M. D., Samelis, J., Metaxopoulos, J. & Wallbanks, S. (1993). Taxonomic studies on some Leuconostoc-like organisms from fermented sausages: description of a new genus *Weissella* for the *Leuconostoc paramesenteroides* group of species. *J Appl Bacteriol* **75**, 595–603.

De Bruyne, K., Schillinger, U., Caroline, L., Boehringer, B., Cleenwerck, I., Vancanneyt, M., De Vuyst, L., Franz, M. A. P. & Vandamme, P. (2007). *Leuconostoc holzapfelii* sp. nov., isolated from Ethiopian coffee fermentation and assessment of sequence analysis of housekeeping genes for delineation of *Leuconostoc* species. *Int J Syst Evol Microbiol* **57**, 2952–2959.

de Vries, M. C., Siezen, R. J., Wijman, J. G. E., Zhao, Y., Kleerebezem, M., de Vos, W. M. & Vaughan, E. E. (2005). Comparative and functional analysis of the rRNA-operons and their tRNA gene complement in different lactic acid bacteria. *Syst Appl Microbiol* **29**, 358–367.

Eisen, J. A., Heidelberg, J. F., White, O. & Salzberg, S. L. (2000). Evidence for symmetric chromosomal inversions around the replication origin in bacteria. *Genome Biol* **1**, RESEARCH0011.

Endo, A. & Okada, S. (2006). *Oenococcus kitaharae* sp. nov., a non-acidophilic and non-malolactic-fermenting oenococcus isolated from a composting distilled shochu residue. *Int J Syst Evol Microbiol* **56**, 2345–2348.

Endo, A. & Okada, S. (2008). Reclassification of the genus *Leuconostoc* and proposals of *Fructobacillus fructosus* gen. nov., comb. nov., *Fructobacillus durionis* comb. nov., *Fructobacillus ficulneus* comb. nov. and *Fructobacillus pseudoficulneus* comb. nov. *Int J Syst Evol Microbiol* **58**, 2195–2205.

Garvie, E. I. (1983). *Leuconostoc mesenteroides* subsp. *cremoris* (Knudsen and Sorensen) comb. nov. and *Leuconostoc mesenteroides* subsp. *Dextranicum* (Beijerinck) comb. nov. *Int J Syst Bacteriol* **33**, 118–119.

Holzapfel, W. H. & Schilinger, U. (1991). The Genus *Leuconostoc*. In *The Prokaryotes*, vol. 3, pp. 1508–1533. Edited by A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K. Scheifer. Berlin: Springer.

Kandler, O. & Weiss, N. (1986). Regular, nonsporing gram-positive rods. In *Bergey's Manual of Systematic Bacteriology*, vol. 2, pp. 1208–1234. Edited by J. G. Holt, P. H. A. Sneath, N. S. Mair & M. E. Sharpe. Baltimore: Williams & Wilkins.

Kim, J. F., Jeong, H., Lee, J., Choi, S., Ha, M., Hur, C., Kim, J., Lee, S., Park, H. & other authors (2008). The complete genome sequence of *Leuconostoc citreum* KM20. *J Bacteriol* **190**, 3093–3094.

Klaenhammer, T., Altermann, E., Arigoni, F., Bolotin, A., Breidt, F., Broadbent, J., Cano, R., Chaillou, S., Deutscher, J. & other authors (2002). Discovering lactic acid bacteria by genomics. *Antonie Van Leeuwenhoek* **82**, 29–58.

Klaenhammer, T. R., Barrangou, R., Buck, B. L., Azcarate-Peril, M. A. & Alterman, E. (2005). Genomic features of lactic acid bacteria effecting bioprocessing and health. *FEMS Microbiol Rev* **29**, 393–409.

Klappenbach, J. A., Dunbar, J. M. & Schmidt, T. M. (2000). rRNA copy number reflects ecological strategies of bacteria. *Appl Environ Microbiol* **66**, 1328–1333.

Lamoureux, M., Prévost, H., Cavin, J. F. & Diviès, C. (1993). Recognition of *Leuconostoc oenos* strains by the use of DNA restriction profiles. *Appl Microbiol Biotechnol* **39**, 547–552.

Le Bourgeois, P., Daveran-Mingot, M. & Ritzenthaler, P. (2000). Genome plasticity among related *Lactococcus* strains: identification of genetic events associated with macrorestriction polymorphisms. *J Bacteriol* **182**, 2481–2491.

Makarova, K. S. & Koonin, E. V. (2007). Evolutionary genomics of lactic acid bacteria. *J Bacteriol* **189**, 1199–1208.

Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V. & other authors (2006). Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* **103**, 15611–15616.

Marcobal, A. M., Sela, D. A., Wolf, Y. I., Makarova, K. S. & Mills, D. A. (2007). The role of hypermutability in the evolution of the genus *Oenococcus*. *J Bacteriol* **190**, 564–570.

Martinez-Murcia, A. J. & Collins, M. D. (1990). A phylogenetic analysis of the genus *Leuconostoc* based on reverse transcriptase sequencing of 16S rRNA. *FEMS Microbiol Lett* **58**, 73–84.

Pitcher, D. G., Saunders, N. A. & Owen, R. J. (1989). Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett Appl Microbiol* **8**, 151–156.

Tenreiro, R., Santos, M. A., Paveia, H. & Vieira, G. (1994). Inter-strain relationships among wine leuconostocs and their divergence from other *Leuconostoc* species, as revealed by low frequency restriction fragment analysis of genomic DNA. *J Appl Bacteriol* **77**, 271–280.

Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K. & Swings, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* **60**, 407–438.

Yang, D. & Woese, C. R. (1989). Phylogenetic structure of the "Leuconostocs": an interesting case of a rapidly evolving organism. *Syst Appl Microbiol* **12**, 145–149.

Zé-Zé, L., Tenreiro, R., Brito, L., Santos, M. & Paveia, H. (1998). Physical map of the genome of *Oenococcus oeni* PSU-1 and localization of genetic markers. *Microbiology* **144**, 1145–1156.

Zé-Zé, L., Tenreiro, R. & Paveia, H. (2000). The *Oenococcus oeni* genome: physical and genetic mapping of strain GM and comparison with the genome of a 'divergent' strain, PSU-1. *Microbiology* **146**, 3195–3204.

Zé-Zé, L., Chelo, I. M. & Tenreiro, R. (2008). Genome organization in *Oenococus* strains studied by comparison of physical and genetic maps. *Int Microbiol* **11**, 237–244.

Edited by: H.-P. Klenk