Are point mutations or DNA rearrangements responsible for the restriction fragment length polymorphisms that are used to type bacteria?

Lucinda M. C. Hall

Tel: +44 71 377 7244 ext. 3228. Fax: +44 71 375 0518. e-mail: l.hall@uk.ac.lhmc/l.hall@lhmc.ac.uk

Department of Medical
Microbiology, London
Hospital Medical College,
Turner Street, London
E1 2AD, UK

Restriction fragment length polymorphisms (RFLPs) detected in total DNA or in rRNA genes are widely used to differentiate strains of bacteria. The changes accounting for these polymorphisms and the extent of genomic difference that they reflect are generally unknown. In this report, several methods have been used to examine the DNA differences between nine *Enterococcus faecalis* isolates. Restriction fragments from total DNA and from rRNA genes were compared between isolates using four and five different restriction enzymes, respectively. The proportion of polymorphic fragments detected was greater with total DNA than with rRNA gene patterns, but depended considerably on the restriction enzyme used. DNA changes underlying nine RFLPs were investigated by using the polymorphic fragments as probes to test for alteration in the position of recognition sites of other enzymes. Two polymorphisms were deduced to result from point mutation in a restriction site. Six were judged to result from DNA rearrangements, five of which involved deletion/insertion of the entire probe fragment. The results demonstrate that DNA rearrangements may be responsible for a high proportion of RFLPs used to differentiate and type strains of bacteria. While this does not limit the utility of such methods, it does preclude calculation of overall DNA sequence conservation from similarities in restriction pattern between isolates. DNA sequence determination of the 16S–23S rRNA intergenic spacer of three isolates revealed minimal base substitutions (less than 1%), suggesting that overall sequence divergence between the isolates may be low.

Keywords: *Enterococcus faecalis*, RFLPs, DNA fingerprinting, typing, ribosomal RNA

INTRODUCTION

Numerous studies have demonstrated that genomic DNA restriction fragment patterns differ between different strains of a single bacterial species. Such restriction fragment length polymorphism (RFLP) analysis has been used as a typing method to trace the dissemination of particular strains of bacteria in communities and within and between hospitals (Bingen *et al.*, 1991; Kristiansen *et al.*, 1986; Ott *et al.*, 1991; Blumberg *et al.*, 1991; Struelens *et al.*, 1992; Murray *et al.*, 1991). This approach has also been used to identify functionally significant subgroups within species, as for example the subset of *Haemophilus influenzae* biogroup Aegyptius isolates responsible for

Brazilian purpuric fever (Irino *et al.*, 1988). A variety of formats have been used for genomic DNA fingerprinting or RFLP analysis (Owen, 1989). Methods using direct visualization of stained fragments involve digestion of total DNA with a suitable restriction enzyme, followed by electrophoresis in agarose or acrylamide gels or by pulsed-field gel electrophoresis, each method separating fragments of a different size range. The restriction enzyme used should generate a moderate number (ideally 10–20) of well-resolved DNA fragments within the appropriate size range for the electrophoresis method. Alternatively, a subset of the genome is detected by hybridizing Southern blots with labelled probes. In the most general method, often referred to as ribotyping, rRNA genes are detected (Stull *et al.*, 1988; Grimont & Grimont, 1986).

In this laboratory, DNA fingerprinting has been used to type *Enterococcus faecalis* (Hall *et al.*, 1992a, b). Digestion of total DNA with the restriction enzyme *SstI* followed by
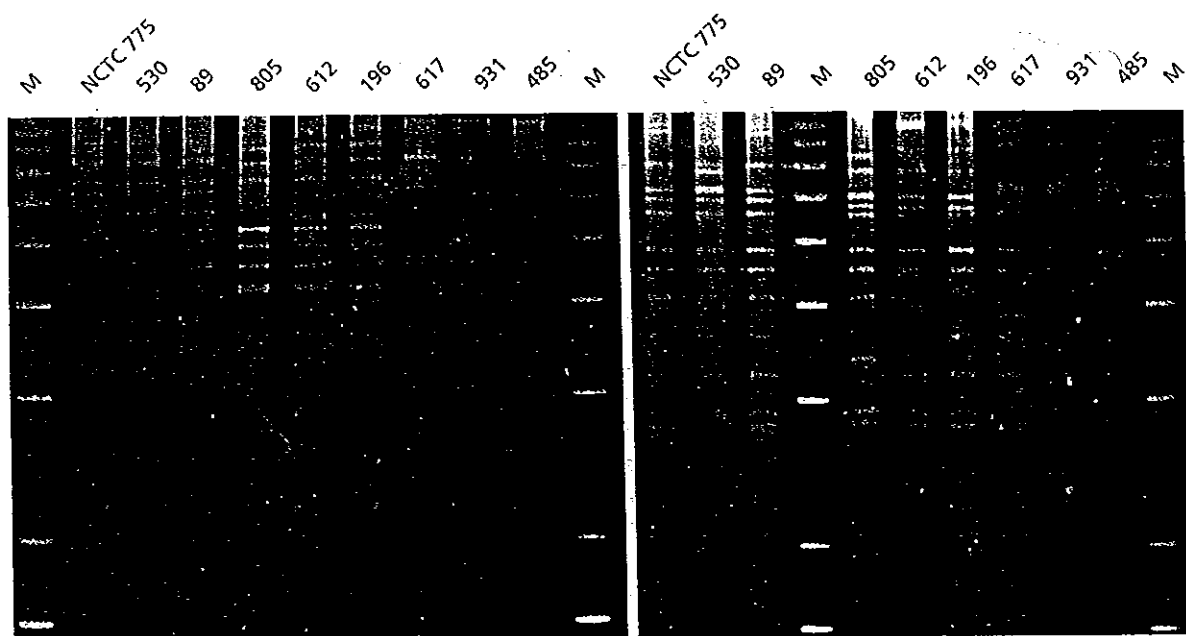
Fig. 1. Total DNA from nine isolates of *E. faecalis* digested with *Sst*I (left panel) and with *Xho*I (right panel). Lanes marked M contain molecular size markers of 1·6, 2, 3, 4, 5, 6, 7, 8, 9 and 10 kb (1 kb ladder; Life Technologies).

separation on conventional agarose gels proved to be highly discriminatory, with, for example, 94 different patterns produced from 135 urinary tract isolates (Hall *et al.*, 1992b). The investigation also highlighted a group of isolates with very similar restriction patterns that was associated with resistance to high-level gentamicin and with *β*-haemolysis. This work has led to further questions about the relationship between isolates with similar restriction patterns: in particular, whether it is valid to relate restriction fragment length conservation with overall DNA sequence conservation in bacteria, as has been done for mitochondrial genomes (Upholt, 1977). The calculation of sequence conservation from fragment patterns described by Upholt (1977) depends, among other factors, on the assumption that fragment changes are due to point mutations rather than DNA rearrangements. It is not clear that this is the case in RFLPs of bacterial DNA.

The present investigation examines nine isolates of *E. faecalis* that fall into three groups of three similar isolates by *Sst*I restriction pattern. Four questions relating to the interpretation of genomic DNA fingerprints are addressed: (1) the extent to which the choice of restriction enzyme alters the discriminatory power of DNA fingerprinting and deduced relationships between isolates; (2) relative discrimination between isolates by total DNA pattern compared with ribotyping; (3) the relative importance of point mutation versus DNA rearrangement in causing restriction fragment length changes; and (4) the extent of DNA sequence divergence between selected isolates in the 16S–23S rRNA intergenic spacer.

METHODS

Bacterial isolates. Eight clinical isolates of *E. faecalis* from urine samples from a collection previously described (Hall *et al.*, 1992b) were used, together with the type strain NCTC 775. The isolates were chosen to represent three different *Sst*I pattern types from each of three groups of related isolates. Isolates 530 and 89 were similar to the type strain. Isolates 805, 612 and 196 were from the previously described group A (associated with a high incidence of antibiotic resistance excluding high-level gentamicin resistance). Isolates 617, 931 and 485 were from the previously described group B (associated with a high incidence of antibiotic resistance, including high-level gentamicin resistance, and with production of haemolysin).

DNA fingerprinting and hybridization. Total DNA was extracted by the method of Pitcher *et al.* (1989), and restriction digests performed and compared exactly as described previously (Hall *et al.*, 1992a). Restriction enzymes used were *Sst*I, *Bgl*II, *Sma*I, *Xho*I and *Sal*I for comparison of total DNA digests, *Hind*III, *Pvu*II, *Eco*RV, *Bst*I and *Eco*RI for ribotyping, and *Sst*I, *Eco*RI and *Hind*III for evaluation of polymorphisms. Extracted DNA was digested with restriction enzyme for at least 4 h, and the fragments separated by electrophoresis in 0·9 % agarose with TBE buffer (0·089 M Tris, 0·089 M boric acid, 0·002 M Na$_2$EDTA) (Sambrook *et al.*, 1989). Plasmid DNA was extracted by alkaline SDS as described by Kado & Liu (1981) and separated by electrophoresis as above. For Southern blots, DNA was transferred to Hybond-N filters (Amersham) using a positive pressure system (Posiblot, Stratagene). DNA was fixed by UV treatment for 2 min. The rRNA probe for ribotyping was labelled, hybridized and detected as described previously (Hall *et al.*, 1992a). DNA fragments to be used as probes were purified from agarose, using Geneclean (Bio 101), and labelled with digoxigenin (DIG DNA labelling and detection kit; Boehringer). Hybridizations were performed at 65 °C and

washed to a final stringency of 0·2 × SSC (1 × SSC is 0·15 M NaCl, 0·015 M sodium citrate) at 65 °C by standard methods. Hybridization was detected by chemiluminescence using AMPPD [4-methoxy-4-(3-phosphate-phenyl)-spiro(1,2-dioxxe-tane-3,2'-adamantane) disodium salt; Boehringer]. Blots were stripped and reused by the method recommended by the supplier.

Polymerase chain reaction and DNA sequence analysis. Primers for amplification of the 16S–23S spacer are related to those described by Barry *et al.* (1991). Primer A was 5′ AGTCGTAACAAGGTAAGCCG 3′, and primer B was 5′ TGCCAAGGCATCCACC 3′. PCR reactions were performed with GeneAmp PCR Core Reagents (Perkin Elmer Cetus) using the standard recommended protocol with 2·5 mM MgCl$_2$. Reaction conditions were 95 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s for 30 cycles, performed on a Hybaid Thermal Reactor. PCR products were rendered blunt-ended by incubation with Klenow polymerase and 0·2 mM dNTPs for 5 min at 37 °C, then recovered by ethanol precipitation. Products were then ligated into the *Sma*I site of M13mp19. DNA sequencing was performed by chain termination using a Sequenase version 2.0 kit and protocol (USB). Sequences were read on a sonic digitizer and processed with PC/gene software (Intelligenetics).

RESULTS

Restriction patterns of total DNA generated with different enzymes

Eight clinical isolates and the type strain of *E. faecalis* were selected for analysis on the basis of their relationships already determined by *Sst*I digestion. All isolates were distinguishable, but they formed three groups of three related isolates (Fig. 1). Groups were defined previously as comprising isolates with > 50% of fragments in common, but in practice the isolates used here have > 80% of *Sst*I fragments in common. Isolates from different groups had less than 50% of fragments in common (only fragments of 1·6–8 kb in length were considered).

DNA samples from the isolates were all digested with four other restriction enzymes, chosen because the fragments generated were well resolved on a conventional agarose gel (example in Fig. 1). For each enzyme, fragments within a defined size range were compared between all isolates, and the isolates placed in types (Table 1). A single band difference was used as the criterion to define a different type. The size-range was chosen to contain about 10–20 bands. However, *Sma*I generated only 3–5 bands in the usable size range and was thus excluded from further analysis. The similarity of fragment pattern between pairs of isolates was calculated by Dice coefficient, i.e. 2 × the number of shared bands divided by the total number of bands (Dice, 1945) (Table 1).

The restriction enzymes used varied in their ability to discriminate between isolates (Table 1). Only *Bgl*II shared with *Sst*I the ability to discriminate all isolates. The other enzymes, *Sal*I and *Xho*I, failed to discriminate between *Sst*I group B isolates. These enzymes also generated identical patterns for some other pairs of isolates, but the pairs comprised different isolates for the two enzymes.

The extent of similarity between isolates also varied with the restriction enzyme used (Table 1). Isolates within *Sst*I groups had highly related restriction patterns for all enzymes (over 70% similar). However, the similarity between isolates from different groups ranged from 24–45% for *Sst*I to 72–88% for *Sal*I. For all four enzymes, *Sst*I, *Bgl*II, *Xho*I and *Sal*I, intra-group pattern similarity was higher than inter-group similarity: the difference was greatest with *Sst*I.

Restriction patterns of ribosomal genes generated with different enzymes

Restriction enzymes used for ribotyping were chosen because they digested total DNA to fragments mostly in the range of 1–20 kb. Again, isolates with patterns differing by one fragment or more were placed in different types (Table 2). Discrimination by ribotyping was lower overall than by total DNA analysis. None of the enzymes used could discriminate between isolates within *Sst*I group B, and only *Eco*R1 differentiated some of group A from group B. The number of types obtained ranged from one with *Hind*III to five with *Eco*R1. Dice coefficients were not calculated as the number of bands was generally low.

DNA probing to investigate RFLPs

Polymorphism in the length of an *Sst*I restriction fragment between isolates of *E. faecalis* could be due to a point mutation, causing loss or gain of an *Sst*I recognition site,

Table 1. Discrimination between *E. faecalis* isolates from total DNA pattern with different restriction enzymes

Types are distinguished by one or more band differences. Similarity between two isolates is calculated as (2 × no. shared bands)/(total no. bands).

Isolate	Group	Type			
		*Sst*I	*Bgl*II	*Xho*I	*Sal*I
NCTC 775	1	I	I	I	I
530	1	II	II	II	I
89	1	III	III	I	I
805	A	IV	IV	III	II
612	A	V	V	IV	III
196	A	VI	VI	IV	IV
617	B	VII	VII	V	V
931	B	VIII	VIII	V	V
485	B	IX	IX	V	V
Range similarity within groups (%)		83–96	73–92	72–100	93–100
Range similarity between groups (%)		24–45	48–69	62–69	72–88
Size range of fragments scored (kb)		1·6–8	1·6–6	1·6–7	1·6–3
No. of fragments per isolate		11–15	17–26	12–16	12–15

199

Table 2. Discrimination between isolates by ribotyping with different restriction enzymes

Types are distinguished by one or more band differences.

Isolate	Group	Type				
		HindIII	PvuII	EcoRV	BscI	EcoRI
NCTC 775	1	I	I	I	I	I
530	1	I	II	II	II	II
89	1	I	I	I	III	III
805	A	I	I	II	IV	IV
612	A	I	I	II	IV	V
196	A	I	I	II	IV	V
617	B	I	I	II	IV	IV
931	B	I	I	II	IV	IV
485	B	I	I	II	IV	IV

or to a DNA rearrangement such as an insertion, deletion or inversion. The causes of polymorphisms were investigated by the following scheme. Five *Ss*I fragments from NCTC 775 and four fragments from isolate 617 were isolated from gels and labelled. All nine fragments were polymorphic in that a fragment of the same length was absent in the DNA of some other isolates. Each was hybridized in turn to panels of DNA from the nine isolates under investigation digested with *Ss*I, HindIII and *Eco*RI. The majority of fragments generated by the latter two enzymes are in the range 1–20 kb.

The resulting fragments detected are tabulated for NCTC 775, isolate 805 and isolate 617, representing the three groups of related isolates (Table 3). In most cases, other isolates in the same group had the same result and occasional exceptions do not alter the interpretation. Two probes, 1 and 3, detected no polymorphism in *Eco*RI or HindIII digests. This indicates that the *Ss*I polymorphism is caused by a point mutation, as other enzyme sites are not affected. Probe 9 detected polymorphism of both *Eco*RI and HindIII fragments between isolate 617, which has a 4·3 kb *Ss*I fragment, and NCTC 775, which does not. As the sites for three different restriction enzymes are affected simultaneously, it is probable that a DNA rearrangement has occurred within this fragment. (There is also an *Eco*RI polymorphism between isolates 617 and 805, though *Ss*I and HindIII fragments are both conserved.) Probe 8 detected polymorphism of HindIII fragments but not of *Eco*RI fragments; no firm conclusion could be drawn as to the change underlying this *Ss*I polymorphism. For five probes (2, 4, 5, 6 and 7) no complementary sequence was detected in isolates that did not share an *Ss*I fragment of the same length. Thus the apparent *Ss*I polymorphism is due to a deletion or insertion event involving the entire *Ss*I fragment (or enough of it to prevent detectable hybridization). (With probes 4 and 5, the two isolates that have an *Ss*I fragment of the same length nonetheless show polymorphic fragments with one or both of the other restriction enzymes.) DNA fragments present in some isolates but not others could derive from plasmids rather than genomic DNA. This possibility was tested by hybridizing the probes to preparations of plasmid DNA from the isolates they were

Table 3. Results of hybridizing total DNA digests of three isolates with polymorphic SstI fragments as probes

Columns S, R and H list the size of fragment (kb) detected by each probe in digests with *Ss*I, *Eco*RI and HindIII, respectively. Probes 1–5 are *Ss*I fragments of the given size from NCTC 775, probes 6–9 are *Ss*I fragments of the given size from isolate 617 (probe fragments shown in bold). Dashes are shown where no band hybridized (in *Ss*I digests this may also correspond to hybridization to a large fragment not resolved on the gel).

Probe	NCTC 775 Enzyme digest			Isolate 805 Enzyme digest			Isolate 617 Enzyme digest		
	S	R	H	S	R	H	S	R	H
1	2·5	8·2	3·9	2·5	8·2	3·9	11·0	8·2	3·9
2	5·2	4·7	5·0	5·2	4·7	5·0	–	–	–
		4·3	2·4		4·3	2·4			
3	5·5	11·0	8·0	5·2	11·0	8·0	5·5	11·0	8·0
4	2·3	5·1	1·75	–	–	–	2·3	6·0	1·75
5	2·2	5·0	5·0	2·2	4·3	5·9	–	–	–
6	–	–	–	–	–	–	2·9	1·5	3·2
								1·3	
7	–	–	–	–	–	–	3·1	8·5	6·6
8	–	2·6	8·0	–	2·6	4·4	3·8	2·6	4·1
			2·2			4·1			2·6
9	–	10·0	2·3	4·3	6·0	2·3	4·3	6·1	2·3
		6·2	1·6		4·5	2·1		2·8	2·1
						1·65			1·65

NCTC 775 | M | bp
1636
1018
517, 506
396
344
298

NCTC 775 | M | bp
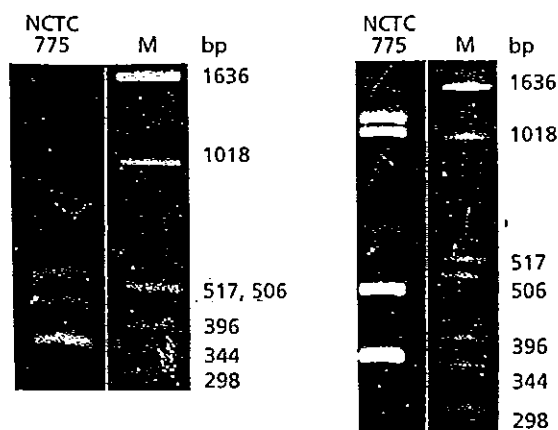1636
1018
517
506
396
344
298

Fig. 2. PCR products of amplification of NCTC 775 DNA with primers A and B run on 1% agarose (left panel) or 4% polyacrylamide gels (right panel). M, size markers.

Sequence of the 16S–23S rRNA intergenic spacer from different strains

To gain a measure of DNA sequence divergence between isolates, the 16S–23S rRNA intergenic spacer was analysed. This was considered to approach the ideal of a selectively neutral DNA sequence that was yet equivalent in all isolates. Primers homologous with conserved sequences at the 3' end of the 16S and the 5' end of the 23S coding regions were used to amplify the spacer region by PCR. Two major bands migrating at approximately 350 and 450 bp were obtained, plus a band which migrated at about 550 bp on agarose gels but resolved to two bands of 1·0 and 1·1 kb on acrylamide gels (Fig. 2). The differential migration indicates that these more slowly migrating bands are not normal double-stranded DNA. The relative intensities of all bands remained constant when annealing temperature or magnesium ion concentration were varied, or DMSO added, indicating that none of the bands was produced by non-specific binding of primers.

The PCR products were cloned into m13 vectors by blunt-end ligation. Sequences corresponding to the 350 bp PCR fragment were determined from two clones of NCTC 775 (type strain), and three clones each of isolates 805 and 617. Sequences corresponding to the 450 bp PCR fragment were determined from two clones of NCTC 775 and one clone each of isolates 805 and 617. (Direct sequencing of gel-purified PCR fragments confirmed which fragments the clones corresponded to.)

derived from. No hybridization to plasmids could be detected.

In summary, out of nine SsrI RFLPs investigated, only two were due to point mutations in a restriction site, while six were due to DNA rearrangements, including deletion/insertion, and one remained unassigned.

```
                              16S rRNA gene
short   AGTCGTAACAAGGTAAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT    50
long    AGTCGTAACAAGGTAAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCTT    50
        primer A          >

        3'>
short   TCTAAGGAATATTACGGAAATACACATTTCGTCTTTACTTTGTTCAGTTT   100
long    TCTAAGGAATATTACGGAAATACACATTTCGTCTTTACTTTGTTCAGTTT   100

short   TGAGAGGTCTACTCTCAA------------------------------   118
long    TGAGAGGTCTACTCTCAATCAAATAARTTCGGGGCCTTAGCTCRGCTGGG   150
S. pn.     tRNAAla                GGGGCCTTAGCTCAGCTGGG    20

short   ------------------------------------------------   118
long    AGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCCGCTAGGCT   200
tRNA    AGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCCGCTAGGCT    70

short   ------------------AACATTCGTTCATTGAAAACTGGATATTGA   148
long    CCATTGATAGCTTTTGCTATCAGATTCGTTCATTGAAMACTGGATATTGA   250
tRNA    CCA---------------------------------------------    73

short   AGTAAAAAGAATCAAAACAAACCGAGAACACCGCGTTGAATGAGTTTTTT   198
long    AGTAAAAAGAATCAARACAAACCGAGAACACCGCGTTGAATGAGTTTTTT   300

short   AATAAGTTCAATTGCTTATTTATTGATTAACCTTCTATCGCTAGAAGAAG   248
long    AATAAGTTCAATTGCTTATTTATTGATTAACCTTCTATCGCTAGAAGAAG   350

                              < 5' 23S rRNA gene
short   TGATCAAGACCCAACCGTAAGGTTGATAAGGTTAAGTGAATAAGGGCGCA   298
long    TGATCAAGACCCAACCGTAAGGTTGATAAGGTTAAGTGAATAAGGGCGCA   400

short   CGGTGGATGCCTTGGCA   315
long    CGGTGGATGCCTTGGCA   417
        < primer B
```

Fig. 3. Alignment of DNA sequences obtained from two forms of the 16S–23S rRNA spacer region of *E. faecalis* and the published tRNA_Ala gene of *S. pneumoniae* (*S. pn.*) (Bacot & Reeves, 1991). Sequences corresponding to the two primers are underlined. Sequences from the 16S and 23S rRNA genes are in bold, the ends of the genes being deduced from comparison with the genes of other eubacteria in the EMBL database. Ambiguities in the long spacer are: R at 127, which is G in 775 clones 1 and 2, and is A in 617 and 805; R at 144, which is G in 775 clone 2, and is A in 775 clone 1, 617 and 805; M at 238, which is A in 775 clones 1 and 2 and 805, and is C in 617; R at 266, which is A in 775 clones 1 and 2 and 805, and is G in 617.

The shorter PCR fragment is 315 bp long (Fig. 3). Sequences in the 16S and 23S coding regions are recognizable by homology with other bacteria. The likely 3' end of 16S and 5' end of 23S were deduced from this homology. The intergenic spacer comprises 225 bp of DNA with a high A + T base composition. No differences were observed in the sequences of eight independent clones from three different isolates. The longer PCR fragment is 417 bp long and contains an intergenic spacer of 327 bp (Fig. 3). The two clones of NCTC 775 contain one sequence heterogeneity. The clone of isolate 617 differs from NCTC 775 at three positions, and the clone of isolate 805 differs from NCTC 775 at one position, but has the same sequence as 617 at that position (Fig. 3).

The longer spacer differs from the shorter in having an extra segment of 102 bp. At the junction with the extra segment there are two base differences between the spacers, but the remainder of the spacer sequence is identical. The extra segment includes a $tRNA_{Ala}$ gene of 73 bp which is identical to the analogous gene in the 16S–23S intergenic spacer in *Streptococcus pneumoniae* (Bacot & Reeves, 1991) (Fig. 3).

To address the possibility that the smaller fragment was generated by an artefact of PCR amplification in which the extra segment was looped out, two further experiments were performed. Reamplification of the purified 417 bp fragment produced an intense band of the same size, with no shorter band. Hybridization of the short PCR fragment to total DNA (from NCTC 775) digested with *Eco*RI detected two bands of 2·1 and 2·2 kb. This confirms that two forms of rRNA operon are present in the genomic DNA (there is no *Eco*RI site in the spacer). The composition of the additional PCR products observed at higher M_r values was not confirmed, but they probably represent heterodimers between the two major bands.

DISCUSSION

This study analyses differences in the DNA of nine isolates of *E. faecalis* by a variety of approaches. Comparison of total DNA fingerprints reveals that the proportion of fragments shared between isolates varies depending on the restriction enzyme used. In other words, the level of discrimination varies. Nevertheless, placing of the isolates into three groups on the basis of *Sst*I pattern was substantiated by results with three other enzymes, in that similarity of pattern between isolates within groups was always higher than between groups. Ribotyping was less discriminatory than total DNA fingerprinting. Relationships between isolates deduced from ribotyping were only partly consistent with *Sst*I groups: similarity within and between groups A and B was greater than similarity within group 1.

The DNA changes underlying nine *Sst*I fragment polymorphisms were studied. To differentiate between point mutation and DNA rearrangement, the fragments were used to probe the size of overlapping fragments generated by other restriction enzymes in isolates with or without an *Sst*I fragment of the same size. A point mutation in an *Sst*I site will not alter the sites of other restriction enzymes. A DNA rearrangement will change the distance between sites for all enzymes. It was found that two of the nine polymorphisms could be interpreted as caused by point mutations in *Sst*I sites, while six were caused by DNA rearrangements and one was inconclusive. Of the six rearrangements, five comprised deletion/insertion of the entire *Sst*I fragment.

Possible sources of error in interpreting these results were considered. When *Eco*RI and *Hind*III fragments are unaltered, it is possible that a rearrangement could coincidentally restore a fragment of the same length, but this seems very unlikely for fragments generated by two different enzymes. A small inversion changing an *Sst*I fragment length could occur within both *Eco*RI and *Hind*III fragments without changing their length, but again the probability seems likely to be low. It is also possible that loss of the *Sst*I site could be due to a change in methylation that would not affect sites for other enzymes: such a change could result from generation of a new modification site overlapping the restriction site – by point mutation – or from alteration of the modification enzymes. The latter would be expected to result in a gross change in the frequency of sites for any affected restriction enzyme and is thus unlikely to account for any of the cases in this study.

When *Eco*RI and *Hind*III fragments detected by a polymorphic *Sst*I fragment do differ between isolates, it is possible that independent point mutations could coincidentally have affected these sites, rather than the change being due to DNA rearrangement. From the results obtained with five restriction enzymes on total DNA, it is found that 58–64 % (mean 60 %) of all fragments compared are shared between representative isolates of different *Sst*I groups. Thus the probability of any fragment being polymorphic would be approximately 0·4, and of any two fragments both being polymorphic would be approximately 0·16. Observation, with three probes, of polymorphism in *Eco*RI or *Hind*III fragments between isolates that do not differ in the corresponding *Sst*I fragment confirms that there is a significant level of independent polymorphisms. Hence, if most polymorphisms were due to point mutation, there would remain an estimated 0·16 probability that point mutations could coincidentally affect both *Eco*RI and *Hind*III fragments detected with a polymorphic *Sst*I fragment, and so lead to misinterpretation of a point mutation as a DNA rearrangement. However, in the present study five of the six polymorphisms considered to be due to DNA rearrangement involve deletion/insertion of a sequence including the whole *Sst*I fragment (or enough of it to prevent detectable hybridization), and are thus not subject to this source of error.

An independent approach to comparing the DNA of different isolates was to sequence the 16S–23S rRNA intergenic spacer. The spacer is considered to be largely non-functional, and therefore not under selective pressure to maintain its sequence. On the other hand, this region is equivalent in all species. It is known that the DNA

sequence of the 16S–23S spacer differs extensively between species of *Clostridium* (Barry *et al.*, 1991) and of *Mycoplasma* (Uemori *et al.*, 1992), and intraspecies variation in spacer length in *Pseudomonas cepacia* has also been observed (Kostman *et al.*, 1992). Further studies will be required to determine how spacer sequence divergence relates to sequence divergence elsewhere in the genome, but it is postulated that this sequence might be useful as an indicator of overall homology. Results in this study showed there to be two forms of the spacer in each isolate, with and without a tRNA gene. Sequences of the shorter spacer from three isolates in different *Ss*1 groups were identical, and from the longer spacer had a maximum of three base changes in 327 bp. Sequence differences could be contributed by different copies of the spacer within the genome, as well as differences between isolates, and also possible PCR artefacts. However, it appears that the extent of base substitution in the spacer region between isolates is under 1 %.

The results of this study lead to the conclusion that differences in *Ss*1 fragment pattern between different isolates of *E. faecalis* are much more frequently due to DNA rearrangements than to point mutations. The rearrangements are likely to include excision and insertion of various transposable elements, insertion sequences, or bacteriophage genomes, although, interestingly, none of the probes appeared to include repeated sequences (hybridization to other *Ss*1 fragments was not observed). The enzyme *Ss*1 was initially selected for typing studies because of the small number of bands produced in the selected size range, arising from the fact that it cuts less frequently than would be expected in a random sequence of the same G + C content. It also produced restriction patterns that were particularly discriminating between isolates. It is possible that use of this enzyme specifically selected for detection of mobile DNA elements. For example restriction–modification systems in this species could result in selection against or modification of *Ss*1 recognition sites (GAGCTC), leaving a very low frequency of sites within essential regions of the genome (analogous to the low frequency of CG dinucleotides in mammalian DNA). Thus the majority of sites might be in DNA segments introduced recently into the genome, such as mobile elements. Intriguingly, a restriction enzyme *Sdu*1, with a specificity that would include all *Ss*1 sites (G[A/G/T]GC[A/C/T]C), has been reported from an isolate of *Enterococcus durans* (Janulaitis *et al.*, 1981). On the other hand, use of pulsed-field gel electrophoresis as an alternative approach to fingerprinting would also maximize detection of DNA deletion/insertion events (above a certain size) because fragment lengths of a large proportion of the genome are compared.

The finding that much of the polymorphism between restriction patterns is due to DNA rearrangements rather than point mutations precludes the use of Upholt's equations to extrapolate pattern similarity to overall sequence similarity between isolates. It indicates that the frequency of base substitution would be lower than such a calculation would predict. This also concords with the results of 16S–23S rRNA spacer sequence analysis that

demonstrate very limited sequence divergence between the isolates studied (although more work is needed to determine how sequence divergence in the rRNA spacer relates to divergence elsewhere in the genome).

It is not known how generally the results of this study may apply amongst other species of bacteria. However, it seems likely that DNA rearrangements, probably involving mobile elements of various kinds, could contribute significantly to RFLPs in many species. Thus, deductions of sequence similarity from pattern similarity are unlikely to be valid. Indeed, detailed investigations of the evolution of sections of the *Escherichia coli* genome demonstrate some large-scale DNA rearrangements due to prophages and insertion sequences among natural isolates of this species (Stoltzfus *et al.*, 1988; Milkman & Bridges, 1990). It is of interest to speculate whether transposition and recombination events are much less frequent in species in which DNA fingerprinting has been less useful in discriminating between isolates such as *Mycobacterium leprae* (Williams *et al.*, 1990).

ACKNOWLEDGEMENTS

REFERENCES

Bacot, C. M. & Reeves, R. H. (1991). Novel tRNA gene organization in the 16S-23S intergenic spacer of the *Streptococcus pneumoniae* rRNA gene cluster. *J Bacteriol* 173, 4234–4236.

Barry, T., Colleran, G., Glennon, M., Dunican, L. K. & Gannon, F (1991). The 16S/23S ribosomal spacer region as a target for DNA probes to identify eubacteria. *PCR Methods Applic* 1, 51–56.

Bingen, E. H., Denamur, E., Lambert-Zechovsky, N. Y., Bourdois A., Mariani-Kurkdjian, P., Cezard, J.-P., Navarro, J. & Elion, J (1991). DNA restriction fragment length polymorphism differentiates crossed from independent infections in nosocomia *Xanthomonas maltophilia* bacteremia. *J Clin Microbiol* 29, 1348–1350

Blumberg, H. M., Kiehlbauch, J. A. & Wachsmuth, I. K. (1991) Molecular epidemiology of *Yersinia enterocolitica* O:3 infections: use of chromosomal DNA restriction fragment length polymorphism of rRNA genes. *J Clin Microbiol* 29, 2368–2374.

Dice, L. R. (1945). Measures of the amount of ecological association between species. *Ecology* 26, 297–302.

Grimont, F. & Grimont, P. A. D. (1986). Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. *Ann Ins Pasteur Microbiol* 137B, 165–175.

Hall, L. M. C., Duke, B., Guiney, M. & Williams, R. (1992a). Typing of *Enterococcus* species by DNA restriction fragment analysis. *J Cli Microbiol* 30, 915–919.

Hall, L. M. C., Duke, B., Urwin, G. & Guiney, M. (1992b) Epidemiology of *Enterococcus faecalis* urinary tract infection in teaching hospital in London, United Kingdom. *J Clin Microbiol* 30 1953–1957.

Irino, K., Grimont, F., Casin, I., Grimont, P. A. D. & The Braziliar Purpuric Fever Study Group. (1988). rRNA gene restriction patterns of *Haemophilus influenzae* biogroup aegyptius strains associated with Brazilian purpuric fever. *J Clin Microbiol* 26, 1535–1538

Janulaitis, A. A., Marcinkeviciene, L., Petrusyte, M. & Mironov, A. (1981). A new sequence-specific endonuclease from *Streptococcus durans*. *FEBS Lett* 134, 172–174.

Kado, C. I. & Liu, S. T. (1981). Rapid procedure for detection and isolation of large and small plasmids. *J Bacteriol* 145, 1365–1373.

Kostman, J. R., Edlind, T. D., Lipuma, J. J. & Stull, T. L. (1992). Molecular epidemiology of *Pseudomonas cepacia* determined by polymerase chain reaction ribotyping. *J Clin Microbiol* 30, 2084–2087.

Kristiansen, B.-E., Sorensen, B., Bjorvatn, B., Falk, E. S., Fosse, E., Bryn, K., Froholm, L. O., Gaustad, P. & Bovre, K. (1986). An outbreak of group B meningococcal disease: tracing the causative strain of *Neisseria meningitidis* by DNA fingerprinting. *J Clin Microbiol* 23, 764–767.

Milkman, R. & Bridges, M. M. (1990). Molecular evolution of the *Escherichia coli* chromosome. III. Clonal frames. *Genetics* 126, 505–517.

Murray, B. E., Singh, K. V., Markowitz, S. M., Lopardo, H. A., Patterson, J. E., Zervos, M. J., Rubeglio, E., Eliopoulos, G. M., Rice, L. B., Goldstein, F. W., Jenkins, S. G., Caputo, G. M., Nasnas, R., Moore, L. S., Wong, E. S. & Weinstock, G. (1991). Evidence for clonal spread of a single strain of β-lactamase-producing *Enterococcus* (*Streptococcus*) *faecalis* to six hospitals in five states. *J Infect Dis* 163, 780–785.

Ott, M., Bender, L., Marre, R. & Hacker, J. (1991). Pulsed field electrophoresis of genomic restriction fragments for the detection of nosocomial *Legionella pneumophila* in hospital water supplies. *J Clin Microbiol* 29, 813–815.

Owen, R. J. (1989). Chromosomal DNA fingerprinting – a new method of species and strain identification applicable to microbial pathogens. *J Med Microbiol* 30, 89–99.

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

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Stoltzfus, A., Leslie, J. F. & Milkman, R. (1988). Molecular evolution of the *Escherichia coli* chromosome. I. Analysis of structure and natural variation in a previously uncharacterized region between *trp* and *tonB*. *Genetics* 120, 345–358.

Struelens, M. J., Deplano, A., Godard, C., Maes, N. & Serruys, E. (1992). Epidemiologic typing and delineation of genetic relatedness of methicillin-resistant *Staphylococcus aureus* by macrorestriction analysis of genomic DNA by using pulsed-field gel electrophoresis. *J Clin Microbiol* 30, 2599–2605.

Stull, T. L., Lipuma, J. J. & Edlind, T. D. (1988). A broad spectrum probe for molecular epidemiology of bacteria: ribosomal RNA. *J Infect Dis* 157, 280–286.

Uemori, T., Asada, K., Kato, I. & Harasawa, R. (1992). Amplification of the 16S-23S spacer region in rRNA operons of Mycoplasmas by the polymerase chain reaction. *Syst Appl Microbiol* 15, 181–186.

Upholt, W. B. (1977). Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. *Nucleic Acids Res* 4, 1257–1265.

Williams, D. L., Gillis, T. P. & Portaels, F. (1990). Geographically distinct isolates of *Mycobacterium leprae* exhibit no genotypic diversity by restriction fragment-length polymorphism analysis. *Mol Microbiol* 4, 1653–1659.