Sequencing the Gene Encoding Manganese-Dependent Superoxide Dismutase for Rapid Species Identification of Enterococci

CLAIRE POYART, GILLES QUESNES, AND PATRICK TRIEU-CUOT*

Laboratoire Mixte Pasteur-Necker de Recherche sur les Streptocoques et Streptococcies and Unité INSERM 411, Faculté de Médecine Necker-Enfants Malades, 75730 Paris Cedex 15, France

Simple PCR and sequencing assays that utilize a single pair of degenerate primers were used to characterize a 438-bp-long DNA fragment internal ($sodA_{int}$) to the $sodA$ gene encoding the manganese-dependent superoxide dismutase in 19 enterococcal type strains (*Enterococcus avium, Enterococcus casseliflavus, Enterococcus cecorum, Enterococcus columbae, Enterococcus dispar, Enterococcus durans, Enterococcus faecalis, Enterococcus faecium, Enterococcus flavescens, Enterococcus gallinarum, Enterococcus hirae, Enterococcus malodoratus, Enterococcus mundtii, Enterococcus pseudoavium, Enterococcus raffinosus, Enterococcus saccharolyticus, Enterococcus seriolicida, Enterococcus solitarius,* and *Enterococcus sulfureus*). Sequence analysis of the $sodA_{int}$ fragments enabled reliable identification of 18 enterococcal species, including *E. casseliflavus-E. flavescens* and *E. gallinarum*. The $sodA_{int}$ fragments of *E. casseliflavus* and *E. flavescens* were almost identical (99.5% sequence identity), which suggests that they should be associated in a single species. Our results confirm that the $sodA$ gene constitutes a more discriminative target sequence than 16S rRNA gene in differentiating closely related bacterial species.

Enterococci, although not highly virulent microorganisms, have emerged worldwide in the last decade as one of the leading causes of nosocomial bacteremia, surgical wound infections, and urinary tract infections (9, 10, 13, 24). This evolution is mainly due to the appearance of multiresistant strains of enterococci that are resistant to most antibiotics used in treatment (e.g., ampicillin, aminoglycosides, and glycopeptides). Most human enterococcal infections ($\geq 90\%$) are caused by *Enterococcus faecalis* and *Enterococcus faecium*; however, the incidence of other species, such as *Enterococcus casseliflavus* and *Enterococcus gallinarum*, could be underestimated because of bacterial misidentification. In clinical laboratories, accurate identification of enterococcal species is required to carry out a proper epidemiologic surveillance and may help in the management of infected patients in case of relapse. This is usually done by testing tolerance to bile esculine and tellurite, growth in 6.5% NaCl broth, and specific carbohydrate utilization (2, 6); by characterizing bacterial motility and pigment production (1); and by using commercial biochemical test systems, such as the API 20STREP or Rapid ID 32 Strep systems. However, these phenotypic methods are often not reliable and the automated systems, such as the Vitek and MicroScan systems, do not properly identify enterococci other than *E. faecalis* and *E. faecium* in the absence of additional tests (11). Consequently, several genotypic methods based on the analysis of PCR products derived from selected target DNA have been developed for the species identification of enterococci (3, 14, 22). This includes the determination of the 16S ribosomal DNA (rDNA) sequence (18), a strategy which is now greatly facilitated by the use of universal 16S PCR primers associated with the development of simplified, par-

tially automated, and cost-effective sequencing technologies. However, the interpretation of these data may be complicated by the fact that divergent 16S rDNA sequences may exist within a single organism (23) or, alternatively, that closely related species may have identical 16S rDNA sequences (8), as recently shown in the genera *Enterococcus* for *E. casseliflavus* and *E. gallinarum* (18). To solve this problem, it is possible to use alternative monocopy target sequences which exhibit a higher divergence than that of the 16S rDNA. The $sodA$ gene of the gram-positive cocci which encodes the manganese-de-

TABLE 1. Enterococcal type strains used in this study

Strain[a]	Other designation[b]	$sodA_{int}$ accession no.
E. avium CIP 103019 T	ATCC 14025	AJ387906
E. casseliflavus CIP 103018 T	ATCC 25788	AJ387907
E. cecorum CIP 103676 T	ATCC 43198	AJ387908
E. columbae CIP 103675 T	ATCC 51263	AJ387909
E. dispar CIP 103664 T	ATCC 51266	AJ387910
E. durans CIP 55.125 T	ATCC 19432	AJ387911
E. faecalis CIP 103015 T	ATCC 19433	AJ387912
E. faecium CIP 103014 T	ATCC 19434	AJ387913
E. flavescens CIP 103525 T	ATCC 49996	AJ387914
E. gallinarum CIP 103013 T	ATCC 49573	AJ387915
E. hirae CIP 53.48 T	ATCC 8043	AJ387916
E. malodoratus CIP 103012 T	ATCC 43197	AJ387917
E. mundtii CIP 103010 T	ATCC 43186	AJ387918
E. pseudoavium CIP 103647 T	ATCC 49372	AJ387919
E. saccharolyticus CIP 103246 T	ATCC 43076	AJ387920
E. solitarius CIP 103330 T	NCTC 12193	AJ387921
E. raffinosus CIP 103329 T	ATCC 49427	AJ387922
E. seriolicida CIP 104369 T	ATCC 49156	AJ387923
E. sulfureus CIP 104373 T	DSM 6905	AJ387924
L. garvieae CIP 102507 T	DSM 20684	AJ387925

[a] CIP, Collection de l'Institut Pasteur.
[b] ATCC, American Type Culture Collection; DSM, Deutsche Sammlung von Mikroorganismen und Zelkulturen; NCTC, National Collection of Type Cultures.

* Corresponding author. Mailing address: Laboratoire Mixte Pasteur-Necker de Recherche sur les Streptocoques et Streptococcies, Faculté de Médecine Necker-Enfants Malades, 75730 Paris Cedex 15, France. Phone: (33) (1) 40 61 56 79. Fax: (33) (1) 40 61 55 92. E-mail: ptrieu@pasteur.fr.

TABLE 2. Identification of various enterococcal strains by sequencing the $sodA_{int}$ fragment

Strain	Relevant characteristic(s)[a]	Bacterial species[b]	Accession no.
NEM1616	E. faecalis; vanA	E. faecalis (99.5)	AJ387927
NEM1617	E. faecalis; vanA	E. faecalis (98.6)	AJ387928
NEM1618	Enterococcus sp.	E. durans (99.3)	AJ387929
NEM1619	E. hirae	E. durans (99.5)	AJ387930
NEM1620	Enterococcus sp.	E. durans (99.1)	AJ387931
NEM1621	E. hirae	E. hirae (99.8)	AJ387932
NEM1622	Enterococcus sp.	E. hirae (99.5)	AJ387933
NEM1623	E. casseliflavus	E. casseliflavus (99.1)	AJ387934
NEM1624	E. faecium; vanB	E. faecium (99.5)	AJ387935
NEM1625	E. faecium; vanA	E. faecium (100)	AJ387936
NEM1626	E. faecium; vanB	E. faecium (99.8)	AJ387937
NEM1627	E. faecium; multiply resistant strain	E. faecium (99.8)	AJ387938
NEM1628	E. faecium; multiply resistant strain	E. faecium (99.8)	AJ387939
NEM1629	Enterococcus sp.	E. gallinarum (98.6)	AJ387940
NEM1630	E. avium	E. avium (100)	AJ387941

[a] Bacterial strains were all clinical isolates from our collection which were identified by using conventional microbiological tests and the ID 32 Strep System (API-bioMérieux). Presence of vanA (NEM1616, NEM1617, and NEM1625) and vanB (NEM1624 and NEM1626) was determined by PCR with specific primers (3).

[b] The species identification was based on the phylogenic position of the $sodA_{int}$ fragment of the strain studied relative to those of the type strains, as shown in Fig. 1. Numbers in parentheses indicate the percentage of identity of the $sodA_{int}$ fragment with that of the corresponding type strains.

pendent superoxide dismutase fulfills these criteria and we recently reported that sequencing of the sodA PCR product with the use of a single pair of degenerate primers constitutes a valuable approach to the genotypic identification of the 29 streptococcal species (20). In this work, the same universal primers (19) were used to construct a sodA database of 19 enterococcal species, including E. casseliflavus and E. gallinarum. We show the usefulness of this library for a rapid sequence-based identification method of enterococcal isolates.

The main characteristics of the bacterial strains used in this study, including the type strains, are listed in Table 1 and 2. Rapid extraction of bacterial genomic DNA was carried out by using the InstaGene Matrix (Bio-Rad, Hercules, Calif.) on cells collected from 2 ml of an overnight culture. The sodA degenerate primers d1 (5′-CCITAYICITAYGAYGCIYTIGA RCC-3′) and d2 (5′-ARRTARTAIGCRTGYTCCCAIACR TC-3′) were used to amplify an internal fragment, designated $sodA_{int}$, representing approximately 85% of their sodA genes. PCRs were performed on a Gene Amp System 9600 instrument (Perkin-Elmer Cetus, Roissy, France) in a final volume of 50 μl containing 250 ng of DNA as template, 0.5 μM each primer, 200 μM each deoxynucleoside triphosphate, and 1 U of AmpliTaq Gold DNA polymerase (Perkin-Elmer) in a 1× amplification buffer (10 mM Tris-HCl [pH 8.3], 50 mM KCl, 1.5 mM $MgCl_2$). The PCR mixtures were denatured (3 min at 95°C) and then subjected to 30 cycles of amplification (60 s of annealing at 37°C, 60 s of elongation at 72°C, and 30 s of denaturation at 95°C) and 7 min at 72°C for the last elongation cycle. A single DNA fragment corresponding to the expected 480-bp amplification product, $sodA_{int}$, was observed in all cases following agarose gel electrophoresis and ethidium bromide staining (data not shown). PCR products were purified on an S-400 Sephadex column (Pharmacia, Uppsala, Sweden) and directly sequenced on both strands with the oligonucleotides d1 and d2 by using the ABI-PRISM Big Dye terminator sequencing kit on a Genetic ABI-PRISM 310 Sequencer Analyzer (Perkin-Elmer). The cycle sequencing protocol was optimized as follows: the sequencing mixtures were subjected to 40 cycles of amplification consisting of 10 s of denaturation at 96°C, 5 s of annealing at 40°C, and 4 min of elongation at 60°C.

The nucleotide sequences of the $sodA_{int}$ fragments from the type strains of Enterococcus avium, E. casseliflavus, Enterococcus cecorum, Enterococcus columbae, Enterococcus dispar, Enterococcus durans, E. faecalis, E. faecium, Enterococcus flavescens, E. gallinarum, Enterococcus hirae, Enterococcus malodoratus, Enterococcus mundtii, Enterococcus pseudoavium, Enterococcus raffinosus, Enterococcus saccharolyticus, Enterococcus seriolicida, Enterococcus solitarius, Enterococcus sulfureus, and Lactococcus garvieae were determined (Table 1). We assumed that the PCR products sequenced were actual $sodA_{int}$

TABLE 3. Identity matrix based on pairwise comparisons of $sodA_{int}$ fragments of enterococcal type strains

Strain no.	Species[a]	% Identity with:																	
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	E. avium	74.0	67.4	71.9	70.3	70.3	73.7	69.9	74.2	75.1	75.6	71.0	80.1	72.6	74.0	85.4	87.9	60.7	67.1
2	E. casseliflavus		66.4	70.8	72.6	72.4	77.9	72.4	99.5	83.1	78.5	77.4	71.5	73.5	74.0	76.7	76.7	66.4	75.6
3	E. cecorum			78.8	72.4	66.0	68.7	66.2	66.9	67.4	70.3	64.2	65.5	71.7	65.8	67.6	66.0	62.8	68.7
4	E. colombae				69.4	68.9	71.7	69.2	70.8	72.6	73.1	68.7	69.6	72.1	72.8	73.3	71.9	67.1	69.2
5	E. dispar					70.3	77.4	68.7	72.8	72.1	73.5	72.8	70.5	71.0	69.4	72.1	70.3	62.6	74.9
6	E. durans						73.1	81.3	72.4	76.3	84.9	80.1	69.6	72.8	70.5	71.5	73.3	62.1	73.7
7	E. faecalis							72.6	78.3	77.6	77.9	72.4	71.2	78.8	73.5	77.9	75.1	67.1	76.5
8	E. faecium								72.4	77.2	83.1	81.7	67.4	72.6	69.4	71.9	71.9	62.1	72.1
9	E. flavescens									83.1	78.3	77.2	72.1	72.8	74.4	77.2	76.9	65.8	74.9
10	E. gallinarum										80.8	78.5	73.3	76.0	73.5	77.4	75.1	66.2	76.7
11	E. hirae											83.6	71.5	73.3	73.5	77.4	76.7	63.5	75.3
12	E. mundtii												70.8	69.4	69.6	71.9	73.5	63.9	74.4
13	E. pseudoavium													70.5	69.2	81.7	80.4	62.8	65.3
14	E. saccharolyticus														72.4	75.1	71.0	62.1	76.7
15	E. solitarius															74.2	72.8	64.8	72.1
16	E. malodoratus																87.9	63.5	70.5
17	E. raffinosus																	64.6	67.6
18	E. seriolicida																		61.6
19	E. sulfureus																		

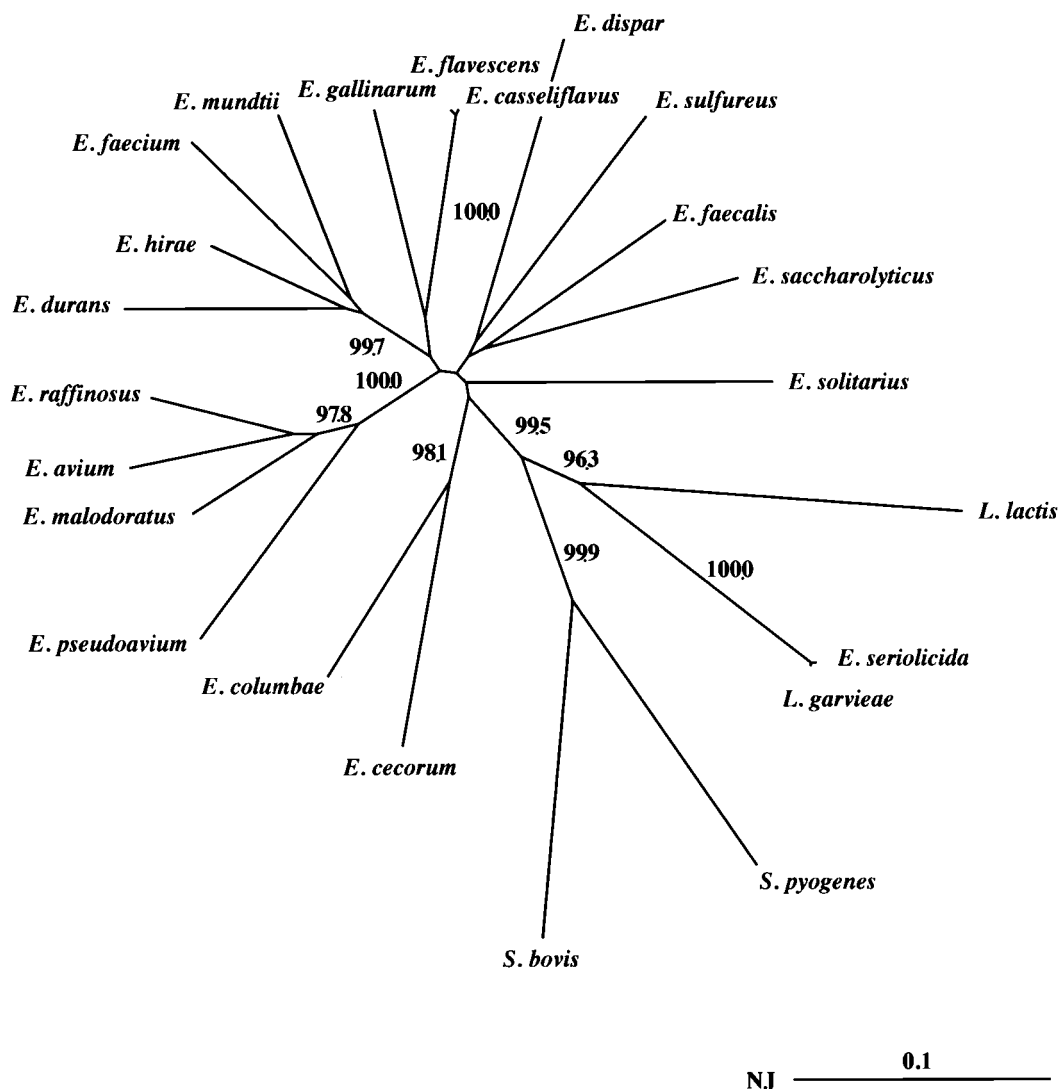[a] The main characteristics of the strains are listed in Table 1.

FIG. 1. Phylogenetic unrooted tree showing relationships among the $sodA_{int}$ fragments from various enterococcal type strains. The tree was established from an analysis of the sequences listed in Table 1 by using the NJ method. The $sodA_{int}$ sequences of *L. lactis*, *L. garvieae*, *S. bovis*, and *S. pyogenes* type strains were included in this work. The value on each branch is the estimated confidence limit (expressed as a percentage) for the position of the branch as determined by bootstrap analysis. Only bootstrap values greater than 95%, which were considered significant, are indicated. The scale bar (NJ distance) represents a 10% difference in nucleotide sequences.

fragments, since the corresponding deduced polypeptides all contained the amino acids characteristic of the manganese-dependent superoxide dismutase (16, 17) at the expected positions (data not shown). Multiple alignment of these $sodA_{int}$ DNA sequences plus those of *L. garvieae* (Table 1), *Lactococcus lactis* (19), *Streptococcus bovis* (20), and *Streptococcus pyogenes* (20) was carried out with the Clustal X program (12), and an unrooted phylogenetic tree was constructed by the neighbor-joining (NJ) method (21). The sequences of the degenerate oligonucleotides *d1* and *d2* and alignment gaps were not taken into consideration for calculations. The reliability of the tree nodes was evaluated by calculating the percentage of 1,000 bootstrap resamplings that support each topological element. Only the nodes having a bootstrap value greater than 95% are shown in Fig. 1, since this critical value could be used to define the monophyly of a clade of related organisms (7). This analysis revealed that, as expected, the members of the genus *Enterococcus*, with the exception of *E. seriolicida*, were clus-

tered within a clade supported by 99.5% of the bootstrap replicates. The $sodA_{int}$ sequences of *E. seriolicida* and of *L. garvieae* were almost identical (99.5% sequence identity) and were clustered with that of *L. lactis* within a clade supported by 96.3% of the bootstrap confidence limit (Table 3 and Fig. 1). These results are consistent with the redesignation of *E. seriolicida* as *L. garvieae* (4). The phylogenetic tree representing the enterococcal $sodA_{int}$ sequences (Fig. 1) has the same topology as the NJ tree constructed from the analysis of their 16S rDNA sequences (18). It is worth noting that the $sodA_{int}$ sequences of the *E. casseliflavus* and *E. gallinarum* type strains displayed 16.9% sequence divergence, a value similar to the 19.7% sequence divergence observed between the *ddl* genes encoding the D-Ala–D-Ala ligases in these species (5). These results do not support the suggestion that *E. casseliflavus* and *E. gallinarum* comprise a single species (18). By contrast, the fact that the 16S rDNA sequence (18) and the *ddl* (15), *vanC* (3), and $sodA_{int}$ (Table 3) genes of the *E. casseliflavus* and *E.*

flavescens type strains were almost identical (99.9, 99.5, 96, and 98% sequence identity, respectively) suggests that they should be associated in a single species.

The phylogenetic tree shown in Fig. 1 revealed the presence of two major clusters within the enterococcal species which we have designated the *faecium* group (*E. faecium*, *E. durans*, *E. hirae*, and *E. mundtii*) and the *avium* group (*E. avium*, *E. malodoratus*, *E. pseudoavium*, and *E. raffinosus*). Within each group, the 16S rDNA sequences exhibited more than 99% sequence identity (18), whereas the highest percentage of similarity found between two $sodA_{int}$ sequences was 87.9% (Table 3). These results confirm that the gene *sodA* constitutes a more discriminative target sequence than 16S RNA in differentiating closely related bacterial species.

Fifteen enterococcal isolates were identified by using conventional microbiological tests, Rapid ID 32 Strep, and the $sodA_{int}$ systems (Table 2). In all cases, the $sodA_{int}$ sequences of the isolates displayed less than 1.5% divergence from the corresponding type strain. For 10 strains (NEM1616, NEM1617, NEM1621, NEM1623, NEM1624, NEM1625, NEM1626, NEM1627, NEM1628, and NEM1630), the two methods gave the same results. Four isolates (NEM1618, NEM1620, NEM1622, and NEM1629) were identified at the species level with the $sodA_{int}$ system but not with conventional microbiological tests or the Rapid ID 32 Strep system. The remaining isolate, NEM1619, was identified with the Rapid ID 32 Strep system as *E. hirae* but was identified with the $sodA_{int}$ system as *E. durans* (Table 2). The reliability of the molecular identification of NEM1164 was based on the fact that its $sodA_{int}$ fragment displays 99.5 and 85% sequence identity with the type strains of *E. durans* and *E. hirae*, respectively.

In conclusion, we have determined the $sodA_{int}$ sequences of the type strains of *E. avium*, *E. casseliflavus*, *E. flavescens*, *E. cecorum*, *E. columbae*, *E. dispar*, *E. durans*, *E. faecalis*, *E. faecium*, *E. gallinarum*, *E. hirae*, *E. malodoratus*, *E. mundtii*, *E. pseudoavium*, *E. raffinosus*, *E. saccharolyticus*, *E. seriolicida*, *E. solitarius*, and *E. sulfureus* and demonstrated the usefulness of this database for the species identification of enterococcal isolates. The identification method presented in this study is not accessible to routine clinical microbiology laboratories, but it may become the "gold standard" technique in reference and large research hospital laboratories for epidemiologic purposes and/or identifying problematic strains.

Nucleotide sequence accession number. Representative nucleotide sequences have been submitted to the EMBL database and have been given accession no. AJ387906 to AJ387925 and AJ387927 to AJ387941.

REFERENCES

1. **Cartwright, C. P., F. Stock, G. A. Fahle, and V. J. Gill.** 1995. Comparison of pigment production and motility tests with PCR for reliable identification of intrinsically vancomycin-resistant enterococci. J. Clin. Microbiol. **33:**1931–1933.
2. **Devriese, L. A., B. Pot, K. Kersters, S. Lauwers, and F. Haesebrouck.** 1996. Acidification of methyl-alpha-D-glucopyranoside: a useful test to differentiate *Enterococcus casseliflavus* and *Enterococcus gallinarum* from *Enterococcus faecium* species group and from *Enterococcus faecalis*. J. Clin. Microbiol. **34:**2607–2608.
3. **Dutka-Malen, S., S. Evers, and P. Courvalin.** 1995. Detection of glycopeptide resistance genotypes and identification to the species level of clinically relevant enterococci by PCR. J. Clin. Microbiol. **33:**24–27. (Erratum, **33:**1434.)
4. **Eldar, A., C. Ghittino, L. Asanta, E. Bozzetta, M. Goria, M. Prearo, and H. Bercovier.** 1996. *Enterococcus seriolicida* is a junior synonym of *Lactococcus garvieae*, a causative agent of septicemia and meningoencephalitis in fish. Curr. Microbiol. **32:**85–88.
5. **Evers, S., B. Casadewall, M. Charles, S. Dutka-Malen, M. Galimand, and P. Courvalin.** 1996. Evolution of structure and substrate specificity in D-alanine: D-alanine ligases and related enzymes. J. Mol. Evol. **42:**706–712.
6. **Facklam, R. R., and M. D. Collins.** 1989. Identification of *Enterococcus* species isolated from human infections by a conventional test scheme. J. Clin. Microbiol. **27:**731–734.
7. **Felsenstein, J.** 1985. Confidence limits on phylogeny and approach using the bootstrap. Evolution **39:**783–791.
8. **Fox, G. E., J. D. Wisotzkey, and P. Jurtshuk, Jr.** 1992. How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int. J. Syst. Bacteriol. **42:**166–170.
9. **Hunt, C. P.** 1998. The emergence of enterococci as a cause of nosocomial infection. Br. J. Biomed. Sci. **55:**149–156.
10. **Huycke, M. M., D. F. Sahm, and M. S. Gilmore.** 1998. Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future. Emerg. Infect. Dis. **4:**239–249.
11. **Iwen, P. C., D. M. Kelly, J. Linder, and S. H. Hinrichs.** 1996. Revised approach for identification and detection of ampicillin and vancomycin resistance in *Enterococcus* species by using MicroScan panels. J. Clin. Microbiol. **34:**1779–1783.
12. **Jeanmougin, F., J. D. Thompson, M. Gouy, D. G. Higgins, and T. J. Gibson.** 1998. Multiple sequence alignment with Clustal X. Trends Biochem. Sci. **23:**403–405.
13. **Moellering, R. C., Jr.** 1998. Vancomycin-resistant enterococci. Clin. Infect. Dis. **26:**1196–1199.
14. **Monstein, H. J., M. Quednau, A. Samuelsson, S. Ahrne, B. Isaksson, and J. Jonasson.** 1998. Division of the genus *Enterococcus* into species groups using PCR-based molecular typing methods. Microbiology **144:**1171–1179.
15. **Navarro, F., and P. Courvalin.** 1994. Analysis of genes encoding D-alanine–D-alanine ligase-related enzymes in *Enterococcus casseliflavus* and *Enterococcus flavescens*. Antimicrob. Agents Chemother. **38:**1788–1793.
16. **Parker, M. W., and C. C. F. Balke.** 1988. Crystal structure of manganese superoxide dismutase from *Bacillus stearothermophilus* at 2.4 A resolution. J. Mol. Biol. **199:**649–661.
17. **Parker, M. W., and C. C. F. Blake.** 1988. Iron- and manganese-containing superoxide dismutases can be distinguished by analysis of their primary structures. FEBS Lett. **229:**377–382.
18. **Patel, R., K. E. Piper, M. S. Rouse, J. M. Steckelberg, J. R. Uhl, P. Kohner, M. K. Hopkins, F. R. Cockerill III, and B. C. Kline.** 1998. Determination of 16S rRNA sequences of enterococci and application to species identification of nonmotile *Enterococcus gallinarum* isolates. J. Clin. Microbiol. **36:**3399–3407.
19. **Poyart, C., P. Berche, and P. Trieu-Cuot.** 1995. Characterization of superoxide dismutase genes from Gram-positive bacteria by polymerase chain reaction using degenerate primers. FEMS Microbiol. Lett. **131:**41–45.
20. **Poyart, C., G. Quesne, S. Coulon, P. Berche, and P. Trieu-Cuot.** 1998. Identification of streptococci to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase. J. Clin. Microbiol. **36:**41–47.
21. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
22. **Tyrrell, G. J., R. N. Bethune, B. Willey, and D. E. Low.** 1997. Species identification of enterococci via intergenic ribosomal PCR. J. Clin. Microbiol. **35:**1054–1060. (Erratum, **35:**2434.)
23. **Ueda, K., T. Seki, T. Kudo, T. Yoshida, and M. Kataoka.** 1999. Two distinct mechanisms cause heterogeneity of 16S rRNA. J. Bacteriol. **181:**78–82.
24. **Woodford, N.** 1998. Glycopeptide-resistant enterococci: a decade of experience. J. Med. Microbiol. **47:**849–862.