A Novel Putative Enterococcal Pathogenicity Island Linked to the *esp* Virulence Gene of *Enterococcus faecium* and Associated with Epidemicity

Helen Leavis,[1,2] Janetta Top,[1] Nathan Shankar,[3] Katrine Borgen,[1] Marc Bonten,[2] Jan van Embden,[4] and Rob J. L. Willems[1,5]

[1]Diagnostic Laboratory for Infectious Diseases and Perinatal Screening, and
[4]Laboratory for Vaccine Preventable Diseases, National Institute for Public Health and the Environment, Bilthoven,
[2]Department of Internal Medicine, Division of Acute Internal Medicine and Infectious Diseases, and
[5]Eijkman-Winkler Institute for Microbiology, Infectious Diseases and Inflammation, University Medical Center Utrecht, Utrecht, The Netherlands,
[3]Department of Pharmaceutical Sciences, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma

Journal of Bacteriology, 2004, 186:672-682

Abstract

Enterococcus faecalis harbors a virulence-associated surface protein encoded by the *esp* gene. This gene has been shown to be part of a 150-kb putative pathogenicity island. A gene similar to *esp* has recently been found in *Enterococcus faecium* isolates recovered from hospitalized patients. In the present study we analyzed the polymorphism in the *esp* gene of *E. faecium*, and we investigated the association of *esp* with neighboring chromosomal genes. The *esp* gene showed considerable sequence heterogeneity in the regions encoding the nonrepeat N- and C-terminal domains of the Esp protein as well as differences in the number of repeats. DNA sequencing of chromosomal regions flanking the *esp* gene of *E. faecium* revealed seven open reading frames, representing putative genes implicated in virulence, regulation of transcription, and antibiotic resistance. These flanking regions were invariably associated with the presence or absence of the *esp* gene in *E. faecium*, indicating that *esp* in *E. faecium* is part of a distinct genetic element. Because of the presence of virulence genes in this gene cluster, the lower G+C content relative to that of the genome, and the presence of *esp* in *E. faecium* isolates associated with nosocomial outbreaks and clinically documented infections, we conclude that this genetic element constitutes a putative pathogenicity island, the first one described in *E. faecium*. Except for the presence of *esp* and *araC*, this pathogenicity island is completely different from the *esp*-containing pathogenicity island previously disclosed in *E. faecalis.*

Introduction

Enterococci are common inhabitants of the gastrointestinal tracts of humans and animals, and although they have been recognized as pathogens able to cause endocarditis, they were generally considered second-rate pathogens. Recent estimates, however, indicate that enterococci are now among the leading causes of nosocomial infections (57). Of all enterococcal species, *Enterococcus faecalis* accounted for the most infections in humans (26). However, during the past decade, the incidence of bloodstream infections caused by *Enterococcus faecium* increased, an increase which has been linked to the emergence of antibiotic resistance in this species (26, 40).

Little is known about virulence determinants in *E. faecium* (20). Recently, however, three potential virulence genes, *esp*, *hyl*, and *acm*, have been described for *E. faecium*. They were all found more frequently in clinical isolates than in fecal isolates or nonhuman isolates (13, 41, 44, 65).

Of these three putative virulence genes, only the *esp* gene is also found in *E. faecalis* (51). The Esp protein in *E. faecalis* is expressed as a large surface-exposed

protein with a molecular mass of approximately 202 kDa. In *E. faecalis*, Esp is thought to be an adhesin contributing to colonization of urinary tract epithelial cells and biofilm formation (50, 59). Although detailed experimental evidence is not yet available, the higher prevalence of the *E. faecium esp* gene in clinical isolates suggests a role of Esp in the pathogenesis of *E. faecium* infections (3, 7, 12, 13, 30, 65, 68). Furthermore, the presence of the *esp* gene in *E. faecium* was also strongly associated with hospital outbreaks of vancomycin-resistant *E. faecium*, suggesting a role for Esp in nosocomial transmission (65).

Recently, the *esp* gene of *E. faecium* strain P61 was cloned and sequenced (13). Analysis of the sequence revealed that the enterococcal Esp (13, 51) belongs to a family of gram-positive surface-exposed proteins with repetitive structures such as the alpha C (38) and Rib (55) proteins of *Streptococcus agalactiae*, the R28 protein of *Streptococcus pyogenes* (54), and the Bap protein of *Staphylococcus aureus* (8), all of which are involved in virulence and in conferring protective immunity. Sequence similarity between these surface proteins is found predominantly in the repeat regions.

In *E. faecalis*, the *esp* gene is contained on a large (150-kb) genetic element (49). This element has all the characteristics of a pathogenicity island (PAI), with a GC content of 32.2%, which is significantly different from that of the rest of the *E. faecalis* chromosome, and the presence of genes encoding transposases, transcriptional regulators, and virulence determinants.

In this study we demonstrate considerable sequence heterogeneity among the *E. faecium esp* genes of various isolates. We also show that *E. faecium esp* is contained on a putative PAI and that the presence of this putative PAI is associated with nosocomial outbreaks of *E. faecium*.

(Part of this study was presented as a poster at the 42nd Interscience Conference on Antimicrobial Agents and Chemotherapy, San Diego, Calif., 27 to 30 September 2002 [abstr. B-803].)

Materials and methods

Bacterial strains

E. faecium isolate E300 from hospital outbreak US-1 (11, 65) was used to clone and sequence the *esp* gene and the DNA region encompassing the putative PAI. *E. faecium* isolate E734 from hospital outbreak NL-1-1 (64, 65) and strain E470 from hospital outbreak NL-3-1 (58, 65) were used to determine sequence heterogeneity in the N- and C-terminal domains of the *esp* gene. Sequencing of the frameshift mutation at positions 12830 to 12832 and the stop codon at position 13719, originally found in strain E300, was performed for isolates E155 from outbreak US-2-6 (5, 65) and E734 from outbreak NL-1-1 (64, 65). Bacteria were grown on

Table 1. Oligonucleotides used in this study.

Name	Sequence	Strand	Start position	Target or function
esp$_{fs}$1R	5'-ACT ATC AAC CTC TCC TGT TTT AG	-	5616[a]	sequencing of the *esp* gene
esp$_{fs}$2R	5'-GAA GAG ACT TCT TCC TCT TTT C	-	5750[a]	sequencing of the *esp* gene
esp$_{fs}$3R	5'-TTC GGC GCT TTT TTA TC	-	5273[a]	sequencing of the *esp* gene; amplification C-repeat region
esp$_{fs}$4F	5'-GGA ACG CCT TGG TAT G	+	1000[a]	sequencing of the *esp* gene
esp$_{fs}$4R	5'-GAA TAT GTC ACT ACA ACC GTA C	-	3254[a]	sequencing of the *esp* gene
esp$_{fs}$5F	5'-GGA AAC CTG AAT TAG AAG AAG	+	1490[a]	sequencing of the *esp* gene
esp$_{fs}$5R	5'-TAC TGC TAA ATC GGT CGT G	-	2295[a]	sequencing of the *esp* gene
esp$_{fs}$6F	5'-ACG TGG ATG TAG AGT TTG C	+	1973[a]	sequencing of the *esp* gene
esp$_{fs}$6R	5'-CCG CTT TTG GTG ATT C	-	1798[a]	sequencing of the *esp* gene
esp$_{fs}$7F	5'-CGA CCG ATT TAG CAG TAA C	+	2279[a]	sequencing of the *esp* gene; amplification A-repeat region
esp$_{fs}$7R	5'-CCG CTG CTT TCA TTT C	-	1309[a]	sequencing of the *esp* gene
esp$_{fs}$8F	5'-GGT AGA GGT TGT TAT TTC TGT AGT AG	+	5233[a]	sequencing of the *esp* gene
esp$_{fs}$8R	5'-TAT AGA AAT CAT CTT GAT CTG TC	-	817[a]	sequencing of the *esp* gene
esp$_{fs}$9F	5'-AAA AGG TAC GGT TGT AGT GAC	+	3228[a]	sequencing of the *esp* gene; PCR
esp$_{fs}$9R	5'-TTC TTC GTA TAT CCC GG	-	305[a]	sequencing of the *esp* gene
esp$_{fs}$10F	5'-GAA ACA ACT GAT ACA CAA ACT G	+	610[a]	sequencing of the *esp* gene
esp$_{fs}$10R	5'-CAG TTT GTG TAT CAG TTG TTT C	-	631[a]	sequencing of the *esp* gene
esp$_{fm}$1F	5'-ATA ACT ATT AAG GGA GTT GAT TTG	+	3063[b]	sequencing of the *esp* gene
esp$_{fm}$2F	5'-TGG TTA GCA AGA ATA ATA AGA GAG	+	3092[b]	sequencing of the *esp* gene
esp$_{fm}$3R	5'-GGA CTT GCA TTA GCA AAA TC	-	4277[b]	sequencing of the *esp* gene
esp$_{fm}$4F	5'-AAG TAG AGG TTA CTA TTT CTG TAG AAG	+	8288[b]	sequencing of the *esp* gene
esp$_{fm}$5R	5'-CAG CTG CGC TAA CAT CTA C	-	6769[b]	amplification A-repeat region
esp$_{fm}$5F	5'-AAA GAA GAT TTA CCA AAA GAT ACT AAG	+	6706[b]	amplification C-repeat region
esp$_{fm}$4R	5'-AAT ACT CTC TTA TTA TTC TTG CTA ACC	-	3119[b]	inverse PCR, cloning and sequencing of PAI
nox 1F	5'-GTA ATT ATT GGA TCA AAC CAT TC	+	8939[b]	inverse PCR, cloning and sequencing of PAI
nox 2R	5'-GAA TGG TTT GAT CCA ATA ATT AC	-	8961[b]	sequencing of PAI; PCR
PAI-1R	5'-ACG TTC ATG TAT GGG AAA G	-	13814[b]	sequencing of PAI
PAI-2F	5'-GTA TTA GCG GTG TTC AAA ATG	+	49[b]	sequencing of PAI; PCR
PAI-2R	5'-TTC CTC TGT CAA AAT AAG CTA AC	-	755[b]	sequencing of PAI
PAI-3R	5'-CGA TAG GTG ACA GAA CTC ATA AC	-	14575[b]	sequencing of PAI; PCR
PAI-4R	5'-GTT CCA AAA AGG CTG ATA ATC	-	14059[b]	sequencing of PAI
PAI-5F	5'-CCT TGT TCC AGT CCC C	+	14524[b]	sequencing of PAI
PAI-5R	5'-AAA ATC AAG CCG CCA AG	-	13891[b]	sequencing of PAI
PAI-6F	5'-GAA GAA GGA ATT TGA AGT CAC	+	2697[b]	sequencing of PAI
PAI-6R	5'-CTA ATG ATC GTG TAG CTA AGA AC	-	13385[b]	sequencing of PAI
PAI-7F	5'-CGG ATC ATA ATA ATT ATT GTC TTT G	+	587[b]	sequencing of PAI
PAI-7R	5'-GAT ATT TGT CAA TCA AAG GTT G	-	12937[b]	sequencing of PAI; PCR
PAI-8R	5'-TTT AGA AGT CGC TTT GCC	-	12492[b]	sequencing of PAI
PAI-9R	5'-ATC AAA GGT CTA AGA ATC CAA C	-	11891[b]	sequencing of PAI
PAI-10R	5'-CAT AGG TTT TAA TTA ATT CAT TTA GC	-	11396[b]	sequencing of PAI
PAI-11R	5'-CGC AGA CTC ACC AAT TTT C	-	10987[b]	sequencing of PAI
PAI-12R	5'-CAG TCG TCT CGG TTC TTT C	-	10511[b]	sequencing of PAI
PAI-13R	5'-CAA AGC TAA TTC TTA ATT TTA CAC G	-	10067[b]	sequencing of PAI
PAI-14R	5'-CTT ATT ATT CTT GCT AAC CAT TAT TC	-	3111[b]	sequencing of PAI
PAI-15R	5'-ATT GGA GTT ATC AAC ATT TTT TC	-	2613[b]	sequencing of PAI
PAI-16R	5'-GTC ATA TTC ATT TAA CAC ACT ATT ATT ACC	-	2186[b]	sequencing of PAI
PAI-17R	5'-CGA TTT CCT TAG TAT AAT AAA CAA TC	-	1665[b]	sequencing of PAI; PCR
PAI-18R	5'-TTT GCA ATG AAT TAT AGA GTC G	-	1170[b]	sequencing of PAI
PAI-19R	5'-AAT CTA TAC ACG AAT AAG AAT ATT ATC C	-	438[b]	sequencing of PAI
PAI-20R	5'-GAG AAA ACA TTG ATA ATA GTC CAG	-	10145[b]	sequencing of PAI
PAI-21R	5'-ATG TAT TCC ATT TTT TGA TAG TAT TTC	-	9547[b]	sequencing of PAI
PAI-2F-Biotin	5'-GTA TTA GCG GTG TTC AAA ATG	+	49[b]	Southern hybridisation; detection *orf1*
esp$_{fm}$1R-Biotin	5'-GTA ATT AGC ATA CCA AGG CG	-	4068[b]	Southern hybridisation; detection *esp*
PAI-10F-Biotin	5'-AAA ATA GTC ACT ACA AGT GGT ACC C	+	9599[b]	Southern hybridisation; detection *orf4*
PAI-11F-Biotin	5'-TTG CAT CAG CAG TTA TAT TAA TG	+	10401[b]	Southern hybridisation; detection *orf5*; PCR
PAI-8F-Biotin	5'-ACC GAA AAA TAA TAC AAG TGG	+	12760[b]	Southern hybridisation; detection *orf6*
PAI-4F-Biotin	5'-CTA GTA TGA CTA TGG CTA CAA ATG C	+	13753[b]	Southern hybridisation; detection *orf7*

[a] Nucleotide reference positions relative to the *E. faecalis esp* sequence (GenBank, EMBL database accession no. AF034779)

[b] Nucleotide reference positions relative to the *E. faecium* PAI sequence deposited in the GenBank/EMBL database under accession no. AY322150

blood agar plates at 37°C for further use.

The following isolates were used to determine the presence of the putative PAI in *E. faecium*: isolates from hospital outbreaks Australia-1, NL-1-1, NL-2-1, NL-2-3, NL-3-1, UK-1, US-1, US-2-1, US-2-2, US-2-3, US-2-4, US-2-5, US-2-6, and US-2-7 (4, 5, 11, 28, 35, 58, 64, 65); 68 clinical isolates (44 from blood, 9 from pus, 7 from urine, 5 from peritoneal fluid, 1 from bile, 1 from lungs, and 1 from skin) (4, 11, 16, 48, 64, 67) from the SENTRY Antimicrobial Surveillance Program, originating from hospitals in 15 different countries (Australia, Austria, Belgium, France, Germany, Israel, Italy, The Netherlands, Poland, Portugal, Spain, Switzerland, Turkey, the United Kingdom, and the United States); 6 hospital surveillance isolates (feces isolates with no link to a hospital outbreak) from three different countries (France, The Netherlands, and the United Kingdom) (28, 48, 58, 64, 66); 3 community surveillance isolates from The Netherlands (feces isolates with no hospital link) (16, 62, 67); and 10 animal feces isolates from The Netherlands (2 each from cats, dogs, calves, swine, and poultry) (61-63, 66).

PCR and sequencing of the *E. faecium esp* gene

The nonrepeat regions of the *E. faecium esp* gene were amplified and sequenced by using a combination of 17 primers based on the published *E. faecalis esp* sequence (GenBank/EMBL accession no. AF034779) (51) and 4 primers based on the *E. faecium* sequence determined in this study (Table 1). Chromosomal DNA was purified as described elsewhere (66, 67). PCR conditions for all amplification reactions were as follows: initial denaturation at 95°C for 15 min, followed by 35 cycles of 30 s at 94°C, 30 s at 52°C, and 30 s at 72°C, and a final 5-min extension at 72°C. Reactions were performed in 25-µl volumes with HotStar *Taq* polymerase and HotStar Master Mix buffers (Qiagen Inc., Valencia, Calif.). PCR products were purified with a PCR purification kit (Qiagen Inc.) and sequenced by using the BigDye Terminator reaction kit and an ABI PRISM 3700 DNA analyzer (both from Applied Biosystems, Foster City, Calif.).

For sequencing of the region encompassing the A and C repeats, a slightly different approach was followed. First the A- and C-repeat regions were amplified with the primer combinations esp_{fs}6F-esp_{fs}4R and esp_{fs}9F-esp_{fs}2R, respectively, and were subsequently cloned into pCR2.1-TOPO by using the TOPO TA cloning kit (Invitrogen Life Technologies, Carlsbad, Calif.) in accordance with the manufacturer's instructions. This resulted in pJT1 and pJT2, harboring the A- and C-repeat regions, respectively. To generate subclones suitable for sequencing, overlapping deletions were constructed with the Erase-a-base system (Promega Corporation, Madison, Wis.). Subclones were sequenced by using the M13 reverse primer, the BigDye Terminator reaction kit, and an ABI PRISM 3700 DNA analyzer (all from Applied Biosystems).

The 5' end of the *esp* gene was amplified by a combination of primer esp_{fs}10R and an 18-mer primer consisting of thymidines only. This PCR fragment was cloned into pCR2.1-TOPO by using the TOPO TA cloning kit (Invitrogen Life Technologies) in accordance with the manufacturer's instructions, and the resulting plasmid, designated pJT3, was sequenced using primers esp_{fs}6R, esp_{fs}10R, and esp_{fm}1F. Clone pJT2 was also used to determine the nucleotide sequence of the 3' end of the *esp* gene, since primer esp_{fs}2R is located just downstream of the *esp* gene.

Determination of variation in the *esp* A and C repeats

Two different primer combinations were used to assess repeat number variation by PCR. Primer sets esp_{fs}7F-esp_{fm}5R and esp_{fm}5F-esp_{fs}3R (Table 1) were used to amplify across the A- and C-repeat regions of the *esp* gene, respectively, in a set of 36 *E. faecium* isolates. Amplification conditions were identical to those described above. Subsequently, the amplicons were subjected to agarose gel electrophoresis (1%) in order to determine their sizes. From the sizes of the amplicons the numbers of repeats were deduced. Amplicon size differences corresponded to multiples of either 252 bp (A repeats) or 246 bp (C repeats).

Cloning and sequencing of the putative PAI

The DNA region adjacent to *esp* was cloned by an inverse-PCR strategy. Approximately 10 μg of chromosomal DNA was digested with *Eco*RI or *Bcl*I, and the resulting fragments were self-ligated. Ligated DNA was amplified with primer esp_{fm}4R, located in the 5' end of the *esp* gene, and primer *nox*1F, located just downstream of the *esp* gene, by using the Expand Long Template PCR system (Roche Diagnostics Nederland B.V., Almere, The Netherlands). Six-kilobase *Bcl*I and 7.9-kb *Eco*RI inverse-PCR products were cloned into pCR2.1-TOPO by using the TOPO TA cloning kit (Invitrogen Life Technologies) in accordance with the manufacturer's instructions, producing plasmids pJT4 (*Eco*RI digest) and pJT5 (*Bcl*I digest). Overlapping deletions of pJT4 and pJT5 were constructed with the Erase-a-base system (Promega) to generate subclones suitable for sequencing. One strand of the pJT4 and pJT5 subclones was sequenced with the M13 forward primer in combination with the BigDye Terminator reaction kit by using an ABI PRISM 3700 DNA analyzer (all from Applied Biosystems). Gaps in the DNA sequence of the first strand and sequence information of the second strand were obtained by direct sequencing of PCR products with primers based on the emerging nucleotide sequence of the first strand. Primers that were used for PCR and sequencing of this DNA region are listed in Table 1. PCR conditions were the same as those described above.

Detection of the putative PAI in *E. faecium* isolates

Southern hybridization was used to determine the presence of six open reading frames (ORFs) contained in the putative PAI in a set of 105 *E. faecium* isolates. For this purpose, chromosomal DNA preparations were digested with *Hae*III, separated by agarose gel electrophoresis (0.7% agarose gels), transferred onto a Hybond N$^+$ nylon membrane (Nycomed Amersham plc, Little Chalfont, Buckingham, United Kingdom), and subsequently hybridized to six biotin-labeled oligonucleotide probes specific for the six ORFs according to the protocol developed by Schouls and coworkers (47a). The oligonucleotides used as probes for hybridization are listed in Table 1.

Nucleotide sequence accession numbers

The DNA sequences reported in this article have been deposited in the GenBank/EMBL/DDBJ nucleotide sequence databases under accession no. AY322150 (*E. faecium* E300 putative PAI), AY322497 (*E. faecium* E155 hypothetical phage gene), AY322498 (*E. faecium* E734 permease gene), AY322499 (*E. faecium* E734 *esp* 5' end), AY322501 (*E. faecium* E734 *esp* 3' end), AY322500 (*E. faecium* E470 *esp* 5' end), and AY322502 (*E. faecium* E470 *esp* 3' end).

Results

Sequence analysis of the *E. faecium esp* gene

In an attempt to obtain the DNA sequence of the *esp* gene in *E. faecium*, a DNA region of strain E300, which was recovered during a hospital outbreak (11, 65), was amplified by PCR using primers based on the *E. faecalis esp* sequences (Table 1), followed either by direct sequencing of the PCR products or by making overlapping deletions of a cloned amplicon followed by sequencing of the deletion mutants. Sequence analysis revealed one ORF of 5,703 nucleotides that is predicted to encode a polypeptide of 1,900 amino acid residues with a calculated molecular mass of 205 kDa.The deduced amino acid sequence of the *E. faecium* Esp protein revealed a high degree of similarity to, but appeared not to be identical with, the recently described Esp of *E. faecium* strain P61 and the *E. faecalis* Esp protein (51). The *E. faecium* E300 Esp is predicted to be synthesized as a precursor with a 49-amino-acid signal peptide that precedes an N-terminal region of 706 amino acids, a central repeat region, and a C-terminal domain (Figure 1). The N-terminal domain has 99 and 91% amino acid sequence identities with *E. faecium* P61 Esp and *E. faecalis* Esp, respectively. Remarkably, the first 23 amino acid residues of the processed protein of E300 are highly different from those of *E. faecalis* Esp. The central repeat region of the variant E300 Esp protein contains five A repeats of 84 amino acids, followed by one B1 repeat (79 amino acids), five C

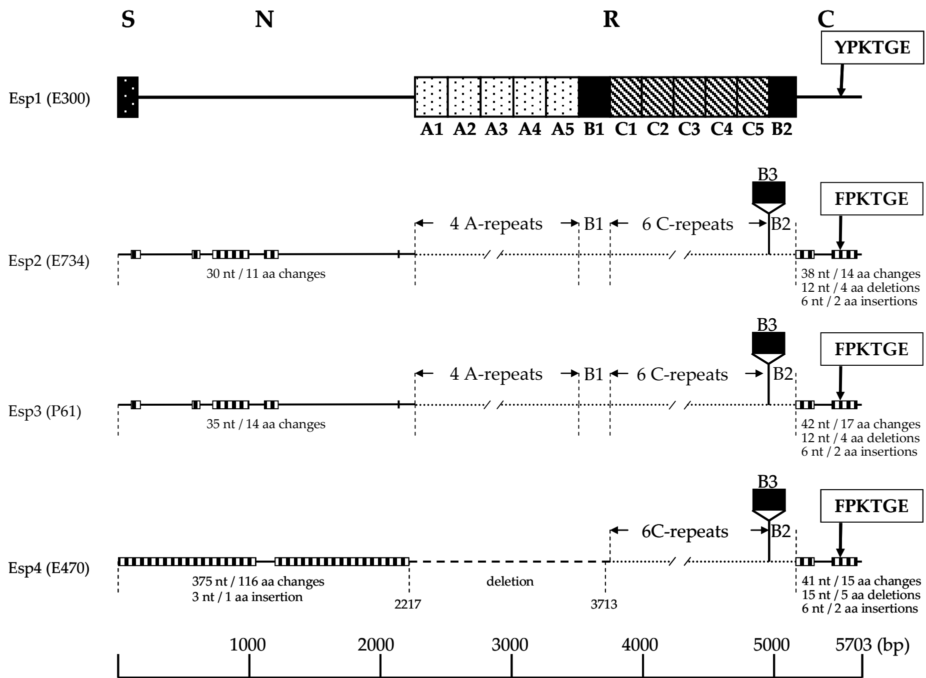

Figure 1. Schematic representation of the inferred *E. faecium* Esp protein and comparison of four *E. faecium* Esp variants. Esp1, deduced sequence of the Esp protein of strain E300, comprising the signal sequence (S) (solid box with white dots), N-terminal region (N), repeat region (R), and C-terminal region (C). The A-, B-, and C-repeat units are indicated (dotted, solid, and crosshatched boxes, respectively). YPKTGE and FPKTGE, anchor motifs in the C-terminal region. Solid lines in Esp2 (accession no. AF444000, AY322499, and AY322501), Esp3 (accession no. AJ487981), and Esp4 (accession no. AY322500 and AY322502) represent regions for which the DNA and amino acid sequences were compared to those of Esp1 (accession no. AF443999 and AY322150), while dotted lines represent regions that were not compared. Striped boxes in Esp2, -3, and -4 indicate locations of nucleotide (nt) and amino acid (aa) changes, with numbers of nucleotide and amino acid changes, insertions, and deletions relative to Esp1 indicated below. The dashed line in Esp4 represents the deletion in the *esp* gene of E470. The start and end points of this deletion, positions 2217 and 3713, respectively, relative to the E300 *esp* sequence, are indicated.

repeats (82 amino acids), and one B2 repeat (68 amino acids) (Figure 1). The beginning and end of the B and C repeats were chosen slightly differently from those published by Shankar et al. (51) and Eaton and Gasson (13), so that only complete instead of truncated copies of C repeats are present in the central part of the *esp* gene (Figure 2). The repeats in *E. faecium* E300 are highly similar to those of *E. faecium* P61 Esp and *E. faecalis* Esp, with amino acid identities of 98 to 99% for the A repeats, 96 to 98% for the B1 repeat, 97 to 98% for the C repeats, and 87 to

```
A-repeats
Efm Esp1-A1  IYENPGENIPAGYHKVTFTAGEGTSIESGTTVFAVKDGVSLPEDKLPVLKAKDGYTDAKWPEEATQPITADDTEFVSSATKLDD
Efm Esp1-A2  .I..................................................................................
Efm Esp1-A3  .I..................................................................................
Efm Esp1-A4  .I..............I...................................................................
Efm Esp1-A5  .I..............I...................................................................
Efm Esp3-A1  ....................................................................................
Efm Esp3-A2  .I.D..............................................T.................................
Efm Esp3-A3  .I....DS.............G................................F.................A...........
Efm Esp3-A4  .I............................................S.....................................
Efs Esp-A1   ....................................................................K...............
Efs Esp-A2   .I....D.............................................................K...............
Efs Esp-A3   .I....D......................................................G......K...............

B-repeats
Efm Esp1-B1  SDADKYNPEGQKVTTELNKEPDASEGIKNKEDLPKDTKYTWKEKVDVSAAGNKKGTVVVTYSDGSSDEVEVDVTVTDNR
Efm Esp1-B2  ...............D...................G.T.................I.....P...KE....TIS.E.K
Efm Esp3-B1  .......................F.....................................P.................
Efm Esp3-B2  ...............D...................G.T.......................P...KE....TIS.V.K
Efm Esp3-B3  .N.............D...................G.T.A............E...I....P...............T.
Efm Esp4-B3  .N.............D...................G.T.A.....................P...............T.
Efs Esp-B1   ......T.......................K.....A.........I.T..............................
Efs Esp-B2   ......T.K............E..D.....S....G.M.F.......GIP....A....I.P...KE.....IS.V.K

C-repeats
Efm Esp1-C1    SDADKYEPTVEGEKVEVGGTVDLTDNVTNLPTLPEGTTVTDVTPDGTIDTNTPGNYEGVIEVTYPDGTKDTVKVPVEVTDNR
Efm Esp1-C2    ......T.M.........................................................................
Efm Esp1-C3    ......T.M...................................G.....................................
Efm Esp1-C4    ............................................G.....................................
Efm Esp1-C5    ......T.M.........................................................................
Efm Esp3-C1    ............................................G.....................................
Efm Esp3-C2    ............................................G.....................................
Efm Esp3-C3    ......T.M...................................G.....................................
Efm Esp3-C4    ......T.M...................................G.....................................
Efm Esp3-C5    ...............................A............G.....................................
Efm Esp3-C6    ................I..............A............G.....................................
Efs Esp-C1     ................I..K..............Q.........G.....................................
Efs Esp-C2     ................I..K..............Q.........G.....................................
Efs Esp-C3     ................I..K..............Q.........G.....................................
Efs Esp-C4     ......T.M.......I..K..............Q.........G.....................................
Efs Esp-C5     ................I..K..............Q.........G.....................................
Efs Esp-C6     ................I..K..............Q...I.....G.....................................
Efs Esp-C7     ................I..K..............Q...I.....G.....................................
```

Figure 2. Comparison of the primary structures of the A, B, and C repeats of *E. faecium* Esp1 (E300), *E. faecium* Esp3 (P61) (13), and *E. faecalis* Esp (51) and the B3 repeat of *E. faecium* Esp4 (E470). Dots indicate identical amino acid residues. Only those amino acid residues of Esp3, Esp4, and *E. faecalis* Esp that differ from the repeats of *E. faecium* Esp1 are represented by letters. Efs, *E. faecalis*; Efm, *E. faecium*.

99% for the B2 repeat. *E. faecium* E300 Esp lacked the third B repeat (B3) that was reported for the P61 Esp (Figure 1 and 2).

The C-terminal domain of 167 amino acid residues contains a membrane-spanning hydrophobic region, the YPKTGE cell wall anchor motif, and a charged tail presumably extending into the cytoplasm, ending with a glutamic acid. This domain is also highly similar to those of *E. faecium* P61 Esp and *E. faecalis* Esp, with 87 and 84% amino acid identities, respectively. The overall similarities of *E. faecium* E300 Esp with the *E. faecium* P61 and *E. faecalis* Esp proteins, disregarding the number of repeats, are 96 and 92%, respectively.

Sequence heterogeneity in the *E. faecium esp* gene

In a previous study, sequence heterogeneity was identified in an internal fragment of the *E. faecium esp* gene (65). To determine sequence heterogeneity in *E. faecium esp* genes in more detail, the DNA regions encoding the N- and C-terminal domains of two additional *esp* genes from two outbreak-related vancomycin-resistant *E. faecium* isolates (E734 and E470) were amplified and sequenced, and the DNA sequences were compared to the corresponding E300 and P61 *esp* sequences. These comparisons revealed considerable polymorphism in the DNA regions encoding the N- and C-terminal domains, resulting in four different copies of the *E. faecium esp* gene, designated *esp1* to *esp4*, tentatively encoding four different Esp proteins, Esp1 to Esp4; *esp1* is the sequenced *esp* gene of strain E300, and *esp3* is the sequenced *esp* gene of *E. faecium* P61 (13) (Figure 1). The *esp2* gene was found in strain E734 from outbreak NL-1 and harbored 70 nucleotide differences from *esp1*, resulting in 26 amino acid changes. Furthermore, a 4-amino-acid deletion and a 2-amino-acid insertion, relative to the Esp1 protein, were found in the C-terminal domain of Esp2, as well as a third copy of the B3

Table 2. Variations in A-, B-, and C-repeats in *esp* analyzed in 36 *E. faecium* isolates

Source	No. of strains analyzed	Country[a]	No. of repeats			*esp* repeat profile
			A	C	B3	
Clinical isolate	1	PO	6	5	1	1
Clinical isolate	1	AU	6	5	1	1
Clinical isolate	1	GR	4	6	1	2
Clinical isolate	1	GR	5	6	1	3
Clinical isolate	1	GB	6	6	1	4
Clinical isolate	1	IT	4	6	1	2
Clinical isolate	1	FR	4	6	1	2
Clinical isolate	1	GB	5	6	1	3
Clinical isolate	1	AU	5	4	1	5
Hospital outbreak NL-1-1	10	NL	4	6	1	2
Hospital outbreak NL-2-1	2	NL	4	6	1	2
Hospital outbreak NL-3	1	NL	0	6	1	6
Hospital outbreak UK-1	4	UK	6	6	1	4
Hospital outbreak US-1	2	US	3	5	0	7
Hospital outbreak US-2-1	1	US	0	6	1	6
Hospital outbreak US-2-2	1	US	0	6	1	6
Hospital outbreak US-2-3	1	US	0	6	1	6
Hospital outbreak US-2-4	1	US	3	6	0	8
Hospital outbreak US-2-5	1	US	3	6	1	9
Hospital outbreak US-2-6	1	US	3	6	1	9
Hospital outbreak US-2-7	1	US	3	6	1	9
Hospital survey	1	GB	6	7	1	10

[a] PO, Poland; AU, Austria; GR, Greece; IT, Italy; FR, France; NL, The Netherlands; UK, United Kingdom; US, United States.

repeat. Also, the FPKTGE cell wall anchor motif in the C-terminal domain of Esp2 was different from that in Esp1 but identical to the anchor motif found in the P61 Esp3 protein (13). In general, the *esp*2 gene closely resembled the P61 *esp3* gene:the sequenced regions of *esp2* differed by only 11 nucleotides from *esp3*. The *esp4* gene, found in strain E470 from outbreak NL-3, contained 416 nucleotide differences in the regions encoding the N- and C-terminal domains relative to *esp1*, resulting in 131 amino acid changes, with most of the differences found in the region encoding the N-terminal domain. In addition to nucleotide changes, the *esp4* gene contained a large deletion in the region encoding the N-terminal domain, which also included the entire A-repeat region.

In addition to the observed nucleotide differences, the repeat regions of Esp appeared to be highly polymorphous, with variations in the numbers of A, B, and C repeats. This is not unexpected, since polymorphisms in these regions have been reported before in *E. faecalis* and *E. faecium* (13, 51). Thirty-six *E. faecium* isolates were analyzed for the numbers of A and C repeats. The number of A repeats varied from 0 to 6, while the number of C-repeats varied from 4 to 7, resulting in 10 different *esp* repeat profiles (Table 2). All strains originating from a single outbreak had identical repeat regions. Ten of the isolates, from outbreak NL-1-1, were collected during a 2-year period between April 2000 and April 2003, and they were all indistinguishable with respect to the number of repeats. In addition, the Esp repeat profile of these isolates was identical to that of the two isolates from outbreak NL-2-1, which previously had been shown to be epidemiologically linked to outbreak NL-1-1 (35). This Esp repeat profile of the Dutch outbreak strains was also found in epidemiologically unrelated clinical isolates from Greece, Italy, and France. This finding suggests that Esp repeat profiles are relatively stable, at least among strains associated with a single outbreak.

A cluster of genes adjacent to the *E. faecium esp* gene

Recently, it was reported that the *esp* gene of *E. faecalis* was part of a large (150-kb) PAI (49). To examine whether the *esp* gene in *E. faecium* was also present on a PAI, an inverse-PCR strategy was used on *Bcl*I- and *Eco*RI-digested chromosomal DNA to obtain sequence information for a 14-kb DNA fragment. Sequencing of this DNA fragment revealed seven ORFs including the *esp* gene (Figure 3). A search for homology using the GenBank/EMBL database showed that the predicted amino acid sequence of ORF1 (41 amino acids) had 42% similarity to the N-terminal part of the Uve2 protein encoded by the *vanE* gene cluster (Table 3). From this similarity it was also clear that only a part of this putative gene was cloned and sequenced. The *uve2* gene contained in the *vanE* gene cluster is 26% identical to the sigma factor SpoIIG of *Bacillus subtilis* (6). The *orf*2 gene is

Table 3. GC contents of the 7 ORFs contained in the putative *E. faecium* PAI and maximum predicted amino acid similarities

Locus	GC content (%)	Homologue	Maximum amino acid similarity of predicted protein to homologue (%)
orf1	32.5	*E. faecalis* Uve2	42
orf2	27.9	*E. faecalis* AraC	61
orf3	38.1	*E. faecalis* Esp	92
orf4	28.9	*E. faecalis* NADH Oxidase	35
orf5	35.6	*S. pyogenes* muramidase	27
orf6	32.8	*L. monocytogenes* bacteriophage protein	33
orf7	43.6	*L. lactis* MDR protein	45

predicted to encode a 401-amino-acid protein. This putative gene exhibited similarity with the *araC* gene found in the *E. faecalis* PAI (Table 3).

The *esp1* gene, the third ORF in this gene cluster, which is described in detail above, is present downstream of the *araC* gene (Figure 3). The *orf4* gene is predicted to encode a 447-amino-acid protein. This putative gene is located just downstream of *esp* and displays the highest similarity with the *nox* gene of *E. faecalis*, encoding an NADH oxidase. Although the overall similarities with homologous proteins in *Streptococcus pneumoniae*, *E. faecalis*, and *S. pyogenes* are relatively low (29.3, 34.8, and 33.8%, respectively), the three sequence motifs representing the flavin adenine dinucleotide (FAD) binding region, the NADH contact region, and a cysteine residue essential for redox activity are conserved in the putative *E. faecium* NADH oxidase present in this gene cluster (2, 19, 45) (Figure 3C).

The *orf5* product is predicted to be synthesized as a 483-amino-acid precursor with an amino-terminal signal sequence of 27 amino acid residues and shows similarity with peptidoglycan hydrolases, *N*-acetylmuramidases, and autolysins of *E. faecalis*, *Enterococcus hirae*, *Lactococcus lactis*, and *S. pyogenes* (Table 3; Figure 3). Alignment of the *E. faecium* putative muramidase polypeptide with the muramidase-2 gene product of *E. hirae* and the *E. faecalis* autolysin reveals that similarity is restricted to the N-terminal enzymatically active domain and that the *E. faecium* putative muramidase protein lacks the C-terminal peptidoglycan anchor domain (29). These findings make it unclear whether this gene encodes a functional muramidase or autolysin. In addition, the putative muramidase also contains the S144SKK, S178GN, D258/E282, and K354TG motifs found in serine ß-lactamases and penicillin-binding proteins (18). This could mean that this *E.*

faecium protein also displays penicillin binding properties comparable to those of the muramidase-2 protein of *E. hirae* (10, 29).

Downstream of the putative muramidase gene are two small ORFs displaying similarity with phage-associated hypothetical proteins of *Lactobacillus* spp., *Listeria monocytogenes*, *S. pyogenes*, and *Pseudomonas aeruginosa*. Detailed examination of the sequence at positions 12830 to 12832 suggests the presence of a frameshift in isolate E300. Sequencing of this region in the epidemic *E. faecium* isolate E155 (US-2-6) (65) showed the presence of an extra nucleotide ("A") at position 12832 and confirmed that in this isolate the two ORFs in fact belong to one single ORF, *orf6*, which is predicted to encode a 256-amino-acid protein with a calculated molecular weight of 29,369. The deduced amino acid sequence of ORF6 revealed the highest similarity, 33%, with an unknown bacteriophage protein of *L.*

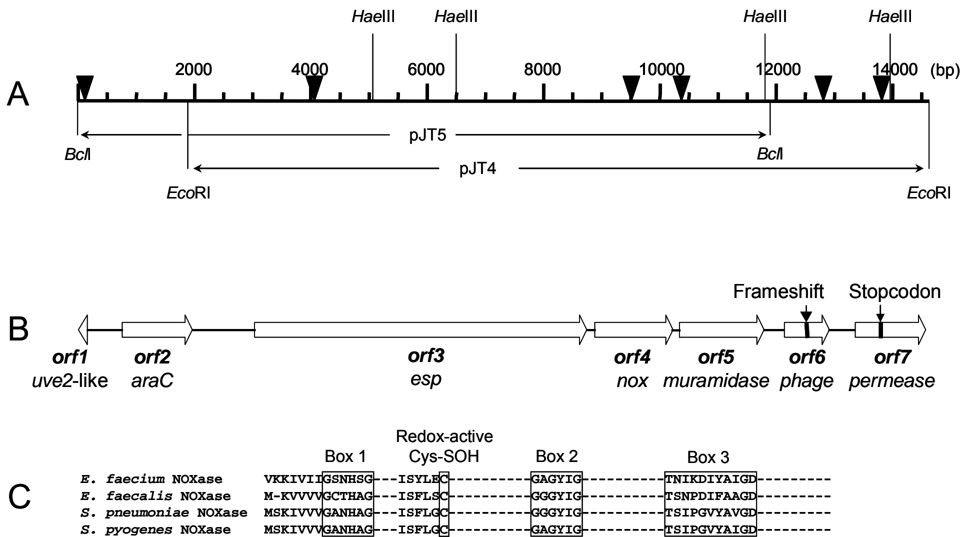

Figure 3. Schematic representation of the putative PAI in *E. faecium* and alignment of NADH oxidase regions. (A) Genetic map of the PAI. Numbers correspond to base pair positions relative to accession no. AY322150. Only restriction enzyme recognition sites relevant for this study are shown. The locations of the two clones that were constructed to derive the entire sequence are indicated. Arrowheads indicate the positions of the oligonucleotide probes used in the Southern hybridization. (B) Physical map of the PAI. Large open arrows with proposed names below indicate sizes, locations, and orientations of predicted ORFs. The positions of the frameshift and stop codon in the PAI of strain E300 are indicated. (C) Comparison of sequence fingerprints of the FAD binding region (boxes 1 and 3), the NADH contact region (box 2), and the cysteine-sulfenic acid redox center of the NADH oxidases (NOXase) of *E. faecium* (this study) with three previously identified homologues: *E. faecalis* (GenBank accession no. X68847) (45), *S. pneumoniae* (GenBank accession no. AF014458) (2), and *S. pyogenes* (GenBank accession no. AF101442) (19).*monocytogenes* strain EGD-e (Table 3; Figure 3) (21).

The last ORF, *orf7*, exhibited amino acid sequence identity with multidrug resistance permeases of *Clostridium perfringens* and *L. lactis*, suggesting that the *orf7* gene may encode a multidrug resistance efflux pump (Table 3; Figure 3). In E300, *orf7* was interrupted by a stop codon at position 13719. Again, repeated sequencing of this region in *E. faecium* isolate E734, which belonged to hospital outbreak NL-1-1 (65), demonstrated that in this isolate the sequence TTA, encoding a leucine, was present instead of the TAA stop codon observed in strain E300. The fact that no stop codon was found in the uninterrupted *orf7* gene suggests that only a part of this gene is present in the cloned and sequenced copy of the *E. faecium esp* gene cluster.

The *E. faecium esp* gene cluster is part of a putative PAI

To investigate whether there was a physical link between *esp* and the other ORFs in this gene cluster, *E. faecium* isolates carrying *esp* and *esp*-deficient strains were analyzed for the presence of the other six ORFs by Southern hybridization. Chromosomal DNAs of 105 *E. faecium* isolates were digested with *Hae*III and hybridized to six oligonucleotide probes derived from internal parts of, and specific for, *orf1* and *orf3* to *orf7* (Figure 3). The selection included 50 vancomycin-susceptible and 55 vancomycin-resistant (*vanA*-positive) isolates (Table 4). All 26 isolates that were *esp* positive also reacted with all six oligonucleotide probes, while all 79 *esp*-negative isolates failed to react with any of the six oligonucleotides. This shows that in this set of isolates, the entire *esp* gene cluster was either present or absent. The 26 isolates carrying this gene cluster included epidemic and clinical isolates, while this cluster was absent in all surveillance and animal isolates. Furthermore, the hybridization results showed that *orf1* and *orf3* were located on identical-sized DNA fragments, as were *orf4* and *orf5*, and *orf6* and *orf7* (data not shown). To further examine whether ORFs 1 to 7 are located in proximity to each other, PCRs were performed with forward and reverse primers specific for different ORFs. PCRs with the primer combinations PAI2F (*orf1*)-PAI17R (*orf2*), esp_{fs}9F (*orf3*)-*nox*2R (*orf4*), PAI11F (*orf5*)-PAI7R (*orf6*), and PAI11F (*orf5*)-PAI3R (*orf7*) (Table 1) demonstrated that ORFs 1 and 2, ORFs 3 and 4, ORFs 5 and 6, and ORFs 6 and 7 are located adjacent to each other (data not shown).

The observation that all six of the putative genes are either present or absent and are physically linked on the genome strongly suggests that this cluster of genes is part of a PAI. This is also supported by the fact that several of the putative genes in this gene cluster have a deviant GC content, ranging from 27.9 to 43.6%, compared to the average GC content of 37.8% in *E. faecium* (Table 3; http://www.jgi.doe.gov/JGI_microbial/html/enterococcus/enterococcus_homepag e.html).

Discussion

In this study we have identified a cluster of six genes in *E. faecium* that are potentially involved in regulation, virulence, and antibiotic resistance and are linked to the *esp* virulence gene. We have designated this gene cluster a putative PAI because of (i) the Southern blot analysis of *esp* positive and -negative *E. faecium* isolates, which demonstrated the presence or absence of the entire gene cluster and a physical link among the seven putative genes in this cluster; (ii) the deviant GC content of this gene cluster (34.4%) compared to the GC content of the

Table 4. Presence of a putative PAI among vancomycin-susceptible and -resistant *E. faecium* isolates from different sources

Strain origin	No. of isolates			
	PAI-		PAI+	
	VSEF[a]	VREF[b]	VSEF	VREF
Hospital outbreak isolates:				
AU-1				1
NL-1-1				1
NL-2-1				2
NL-2-3		1		
NL-3-1				1
UK-1				2
US-1				3
US-2-1				1
US-2-2				1
US-2-3				1
US-2-4				1
US-2-5				1
US-2-6				1
US-2-7				1
Clinical isolates	42	17	8	1
Hospital surveillance isolates		6		
Community surveillance isolates		3		
Pet isolates (dogs, cats)		4		
Swine isolates		2		
Poultry isolates		2		
Calf isolates		2		
Total	42	37	8	18

[a] Vancomycin-susceptibel *E. faecium*

[b] Vancomcyin-resistant *E. faecium*

E. faecium strain DO chromosome (37.8%); (iii) the presence of putative virulence genes; and (iv) the fact that this island is absent in all human surveillance and animal isolates but present in epidemic and clinical isolates. The large variation in GC content of the seven ORFs present in the putative PAI, ranging from 27.9 up to 43.6%, suggests that this island was generated not as a result of one single event but as a result of a complex evolution involving multiple steps in different bacterial ancestors, and that it was finally acquired by *E. faecium* through horizontal gene transfer.

The presence of the putative PAI seems to be associated with epidemicity, since 13 of the 14 clones analyzed from different hospital outbreaks contained this PAI. This finding is in line with previous findings that suggested the existence of an epidemic *E. faecium* subpopulation with specific genetic characteristics (65). The fact that such a subpopulation is characterized not only by the presence of *esp* but also by the acquisition of a large genomic island may improve rapid detection of potential epidemic *E. faecium* strains, thus facilitating rapid implementation of infection control strategies. Furthermore, proteins encoded by the putative PAI may be potential targets for specific therapies, for example, to eradicate or prevent gastrointestinal colonization by potentially epidemic *E. faecium*.

A homologue of *E. faecium esp* contained in this putative PAI was first described in *E. faecalis*, where it was found in a high proportion of clinical strains (51). Recently it was shown that the *E. faecalis esp* gene is part of a large (150-kb) PAI (49). The *E. faecalis esp* gene encodes a surface-exposed protein and is thought to be involved in colonization of the urinary tract (50) and biofilm formation (59). In *E. faecium*, the *esp* gene was initially found in vancomycin-resistant outbreak-related isolates (65); later, it was also found in vancomycin-susceptible clinical isolates (3, 12, 13, 68). Recently, the *E. faecium* gene from a clinical isolate, P61, was cloned and sequenced by Eaton and Gasson (13). It displayed 89% similarity with the *E. faecalis esp* gene. It also exhibited global structural similarity to the *S. agalactiae* Rib and alpha C proteins, the R28 protein of *S. pyogenes*, and the biofilm-associated protein (Bap) of *S. aureus*, all of which are known virulence factors conferring protective immunity (8, 33, 37, 54, 55). All these proteins contain a repeat region in which amino acid similarities are most prominent (13, 51). The *E. faecium esp* genes analyzed in this study were highly similar but not identical to the P61 *esp* gene. In addition to variations in the numbers of A, B, and C repeats, extensive polymorphism was found in the N- and C-terminal nonrepeat regions. This may suggest that the *esp* gene was not acquired recently by *E. faecium* or that the *esp* gene is a relatively "ancient" gene acquired by *E. faecium* during multiple occasions. In addition, heterogeneity, especially in the surface-exposed N-terminal region, may correspond to different functions or specificities of different Esp variants. Differences in repeat numbers in *esp*, both in *E. faecalis* and in *E.*

faecium, have been reported previously (13, 51). It is questionable whether this heterogeneity in repeat numbers can be used as an epidemiological tool. Comparison of the *esp* repeat profiles of epidemiologically linked and unrelated strains suggests that *esp* repeat profiling may be used to study local outbreaks but probably does not discriminate sufficiently to serve as a major tool for global epidemiology unless it is used in combination with genotyping schemes such as multilocus sequence typing or pulsed-field gel electrophoresis.

The presence of the *esp* gene in isolates from epidemiologically distinct sources seems to differ between *E. faecalis* and *E. faecium*. While the presence of the *esp* gene in *E. faecium* is confined to clinical and epidemic isolates, in *E. faecalis* the *esp* gene is also found in isolates from farm animals and food (12, 17, 24). This could be related to differences in the frequency of horizontal transmission of the *esp* gene in *E. faecalis* and *E. faecium*.

In addition to *esp*, two other putative virulence genes were found on this genetic island: the *nox* and muramidase genes, encoding a NADH oxidase and muramidase or autolysin, respectively. NADH oxidases are enzymes that can catalyze the four-electron reduction of O_2 to H_2O and are considered to perform normal household functions. In *E. faecalis*, NADH oxidase is involved in glycolytic metabolism (47). However, similar enzymes in *S. pyogenes*, *Streptococcus mutans*, and *S. pneumoniae* are considered virulence factors involved in adaptive responses to O_2, enabling these bacteria to grow in O_2-rich environments (2, 19, 25, 69). Furthermore, the NADH oxidase of *S. pneumoniae* is also involved in natural competence for genetic exchange (2, 14). It is not yet known whether the NADH oxidase found on the *E. faecium* putative PAI is involved in virulence. One can speculate that *E. faecium* isolates harboring this enzyme are better equipped to leave the anaerobic conditions in the gut and grow in more oxygen rich niches such as the urinary tract or the bloodstream.

The muramidase gene is predicted to encode an enzyme with important physiological functions during cell growth and division (52, 53, 56). Most of these enzymes have a domain structure (42). The *E. faecium* muramidase encoded by the putative PAI displayed similarity only with the N-terminal enzymatically active domain of the *E. hirae* muramidase-2 and seems to lack the C-terminal peptidoglycan binding domain. It was shown previously that the muramidase-2 enzyme of *E. hirae* covalently binds penicillin (10). It is not known whether the *E. faecium* muramidase described here is able to bind ß-lactam antibiotics, but the characteristic motifs present in serine ß-lactamases and penicillin-binding proteins are also conserved in this protein. In addition to basic cell functions, some bacterial peptidoglycan hydrolases, muramidase or autolysin, have been implicated in virulence by contributing to primary adhesion, biofilm formation, or other, yet unknown processes (1, 22, 27, 31, 36, 39, 46). Some other murein

hydrolases, such as the lysostaphin of *Staphylococcus simulans*, may also act as bacteriocins (70). The production of bacteriocins may provide a competitive advantage in specific niches, thus promoting intestinal colonization. Furthermore, a peptidoglycan hydrolase gene of *Neisseria gonorrhoeae*, *atlA*, was also found on a PAI (9), and it was demonstrated that this *atlA* gene was required for DNA secretion during growth. This suggests that peptidoglycan hydrolases may also play a role in DNA transfer events. Further characterization of the peptidoglycan hydrolase encoded by the muramidase-like gene on the putative PAI is needed to establish a potential role in penicillin binding, pathogenesis, or intestinal colonization.

The first two ORFs of this putative PAI encode putative transcriptional regulators. *orf1*, which was cloned and sequenced only partially, may encode a sigma-like factor, while *orf2* most likely encodes a protein that belongs to the AraC family of global regulators. Both AraC and alternate sigma factors are often found on PAIs (reviewed by Hacker and Kaper [23] and Egan [15]). Interestingly, an *araC*-like gene was also found on the recently described *E. faecalis* PAI, downstream of *esp*, while the *E. faecium araC* is located upstream of *esp* (49). Transcriptional regulators contained on PAIs may regulate virulence genes located on the same island or genes located outside the PAI. At this moment the role of these regulators in *E. faecium* is the subject of ongoing research.

The last two ORFs were disrupted in isolate E300 but were found intact in other isolates. They may encode a hypothetical bacteriophage protein and a multidrug resistance efflux pump. Bacteriophages have been implicated in the mobilization of PAIs, and several PAIs contain sequences with homology to bacteriophage integrase genes (reviewed by Hacker and Kaper [23]). The exact function of this putative phage protein remains to be elucidated. Sequence analysis and alignment of the last ORF suggested that *orf7* may encode a putative multidrug resistance efflux pump that was only partially present on the cloned and sequenced copy of the putative PAI. Although virulence and antibiotic resistance may often be linked (34), antibiotic resistance genes are rarely found on PAIs. Recently, a PAI carrying a resistance locus conferring resistance to streptomycin, ampicillin, chloramphenicol, and tetracycline was found in *Shigella flexneri* (32, 60).

Comparison of the putative *E. faecium* PAI and the recently published *E. faecalis* PAI revealed that these two enterococcal PAIs are different, although they share at least two genes: *araC* and *esp* (49). It is intriguing that these two related enterococcal species, which are often found in the same niche, carry different PAIs. On the other hand, the epidemiology of the two species seems to be different. While *E. faecalis* is more frequently encountered among clinical isolates, *E. faecium*, mainly ampicillin- and vancomycin-resistant isolates, is more often

associated with epidemic spread in hospitals (26, 40, 43). It is possible that differences in PAI sequences between *E. faecium* and *E. faecalis*, in addition to differences in antibiotic susceptibility, may account for these epidemiological differences.

Since PAIs may provide a rapid and flexible means of evolution of virulence by generating new pathogenic variants, it is not unlikely that the acquisition of a PAI by *E. faecium* has played an important role in the rapid emergence of *E. faecium* as a nosocomial pathogen.

Acknowledgments

We thank Michael Dunne, Jr., J. Zoe Jordens, Christina Vandenbroucke-Grauls, David Tribe, Ellen Mascini, Ad Fluit, Ellen Stobberingh, and Dik Mevius for providing isolates. We also thank Han de Neeling for helpful discussions and critical reading of the manuscript.

This work was supported by a Marie Curie Fellowship of the European Community program "Quality of Life and Management of Living Resources" under contract QLK2-CT-2001-50991

References

1. Allignet, J., S. Aubert, K. G. Dyke, and N. El Solh. 2001. *Staphylococcus caprae* strains carry determinants known to be involved in pathogenicity: a gene encoding an autolysin-binding fibronectin and the *ica* operon involved in biofilm formation. Infect. Immun. 69:712-718.

2. Auzat, I., S. Chapuy-Regaud, G. Le Bras, D. Dos Santos, A. D. Ogunniyi, I. Le Thomas, J. R. Garel, J. C. Paton, and M. C. Trombe. 1999. The NADH oxidase of *Streptococcus pneumoniae*: its involvement in competence and virulence. Mol. Microbiol. 34:1018-1028

3. Baldassarri, L., L. Bertuccini, M. G. Ammendolia, G. Gherardi, and R. Creti. 2001. Variant *esp* gene in vancomycin-sensitive *Enterococcus faecium*. Lancet 357:1802.

4. Bell, J. M., J. C. Paton, and J. Turnidge. 1998. Emergence of vancomycin-resistant enterococci in Australia: phenotypic and genotypic characteristics of isolates. J. Clin. Microbiol. 36:2187-2190.

5. Bonten, M. J. M., M. K. Hayden, C. Nathan, T. W. Rice, and R. A. Weinstein. 1998. Stability of vancomycin-resistant enterococcal genotypes isolated from long-term-colonized patients. J. Infect. Dis. 177:378-382

6. Boyd, D. A., T. Cabral, P. Van Caeseele, J. Wylie, and M. R. Mulvey. 2002. Molecular characterization of the *vanE* gene cluster in vancomycin-resistant *Enterococcus faecalis* N00-410 isolated in Canada. Antimicrob. Agents Chemother. 46:1977-1979.

7. Coque, T. M., R. Willems, R. Canton, R. Del Campo, and F. Baquero. 2002. High occurrence of *esp* among ampicillin-resistant and vancomycin-susceptible *Enterococcus faecium* clones from hospitalized patients. J. Antimicrob. Chemother. 50:1035-1038.

8. Cucarella, C., C. Solano, J. Valle, B. Amorena, I. Lasa, and J. R. Penades. 2001. Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. J. Bacteriol. 183:2888-2896.

9. Dillard, J. P., and H. S. Seifert. 2001. A variable genetic island specific for *Neisseria gonorrhoeae* is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. Mol. Microbiol. 41:263-277

10. Dolinger, D. L., L. Daneo Moore, and G. D. Shockman. 1989. The second peptidoglycan hydrolase of *Streptococcus faecium* ATCC 9790 covalently binds penicillin. J. Bacteriol. 171:4355-4361.

11. Dunne, W. M., and W. Wang. 1997. Clonal dissemination and colony morphotype variation of vancomycin-resistant *Enterococcus faecium* isolates in metropolitan Detroit, Michigan. J. Clin. Microbiol. 35:388-392.

12. Eaton, T. J., and M. J. Gasson. 2001. Molecular screening of *Enterococcus* virulence determinants and potential for genetic exchange between food and medical isolates. Appl. Environ. Microbiol. 67:1628-1635.

13. Eaton, T. J., and M. J. Gasson. 2002. A variant enterococcal surface protein Esp_{fm} in *Enterococcus faecium*; distribution among food, commensal, medical, and environmental isolates. FEMS Microbiol. Lett. 216:269-275

14. Echenique, J. R., S. Chapuy-Regaud, and M. C. Trombe. 2000. Competence regulation by oxygen in *Streptococcus pneumoniae*: involvement of *ciaRH* and *comCDE.* Mol. Microbiol. 36:688-696

15. Egan, S. M. 2002. Growing repertoire of AraC/XylS activators. J. Bacteriol. 184:5529-5532]

16. Endtz, H., N. van den Braak, A. van Belkum, J. A. J. W. Kluytmans, J. G. M. Koeleman, L. Spanjaard, A. Voss, A. J. L. Weersink, C. M. J. E. Vandenbroucke-Grauls, A. G. M. Buiting, A. van Duin, and H. A. Verbrugh. 1997. Fecal carriage of vancomycin-resistant enterococci in hospitalized patients and those living in the community in The Netherlands. J. Clin. Microbiol. 35:3026-3031.

17. Franz, C. M. A. P., A. B. Muscholl Silberhorn, N. M. K. Yousif, M. Vancanneyt, J. Swings, and W. H. Holzapfel. 2001. Incidence of virulence factors and antibiotic resistance among enterococci isolated from food. Appl. Environ. Microbiol. 67:4385-4389.

18. Ghuysen, J. M. 1991. Serine beta-lactamases and penicillin-binding proteins. Annu. Rev. Microbiol. 45:37-67

19. Gibson, C. M., T. C. Mallett, A. Claiborne, and M. G. Caparon. 2000. Contribution of NADH oxidase to aerobic metabolism of *Streptococcus pyogenes.* J. Bacteriol. 182:448-455.

20. Gilmore, M. S., P. S. Coburn, S. R. Nallapareddy, and B. E. Murray. 2002. Enterococcal virulence, p. 301-354. *In* M. S. Gilmore, D. B. Clewell, P. Courvalin, G. M. Dunny, B. E. Murray, and L. B. Rice (ed.), The enterococci: pathogenesis, molecular biology, and antibiotic resistance. American Society for Microbiology, Washington, D.C.

21. Glaser, P., L. Frangeul, C. Buchrieser, C. Rusniok, A. Amend, F. Baquero, P. Berche, H. Bloecker, P. Brandt, T. Chakraborty, A. Charbit, F. Chetouani, E. Couve, A. de Daruvar, P. Dehoux, E. Domann, G. Dominguez-Bernal, E. Duchaud, L. Durant, O. Dussurget, K.-D. Entian, H. Fsihi, F. G.-D. Portillo, P. Garrido, L. Gautier, W. Goebel, N. Gomez-Lopez, T. Hain, J. Hauf, D. Jackson, L.-M. Jones, U. Kaerst, J. Kreft, M. Kuhn, F. Kunst, G. Kurapkat, E. Madueno, A. Maitournam, J. M. Vicente, E. Ng, H. Nedjari, G. Nordsiek, S. Novella, B. de Pablos, J.-C. Perez-Diaz, R. Purcell, B. Remmel, M. Rose, T. Schlueter, N. Simoes, A. Tierrez, J.-A. Vazquez-Boland, H. Voss, J. Wehland, and P. Cossart. 2001. Comparative genomics of *Listeria* species. Science 294:849-852.

22. Gotz, F. 2002. *Staphylococcus* and biofilms. Mol. Microbiol. 43:1367-1378

23. Hacker, J., and J. B. Kaper. 2000. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. 54:641-679

24. Hammerum, A. M., and L. B. Jensen. 2002. Prevalence of *esp*, encoding the enterococcal surface protein, in *Enterococcus faecalis* and *Enterococcus faecium* isolates from hospital patients, poultry, and pigs in Denmark. J. Clin. Microbiol. 40:4396.

25. Higuchi, M. 1984. The effect of oxygen on the growth and mannitol fermentation of *Streptococcus mutans.* J. Gen. Microbiol. 130:1819-1826

26. Iwen, P. C., D. M. Kelly, J. Linder, S. H. Hinrichs, E. A. Dominguez, M. E. Rupp, and K. D. Patil. 1997. Change in prevalence and antibiotic resistance of *Enterococcus* species isolated from blood cultures over an 8-year period. Antimicrob. Agents Chemother. 41:494-495.

27. Jedrzejas, M. J. 2001. Pneumococcal virulence factors: structure and function. Microbiol. Mol. Biol. Rev. 65:187-207.

28. Jordens, J. Z., J. Bates, and D. T. Griffiths. 1994. Faecal carriage and nosocomial spread of vancomycin-resistant *Enterococcus faecium.* J. Antimicrob. Chemother. 34:515-528.

29. Joris, B., S. Englebert, C. P. Chu, R. Kariyama, L. Daneo Moore, G. D. Shockman, and J. M. Ghuysen. 1992. Modular design of the *Enterococcus hirae* muramidase-2 and *Streptococcus faecalis* autolysin. FEMS Microbiol. Lett. 70:257-264

30. Leavis, H. L., R. J. L. Willems, J. Top, E. Spalburg, E. M. Mascini, A. C. Fluit, A. Hoepelman, A. J. de Neeling, and M. J. M. Bonten. 2003. Evolutionary insights in the emergence of epidemic and non-epidemic multi-resistant *Enterococcus faecium.* Emerg. Infect. Dis. 9:1108-1115

31. Loessner, M. J., K. Kramer, F. Ebel, and S. Scherer. 2002. C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. Mol. Microbiol. 44:335-349

32. Luck, S. N., S. A. Turner, K. Rajakumar, H. Sakellaris, and B. Adler. 2001. Ferric dicitrate transport system (Fec) of *Shigella flexneri* 2a YSH6000 is encoded on a novel pathogenicity island carrying multiple antibiotic resistance genes. Infect. Immun. 69:6012-6021.

33. Madoff, L. C., J. L. Michel, E. W. Gong, D. E. Kling, and D. L. Kasper. 1996. Group B streptococci escape host immunity by deletion of tandem repeat elements of the alpha C protein. Proc. Natl. Acad. Sci. USA 93:4131-4136.

34. Martinez, J. L., and F. Baquero. 2002. Interactions among strategies associated with bacterial infection: pathogenicity, epidemicity, and antibiotic resistance. Clin. Microbiol. Rev. 15:647-679.

35. Mascini, E. M., A. C. Gigengack-Baars, R. J. Hene, T. E. Kamp-Hopmans, A. J. Weersink, and M. J. Bonten. 2000. Epidemiologic increase of various genotypes of vancomycin-resistant *Enterococcus faecium* in a university hospital. Ned. Tijdschr. Geneeskd. 144:2572-2576. (In Dutch.)

36. Mercier, C., C. Durrieu, R. Briandet, E. Domakova, J. Tremblay, G. Buist, and S. Kulakauskas. 2002. Positive role of peptidoglycan breaks in lactococcal biofilm formation. Mol. Microbiol. 46:235-243

37. Michel, J. L., L. C. Madoff, D. E. Kling, D. L. Kasper, and F. M. Ausubel. 1991. Cloned alpha and beta C-protein antigens of group B streptococci elicit protective immunity. Infect. Immun. 59:2023-2028.

38. Michel, J. L., L. C. Madoff, K. Olson, D. E. Kling, D. L. Kasper, and F. M. Ausubel. 1992. Large, identical, tandem repeating units in the C protein alpha antigen gene, *bca*, of group B streptococci. Proc. Natl. Acad. Sci. USA 89:10060-10064.

39. Milohanic, E., R. Jonquieres, P. Cossart, P. Berche, and J. L. Gaillard. 2001. The autolysin Ami contributes to the adhesion of *Listeria monocytogenes* to eukaryotic cells via its cell wall anchor. Mol. Microbiol. 39:1212-1224

40. Murdoch, D. R., S. Mirrett, L. J. Harrell, J. S. Monahan, and L. B. Reller. 2002. Sequential emergence of antibiotic resistance in enterococcal bloodstream isolates over 25 years. Antimicrob. Agents Chemother. 46:3676-3678.

41. Nallapareddy, S. R., G. M. Weinstock, and B. E. Murray. 2003. Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. Mol. Microbiol. 47:1733-1747

42. Navarre, W. W., and O. Schneewind. 1999. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. Microbiol. Mol. Biol. Rev. 63:174-229.

43. Rice, L. B. 2001. Emergence of vancomycin-resistant enterococci. Emerg. Infect. Dis. 7:183-187

44. Rice, L. B., L. Carias, S. Rudin, C. Vael, H. Goossens, C. Konstabel, I. Klare, S. R. Nallapareddy, W. Huang, and B. E. Murray. 2003. A potential virulence gene, hyl_{Efm}, predominates in *Enterococcus faecium* of clinical origin. J. Infect. Dis. 187:508-512

45. Ross, R. P., and A. Claiborne. 1992. Molecular cloning and analysis of the gene encoding the NADH oxidase from *Streptococcus faecalis* 10C1. Comparison with NADH peroxidase and the flavoprotein disulfide reductases. J. Mol. Biol. 227:658-671

46. Rupp, M. E., P. D. Fey, C. Heilmann, and F. Gotz. 2001. Characterization of the importance of *Staphylococcus epidermidis* autolysin and polysaccharide intercellular adhesin in the pathogenesis of intravascular catheter-associated infection in a rat model. J. Infect. Dis. 183:1038-1042

47. Schmidt, H. L., W. Stocklein, J. Danzer, P. Kirch, and B. Limbach. 1986. Isolation and properties of an H_2O-forming NADH oxidase from *Streptococcus faecalis.* Eur. J. Biochem. 156:149-155.

48. Schouls, L. M., C. S. Schot, and J. A. Jacobs. 2003. Horizontal transfer of segments of the 16S rRNA genes between species of the *Streptococcus anginosus* group. J. Bacteriol. 185:7241-7246.

49. Schouten, M. A., R. J. Willems, W. A. Kraak, J. Top, J. A. Hoogkamp-Korstanje, and A. Voss. 2001. Molecular analysis of Tn*1546*-like elements in vancomycin-resistant enterococci isolated from patients in Europe shows geographic transposon type clustering. Antimicrob. Agents Chemother. 45:986-989.

50. Shankar, N., A. S. Baghdayan, and M. S. Gilmore. 2002. Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis.* Nature 417:746-750

51. Shankar, N., C. V. Lockatell, A. S. Baghdayan, C. Drachenberg, M. S. Gilmore, and D. E. Johnson. 2001. Role of *Enterococcus faecalis* surface protein Esp in the pathogenesis of ascending urinary tract infection. Infect. Immun. 69:4366-4372.

52. Shankar, V., A. S. Baghdayan, M. M. Huycke, G. Lindahl, and M. S. Gilmore. 1999. Infection-derived *Enterococcus faecalis* strains are enriched in *esp*, a gene encoding a novel surface protein. Infect. Immun. 67:193-200.

53. Shockman, G. D. 1992. The autolytic ('suicidase') system of *Enterococcus hirae*: from lysine depletion autolysis to biochemical and molecular studies of the two muramidases of *Enterococcus hirae* ATCC 9790. FEMS Microbiol. Lett. 79:261-267

54. Shockman, G. D., L. Daneo-Moore, R. Kariyama, and O. Massidda. 1996. Bacterial walls, peptidoglycan hydrolases, autolysins, and autolysis. Microb. Drug Resist. 2:95-98

55. Stalhammar-Carlemalm, M., T. Areschoug, C. Larsson, and G. Lindahl. 1999. The R28 protein of *Streptococcus pyogenes* is related to several group B streptococcal surface proteins, confers protective immunity and promotes binding to human epithelial cells. Mol. Microbiol. 33:208-219

56. Stalhammar-Carlemalm, M., L. Stenberg, and G. Lindahl. 1993. Protein Rib: a novel group B streptococcal cell surface protein that confers protective immunity and is expressed by most strains causing invasive infections. J. Exp. Med. 177:1593-1603.

57. Takahashi, J., H. Komatsuzawa, S. Yamada, T. Nishida, H. Labischinski, T. Fujiwara, M. Ohara, J. Yamagishi, and M. Sugai. 2002. Molecular characterization of an *atl* null mutant of *Staphylococcus aureus.* Microbiol. Immunol. 46:601-612

58. Tannock, G. W., and G. Cook. 2002. Enterococci as members of the intestinal microflora of humans, p. 101-132. *In* M. S. Gilmore, D. B. Clewell, P. Courvalin, G. M. Dunny, B. E. Murray, and L. B. Rice (ed.), The enterococci: pathogenesis, molecular biology, and antibiotic resistance. American Society for Microbiology, Washington, D.C.

59. Timmers, G. J., W. C. van Der Zwet, I. M. Simoons-Smit, P. H. Savelkoul, H. H. Meester, C. M. Vandenbroucke-Grauls, and P. C. Huijgens. 2002. Outbreak of vancomycin-resistant *Enterococcus faecium* in a haematology unit: risk factor assessment and successful control of the epidemic. Br. J. Haematol. 116:826-833

60. Toledo Arana, A., J. Valle, C. Solano, M. J. Arrizubieta, C. Cucarella, M. Lamata, B. Amorena, J. Leiva, J. R. Penades, and I. Lasa. 2001. The enterococcal surface protein, Esp, is involved in *Enterococcus faecalis* biofilm formation. Appl. Environ. Microbiol. 67:4538-4545.

61. Turner, S. A., S. N. Luck, H. Sakellaris, K. Rajakumar, and B. Adler. 2003. Molecular epidemiology of the SRL pathogenicity island. Antimicrob. Agents Chemother. 47:727-734.

62. van Belkum, A., N. van den Braak, R. Thomassen, H. Verbrugh, and H. Endtz. 1996. Vancomycin-resistant enterococci in cats and dogs. Lancet 348:1038-1039.

63. van den Bogaard, A. E., P. Mertens, N. H. London, and E. E. Stobberingh. 1997. High prevalence of colonization with vancomycin- and pristinamycin-resistant enterococci in healthy humans and pigs in The Netherlands: is the addition of antibiotics to animal feeds to blame? J. Antimicrob. Chemother. 40:454-456.

64. van den Bogaard, A. E., R. Willems, N. London, J. Top, and E. E. Stobberingh. 2002. Antibiotic resistance of faecal enterococci in poultry, poultry farmers and poultry slaughterers. J. Antimicrob. Chemother. 49:497-505.

65. van der Steen, L. F., M. J. Bonten, E. van Kregten, J. J. Harssema-Poot, R. Willems, and C. A. Gaillard. 2000. Vancomycin-resistant *Enterococcus faecium* outbreak in a nephrology ward. Ned. Tijdschr. Geneeskd. 144:2568-2572. (In Dutch.)

66. Willems, R. J., W. Homan, J. Top, M. van Santen-Verheuvel, D. Tribe, X. Manzioros, C. Gaillard, C. M. Vandenbroucke-Grauls, E. M. Mascini, E. van Kregten, J. D. van Embden, and M. J. Bonten. 2001. Variant *esp* gene as a marker of a distinct genetic lineage of vancomycin-resistant *Enterococcus faecium* spreading in hospitals. Lancet 357:853-855

67. Willems, R. J. L., J. Top, N. van den Braak, A. van Belkum, H. Endtz, D. Mevius, E. Stobberingh, A. van den Bogaard, and J. D. A. van Embden. 2000. Host specificity of vancomycin-resistant *Enterococcus faecium.* J. Infect. Dis. 182:816-823

68. Willems, R. J. L., J. Top, N. van den Braak, A. van Belkum, D. J. Mevius, G. Hendriks, M. Van Santen Verheuvel, and J. D. A. van Embden. 1999. Molecular diversity and evolutionary relationships of Tn*1546*-like elements in enterococci from humans and animals. Antimicrob. Agents Chemother. 43:483-491.

69. Woodford, N., M. Soltani, and K. J. Hardy. 2001. Frequency of *esp* in *Enterococcus faecium* isolates. Lancet 358:584.

70. Yu, J., A. P. Bryant, A. Marra, M. A. Lonetto, K. A. Ingraham, A. F. Chalker, D. J. Holmes, D. Holden, M. Rosenberg, and D. McDevitt. 2001. Characterization of the *Streptococcus pneumoniae* NADH oxidase that is required for infection. Microbiology 147:431-438.

71. Zygmunt, W. A., H. P. Browder, and P. A. Tavormina. 1967. Lytic action of lysostaphin on susceptible and resistant strains of *Staphylococcus aureus.* Can. J. Microbiol. 13:845-853