

Allelic Diversity and Population Structure in *Oenococcus oeni* as Determined from Sequence Analysis of Housekeeping Genes

Blanca de las Rivas, Ángela Marcobal, and Rosario Muñoz*

Instituto de Fermentaciones Industriales, CSIC, Madrid, Spain

Received 8 June 2004/Accepted 20 July 2004

***Oenococcus oeni* is the organism of choice for promoting malolactic fermentation in wine. The population biology of *O. oeni* is poorly understood and remains unclear. For a better understanding of the mode of genetic variation within this species, we investigated by using multilocus sequence typing (MLST) with the *gyrB*, *pgm*, *ddl*, *recP*, and *mleA* genes the genetic diversity and genetic relationships among 18 *O. oeni* strains isolated in various years from wines of the United States, France, Germany, Spain, and Italy. These strains have also been characterized by ribotyping and restriction fragment length polymorphism (RFLP) analysis of the PCR-amplified 16S-23S rRNA gene intergenic spacer region (ISR). Ribotyping grouped the strains into two groups; however, the RFLP analysis of the ISRs showed no differences in the strains analyzed. In contrast, MLST in oenococci had a good discriminatory ability, and we have found a higher genetic diversity than indicated by ribotyping analysis. All sequence types were represented by a single strain, and all the strains could be distinguished from each other because they had unique combinations of alleles. Strains assumed to be identical showed the same sequence type. Phylogenetic analyses indicated a panmictic population structure in *O. oeni*. Sequences were analyzed for evidence of recombination by split decomposition analysis and analysis of clustered polymorphisms. All results indicated that recombination plays a major role in creating the genetic heterogeneity of *O. oeni*. A low standardized index of association value indicated that the *O. oeni* genes analyzed are close to linkage equilibrium. This study constitutes the first step in the development of an MLST method for *O. oeni* and the first example of the application of MLST to a nonpathogenic food production bacteria.**

Oenococcus oeni, formerly *Leuconostoc oenos* (7), is the species of lactic acid bacteria (LAB) most frequently associated with malolactic fermentation (MLF) in wine. MLF, which occurs after alcoholic fermentation during wine making, is induced by the growth of LAB. Normally, spontaneous MLF takes place when LAB develop in wine after alcoholic fermentation. However, when fermentation by indigenous bacteria is relied upon, the fitness of the bacteria present to carry out MLF may be highly variable, and consequently their wine-making properties are unpredictable. In order to exercise greater control over wine-making processes, common wine-making practices therefore involve inoculation of wine with either commercially prepared strains or in-house winery strains of malolactic bacteria (4). Therefore, the differentiation of *O. oeni* strains at the strain level becomes a major concern, since their adaptation to wine and influence on organoleptic quality are strain specific (2). Moreover, manufacturers of malolactic starters need accurate control of their products, and wine-makers have to be able to recognize the inoculated strains during vinification. The need for positive identification of different isolates is also acknowledged by research workers in the field, since many strains from diverse origins are often exchanged between laboratories, and no reliable phenotypic method for certifying their identities is available (20). Because strains of *O. oeni* are not readily typed with phenotypic tests (20), the study of the genome structure can constitute a useful tool for the taxonomic characterization of strains and their

typing. Therefore, unambiguous, discriminatory isolate characterization schemes are essential for population genetic and evolutionary studies.

The genus *Oenococcus* (7) contains a sole species, *O. oeni*, with a restricted ecological niche (wine and related habitats); although its dissimilarity to the genus *Leuconostoc* is generally accepted, there is some controversy regarding both the evolution rate of the species and its diversity (6, 46). Despite the exhaustive phenetic and molecular studies that have been performed on *O. oeni*, little is known about its population genetics. Studies carried out by using different molecular techniques such as chromosomal DNA-DNA hybridizations (7), 16S and 23S rRNA sequence analysis (39), and 16S-23S rRNA gene (rDNA) intergenic spacer region (ISR) sequencing (30, 46) and also by DNA fingerprinting (29, 44), pulsed-field gel electrophoresis (26), and randomly amplified polymorphic DNA (RAPD) analysis (45) suggest that this species is homogeneous. However, metabolic or physiological criteria, such as lactate dehydrogenases (17), carbohydrate fermentation (18), and cellular fatty acid patterns (42), have shown considerable diversity among strains of *O. oeni*. Therefore, *O. oeni* has often been referred to as a highly heterogeneous species, and splitting it into two species (38) or subspecies (41) has been proposed.

Multilocus sequence typing (MLST), a method that is based on partial nucleotide sequences of multiple housekeeping genes, has recently been shown to be a powerful technique for bacterial typing (11). Housekeeping genes are preferred because an analysis of mutations in such genes is more likely to properly reflect the phylogeny of strains. MLST uses variation that accumulates slowly, which is expected to be selectively neutral, and achieves very high resolution by analyzing multi-

* Corresponding author. Mailing address: Departamento de Microbiología, Instituto de Fermentaciones Industriales, CSIC, Juan de la Cierva 3, 28006 Madrid, Spain. Phone: 34 91 562 2900. Fax: 34 91 564 4853. E-mail: rmunoz@ifi.csic.es.

TABLE 1. Properties of *O. oeni* isolates analyzed and their allele profiles at each locus

Strain no.	Strain	RT ^a	ST	Allele no. at locus:					Source of isolate	
				<i>gyrB</i>	<i>ddl</i>	<i>pgm</i>	<i>recP</i>	<i>mleA</i>	Country	Year
1	MCW (PSU-1)	ND ^b	1	1	1	1	1	1	California, United States	1958
2	CECT 4028 (DSMZ 20255)	B	2	2	1	2	2	1	Bordeaux, France	1961
3	CECT 4029 (DSMZ 20257)	B	3	3	2	2	2	1	CSIRO, Australia	1958
4	CECT 4100 ^T (ATCC 23279)	B	4	4	1	3	2	1	Bordeaux, France	1961
5	CECT 4721 (ATCC 23278)	B	5	4	3	4	2	1	Merbein, Australia	1958
6	CECT 4725 (ATCC 23277)	B	6	4	1	2	3	1	Bordeaux, France	1961
7	CECT 4728 (ML27)	B	7	5	1	5	3	1	California, United States	1965
8	CECT 4758	A	8	6	4	6	1	2	Valladolid, Spain	1995
9	BIFI-1	A	9	6	5	7	2	1	Valladolid, Spain	1999
10	BIFI-9	A	10	6	5	8	1	2	Logroño, Spain	1999
11	BIFI-21	A	11	6	1	8	4	2	Madrid, Spain	2001
12	BIFI-26	A	12	7	7	9	5	1	Valladolid, Spain	2001
13	BIFI-86	A	13	6	1	8	6	2	Logroño, Spain	2003
14	5001	A	14	8	8	8	7	1	Italy	ND
15	Uvaferm ALPHA	A	15	6	9	2	2	1	Bordeaux, France	ND
16	Uvaferm MLD	A	16	6	1	10	1	2	ND	ND
17	Viniflora OENOS	A	17	6	1	11	8	2	ND	1993
18	Viniflora CH35	A	18	6	10	12	5	2	Burgundy, France	ND

^a RT, ribotype. Ribotypes are as defined by Zavaleta et al. (47).

^b ND, no data available.

ple loci. The MLST method was introduced in 1998 and since then has been used for phylogenetic analysis of many bacterial pathogens such as *Neisseria meningitidis* (31), *Streptococcus pneumoniae* (10), *Haemophilus influenzae* (35), etc. MLST has led to a better understanding of the mode of genetic variation within a bacterial species. Since 1998, the MLST scheme has been applied to important bacterial pathogens including several food-borne human pathogens such as *Campylobacter jejuni* (8), *Vibrio cholerae* (14), and *Bacillus cereus* (23); however, until now, MLST had not been applied to a nonpathogenic food production bacteria used in the food industry, where precise methods for characterizing isolates are required as well.

The present study was undertaken to evaluate the discriminatory power of MLST in *O. oeni*. The results of our analysis indicate that the *O. oeni* population exhibits high genetic diversity close to linkage equilibrium with a panmictic population structure. This study also constitutes the first step for the development of an MLST method for *O. oeni*. The sequence diversity of five genes, possible candidates to be included in a future MLST scheme, was analyzed.

MATERIALS AND METHODS

Strains and DNA preparation. In this study, we examined a total of 18 *O. oeni* isolates (Table 1), including reference strains provided by the Spanish Type Culture Collection (CECT) and strains belonging to the bacterial culture collection of the Instituto de Fermentaciones Industriales (BIFI), CSIC, Madrid, Spain (36). Four additional *O. oeni* strains were isolated from commercial malolactic starter preparations provided by Lallemand Inc. (Montreal, Canada) and Christian Hansen A/S (Horsholm, Denmark). Strain 5001 was kindly provided by E. García-Moruno from the Instituto Sperimentale per l'Enologia, Asti, Italy. Strain *O. oeni* 51 was isolated from the same wine fermentation and showed a RAPD profile identical to that of *O. oeni* BIFI-21 (data not shown); therefore, we assumed that the strains are identical. To check the stability of the gene sequences, BIFI-21 was subcultured during approximately 250 generations in culture medium, and its DNA was extracted. Isolates were grown on medium for *L. oenos* (3) supplemented with 10% tomato juice. Strains were incubated at 30°C in a 5% CO₂ atmosphere. Chromosomal DNA was prepared as described previously (46).

RFLP of the PCR-amplified 16S-23S rDNA ISR. Restriction fragment length polymorphism (RFLP) analysis of the ISRs was performed as described previ-

ously (46). Briefly, the *O. oeni* ISR PCR-specific reactions were performed by using the primers 16S14f (5'-CTGTACACACCGCCCGTC) and 23S1R (5'-GGTTTCCCCATTCGGAAATC) described by Zavaleta et al. (46). These primers amplified a 650-bp fragment in all the *O. oeni* strains tested. The PCR was performed with approximately 100 ng of total DNA in 0.2-ml microcentrifuge tubes in a total volume of 50 µl as described previously (46). The amplified 16S-23S ISRs from *O. oeni* strains were digested with the restriction enzymes DdeI, CfoI, NdeI, and TaqI (Roche). The digested products were separated by electrophoresis in 4.5% MS-8 agarose gels (Hispanlab).

Ribotyping. Chromosomal DNA was digested with HindIII and EcoRI (Roche), and the products were separated by electrophoresis in 0.7% agarose gels in 1× Tris-acetate-EDTA buffer.

Digested DNA was transferred onto positively charged nylon membranes (Roche) by the Southern method. Probe 16S rDNA was obtained from *O. oeni* CECT 4100^T by PCR by using the eubacterial universal pair of primers 63f and 1387r (32). The 16S rDNA probe was digoxigenin labeled and detected by chemiluminescence by using a DIG-High Prime DNA Labeling and Detection Starter Kit (Roche) according to the manufacturer's instructions.

PCR amplification and DNA sequencing. Since the genome sequence of *O. oeni* had not been published when this study was performed, we identified a number of candidate loci by searching the *O. oeni* genome database (http://genome.jgi-psf.org/draft_microbes/oenoe/oenoe.home.html) with gene sequences from other bacteria. The following four loci were chosen for the sequence analysis scheme: *gyrB* (coding for the B subunit of DNA gyrase and located on contig NZ_AABJ02000001.1), *pgm* (coding for phosphoglucomutase and located on contig NZ_AABJ02000004.1), *ddl* (coding for D-alanine-D-alanine ligase, located on contig NZ_AABJ02000021.1), and *recP* (coding for transketolase and located on contig NZ_AABJ02000005.1). These genes were selected based on the criterion that they are presumptively widely separated on the chromosome. A gene (*mleA*) coding for the malolactic enzyme, an enzyme of technological interest in *O. oeni*, was also included in the study (Table 2). The *mleA* gene is located on contig NZ_AABJ02000005.1 but 85 kb apart from the *recP* gene in the *O. oeni* strain MCW chromosome. Primers were designed by using highly conserved DNA regions of these genes for *O. oeni* obtained from the GenBank database, with the exception of *mleA*, for which we used primers previously described by Divol et al. (9).

PCR was performed to amplify gene fragments from chromosomal DNA of the *O. oeni* strains. Each 25-µl amplification reaction mixture contained 10 ng of template DNA, 20 mM Tris-HCl (pH 8.0), 50 mM KCl, 2.5 mM MgCl₂, a 200 µM concentration of each deoxynucleoside triphosphate, a 1 µM concentration of each primer, and 1 U of AmpliTaq Gold DNA polymerase. The reaction was performed in a GeneAmp PCR System 2400 (Perkin-Elmer) by using the following cycling parameters: 10 min of enzyme activation at 95°C, followed by 30 cycles of 30 sec at 95°C, 1 min at 48°C, and 1 min at 72°C. Amplified products were resolved on a 1.5% agarose gel. The amplification products were purified

TABLE 2. Primers used for MLST of *O. oeni* strains

Protein	Accession no. ^a	Gene	Primers	Sequence 5'→3'	5' start position ^a	PCR product length (bp)	Accession no.
B subunit of DNA gyrase	ZP_00069208	<i>gyrB</i>	GYRB1	TGGGCTTCATGGTGTGGC	375	947	AJ618989–AJ618996
			GYRB2	CCCTCGACGATAAACAATTC	1322		
D-Alanine-D-alanine ligase	ZP_00069716	<i>ddl</i>	DDL1	CGATGTTAGCAAGCGTTTCG	54	911	AJ618997–AJ619006
			DDL2	TTCGTATTTCCCGTAGTG	965		
Phosphoglucosyltransferase	ZP_00070472	<i>pgm</i>	PGM1	CATCCGACTCCGGAATTGAC	370	826	AJ619007–AJ619018
			PGM2	CCGTAGGATTCTTCAAACC	1196		
Transketolase	ZP_00069744	<i>recP</i>	RECP1	GGCGATGGGACTTAATGG	550	926	AJ619666–AJ619673
			RECP2	CCCCTCTTCGCCGACAGC	1476		
Malolactate dehydrogenase	ZP_00069821	<i>mleA</i>	OO1 ^b	GTGCCGCTTTTTTGGATATTA	332	430	AJ619674–AJ619675
			OO2 ^b	AGCAATTTTATCTTTATAGCT	762		

^a The protein accession number and the indication of the gene position refer to the uncompleted genome sequence of *O. oeni* MCW (formerly, PSU-1) by the DOE Joint Genome Institute (http://genome.jgi-psf.org/draft_microbes/oenoe/oenoe.home.html).

^b The *mleA* primers have been described by Divol et al. (9).

on QIAquick spin columns (QIAGEN) for direct sequencing with the same PCR primers by the Sanger dideoxynucleotide chain termination method by using an ABI Prism dRhodamine Terminator Cycle Sequencing Ready Reaction Kit (PE Biosystems, Warrington, England), in accordance with the manufacturer's protocol, and an ABI Prism 377 DNA sequencer (Applied Biosystems, Inc.).

Phylogenetic analysis. For each locus, the sequences obtained for all isolates were compared, and the different sequences were assigned arbitrary allele numbers. For each isolate, the combination of alleles obtained at each locus defined its allelic profile (Table 1). Each isolate was therefore designated by five numbers, constituting an allelic profile or sequence type (ST). We refer to a unique combination of alleles as an ST. The STs were identified by arbitrary numbers assigned in order of description.

Sequence alignments and comparison were done with the program BioEdit, version 4.8.10 (<http://jwbrown.mbio.ncsu.edu/BioEdit/bioedit.html>) (21), and converted into MEGA and NEXUS files with START (sequence type analysis and recombinatorial tests) (<http://outbreak.ceid.ox.ac.uk/software.html>). Phylogenetic trees were compiled with MEGA version 2.1 software (<http://www.megasoftware.net>) by using the unweighted pair group method with arithmetic averages (UPGMA) (28).

The method of split decomposition was used to assess the degree of tree-like structure present in the alleles found for each locus in the complete set of 18 isolates (24). The sequence alignments were converted to NEXUS files, and the split decomposition was performed with SplitsTree 2.0 (<http://bibiserv.techfak.uni-bielefeld.de/splits/>).

Three types of statistical analysis were applied to the data: the index of association (I_A), homoplasy test, and Sawyer's run test. Multilocus linkage disequilibrium was estimated by measuring I_A (34). The homoplasy test (33) and the Sawyer's run test (40) were performed according to methods described previously with the START program. Using the START program, we also performed a test to detect selection in our population, and we calculated the d_N/d_S ratio as described by Nei and Gojobori (37).

Nucleotide sequence accession numbers. The sequences of all alleles have been deposited in the EMBL and GenBank databases under the accession numbers AJ618989 to AJ618996 (*gyrB*[r] fragment), AJ618997 to AJ619006 (*ddl* fragment), AJ619007 to AJ619018 (*pgm* fragment), AJ619666 to AJ619673 (*recP* fragment), and AJ61974 to AJ619675 (*mleA* fragment) (Table 2).

RESULTS

Characterization of *O. oeni* strains by ribotyping and RFLP analysis of the ISRs. The *O. oeni* strains recovered from various years and geographic locations were expected to be diverse. Therefore, PCR products of the 16S-23S rDNA ISRs from all *O. oeni* strains were digested with endonucleases TaqI, CfoI, DdeI, and NdeI. The 18 strains analyzed showed identical RFLP patterns for each enzyme. The TaqI and CfoI RFLP patterns were coincident with those previously described by Zavaleta et al. (46), with six bands (40, 75, 90, 100, 130, and 220 bp) and three bands (110, 230, and 330 bp), respectively. However, DdeI digestion originates a three-band RFLP pattern

(75, 110, and 400 bp), and, unexpectedly, the 16S-23S rDNA ISR PCR fragments remain uncut by NdeI.

Ribopattern analysis with EcoRI or HindIII revealed three bands for all *O. oeni* strains. Two different groups of strains were found, corresponding to groups defined previously by Zavaleta et al. (47). Ribotype A showed bands of 1.6, 2.5, and 4 kb for EcoRI and 1, 5, and 12 kb for HindIII; ribotype B showed bands of 1.6, 2.5, and 8 kb for EcoRI and 1, 5, and 6 kb for HindIII. Table 1 lists the assignment of ribotype group A or B to each *O. oeni* strain. The ribotyping patterns of the EcoRI digest are shown in Fig. 1.

Diversity of gene fragments. We sequenced nucleotide fragments of five genes for each of the 18 *O. oeni* isolates. Table 3 shows the fragment size for each locus used for this analysis. For all loci, the 18 sequences aligned without gaps or insertions. As shown in Table 3, all loci were polymorphic, and the number of polymorphic nucleotide sites varied between 1 (*mleA*) and 36 (*recP*). The proportion of variable sites present in the alleles ranged from 0.3% in the *mleA* locus to 6.6% in the *recP* locus, with 1.7, 2.2, and 2.5% for the *pgm*, *ddl*, and *gyrB* loci, respectively. Figure 2 shows the position of the polymorphic sites within sequenced fragments for all loci. Nucleotide substitutions occurring at the third base of a codon ranged from 64% in *gyrB* to 20% in *ddl* (Fig. 2). Phylogenetically

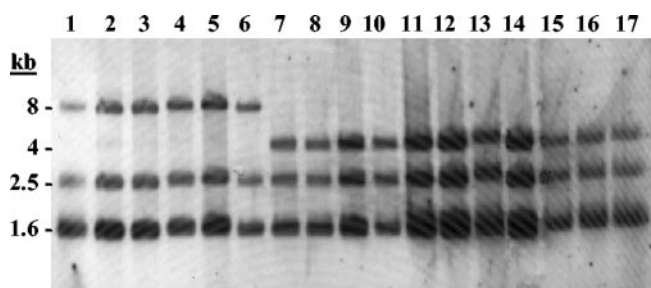


FIG. 1. Ribotyping patterns of EcoRI digest of chromosomal DNAs from *O. oeni* strains. Lane 1, CECT 4028; lane 2, CECT 4029; lane 3, CECT 4100; lane 4, CECT 4721; lane 5, CECT 4725; lane 6, CECT 4728; lane 7, CECT 4758; lane 8, BIFI-1; lane 9, BIFI-9; lane 10, BIFI-21; lane 11, BIFI-26; lane 12, BIFI-86; lane 13, 5001; lane 14, Uvaferm ALPHA; lane 15, Uvaferm MLD; lane 16, Viniflora OENOS; lane 17, Viniflora CH35. The molecular sizes (in kilobases) of the labeled fragments are indicated on the left.

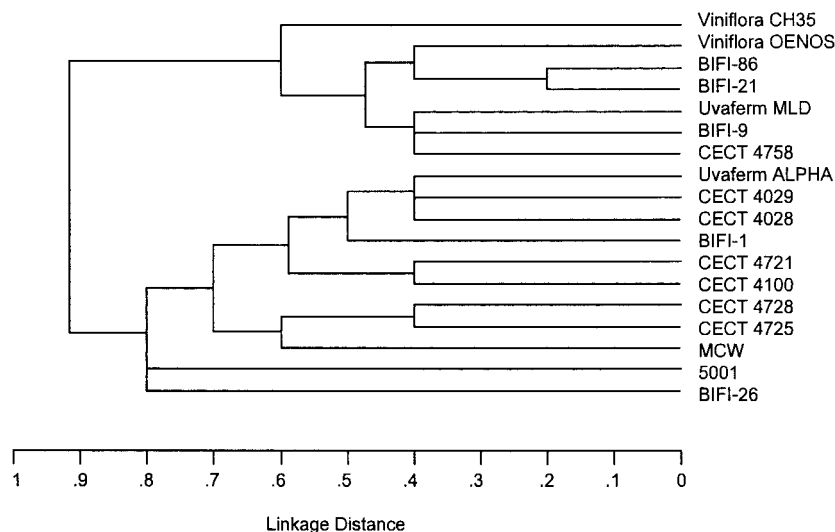


FIG. 3. UPGMA dendrogram showing the genetic relatedness of the 18 *O. oeni* strains examined in this study. The dendrogram was constructed from a matrix of the pairwise distances between the allelic profiles of 18 *O. oeni* strains.

related with the geographic origin of the strains. Concatenated *gyrB*, *ddl*, *pgm*, *recP*, and *mleA* gene sequence fragments were analyzed and graphically displayed with SplitsTree (Fig. 4). The relationship of *O. oeni* isolates based on this analysis is depicted as a star-like structure with rays of different lengths. This star phylogeny is also consistent with a recombinational population structure and with the observation that each strain carries its own allelic combination. Isolates *O. oeni* CECT 4028, CECT 4721, CECT 4725, CECT 4728, and 5001 are more closely related than the other strains and their tree branches are interconnected, suggesting recombinational events between them. Consistent with the lineage assignment by BURST, this placed the predicted *O. oeni* founder strains (CECT 4725, Uvaferm ALPHA, and Viniflora OENOS) at a central position of the split graph. The relationships among other members of the group were

assessed by examining the number of nodes between two isolates.

The d_N/d_S ratio, where d_N indicates the number of nonsynonymous substitutions per nonsynonymous site and d_S indicates the number of synonymous substitutions per synonymous site, was calculated as a measure of the degree of selection in the population. Investigating the number of synonymous and nonsynonymous substitutions may, therefore, provide information about the degree of selection operating on a population. The d_N/d_S ratio was calculated for all five loci, and in *gyrB* and *mleA* it was <10% (Table 3). However, the d_N/d_S ratios of the *recP* (22.46%), *pgm* (43.94%), and *ddl* (62.01%) loci could indicate that the degree of selection is not as strong as in typical housekeeping genes. In the *mleA* locus, the only nucleotide substitution was synonymous and did not affect the amino acid composition.

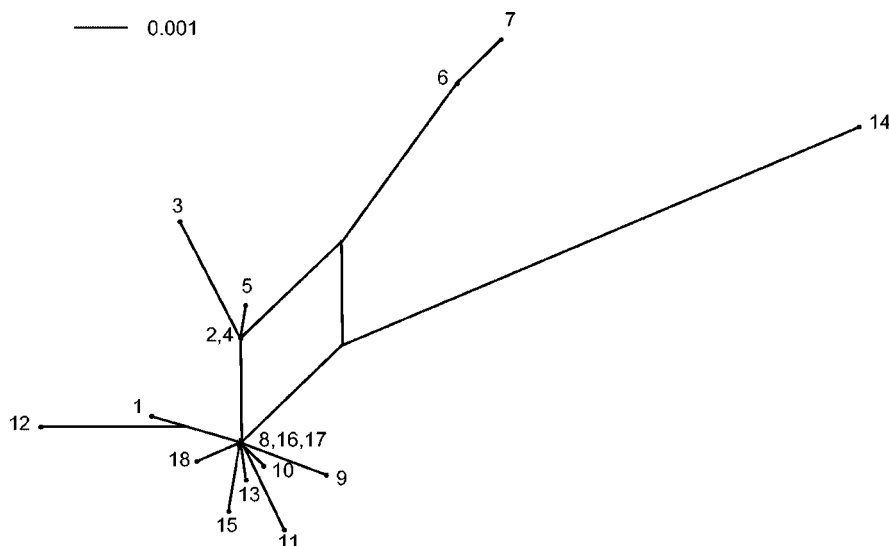


FIG. 4. Split decomposition analysis based on the allelic profiles of the 18 *O. oeni* strains examined in this study. The numbering refers to strain numbers.

TABLE 4. Sawyer's test analysis for evidence of intragenic recombination^a

Locus	SSCF ^b (<i>P</i>)	MCF ^c (<i>P</i>)
<i>gyrB</i>	275 (0.61)	7 (1.00)
<i>ddl</i>	0 (1.00)	0 (1.00)
<i>pgm</i>	1 (1.00)	1 (1.00)
<i>recP</i>	5,081 (0.04)	18 (1.00)
<i>mleA</i>	0 (1.00)	0 (1.00)

^a The Sawyer's test analyses were carried out by using the START program. Results were obtained from 10,000 random resamplings.

^b SSCF, sum of the square of condensed fragments.

^c MCF, maximum condensed fragment.

Comparative results of ribotyping and RFLP-ISR versus MLST. Only one RFLP-ISR type was identified among the 18 *O. oeni* strains analyzed in our study; similarly, the same strains were included in only two ribotypes. However, the number of sequence types (18) was larger (Table 1). This observation suggests that the discriminatory ability of MLST is better than that observed by the use of ribotyping and RFLP analysis of the ISR. A specific ribotype was not restricted to specific STs, and, for example, 11 STs contained isolates with the same ribotype B. These results show clearly that ribotype does not correlate with sequence type (Table 1).

Evidence for recombination in *O. oeni*. Bacteria can differ widely in their population structures. While some of them have a clonal population structure, in which all sequence diversity has arisen from the sequential accumulation of point mutations, in many other sequences diversity is greatly increased by intraspecies recombination. The sets of sequences were tested with the homoplasmy test (33). The homoplasmy test requires a sufficient number of informative sites to yield interpretable results, and if a locus has fewer than 10 informative sites, the analysis will not be performed on it. In all of our data sets, there was insufficient sequence diversity to perform the homoplasmy test.

There were two possible examples of a recombinational event in the *gyrB* and *recP* genes from unknown sources. The mean divergence between allele 7 of *recP* (6.66%) is much higher than the mean diversity within the other *recP* alleles (0.37%), with the exception of allele 3 (3.61%). Alleles 7 and 3 of *recP* seem to have originated from a different source. In *gyrB*, allele 3 showed a divergence of 6.42%, and alleles 2, 4, and 5 showed a mean diversity of 4.28%. Therefore, in *gyrB* alleles 1, 6, 7, and 8 (mean diversity, 1.28%) could have originated from a different source from alleles 3, 2, 4, and 5.

Sawyer's test revealed no detectable cases of intragenic recombination in the sample except in the case of the *recP* locus, in which there was evidence against the null hypothesis of no recombination ($P = 0.04$) (Table 4). This evidence disappeared when the maximum condensed fragment value ($P = 1.00$) was considered.

The complete set of isolates was analyzed for linkage disequilibrium. The level of linkage disequilibrium between alleles was assessed by using the I_A value (31). In our study, an I_A of 0.155 was obtained; this value was not significantly greater than the I_A value of 0 expected for a population of linkage equilibrium, implying no departure from the null hypothesis of linkage equilibrium between alleles. To avoid dependence on the number of loci, Hudson described another statistic, the

standardized I_A (I_A^S) (22), which does not depend on the number of loci analyzed and is expected to be zero when alleles are in linkage equilibrium (free recombination). That means probably that the distribution of alleles occurs independently from each other. The I_A^S value was 0.038 for the sample analyzed in this study. This low I_A^S value is indicative of extensive recombination. The I_A^S value calculated for *O. oeni* supports our estimation that the genes investigated are close to linkage equilibrium.

Comparison between gene trees. In order to further analyze the population structure of *O. oeni*, we used split decomposition analysis. The algorithm used in this software is able to display conflicting results in the phylogenetic descent of sequences. A tree-like structure is created when the descent is clonal, but an interconnecting network or a bush-like structure will appear when recombination plays a role in the evolutionary history of *O. oeni* genes. Figure 5 shows the split graphs for all alleles of the five fragments analyzed. The structure of the split graphs varied substantially between the different loci. The SplitsTree graphs obtained with *ddl*, *pgm*, and *recP* loci present network-like structures. This indicates the presence of homoplasies, probably evolved by intragenic recombination. The split graphs of the other two loci showed no evidence of network-like evolution. The split graph of the *gyrB* gene displays a star- or bush-like structure consisting of a single origin in the center of the graph, from which single branches radiate. However, in *gyrB* an additional uncentered edge was observed, suggesting that the evolution of some of the *gyrB* genes has been initiated by a couple of parallel mutations originating from one ancestor. The split graph of the *mleA* gene displays a line because only two alleles are analyzed. The fit parameter was 100 for *gyrB*, *ddl*, *recP*, and *mleA*, indicating that all phylogenetic information in the sequences could be visualized in the graphs. The fit parameter for *pgm* was lower (40.4), indicating that not all information could be integrated into the graph. The differences in structure among the split graphs obtained for the five loci can be explained by recombination, because recombination can lead to the assembly of genes with different evolutionary histories within one strain.

DISCUSSION

One of the main objectives of this study was to investigate global genomic similarity at the intraspecies level and to develop a sensitive method for strain identification in *O. oeni*. Therefore, we performed multilocus sequence analysis with five chromosomal genes. We determined the degree of allelic variation in these genes of *O. oeni* by using a sample of 18 isolates originating in various countries. MLST schemes are based on sequences of multiple (usually seven) loci because the analysis of a single gene provides too little discrimination to be used for molecular typing. As a first step for developing a typing method, we analyzed the sequence diversity of four housekeeping genes and the *mleA* gene in order to know if they are sufficient to have enough typing discrimination, since the number of loci can be increased to improve resolution, but there will come a point when, for epidemiological purposes, little additional information is obtained for the cost and effort involved (43).

The internal fragments of the five loci that were selected

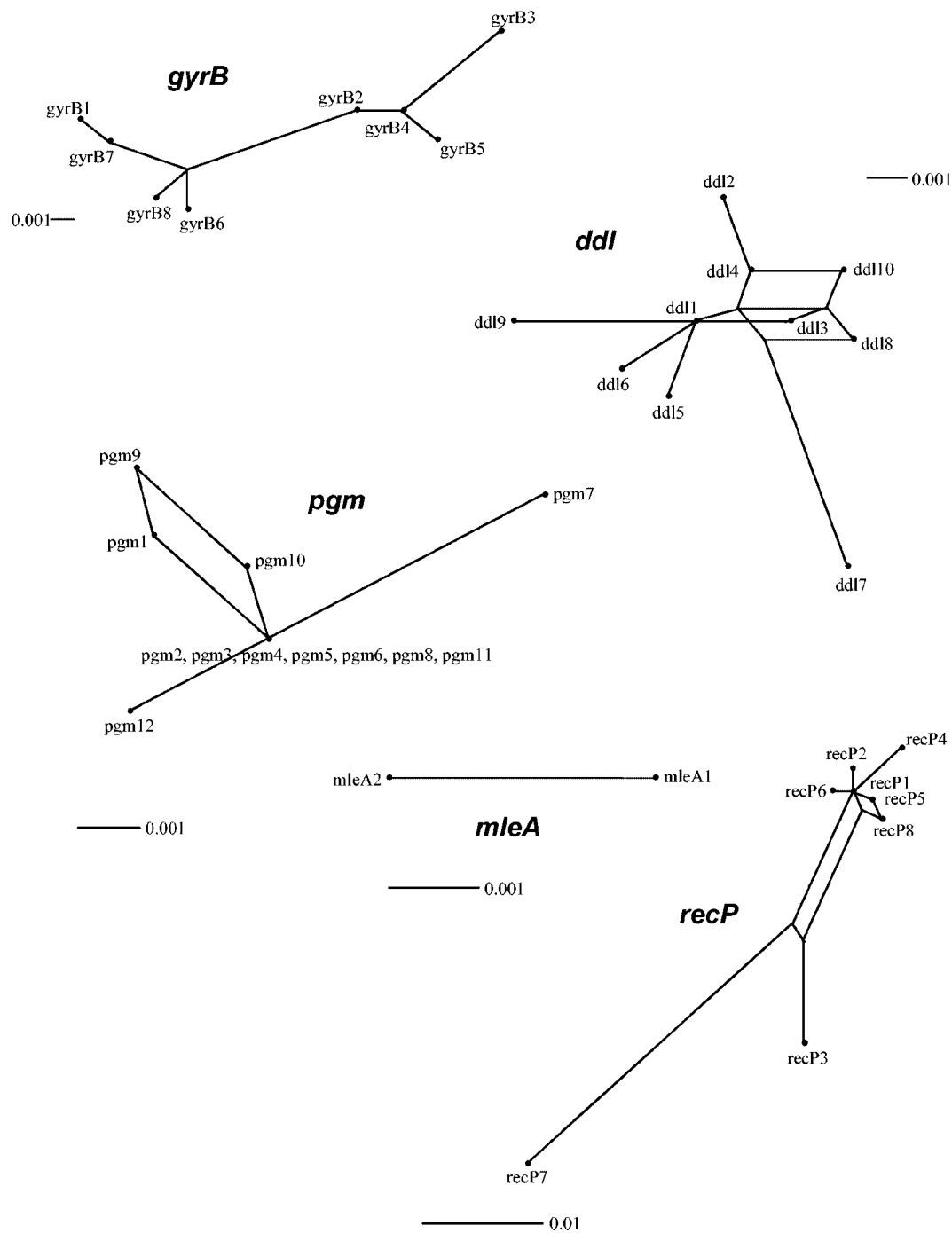


FIG. 5. Split decomposition analysis of alleles obtained from 18 *O. oeni* strains for five loci. All branch lengths are drawn to scale. The observation that in the *ddl*, *pgm*, and *recP* graphs several alleles in the sample are connected to each other by multiple pathways, forming an interconnected network, is suggestive of recombination. The numbering refers to allele numbers.

could be amplified from all strains that we examined. The amplified internal fragments were sequenced, and from these sequences we were able to use fragments between 339 and 554 bp for analyses (Table 3). Only two alleles were found in the *mleA* locus (Tables 1 and 2 and Fig. 2). The lack of diversity among the *mleA* locus sequences results in less discriminatory power than found in the housekeeping loci. It was previously

shown that the partial sequence of the *O. oeni* ATCC 23279^T (CECT 4100) *mleA* gene was also exactly identical to the *mleA* gene sequence of *O. oeni* IOEB 8413. The authors postulated that the malolactic enzyme (MleA) appears to have an identical sequence at species level. They found that the percentage of changes in sequences between *O. oeni* and other LAB was lower for *mleA* than for 16S rRNA (19).

The number of alleles per housekeeping locus ranged from 8 to 12 (Table 3). The four loci were highly polymorphic, and all multilocus types were represented by a single strain. The percentage of variable sites (1.7 to 6.6%) in the housekeeping genes was comparable to that seen by Enright et al. (12) in their analysis of a related gram-positive bacteria, *Streptococcus pyogenes*. The percentage of variable sites was higher than that seen in group B streptococcus (1.2 to 2.5%) (25) and *V. cholerae* (1.1 to 3.5%) (14) and considerably less than that of *C. jejuni* (9.2 to 21.7%) (8) or the *B. cereus* group (9 to 26%) (23).

The percentage of variable sites observed in *O. oeni* indicates a considerably high degree of genetic diversity, and it is not consistent with some data obtained by other experimental methods. The analysis of ribotyping and RAPD patterns has revealed a high level of genetic homogeneity in *O. oeni* (47). This homogeneity is consistent with the results of DNA-DNA hybridization (7), total soluble cell protein analysis (6), and 16S-23S ISR sequences (46). The RAPD profiles discerned two main groups of strains coincident with clusters obtained by macrorestriction typing in previous work (41). Ribotyping of the 18 *O. oeni* strains analyzed in this study indicated that all strains are included in only two ribotypes, A and B, previously described by Zavaleta et al. (47). The study by Zavaleta et al. (47) included 37 culture collection strains, 17 of which fell in group A (which includes CECT 4100^T), and 20 fell in group B (represented by CECT 4028 and CECT 4029). The investigators concluded that the two well-defined and consistent groups of strains shown by all these methods are indicative of two distinct main patterns of DNA arrangement in the genome of the species *O. oeni*. However, by using MLST we could not observe the two well-defined groups of strains previously described.

Although more extensive studies are necessary to assess the population structure of *O. oeni*, the UPGMA tree based on the allelic profiles (Fig. 3), the split decomposition analysis of concatenated gene sequences (Fig. 4), and the I_A values were consistent with the conclusion that *O. oeni* represents a good example of panmictic structure. Panmictic populations (as found in *Neisseria gonorrhoeae*) may be so variable that identical strains are only found among isolates from direct contacts. Even panmictic populations can contain clonal groupings or geographical specialization (34). This was not the case in our study, but the number of isolates analyzed here is too low to allow a definitive conclusion. In previous studies based on RAPD, ribotyping, small-plasmid content, and sequencing of RAPD markers, Zavaleta et al. (47) hypothesized that *O. oeni* follows a closely clonal mode of evolution, similar to the case of some pathogenic bacterial clones such as those of *V. cholerae*. However, the analysis of the population structure of *O. oeni* presented here shows a substantial extent of recombination. In *Porphyromonas gingivalis*, Frandsen et al. (16) and Koehler et al. (27) found I_A^S values of 0.068 and 0.089, respectively, and suggested a nonclonal population structure characterized by recombination. In our study we calculated a similar I_A^S value (0.038) that confirms the importance of recombination in *O. oeni*. Moreover, the I_A^S value calculated is similar to that reported for *N. gonorrhoeae* and, therefore, supports our estimation that the genes investigated in *O. oeni* are close to linkage equilibrium.

The examination of the sequences of housekeeping genes

can provide evidence for the significance of recombination, since the variation within these genes is likely to be selectively neutral. The homoplasy test (33) measures the importance of recombination between members of a population. It is only valid when sequences differ by <5% of the nucleotides, and the test requires a sufficient number of alleles and informative sites to yield interpretable results. In our case, all loci analyzed had insufficient numbers of alleles and informative sites to perform this analysis.

Sawyer's test did not detect clear evidence of recombination in any of the loci analyzed (Table 4). In our analysis this test failed to detect recombination, in a similar way as described previously by Farfán et al. (14) for *V. cholerae*. These authors suggest that the reason for this disagreement may reside in the definition of condensed fragments themselves, which are designed to detect gene conversion in which both source and target sequences are in the sample. From our data *gyrB*, *pgm*, and *recP* showed a net structure in the split graph, suggesting the existence of recombination.

Falush et al. (13) suggested that the panmictic structure in *H. pylori* may result from frequent recombination during mixed colonization by unrelated strains during chronic colonization of the gastric mucosa. Recombination between *O. oeni* strains might occur on the fermentation tanks, where several strains of *O. oeni* may exist simultaneously in a single fermentation and where a favorable environment for horizontal gene transfers could be created. Although it has been shown that *O. oeni* is able to receive foreign DNA by transformation in vitro (5) and by conjugation (49), there is no evidence for horizontal gene transfer in vivo. The existence of recombination is not unexpected, because the presence of plasmids (1, 46), bacteriophages, and insertion sequences (48) was previously shown in *O. oeni*. Zé-Zé et al. (48) reported that small differences observed in two *O. oeni* strains are apparently due to insertion or deletion events related to the presence of insertion sequences.

Recombination can be detected in the aligned sequences by a number of means, the most simple method being the detection either by eye or through statistical analysis of mosaic structure, where different regions of the gene appear to have different evolutionary histories. For example, most of a gene may be identical in sequence for two isolates of a species, whereas a 500-bp region in the middle may differ at 5% of nucleotide sites. Significant mosaic structure is indicative of recombinational exchange, usually among isolates of the same species but occasionally also between closely related species (15). A point mutation will generate a single nucleotide difference, whereas a recombinational exchange is likely to introduce multiple nucleotide differences. Enright and Spratt (10) found in *S. pneumoniae* a possible example of an interspecies recombinational event in *gdh* as the mean divergence between an allele and the other pneumococcal alleles was 4.58%, which was much higher than the mean diversity within the other pneumococcal alleles (0.95%) and similar to the mean divergence between these alleles and that of *Streptococcus mitis* (5.16%). In our study, two possible examples of recombinational events from unknown sources were found. The mean divergence between allele 7 of *recP* (6.66%) is much higher than the mean diversity within the other *recP* alleles (0.37%), with the exception of allele 3 (3.61%) (Fig. 2). This divergence is higher than the divergence observed by Enright and Spratt

(10) in a possible example of an interspecies recombinational event in the *gdh* gene between *S. pneumoniae* and *S. mitis* (5.16%). In *O. oeni*, the *recP* gene of strain 5001 represents a possible example of a recombinational event, but in this case, the possible source remains unknown. As far as we know, this is the first example of a possible recombinational event described in *O. oeni*. The extension of the analysis to a higher number of strains might offer information about the possible source of this allele.

In summary, our data suggest strongly that the four house-keeping loci chosen were a suitable basis for an MLST scheme, as they could be amplified and sequenced from isolates obtained from a wide variety of sources, were unlinked on the *O. oeni* chromosome, and exhibited sufficient diversity to provide a high degree of resolution. This study also demonstrates that MLST discriminates among *O. oeni* isolates effectively and generates data that can be applied to the investigation of the population structure and evolutionary mechanism in this important food organism.

ACKNOWLEDGMENTS

This work was supported by grant 07G/0035/2003 from the Comunidad de Madrid and RM03-002 from the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA). B. de las Rivas and A. Marcobal were the respective recipients of a postdoctoral fellowship and a predoctoral fellowship from the Comunidad de Madrid.

We thank R. González, A. V. Carrascosa, and E. García for their critical reading of the manuscript. The help of A. V. Carrascosa in providing the commercial starter strains was greatly appreciated.

REFERENCES

1. Brito, L., and H. Pavela. 1999. Presence and analysis of large plasmids in *Oenococcus oeni*. *Plasmid* **41**:260–267.
2. Britz, T. J., and R. P. Tracey. 1990. The combination effect of pH, SO₂, ethanol and temperature on the growth of *Leuconostoc oenos*. *J. Appl. Bacteriol.* **68**:23–31.
3. Caspritz, G., and F. Radler. 1983. Malolactic enzyme of *Lactobacillus plantarum*. *J. Biol. Chem.* **258**:4907–4910.
4. Delaquis, P., M. Cliff, M. King, B. Girard, J. Hall, and A. Reynolds. 2000. Effect of two commercial malolactic cultures on the chemical and sensory properties of Chancellor wines vinified with different yeasts and fermentation temperatures. *Am. J. Enol. Vitic.* **51**:42–48.
5. Dicks, L. T. M. 1994. Transformation of *Leuconostoc oenos* by electroporation. *Biotechnol. Techniques* **8**:901–904.
6. Dicks, L. T. M., H. J. J. Van Vuuren, and F. Dellaglio. 1990. Taxonomy of *Leuconostoc* species, particularly *Leuconostoc oenos*, as revealed by numerical analysis of total soluble cell protein pattern, DNA base composition and DNA-DNA hybridization. *Int. J. Syst. Bacteriol.* **40**:83–91.
7. Dicks, L. T. M., F. Dellaglio, and M. D. Collins. 1995. Proposal to reclassify *Leuconostoc oenos* as *Oenococcus oeni* [corrig.] gen. nov., comb. nov. *Int. J. Syst. Bacteriol.* **45**:395–397.
8. Dingle, K. E., F. M. Colles, D. R. A. Wareing, R. Ure, A. J. Fox, F. E. Bolton, H. J. Bootsma, R. J. L. Willems, R. Urwin, and M. C. J. Maiden. 2001. Multilocus sequence typing system for *Campylobacter jejuni*. *J. Clin. Microbiol.* **39**:14–23.
9. Divol, B., T. Tonon, S. Morichon, E. Gindreau, and A. Lonvaud-Funel. 2003. Molecular characterization of *Oenococcus oeni* genes encoding proteins involved in arginine transport. *J. Appl. Microbiol.* **94**:738–746.
10. Enright, M. C., and B. G. Spratt. 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144**:3049–3060.
11. Enright, M. C., and B. G. Spratt. 1999. Multilocus sequence typing. *Trends Microbiol.* **7**:482–487.
12. Enright, M. C., B. G. Spratt, A. Kalia, J. H. Cross, and D. E. Bessen. 2001. Multilocus sequence typing of *Streptococcus pyogenes* and the relationships between *emm* type and clone. *Infect. Immun.* **69**:2416–2427.
13. Falush, D., C. Kraft, N. S. Taylor, P. Correa, J. G. Fox, M. Achtman, and S. Suerbaum. 2001. Recombination and mutation during long-term gastric colonization by *Helicobacter pylori*: estimates of clock rates, recombination size, and minimal age. *Proc. Natl. Acad. Sci. USA* **98**:15056–15061.
14. Farfán, M., D. Miñana-Galbis, M. C. Fusté, and J. G. Lorén. 2002. Allelic diversity and population structure in *Vibrio cholerae* O139 Bengal based on nucleotide sequence analysis. *J. Bacteriol.* **184**:1304–1313.
15. Feil, E. J., M. C. Enright, and B. G. Spratt. 2000. Estimating the relative contributions of mutation and recombination to clonal diversification: a comparison between *Neisseria meningitidis* and *Streptococcus pneumoniae*. *Res. Microbiol.* **151**:465–469.
16. Frandsen, E. V., K. Poulsen, M. A. Curtis, and M. Kilian. 2001. Evidence of recombination in *Porphyromonas gingivalis* and random distribution of putative virulence markers. *Infect. Immun.* **69**:4479–4485.
17. Garvie, E. I. 1969. Lactic acid dehydrogenases of strains of the genus *Leuconostoc*. *J. Gen. Microbiol.* **58**:85–94.
18. Garvie, E. I. 1986. Gram-positive cocci: genus *Leuconostoc*, p. 1071–1075. In P. H. A. Sneath, N. S. Mair, M. E. Sharpe, and J. G. Holt (ed.), *Bergey's manual of systematic bacteriology*, vol. 2. Williams & Wilkins, Baltimore, Md.
19. Groisillier, A., and A. Lonvaud-Funel. 1999. Comparison of partial malolactic enzyme gene sequences for phylogenetic analysis of some lactic acid bacteria species and relationships with the malic enzyme. *Int. J. Syst. Bacteriol.* **49**:1417–1428.
20. Guerrini, S., A. Bastianini, G. Blaiotta, L. Granchi, G. Moschetti, S. Coppola, P. Romano, and M. Vicenzini. 2003. Phenotypic and genotypic characterization of *Oenococcus oeni* strains isolated from Italian wines. *Int. J. Food Microbiol.* **83**:1–14.
21. Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98NT. *Nucleic Acids Symp. Ser.* **41**:95–98.
22. Haubold, B., and R. R. Hudson. 2000. LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics* **16**:847–848.
23. Helgason, E., N. J. Tourasse, R. Meisal, D. A. Caugant, and A. B. Kolsto. 2004. Multilocus sequence typing scheme for bacteria of the *Bacillus cereus* group. *Appl. Environ. Microbiol.* **70**:191–201.
24. Huson, D. H. 1998. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* **14**:68–73.
25. Jones, N., J. F. Bohnsack, S. Takahashi, K. A. Oliver, M. S. Chan, F. Kunst, P. Glaser, C. Rusniok, D. W. M. Crook, R. M. Harding, N. Bisharat, and B. G. Spratt. 2003. Multilocus sequence typing system for group B *Streptococcus*. *J. Clin. Microbiol.* **41**:2530–2536.
26. Kelly, W. J., C. M. Huang, and R. V. Asmundson. 1993. Comparison of *Leuconostoc oenos* strains by pulsed-field gel electrophoresis. *Appl. Environ. Microbiol.* **59**:3969–3972.
27. Koehler, A., H. Karch, T. Beikler, T. F. Flemmig, S. Suerbaum, and H. Schmidt. 2003. Multilocus sequence analysis of *Porphyromonas gingivalis* indicates frequent recombination. *Microbiology* **149**:2407–2415.
28. Kumar, S., K. Tamura, and M. Nei. 1994. MEGA: molecular evolutionary genetics analysis software for microcomputers. *Comput. Appl. Biosci.* **10**:189–191.
29. Lamoureux, M., H. Prévost, J. F. Cavin, and C. Diviès. 1993. Recognition of *Leuconostoc oenos* strains by the use of DNA restriction profiles. *Appl. Microbiol. Biotechnol.* **39**:547–552.
30. Le Jeune, C., and A. Lonvaud-Funel. 1997. Sequence of DNA 16S/23S spacer region of *Leuconostoc oenos* (*Oenococcus oeni*): application to strain differentiation. *Res. Microbiol.* **148**:79–86.
31. Maiden, M. C., J. J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA* **95**:3140–3145.
32. Marchesi, J. R., T. Sato, A. J. Weightman, T. A. Martin, J. C. Fry, S. J. Hiom, and W. G. Wade. 1998. Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Appl. Environ. Microbiol.* **64**:795–799.
33. Maynard Smith, J., and N. H. Smith. 1998. Detecting recombination from gene trees. *Mol. Biol. Evol.* **15**:590–599.
34. Maynard Smith, J., N. H. Smith, M. O'Rourke, and B. G. Spratt. 1993. How clonal are bacteria? *Proc. Natl. Acad. Sci. USA* **90**:4384–4388.
35. Meats, E., E. J. Feil, S. Stringer, A. J. Cody, R. Goldstein, J. S. Kroll, T. Popovic, and B. G. Spratt. 2003. Characterization of encapsulated and non-encapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. *J. Clin. Microbiol.* **41**:1623–1636.
36. Moreno-Arribas, M. V., M. C. Polo, F. Jorganes, and R. Muñoz. 2003. Screening of biogenic amine production by lactic acid bacteria isolated from grape must and wine. *Int. J. Food Microbiol.* **84**:117–123.
37. Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**:418–426.
38. Peynaud, E., and S. Domercq. 1968. Étude de quatre cents souches de coques hétérolactiques isolés de vins. *Ann. Inst. Pasteur* **19**:159–170.
39. Sato, H., F. Yanagida, T. Shinohara, M. Suzuki, K. Suzuki, and K. Yokotsuka. 2001. Intraspecific diversity of *Oenococcus oeni* isolated during red wine-making in Japan. *FEMS Microbiol. Lett.* **202**:109–114.

40. **Sawyer, S.** 1989. Statistical tests for detecting gene conversion. *Mol. Biol. Evol.* **6**:526–538.
41. **Tenreiro, R., M. A. Santos, H. Pavela, and G. Vieira.** 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.
42. **Tracey, R. P., and T. J. Britz.** 1989. Cellular fatty acid composition of *Leuconostoc oenos*. *J. Appl. Bacteriol.* **66**:445–456.
43. **Urwin, R., and M. C. J. Maiden.** 2003. Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol.* **11**:479–487.
44. **Viti, C., L. Giovannetti, L. Granchi, and S. Ventura.** 1996. Species attribution and strain typing of *Oenococcus oeni* (formerly *Leuconostoc oenos*) with restriction endonuclease fingerprints. *Res. Microbiol.* **147**:651–660.
45. **Zapparoli, G., C. Reguant, A. Bordons, S. Torriani, and F. Dellaglio.** 2000. Genomic DNA fingerprinting of *Oenococcus oeni* strains by pulsed-field gel electrophoresis and randomly amplified polymorphic DNA-PCR. *Curr. Microbiol.* **40**:351–355.
46. **Zavaleta, A. I., A. J. Martínez-Murcia, and F. Rodríguez-Varela.** 1996. 16S-23S rDNA intergenic sequences indicate that *Leuconostoc oenos* is phylogenetically homogeneous. *Microbiology* **142**:2105–2114.
47. **Zavaleta, A. I., A. J. Martínez-Murcia, and F. Rodríguez-Varela.** 1997. Intraspecific genetic diversity of *Oenococcus oeni* as derived from DNA fingerprinting and sequence analyses. *Appl. Environ. Microbiol.* **63**:1261–1267.
48. **Zé-Zé, L., R. Tenreiro, and H. Pavela.** 2000. The *Oenococcus oeni* genome: physical and genetic mapping of a strain GM and comparison with the genome of a “divergent” strain, PSU-1. *Microbiology* **146**:3195–3204.
49. **Zúñiga, M., I. Pardo, and S. Ferrer.** 1996. Transposons Tn916 and Tn925 can transfer from *Enterococcus faecalis* to *Leuconostoc oenos*. *FEMS Microbiol. Lett.* **135**:179–185.