

A novel class of small repetitive DNA sequences in *Enterococcus faecalis*

Rossella Venditti¹, Eliana De Gregorio¹, Giustina Silvestro¹, Tullia Bertocco¹, Maria Francesca Salza², Raffaele Zarrilli^{2,3} & Pier Paolo Di Nocera¹

¹Dipartimento di Biologia e Patologia Cellulare e Molecolare, Facoltà di Medicina, Università Federico II, Napoli, Italy; ²Dipartimento di Scienze Mediche Preventive, Sezione di Igiene, Università Federico II, Napoli, Italy; and ³CEINGE Biotecnologie Avanzate, Napoli, Italy

Correspondence: Pier Paolo Di Nocera, Dipartimento di Biologia e Patologia Cellulare e Molecolare, Facoltà di Medicina, Università Federico II, Via S. Pansini 5, 80131 Napoli, Italy. Tel.: +39 81 7462059; fax: +39 81 7703285; e-mail: dinocera@unina.it

Received 26 January 2007; revised 6 March 2007; accepted 6 March 2007.
First published online 10 April 2007.

DOI:10.1111/j.1574-6968.2007.00717.x

Editor: Marco Soria

Keywords

stem-loop structures; palindromic DNA; miniature insertion sequence; genome analysis; microbiological diagnostic.

Introduction

Enterococci are nonspore-forming Gram-positive microorganisms normally considered commensal of the gastrointestinal tracts of humans and animals, and are commonly found in soil, sewage, water and food, frequently through fecal contamination. In recent years, Enterococci have received growing attention as opportunistic pathogens of clinical significance because they are capable of causing serious diseases (Murray, 2000; Malani *et al.*, 2002). *Enterococcus faecalis* is responsible for severe sepsis and endocarditis, and is an important etiological agent of nosocomial infections. The intrinsic antibiotic resistance, the tolerance of adverse environmental conditions, the promiscuity in acquisition and dissemination of genetically mobile antibiotic resistance elements are all factors that present serious challenges to the treatment of enterococcal infections.

Multilocus sequence typing allowed recently to identify two clonal complexes of *E. faecalis*, CC2 and CC9, responsible for outbreaks and life-threatening infections, mostly in the hospital environment (Ruiz-Garbajosa *et al.*, 2006). In turn, CC2 includes the BVE (Bla⁺, Van^r endocarditis; Nallapareddy *et al.*, 2005) complex to which belongs the wholly sequenced V583 strain (Paulsen *et al.*, 2003). BVE

Abstract

The structural organization of *Enterococcus faecalis* repeats (EFAR) is described, palindromic DNA sequences identified in the genome of the *Enterococcus faecalis* V583 strain by *in silico* analyses. EFAR are a novel type of miniature insertion sequences, which vary in size from 42 to 650 bp. Length heterogeneity results from the variable assembly of 16 different sequence types. Most elements measure 170 bp, and can fold into peculiar L-shaped structures resulting from the folding of two independent stem-loop structures (SLSs). Homologous chromosomal regions lacking or containing EFAR sequences were identified by PCR among 20 *E. faecalis* clinical isolates of different genotypes. Sequencing of a representative set of 'empty' sites revealed that 24–37 bp-long sequences, unrelated to each other but all able to fold into SLSs, functioned as targets for the integration of EFAR. In the process, most of the SLS had been deleted, but part of the targeted stems had been retained at EFAR termini.

clones are rarely found among *E. faecalis* isolates. Progression towards hospital adaptation is likely a multi-step cumulative process where genetic exchange or mutation may lead to epidemic, rather than to clonal population structures (see Feil & Spratt, 2001). Not surprisingly, a prominent feature of the V583 genome is the extraordinary abundance (*c.* 25%) of probable mobile and/or foreign DNA including a plethora of insertion elements (IS), multiple transposons, integrated phage regions and plasmid genes (Paulsen *et al.*, 2003).

Herein is reported the organization of a family of repeated DNA sequences identified in the V583 genome by means of bioinformatic approaches (Petrillo *et al.*, 2006), which was called EFAR (for *Enterococcus faecalis* repeat). These elements partly resemble miniature insertion transposable elements or MITEs, small noncodogenic sequences, which also fold into secondary structures. MITEs feature long terminal inverted repeats (TIRs), and their mobilization is mediated by transposases encoded by ISs featuring similar TIRs (Oggioni & Claverys, 1999; Brugger *et al.*, 2002; De Gregorio *et al.*, 2003a, b, 2005). EFAR lack TIRs, and seem to transpose by an unusual cut-and-paste process. It is demonstrated, by means of *in silico* and *in vivo* data, that EFAR have a highly modular structure. Changes in the

organization of the EFAR family among clinical isolates can be easily detectable, making EFAR repeats suitable probes to investigate the epidemiology and population structure of *E. faecalis*.

Materials and methods

Bacterial isolates and growth conditions

Twenty clinical isolates of *E. faecalis* collected from different patients in the Neapolitan area were included in the study (Zarrilli *et al.*, 2005). Relevant background and characteristics of the isolates are detailed in Table 1. All isolates were grown in blood-agar plates at 37 °C and stored at –70 °C in brain heart infusion (BHI) broth plus 10% glycerol. Bacteria were identified by conventional methods (Gram stain, catalase test) and by biochemical tests using API 20 Strep (bioMérieux, France). All isolates were further identified as *E. faecalis* by amplification and sequence analysis of the 16S rRNA gene performed as previously described (Angeletti *et al.*, 2001).

PCR analyses

Genomic DNA was purified from cultures of *E. faecalis* grown at 37 °C in BHI broth by phenol–chloroform extractions as described (Sambrook *et al.*, 1989). EFAR sequences were amplified by standard protocols using 10 ng of genomic DNA and 160 ng of the *for* (5'-GAGCGTGGGA

CAAAAATCAC-3') and *rev* (5'-GAGGTCGGGACAGAA CCGTT-3') primers. One oligonucleotide of the pair had been ³²P-end-labelled at the 5' terminus with the polynucleotide kinase. Amplimers were electrophoresed on 6% acrylamide-urea gels and detected by autoradiography. Amplimers labeled as IX–XV in Fig. 2 were gel-purified and reamplified with the *for* and *rev* EFAR primers. Amplimers were then purified from 1.4% agarose gels, and their nucleotide sequence determined by the dye-terminator method. Specific chromosomal segments of the *E. faecalis* genome were amplified by PCR using pairs of oligonucleotides complementary to coding regions flanking EFAR elements in the V583 genome at the concentrations described above. The amplimers were electrophoresed on 1.4% agarose gels along with a commercial DNA ladder (Ladder 100 plus, MBI) as molecular weight marker. Sequences of the PCR primers used are available upon request.

Slot-blot analyses

0.25, 0.5 and 1 µg of DNA from specific *E. faecalis* isolates were loaded onto a Hybond filter and cross-linked by UV treatment as described (Carlomagno *et al.*, 1988). The filter was hybridized to a ³²P-radiolabeled PCR product spanning a unit length EFAR repeat. Radioactivity signals were quantitated by phosphorimagery. Signals resulting from hybridization to cold probe DNA loaded on the filter were used to estimate the relative abundance of EFAR DNA in each isolate.

Table 1. *Enterococcus faecalis* isolates by source of isolation, resistance phenotype, PFGE type*

Isolate	Clinical source	Month/year of isolation	Resistance phenotype			PFGE type
			Bla	VR	HLAR	
67	Endocarditis	12/1987	–	–	–	24
72	Bronchial aspirate	12/1987	–	–	–	25
75	Urine	12/1987	–	–	–	51
81	Urine	12/1987	–	–	–	19
93	Urine	12/1988	–	–	–	13
183	Bronchial aspirate	1/1990	–	–	+	20
308	Wound swab	12/1989	–	–	+	11
412	Urine	6/1991	–	–	–	17
413	Urine	6/1991	–	–	–	21
595	Wound swab	11/1992	–	–	–	28
617	Urine	12/1992	–	–	–	34
921	Wound swab	4/1995	–	–	+	10
1070	Endocarditis	2/1997	–	+	+	3
1146	Blood	5/1998	–	–	–	1
1185	Bronchial aspirate	1/1999	–	–	+	9d
1226	Endocarditis	7/1999	–	–	–	47
1319	Endocarditis	11/2000	–	–	+	7
1340	Urine	5/2001	–	–	+	39b
1342	Endocarditis	6/2001	+	–	+	8
1423	Urine	2/2003	–	–	+	23a

*As designated in Zarrilli *et al.* (2005).

Bla, resistant to betalactams; VR, vancomycin resistance; HLAR, high-level aminoglycoside resistance.

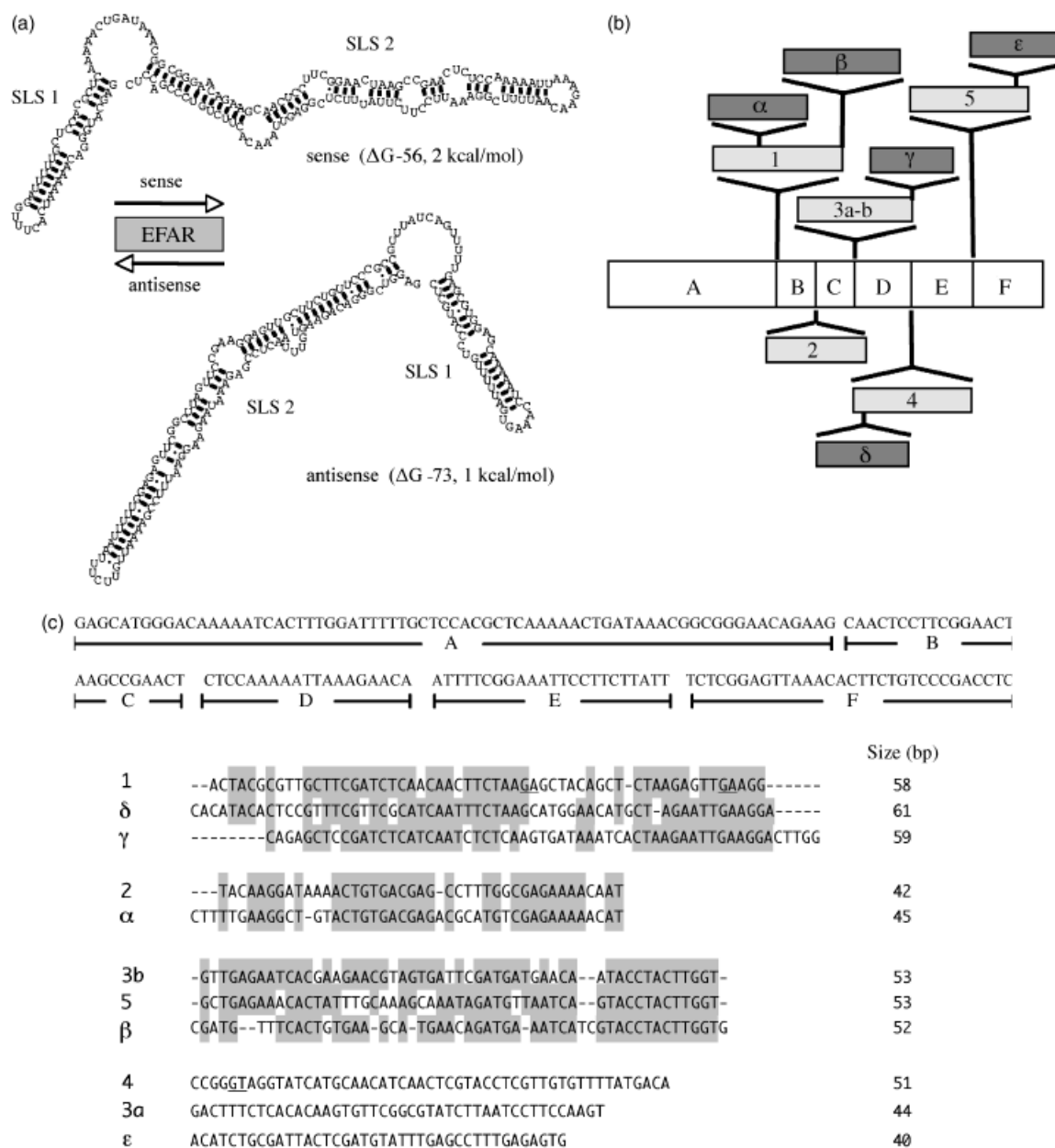


Fig. 1. Structure and organization of EFAR repeats. (a) Secondary structures formed by a consensus 170 bp-long EFAR sequence in the 'sense' and 'antisense' orientation. SLS1 and SLS2 are shown. The Gibbs free energies are indicated. GU pairing is marked by dots. (b) Modular organization of EFAR repeats. The A–F modules, the sites of integration of primary insertions 1–5, and secondary insertions α – ϵ are shown; 3a and 3b are inserted at the same site. (c) Consensus sequences of EFAR modules and insertions. Sequence relatedness is highlighted. Dashes have been introduced to maximize homologies. Underlined residues mark the sites of integration of secondary insertions. Numbers to the right denote the size in bp of each insertion.

Results

Structure and modular organization of EFAR repeats

The interest was in developing systematic searches for prokaryotic sequences able to fold into stem-loop structures

(SLSs; see Petrillo *et al.*, 2006). SLSs were searched as stems measuring at least 12 bp, bordering loops 5–100 nt in length. G-U pairing in the stems was allowed. *In silico* analyses of the chromosome of the *E. faecalis* V583 strain (Paulsen *et al.*, 2003) led to the identification of a novel class of repetitive sequences that was called EFAR in this study. The EFAR family includes 55 members, which vary in size from 42 to

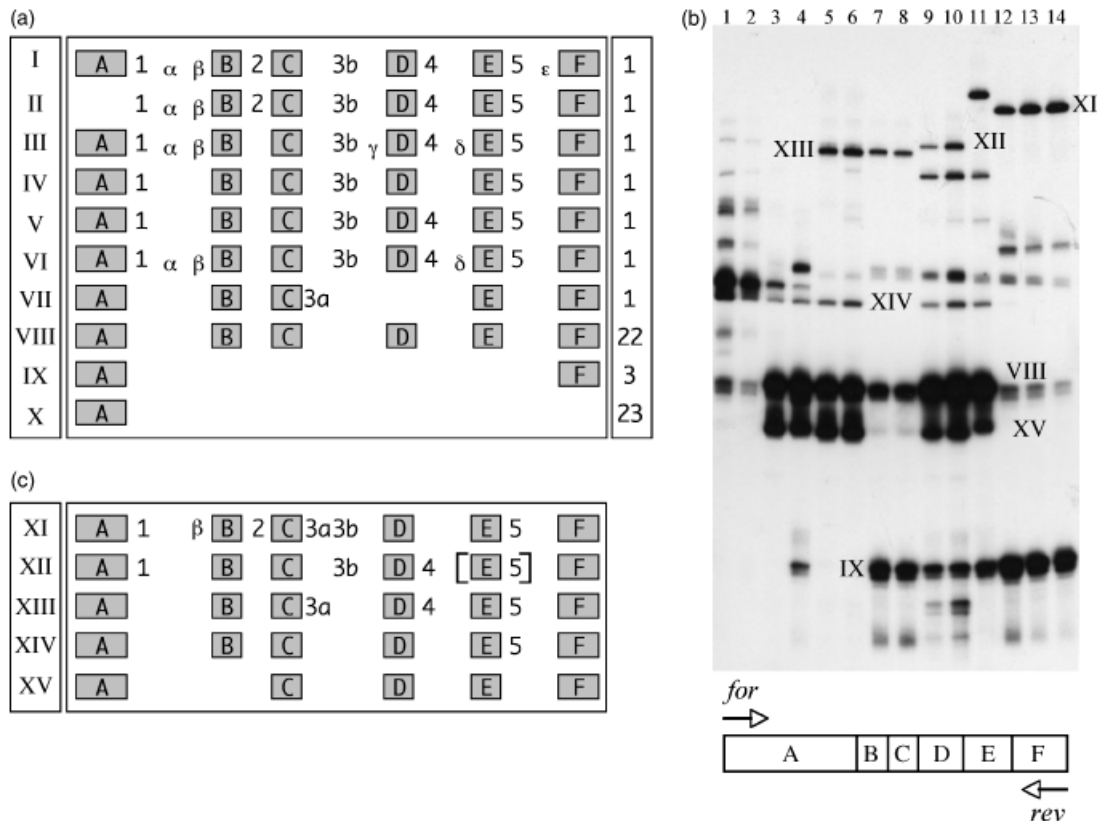


Fig. 2. Genomic organization of EFAR sequences. (a) Modules and insertions defining the 10 (I–X) EFAR subtypes identified in the V583 genome are shown. The number of elements within each subtype is given in the column to the right. (b) DNA derived from 14 isolates of *Enterococcus faecalis* was amplified with the oligomers *for* and *rev* complementary to the EFAR A and F modules. The PCR products were resolved on a 6% polyacrylamide-8M urea gel, and detected by autoradiography. Amplimers labeled VIII–XV were excised from the gel and reamplified using the same primers. The cold PCR products obtained were purified and subjected to sequence analysis. Lanes: 1 (isol. 75), 2 (isol. 617), 3 (isol. 1146), 4 (isol. 595), 5 (isol. 413), 6 (isol. 1226), 7 (isol. 412), 8 (isol. 93), 9 (isol. 1185), 10 (isol. 921), 11 (isol. 67), 12 (isol. 1379), 13 (isol. 1340), 14 (isol. 1342). (c) Modules and insertions in the EFAR subtypes labeled in (b).

650 bp (see Table 2) and exhibit 90–95% sequence homology. The most abundant repeats measure 170 bp. These unit-length elements can fold into L-shaped structures of relatively low free energy, in which a short SLS1 and a long SLS2 are separated by a 20 nt single stranded region (Fig. 1a). The folding of EFAR is peculiar, since elements of comparable size found in other prokaryotes have been shown to potentially fold into single SLSs (see De Gregorio *et al.*, 2003a, b, 2005, 2006). Thirty-five of the EFAR family members listed in Table 2 are at a distance of 30 bp or less from the stop codon of adjacent ORFs. Thus, it is plausible that most EFAR are cotranscribed along with flanking coding sequences.

EFAR have a peculiar modular composition. The presence at specific sites of five DNA sequences (primary insertions 1–5) brought us to subdivide unit-length repeats into the six A–F modules shown in Fig. 1b. Primary insertions may in turn be interrupted by other DNA (secondary insertions α – ϵ in Fig. 1b). Interestingly, insertions 1, δ and γ are 68–70%

homologous to each other. The same holds true for insertions 2 and α , as for insertions 3b, 5 and β (Fig. 1c). Taking into account the presence/absence of all modules and insertions, 10 different element subtypes (Fig. 2a) could be defined. One third of the repeats located in the V583 strain were found lacking one or more modules. Subtype IX repeats keep the terminal A and F modules, subtype X repeats just sequences spanning the A module. Of these, some are heterogeneous in size and clearly represent deletion derivatives of larger EFAR repeats, others measure 42–44 bp, and span just the segment that encodes SLS1 (see Fig. 1a). This supports the hypothesis that EFAR may have originated, as suggested from secondary structure data, from the fusion of independent DNA sequences.

EFAR families in the *E. faecalis* population

To validate knowledge on the structure of the EFAR family emerging from *in silico* analyses, PCR analyses on 14 clinical

Table 2. EFARs identified in the V583 strain and flanking ORFs*

ORF	EFAR	ORF	ORF	EFAR	ORF
89	→ 1	→ 90	1943	→ 29	← 1945
167	← 2	← 168	1980	→ 30	→ 1982
253	→ 3	← 255	2052	← 31	← 2055
401	→ 4	→ 402	2148	← 32	← 2149
672	→ 5	← 673	2150	→ 33	→ 2151
747	→ 6	→ 748	2374	← 34	← 2376
781	→ 7	→ 782	2378	← 35	← 2379
809	→ 8	→ 810	2480	→ 36	← 2481
811	→ 9	→ 812	2495	← 37	→ 2496
822	→ 10	← 824	2504	← 38	← 2505
897	→ 11	← 899	2583	→ 39	→ 2585
921	← 12	← 922	2595	← 40	← 2597
958	→ 13	→ 960	2664	← 41	← 2665
1031	→ 14	← 1032	2698	→ 42	→ 2700
1104	← 15	← 1105	2706	→ 43	← 2708
1116	→ 16	← 1117	2918	← 44	← 2919
1123	→ 17	← 1124	3081	← 45	← 3082
1197	→ 18	→ 1198	3090	→ 46	→ 3091
1313	→ 19	→ 1314	3133	← 47	← 3134
1391	→ 20	→ 1392	3144	→ 48	← 3145
1400	→ 21	← 1402	3206	→ 49	← 3207
1709	→ 22	← 1710	3213	→ 50	→ 3214
1728	→ 23	← 1730	3277	← 51	← 3278
1790	→ 24	← 1791	3281	← 52	← 3282
1809	← 25	← 1810	3283	← 53	← 3284
1811	← 26	← 1812	3292	→ 54	← 3293
1904	→ 27	← 1906	3301	← 55	← 3303
1923	→ 28	← 1925			

*As designated by Paulsen *et al.* (2003).

Arrows denote ORF orientation.

isolates of *E. faecalis* were performed. Since most EFAR elements carry both A and F modules, oligomers complementary to either module were used as primers to monitor the distribution of EFAR repeats among the different *E. faecalis* isolates by PCR. Using unlabeled primers, in several isolates the 170 bp-long EFAR sequences were the predominant amplimers detected. When PCR experiments were performed with 32-P-end labelled primers, a more complex scenario was obtained. A representative electrophoretic profile obtained by this kind of experiment is shown in Fig. 2b. While major patterns of amplification could be distinguished, most clones exhibited a unique PCR pattern. To validate data, several amplimers were gel-purified and reamplified with the same oligonucleotides. The reaction products were electrophoresed onto 1.4% agarose gels, purified and their sequence determined (Fig. 2). Some amplimers were identical to the EFAR subtypes VIII and IX present also in the V583 genome. In contrast, because of changes in the organization of EFAR modules and insertions, other PCR products resulted to be sequence variants not found in the V583 genome. The degree of variability is illustrated by the comparison of subtypes XII and XIII. While similar in size,

the two novel subtypes differ for the presence/absence of insertion 1, and for the alternative presence of insertions 3a and 3b. Moreover, the E module and insertion 5 are partly duplicated in subtype XII (see brackets in Fig. 2c). Several isolates were found to contain unit-length EFARs either selectively decorated by insertion 5 (subtype XIV), or specifically devoided of the B module (subtype XV). Subtype XV resulted in as abundant as subtype VIII in several isolates (lanes 3–6, 9–11). EFARs carrying just the terminal A and F modules (subtype IX repeats) were detected in several isolates, and in some resulted in apparently more abundant than 170 bp-long repeats (see lanes 12–14).

Genomic conservation of EFAR⁺ loci

Next, the extent of conservation of EFAR⁺ loci in the population was assessed. To this end, 12/22 chromosomal regions marked in the V583 strain by the presence of 170 bp-long EFAR sequences were monitored by PCR in 20 *E. faecalis* isolates of different genotypes (Table 1). Genomic DNAs were amplified using oligomers complementary to DNA segments flanking EFAR repeats, located 300–700 bp in the V583 genome. For all tested sites, a PCR product was obtained. The size of the PCR products allowed to easily classify regions analyzed as either 'filled' or 'empty' (i.e. containing or lacking EFAR sequences) sites. PCR results are summarized in Fig. 3a. Data revealed a poor conservation of EFAR⁺ regions on the whole. Only 2/12 elements were retained in all the isolates: EFAR 16, located between the genes encoding the phenylalanyl-tRNA synthetase (ORF 1116) and an ABC transporter permease (ORF 1117), and EFAR 31, located between *ftsK* (ORF 2052) and the gene encoding a pyridine nucleotide-dissulphide oxidoreductase (ORF 2055).

Conservation at other loci varied from 80% to 70% (see loci defined by EFARs 43, 47 and 22) down to 15–10% (see loci defined by EFAR 10 and 29, respectively). In 11/16 filled regions containing EFAR 43, amplification yielded a PCR product larger than expected. As revealed by sequence analysis of the amplimer derived from the isolate 617, size increase is due to a type 1 insertion. The nature of the sequences inserted into EFAR 52 in the 412 and 595 isolates was not investigated.

In view of the results emerging from PCR surveys, the amount of EFAR DNA in each isolate was determined by slot-blot hybridization (Fig. 3b), and shown to vary from a minimum of *c.* 10 copies, as in isolates 183 and 1070, to a maximum of *c.* 30–35 copies, as in isolates 67 and 413.

No correlation could be drawn between the relative abundance and genomic conservation of EFAR repeats. Thus, for example, isolates 921 and 1185, which resulted to be EFAR⁺ at all of the loci tested, had less EFAR DNA than

isolates 67, 413 and 1226, which in contrast seemed lacking EFAR sequences at several of the loci analyzed.

Analysis of EFAR empty sites

On the basis of the sizes of the amplification products, it was postulated that EFAR sequences are missing at several expected chromosomal regions. The analysis of the sequence content of 10 different empty sites (marked in Fig. 3a by asterisks) confirmed the lack of EFAR sequences (Fig. 4a). Sequence data unexpectedly revealed the presence, at the site of EFAR insertions, of short SLSs ranging in size from 24 to 37 bp. These alternative sequence elements did not belong to specific DNA families, as they exhibited poor homology to each other. Furthermore, no related sequences were identified in the V583 genome by BLAST searches. The alternative presence of EFAR and shorter SLSs at the same genomic sites may be interpreted in two different ways. According to one view, EFAR sequences may have been excised from the genome, and replaced at each site by a small SLS (Fig. 4b). A different view suggests, in contrast, that each of the small SLSs functioned as entry sites for the genomic integration of an EFAR repeat. In support of the latter hypothesis, the sequence analysis of the same empty sites from different isolates showed the same alternative SLS in all the specific EFAR regions analyzed.

It is worth noting that sequences at the edges of the alternative SLSs coincided with base-paired regions found at filled sites at the termini of EFAR repeats (highlighted residues in Fig. 4a). This finding can be rationalized by hypothesizing that the enzyme(s) mediating the insertion of EFAR might leave behind part of the AT-rich stems of targeted SLSs upon cleavage.

Discussion

EFAR are relatively large palindromic repeats exhibiting a highly modular structure. Unit-length sequences measure 170 bp and can fold into characteristic L-shaped structures where two distinct hairpins, SLS1 and SLS2, are connected by a 20 nt-long single-stranded region. The identification of elements carrying just sequences spanning SLS1 supports the hypothesis that EFAR may result from the combination of independent sequence types having the ability to fold into SLSs.

None of the mobile DNA sequences found in the *E. faecalis* V583 strain (Paulsen *et al.*, 2003) is related, as shown by BLAST homology searches, to EFAR. All the intergenic regions of V583 were compared by the SEQMATCHALL program of the EMBOS package. Surprisingly, EFAR make up the only family of small (< 300 bp) DNA sequences spread in the genome of enterococci.

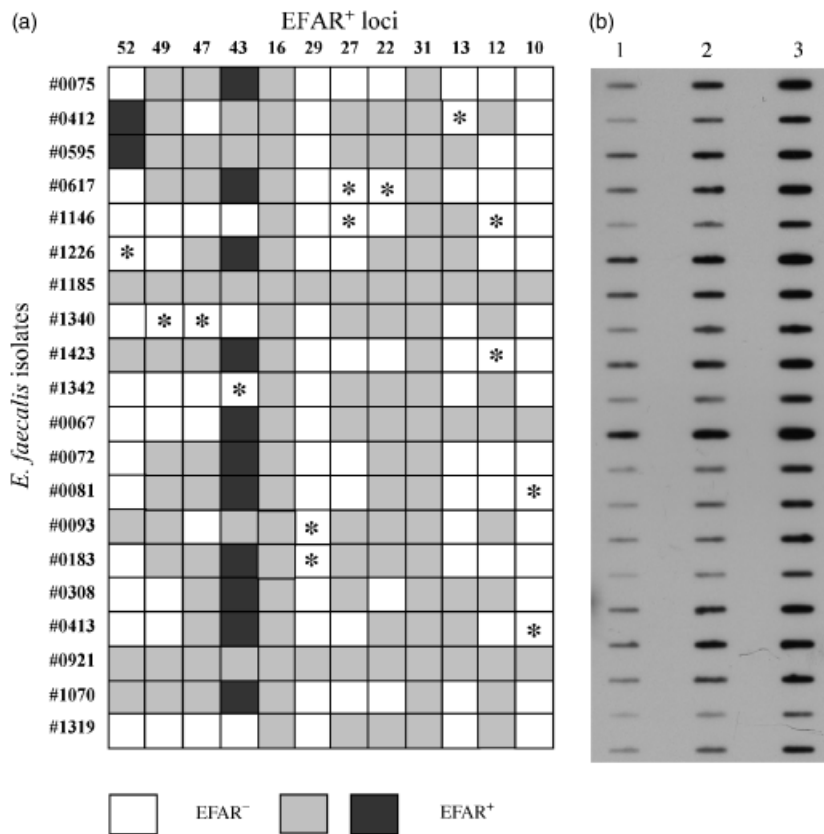


Fig. 3. (a) Analyses of EFAR⁺ loci. The conservation of EFAR sequences among clinical isolates as assessed by PCR is shown. Numbers to the top refer to chromosomal loci defined by unit-length EFARs identified in the V583 strain (see Table 2), numbers to the left denote *Enterococcus faecalis* isolates. White and grey boxes denote empty and filled sites, respectively. Dark grey boxes mark filled sites with a size larger than expected. Asterisks mark amplimers selected for sequence analysis. (b) The genomic abundance of EFAR sequences in the 20 *E. faecalis* isolates analyzed in (a) was evaluated by slot-blot hybridization. 0.25, 0.5 and 1 µg of DNA from each isolate (lanes 1 to 3, respectively) were loaded onto a Hybond filter, and hybridized to a ³²P-radiolabeled EFAR DNA probe (see 'Materials and methods').

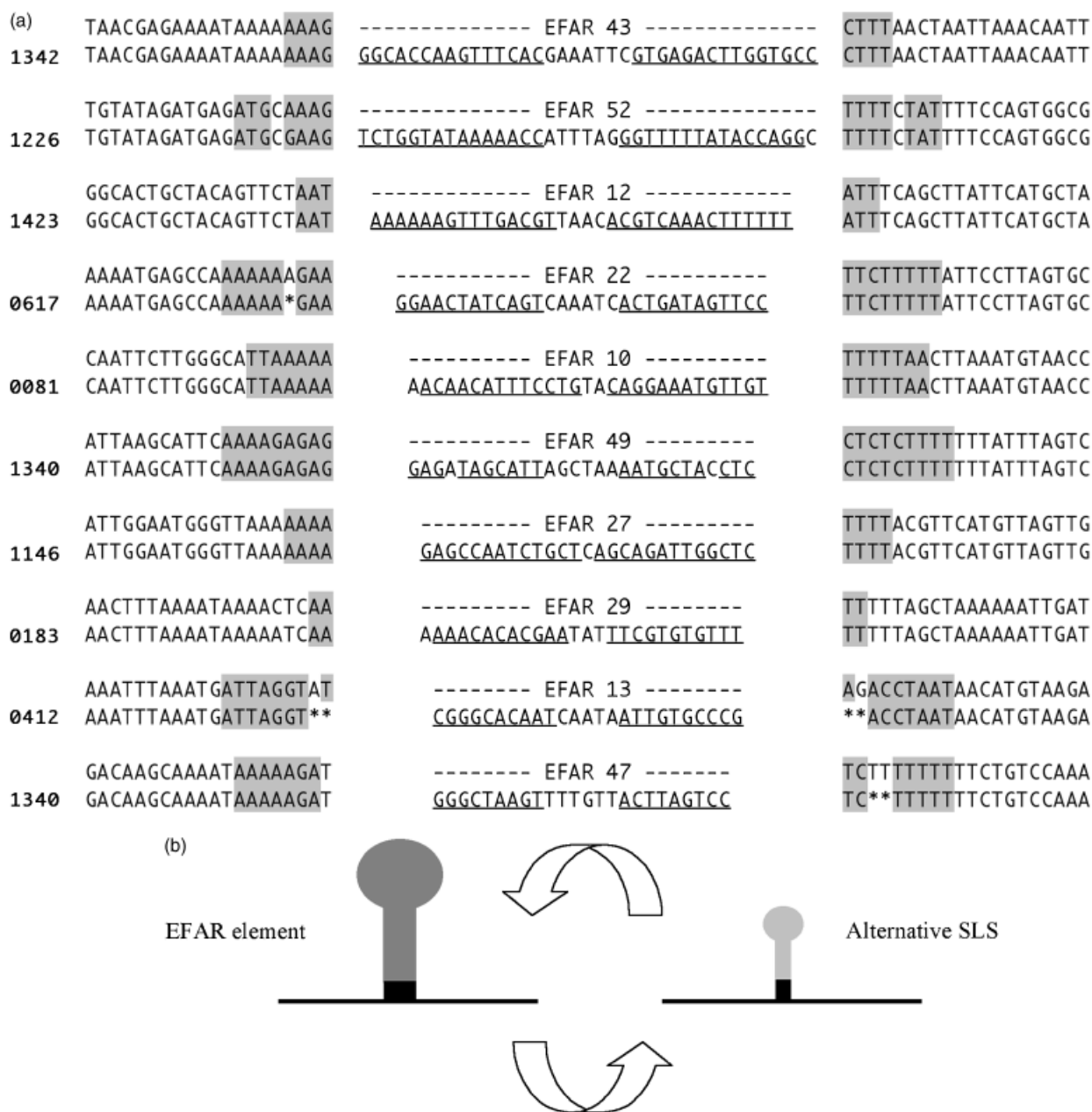


Fig. 4. Analyses of empty EFAR sites. (a) Alignments of homologous DNA regions containing an EFAR repeat in the V583 strain (top sequences) and lacking it in the indicated *Enterococcus faecalis* isolates (bottom sequences). Regions of dyad symmetry are underlined. Asterisks have been introduced to maximize homologies. (b) EFAR and alternative SLSs are sketched not in scale. Black boxes mark regions of base pairing found at the termini of both types of elements, and correspond to the bases highlighted in (a).

EFAR may be interrupted by different types of insertions. All these repeats are inserted in a sequence-specific manner, and most are homologous to each other (Fig. 1c). EFAR insertions seem to have strictly coevolved with EFAR elements, as no homologous sequences were identified outside the mapped EFAR⁺ loci in the V583 genome. However, it cannot be formally ruled out that insertions rather

represent remnants of larger repeats measuring 560–600 bp (see EFAR subfamilies I–II in Fig. 2). Size heterogeneity of EFAR repeats is correlated to the presence/absence of both insertions and modules. Interestingly, most of the clinical isolates analyzed in Fig. 3 exhibited quite distinct different EFAR-PCR patterns. Changes in the distribution of repetitive sequences among bacterial strains are monitored using

PCR primers, which hybridize to the conserved ERIC repeats (Versalovic *et al.*, 1991). A major bias in this type of analyses is data reproducibility, the degeneracy of the primers allowing the detection of amplification patterns also in species lacking ERIC DNA (see Gillings & Holley, 1997). EFAR spread in the genomes of enterococci likely by transposition, and empty and filled homologous chromosomal regions can be distinguished among clinical isolates (Fig. 3). The genomic integration of mobile elements is frequently associated to the generation of target site duplications (TSDs) ranging in size from 2 to 13 bp at the point of insertion. TSDs are not found at the termini of EFAR repeats, and the mechanism of integration of EFAR seems to be indeed rather unusual. The analysis of a representative set of empty sites (Fig. 4a) unequivocally showed that EFAR target sites coincided with 25–40 bp-long DNA regions able to fold into SLSs, which featured AT-rich complementary tracts at their ends. This type of SLS is overrepresented in the genomes of low-GC firmicutes, and may serve multiple functions (Petrillo *et al.*, 2006). Several ISs tend to insert into regions of dyad symmetry (Odaert *et al.*, 1998; Calcutt *et al.*, 1999; Hu *et al.*, 2001; Choi *et al.*, 2003), and rho-independent transcriptional terminator-like sequences are privileged sites of integration for the small (130–170 bp) YPAL repeats in *Yersinia* (De Gregorio *et al.*, 2006). Interestingly, YPAL induce the duplication of 8–25 bp of target sequences, and this results in the formation of long complementary terminal regions (De Gregorio *et al.*, 2006). In contrast, the integration of EFAR was accompanied by the deletion of most of the target. Yet, EFAR were similarly flanked by base-paired residues provided by the targeted SLS (Fig. 4). Complementary termini are crucial for the recognition of SLSs formed by YPAL RNAs by the RNaseIII (De Gregorio *et al.*, 2006), and may plausibly be important for the mobilization of EFAR.

The isolates of *E. faecalis* analyzed in this work feature distinct PFGE patterns (Table 1) and exhibit differences in the organization, or the interspersions of EFAR sequences (Figs 2 and 3). PCR assays similar to those reported in Fig. 2 revealed that isolates with identical PFGE types showed close or identical EFAR profiles (R. Zarrilli, E. De Gregorio and PP. Di Nocera, in preparation). Thus, changes in the structural organization of the EFAR family may be an additional tool to investigate the epidemiology and population structure of *E. faecalis*. The standardization of PCR and electrophoresis conditions should enable different labs to easily obtain validated EFAR-PCR profiles for genotype analysis.

Acknowledgements

M.S. Carlomagno is thanked for critical reading of the manuscript.

References

- Angeletti S, Lorino G, Gherardi G, Battistoni F, De Cesaris M & Dicuozzo G (2001) Routine molecular identification of enterococci by gene-specific PCR and 16S ribosomal DNA sequencing. *J Clin Microbiol* **39**: 794–797.
- Brugger K, Redder P, She Q, Confalonieri F, Zivanovic Y & Garrett RA (2002) Mobile elements in archaeal genomes. *FEMS Microbiol Lett* **206**: 131–141.
- Calcutt MJ, Lavrrar JL & Wise KS (1999) IS1630 of *Mycoplasma fermentans*, a novel IS30-tYPAL insertion element that targets and duplicates inverted repeats of variable length and sequence during insertion. *J Bacteriol* **181**: 7597–7607.
- Carlomagno MS, Chiariotti L, Alifano P, Nappo AG & Bruni CB (1988) Structure and function of the *Salmonella typhimurium* and *Escherichia coli* K-12 histidine operons. *J Mol Biol* **203**: 585–606.
- Choi S, Ohta S & Ohtsubo E (2003) A novel IS element, IS621, of the IS110/IS492 family transposes to a specific site in repetitive extragenic palindromic sequences in *Escherichia coli*. *J Bacteriol* **185**: 4891–4900.
- De Gregorio E, Abrescia C, Carlomagno MS & Di Nocera PP (2003a) Ribonuclease III-mediated processing of specific *Neisseria meningitidis* mRNAs. *Biochem J* **374**: 799–805.
- De Gregorio E, Abrescia C, Carlomagno MS & Di Nocera PP (2003b) Asymmetrical distribution of *Neisseria* miniature insertion sequence DNA repeats among pathogenic and nonpathogenic *Neisseria* strains. *Infect Immun* **71**: 4217–4221.
- De Gregorio E, Silvestro G, Petrillo M, Carlomagno MS & Di Nocera PP (2005) Enterobacterial repetitive intergenic consensus sequence repeats in *Yersinia*: genomic organization and functional properties. *J Bacteriol* **187**: 7945–7954.
- De Gregorio E, Silvestro G, Venditti R, Carlomagno MS & Di Nocera PP (2006) Structural organization and functional properties of miniature DNA insertion sequences in *Yersinia*. *J Bacteriol* **188**: 7876–7884.
- Feil EJ & Spratt BG (2001) Recombination and the population structures of bacterial pathogens. *Annu Rev Microbiol* **55**: 561–590.
- Gillings M & Holley M (1997) Repetitive element PCR fingerprinting (rep-PCR) using enterobacterial repetitive intergenic consensus (ERIC) primers is not necessarily directed at ERIC elements. *Lett Appl Microbiol* **25**: 17–21.
- Hu WY, Thompson W, Lawrence CE & Derbyshire KM (2001) Anatomy of a preferred target site for the bacterial insertion sequence IS903. *J Mol Biol* **306**: 403–416.
- Malani PN, Kauffman CA & Zervos MJ (2002) Enterococcal disease, epidemiology, and treatment. *The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance* (Gillmore MS, Clewell DB, Courvalin PM, Dunny GM, Murray BM & Rice LB, eds), pp. 385–408. ASM Press, Washington, DC.
- Murray BE (2000) Vancomycin-resistant enterococcal infections. *N Engl J Med* **342**: 710–721.
- Nallapareddy SR, Wenxiang H, Weinstock GM & Murray BE (2005) Molecular characterization of a widespread,

- pathogenic, and antibiotic resistance-receptive *Enterococcus faecalis* lineage and dissemination of its putative pathogenicity island. *J Bacteriol* **187**: 5709–5718.
- Odaert M, Devalckenaere A, Trieu-Cuot P & Simonet M (1998) Molecular characterization of IS1541 insertions in the genome of *Yersinia pestis*. *J Bacteriol* **180**: 178–181.
- Oggioni MR & Claverys JP (1999) Repeated extragenic sequences in prokaryotic genomes, a proposal for the origin and dynamics of the RUP element in *Streptococcus pneumoniae*. *Microbiology* **145**: 2647–2653.
- Paulsen IT, Banerjee L, Myers GS *et al.* (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**: 2071–2074.
- Petrillo M, Silvestro G, Di Nocera PP, Boccia A & Paoletta G (2006) Stem-loop structures in prokaryotic genomes. *BMC Genomics* **2006** **7**: 170.
- Ruiz-Garbajosa P, Bonten MJ, Robinson DA *et al.* (2006) Multilocus sequence typing scheme for *Enterococcus faecalis* reveals hospital-adapted genetic complexes in a background of high rates of recombination. *J Clin Microbiol* **44**: 2220–2228.
- Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Versalovic J, Koeuth T & Lupski JR (1991) Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* **19**: 6823–6831.
- Zarrilli R, Tripodi MF, Di Popolo A, Fortunato R, Bagattini M, Crispino M, Florio A, Triassi M & Utili R (2005) Molecular epidemiology of high-level aminoglycoside-resistant enterococci isolated from patients in a university hospital in southern Italy. *J Antimicrob Chemother* **56**: 827–835.