JOURNAL OF CLINICAL MICROBIOLOGY, Aug. 2002, p. 2795–2800 Vol. 40, No. 8
0095-1137/02/$04.00+0 DOI: 10.1128/JCM.40.8.2795–2800.2002

Fluorescent Amplified Fragment Length Polymorphism Probabilistic Database for Identification of Bacterial Isolates from Urinary Tract Infections

Yankuba Kassama,[1] Paul J. Rooney,[2]† and Royston Goodacre[1]*

Institute of Biological Sciences, University of Wales, Aberystwyth, Ceredigion, SY23 3DD,[1] and Bronglais General Hospital, Aberystwyth, Ceredigion, SY23 1ER,[2] United Kingdom

Received 4 February 2002/Returned for modification 24 March 2002/Accepted 12 May 2002

The ability of the fluorescent amplified fragment length polymorphism (FAFLP) technique to identify bacterial isolates from urinary tract infections (UTIs) was investigated. FAFLP was carried out using the single primer combination *Mse*I plus CT and *Eco*RI plus 0, and information-rich FAFLP profiles were generated from all 69 UTI isolates studied, which comprised both gram-negative and gram-positive bacteria encompassing eight genera. The genetic relatedness of these 69 bacteria was determined by cluster analysis, and this revealed eight main groups corresponding to the eight bacterial genera. Finer discrimination on the same dendrogram showed species and subspecies differentiations, thus demonstrating the potential of FAFLP for describing a wide diversity range within microbial populations. The interpretation of FAFLP profiles is often complicated because it relies upon the investigator interpreting dendrograms; this process may be subjective if the tree is complicated, particularly if it includes polytomies (unresolved nodes). Therefore, we have developed a method based on Bayes' theorem for the identification of bacteria against an FAFLP probabilistic identification matrix. Thus, FAFLP is suitable for the objective identification of causal agents of UTI, and the procedure offers great potential in the clinical laboratory.

Urinary tract infections (UTIs) are of significant clinical concern worldwide. For example, the annual incidence in England and Wales is about 2.5 million, with the family doctor consultation rate for women with UTI being 63.5 consultations/1,000 women/year (37). The fact that UTIs mainly afflict women is mirrored in the United States, where 11.3 million women had at least one presumed UTI treated with antibiotics in 1995 (10). Nosocomial UTI is also common (3), as is incidence in the elderly and children, with 2% of boys and 8% of girls showing clinical symptoms by the age of 7 (6). The bacteria typically associated with UTI are members of the family *Enterobacteriaceae*, predominantly *Escherichia coli* (the causative organism of >50% of all cases) and *Klebsiella* spp., while gram-positive bacteria, in particular enterococci, are also a significant problem (29).

Despite such clinical importance, there are considerable variations in the approaches of laboratories and physicians towards UTIs. Routine diagnosis includes a quantitative microbial count of urine. Counts of $>10^5$ organisms/ml are regarded as showing significant bacteriuria (25). Diagnosis is followed by characterization of the causal agent and sometimes antibiotic sensitivity tests to determine appropriate courses of treatment, because the empirical choice of an effective treatment is becoming more difficult as urinary pathogens become increasingly resistant to commonly used antibiotics (15).

Further identification of the isolates may not be undertaken except in complex and acute cases (27), because conventional methods are expensive, time-consuming, and labor-intensive (11). However, recurrent infections are common, particularly in children, and these may lead to further complications, such as renal scarring and hypertension (26). Typing of urine isolates from recurrent cases will distinguish between reinfection with a new causative organism and relapses perhaps due to treatment failures (21). Furthermore, the dynamics of infection within the population can be established.

For routine purposes, the ideal method for UTI characterization would be rapid and automatable and would permit the probabilistic identification of an unknown causal agent against a stable database. Recently, there has been great interest in the amplified fragment length polymorphism (AFLP) technique (35) for the genetic fingerprinting of microorganisms (16, 22, 28).

In the fluorescent AFLP (FAFLP) method, genomic DNA is digested with two different restriction enzymes (one is a frequent cutter, e.g., *Mse*I with a 4-bp recognition site, while the other is a more infrequent cutter, e.g., *Eco*RI recognizing a 6-bp sequence), creating fragments of suitable size to be resolved by polyacrylamide gel electrophoresis. Oligonucleotide linkers (12 to 20 bp long) are then ligated onto the cohesive ends of the resulting restriction fragments, and these serve as primer-binding sites for PCR. The PCR product is fluorescently labeled, and the resultant PCR products are separated and visualized on a polyacrylamide gel by laser excitation. The primers recognize the linker and restriction site and usually contain selective nucleotides to reduce the number of products amplified by PCR and therefore detected.

AFLP has been used to type strains belonging to a wide range of genera, including *Bacillus anthracis* (19), *Campylobacter fetus* (36), *Corynebacterium diphtheriae* (8), *E. coli*

* Corresponding author. Mailing address: Institute of Biological Sciences, University of Wales, Aberystwyth, Ceredigion, SY23 3DD, United Kingdom. Phone: 44 (0)1970 621947. Fax: 44 (0)1970 621947. E-mail: rrg@aber.ac.uk.

† Present address: Belfast City Hospital, Belfast, BT9 7AB, United Kingdom.

O157 (30), *Neisseria meningitidis* (12), *Salmonella enterica* (7), and *Staphylococcus aureus* (14). In order to develop the AFLP technique for each of these different groups, various restriction enzymes are investigated and primers with selective nucleotides are optimized for each bacterial species, often with recourse to whole-genome sequences for in silico AFLP predictions (2, 13).

By contrast to the reports mentioned above, the present study employed FAFLP using a dual restriction digest and a single primer combination set, optimized to characterize UTI isolates belonging to eight different genera. Identification of isolates from patients with bacteriuria was then possible by matching them against a probabilistic database generated for UTI using Bayes' theorem.

MATERIALS AND METHODS

Bacterial strains and growth conditions. Sixty-nine clinical bacterial isolates from patients with UTI were obtained from the local Bronglais hospital. All bacteria were collected from midstream urine samples. All isolates were typed by conventional biochemical tests as belonging to *E. coli* (25), *Proteus mirabilis* (11), enterococci (11), *Klebsiella pneumoniae* (9), *Klebsiella oxytoca* (2), *Citrobacter freundii* (2), *Morganella morganii* (3), *Enterobacter agglomerans* (2), *Enterobacter cloacae* (3), and *Providencia rettgeri* (1) (Table 1 shows details). All isolates were stored at −80°C in 30% glycerol. The samples were grown on nutrient agar, and single colonies were used to inoculate 5 ml of broth. These were grown aerobically at 37°C for 7 h in a shaking incubator. Biomass was harvested by centrifugation for 5 min in a microcentrifuge (Microcentaur, Oxbridge, United Kingdom) at 13,000 rpm.

DNA extraction. Cellular DNA was extracted by the phenol-chloroform method. Briefly, the harvested bacterial cells were suspended in 500 μl of Tris-EDTA buffer and lysed with 30 mg of lysozyme (Sigma)/ml and 10% (wt/vol) sodium dodecyl sulfate. Organic cellular debris was removed with 24:1:1 phenol-chloroform-isoamyl alcohol. DNA was precipitated with 0.6 volume of isopropanol, and the pellets were washed with 70% ethanol. The DNA pellets were then redissolved in 50 μl of Tris-EDTA buffer and incubated with 10 mg of DNase-free RNase (bovine pancreas; Sigma)/ml. The quantity and quality of DNA obtained was determined in 0.8% agarose. Final concentrations were adjusted to 50 ng/μl for FAFLP.

FAFLP analysis. All AFLP reactions were performed in a Progene thermocycler [Techne (Cambridge) Ltd., Cambridge, United Kingdom]. Restriction of genomic DNA and ligation of adapters were performed using GIBCO (Life Technologies, Paisley, United Kingdom) AFLP core reagent kits. A total of 250 ng was simultaneously digested with *Eco*RI and *Mse*I restriction enzymes, 2.5 U each, in a final reaction volume of 25 μl. The reaction mixtures were incubated at 37°C for 2 h and then at 70°C for 15 min to heat inactivate the enzymes. *Eco*RI and *Mse*I adapters were added at final concentrations of 0.2 and 2.2 μM, respectively, and ligated for 2 h at 20°C using 1 U of T4 DNA ligase. The final ligation mixture volume was 50 μl.

PCR amplification of the adapter-ligated fragments was performed with 5′ IRD800-labeled *Mse*I plus CT and unlabeled *Eco*RI plus 0 primers (MWG-Biotech UK Ltd., Milton Keynes, United Kingdom). The PCR mixtures (20 μl) constituted 5 μl of ligated DNA, 40 ng of labeled *Mse*I primer, 4 ng of *Eco*RI primer, 0.2 mM (each) deoxynucleoside triphosphates, 0.4 U of *Taq* polymerase (Promega UK, Southampton, United Kingdom), 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, and 50 mM KCl. Twenty-five amplification cycles were performed with the following cycle profile: cycle 1, 60 s at 94°C, 30 s at 65°C, 60 s at 72°C; cycles 2 to 12, 30 s at 94°C, 30 s at an annealing temperature 0.7°C lower than for each previous cycle (starting at 64.3°C), 60 s at 72°C; cycles 13 to 24, 30 s at 94°C, 30 s at 56°C, 60 s at 72°C. This "touchdown" PCR protocol was used to minimize PCR artifacts (2).

Electrophoresis of PCR products. The PCR-amplified products were separated on 5% denaturing polyacrylamide gels (25 cm^2 by 0.2 mm thick) on a LICOR 4200 automated sequencer (MWG-Biotech). The gels were prepared with 20 ml of monomer solution and 5 ml of complete buffer from the SequaGel XR premixed gel system [National Diagnostics (UK) Ltd., East Riding, Yorkshire, United Kingdom]. These mixtures were briefly degassed, and 1 ml of 10% ammonium persulfate was added prior to casting.

The amplified products were mixed with 10 μl of sequencing loading dye (Sigma) and heated to 95°C for 3 min. These denatured mixtures were rapidly placed in ice to prevent the DNAs from reverting to their secondary structures. Each sample (0.5 μl) was spotted on a CAM48 membrane comb (National Diagnostics) together with an IRD800-labeled molecular size marker of 50 to 350 bp (MWG-Biotech) after every six samples. Electrophoresis was performed at 1,500 V, 25 W, and 45°C using Tris-borate-EDTA buffer.

Fragment analysis. Fragments were sized with RFLPScan version 3.56 software (Scanalytics, Inc., Billeric, Mass.). Electropherograms of all AFLP profiles were visually inspected for polymorphisms, with the presence (1) or absence (0) of fragments from 60 to 350 bp scored in a binary matrix and stored in Microsoft Excel 2000.

From the literature, it was apparent that the majority of AFLP analyses of the same species studied here used *Mse*I plus C and *Eco*RI plus one or two selective nucleotides (2, 23, 30, 40). Therefore, we initially used *Mse*I plus C and *Eco*RI plus 0, but too many bands were produced. Experiments using a second selective nucleotide on the *Mse*I binding site led us to a common primer combination of *Mse*I plus CT and *Eco*RI plus 0, which produced between 40 and 60 bands that could be accurately sized and scored.

Cluster analysis. Cluster analysis used the simple matching genetic distance estimation and unweighted pair group method using arithmetic averages (UPGMA) (31). All calculations and displays were performed using TREECON for Windows version 1.3 (33), which runs under Microsoft Windows NT on an IBM-compatible personal computer.

Probabilistic identification. The approach adopted here uses Bayes' theorem, which is an established method that has been used widely for the identification of bacteria against an identification matrix (9, 20, 39), particularly when the characters are binary in nature (4), such as those in this study. Note that the theoretical requirement for using Bayes' theorem is that all taxa are exclusive and exhaustive, and while this may not always be the case for AFLP (or any other typing method, viz., API and Vitek) in that two different taxa (organisms) may give rise by chance to identical AFLP patterns, the probability of this happening is very small, and it was not encountered in this study (see below). Therefore, the practical implementation using Bayes' theorem for the construction of a probabilistic database from FAFLP data can be considered valid. The probabilistic database comprised two parts, (i) the construction of the probabilistic identification matrix and (ii) the identification of an unknown isolate against the identification matrix. All calculations were performed using Matlab version 5 (The MathWorks, Inc., Natick, Mass.), which runs under Microsoft Windows NT on an IBM-compatible personal computer.

The first stage was the construction of an identification matrix which consisted of 291 columns (1 for each possible FAFLP fragment from 60 to 350 bp), while there were either (i) 8 rows, 1 for each genus, or (ii) 10 rows, 1 for each species. The responses used in each column were the probabilities of observing an FAFLP fragment of 60 to 350 bp for that taxon. The probabilities were calculated as the arithmetic mean of the presence (1) or absence (0) of each fragment for each taxon and were scaled to range between 0.01 and 0.99. Entries of 0.0 and 1.0 were not employed, since an isolate with an atypical or erroneous result would generate an identification score of zero for a strain similar to it (5). Forty-five of the FAFLP patterns were used to construct these matrices, and the rest (24 FAFLP profiles) were used as an independent test set; these patterns were chosen randomly and are detailed in Table 1.

The second stage was the identification of an unknown isolate against the probability identification matrix. The identification was based on the implementation by Willcox et al. of Bayes' theorem for use with bacteria (38, 39) and has been adapted by Lapage et al. (20) to be $L_i^* = L_i/\Sigma L_i$, where L_i is the likelihood of an isolate belonging to taxon i. The identification scores for each taxon are normalized so that $\Sigma L_i^* = 1$; thus, identification of an unknown isolate is achieved when ΣL_i^* for one taxon exceeds 0.999 (5, 20).

RESULTS AND DISCUSSION

AFLP has been confirmed as a rapid, reproducible, and highly discriminatory method of microbial identification to the species level (16, 22, 28). However, there is considerable variation in the choice of primer combinations for subspecies characterization of different bacterial genera. *Eco*RI plus C and *Mse*I plus C primers have been used for *E. coli* (23), although many other primer combinations have been exploited (2, 30, 40): *Eco*RI plus C-*Mse*I plus A for *K. pneumoniae* (34) and *B. anthracis* with *Eco*RI plus C-*Mse*I plus G (19). By contrast, *Enterococcus faecium*, *N. meningitidis*, and *Vibrio cholerae* have

TABLE 1. Summary of 69 UTI isolates analyzed by AFLP

Strain	Organism	Create or test database	FAFLP profile[a]
Cf102	*C. freundii*	Create	A9a
Cf109	*C. freundii*	Test	A9b
Ea98	*E. agglomerans*	Create	A3a
Ea103	*E. agglomerans*	Test	A3b
Et101	*E. cloacae*	Create	A4a
Et107	*E. cloacae*	Create	A4a
Et111	*E. cloacae*	Test	A4a
EntC82	*Enterococcus* spp.	Create	A12a
EntC83	*Enterococcus* spp.	Create	A12g
EntC85	*Enterococcus* spp.	Create	A12h
EntC86	*Enterococcus* spp.	Create	A12d
EntC87	*Enterococcus* spp.	Create	A12b
EntC88	*Enterococcus* spp.	Create	A12e
EntC89	*Enterococcus* spp.	Create	A11a
EntC90	*Enterococcus* spp.	Create	A12f
EntC91	*Enterococcus* spp.	Test	A12i
EntC92	*Enterococcus* spp.	Test	A12c
EntC93	*Enterococcus* spp.	Test	A12j
Eco9	*E. coli*	Create	A10h
Eco11	*E. coli*	Create	A10f
Eco13	*E. coli*	Create	A10j
Eco16	*E. coli*	Create	A10k
Eco19	*E. coli*	Create	A10e
Eco20	*E. coli*	Create	A10j
Eco22	*E. coli*	Create	A10f
Eco33	*E. coli*	Create	A10j
Eco34	*E. coli*	Create	A10i
Eco36	*E. coli*	Create	A10j
Eco38	*E. coli*	Create	A10f
Eco40	*E. coli*	Create	A10k
Eco44	*E. coli*	Create	A10k
Eco47	*E. coli*	Create	A10k
Eco49	*E. coli*	Create	A10k
Eco5	*E. coli*	Test	A10b
Eco6	*E. coli*	Test	A10a
Eco24	*E. coli*	Test	A10g
Eco25	*E. coli*	Test	A10a
Eco31	*E. coli*	Test	A10d
Eco41	*E. coli*	Test	A10g
Eco42	*E. coli*	Test	A10a
Eco46	*E. coli*	Test	A10c
Eco48	*E. coli*	Test	A10e
Eco50	*E. coli*	Test	A10k
kox105	*K. oxytoca*	Create	A6a
kox108	*K. oxytoca*	Test	A6b
kp52	*K. pneumoniae*	Create	A8e
kp53	*K. pneumoniae*	Create	A8c
kp54	*K. pneumoniae*	Create	A7a
kp55	*K. pneumoniae*	Create	A8d
kp56	*K. pneumoniae*	Create	A8e
kp58	*K. pneumoniae*	Create	A8a
kp59	*K. pneumoniae*	Test	A8b
kp60	*K. pneumoniae*	Test	A8f
kp62	*K. pneumoniae*	Test	A8g
Mm95	*M. morganii*	Create	A2a
Mm100	*M. morganii*	Create	A2a
Mm106	*M. morganii*	Test	A2a
pm63	*P. mirabilis*	Create	A1a
pm64	*P. mirabilis*	Create	A1a
pm65	*P. mirabilis*	Create	A1a
pm66	*P. mirabilis*	Create	A1a
pm67	*P. mirabilis*	Create	A1a
pm68	*P. mirabilis*	Create	A1a
pm69	*P. mirabilis*	Create	A1a
pm70	*P. mirabilis*	Create	A1a
pm71	*P. mirabilis*	Test	A1a
pm73	*P. mirabilis*	Test	A1a
pm74	*P. mirabilis*	Test	A1a
Pr110	*P. rettgeri*	Create	A5a

[a] Based on the dendrogram in Fig. 1 using genetic differences of 10% (uppercase letter and number, e.g., A1) and 5% (lowercase letter, e.g., A1a).

all been analyzed with different sets of restriction enzymes (1, 12, 17). Thus, it is necessary to establish the genus of a bacterial isolate prior to nominating the restriction enzymes and primer combination for AFLP.

The common bacterial causal agents of UTI are very varied, and most of the common ones, including *E. coli*, *P. mirabilis*, *K. pneumoniae*, *K. oxytoca*, *C. freundii*, *M. morganii*, *E. agglomerans*, *E. cloacae*, *P. rettgeri*, and *Enterococcus* spp., were encountered and analyzed in this study. For these eight genera, encompassing both gram-positive and gram-negative bacteria, it was found after extensive investigation of different primer combinations that a dual restriction digest with *Eco*RI and *Mse*I followed by *Mse*I plus CT and *Eco*RI plus 0 as a common primer combination was an excellent approach for generating information-rich FAFLP patterns from all of these bacterial isolates. Moreover, all FAFLP reactions were run in triplicate over a period of 4 months, and the same FAFLP profiles for each set of replicates were observed; this highlights the excellent reproducibility of FAFLP, a phenomenon observed by other workers, who showed the robustness of AFLP between different laboratories (18).

The 69 bacterial isolates that were examined by FAFLP showed distinctive AFLP profiles at the genus, species, and subspecies levels, and these are shown digitally in Fig. 1 and recorded in Table 1. These FAFLP patterns typically comprised between 40 and 60 fragments in the selected size range of 50 to 350 bp (Fig. 1); these limits were set by the molecular size marker. All FAFLP profiles were sized using RFLPScan software, and a matrix of ones and zeros representing the presence or absence of amplified fragments, respectively, was generated for all of the UTI isolates. The genetic relatedness of the samples was determined by cluster analysis using the UPGMA algorithm, and the resulting dendrogram of the 69 FAFLP profiles revealed eight main clusters (Fig. 1) corresponding to the eight bacterial genera.

The arbitrary line shown in Fig 1 at 15% genetic difference shows the level of genus-specific clustering, and at ~25%, the first branch in this tree clearly discriminates between the gram-positive enterococci and enteric gram-negative isolates. These discriminations strongly demonstrate the potential of AFLP, contrary to popular opinion (32), for describing a wider diversity range within microbial populations where identification rather than subspecies typing is required.

Clinical cases of UTIs will often require accurate typing of the causative organism in order to distinguish relapses, detect reinfection with different organisms, and identify outbreaks of cross-infection. This higher level of discrimination was obtained from the combined FAFLP data. The *Klebsiella* and *Enterobacter* isolates both show two well-defined subclusters at a genetic-difference threshold of 10% (Fig. 1 and Table 1), corresponding to *K. oxytoca* and *K. pneumoniae* and to *E. agglomerans* and *E. cloacae*, respectively. Furthermore, when the threshold of 5% genetic difference is used, subspecies discrimination among the UTI isolates is also observed (Fig. 1 and Table 1). Finally, to test whether subspecies discrimination is dependent on very different microorganisms being included in tree production, cluster analysis was performed on the 25 *E. coli* isolates alone. The resultant dendrogram in Fig. 2 shows the same discrimination observed when all the data were combined (Fig. 1), indicating that having very different FAFLP

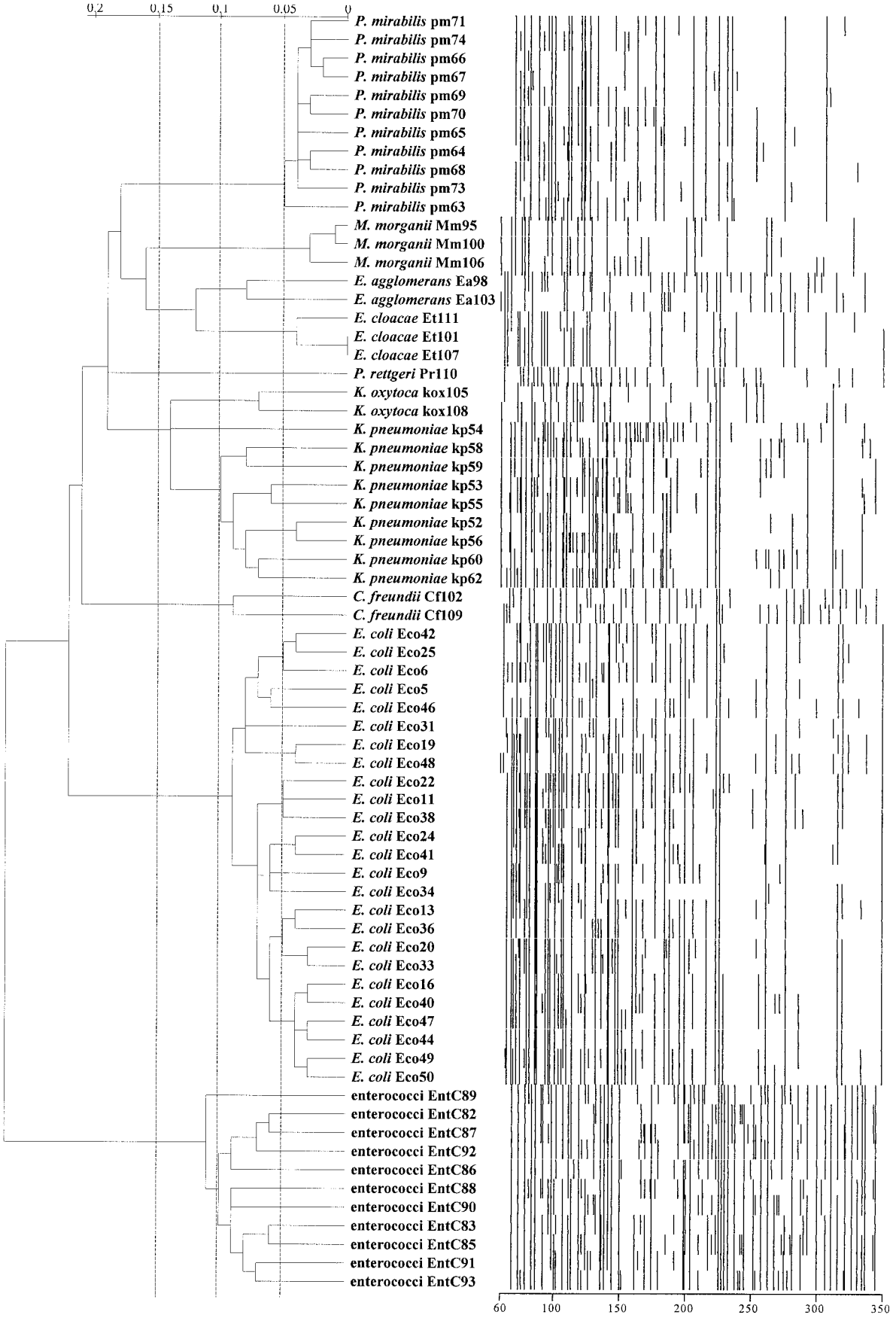

FIG. 1. UPGMA dendrogram derived from FAFLP data showing all 69 isolates analyzed. The scale represents the genetic difference. Also shown are the digitized AFLP patterns with the scale in base pairs.

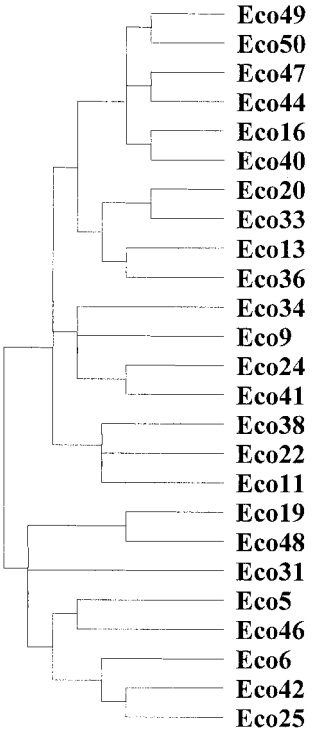

FIG. 2. UPGMA dendrogram derived from FAFLP data showing the relationships among the 25 *E. coli* isolates.

profiles in the cluster analysis does not skew the discrimination observed. In combination, the above-mentioned results clearly suggest that the single primer combination *Eco*RI plus 0 and *Mse*I plus CT was sufficient to characterize and type bacteria from these different genera to the species level and further. It is hoped that this will also be true for other genera and species therein, but this will be a subject for future investigation.

Interpretation of the AFLP profiles has conventionally been done by the application of "unsupervised" pattern recognition methods of cluster analysis (16, 22, 28). With "unsupervised learning" methods of this sort, the relevant algorithm (e.g., UPGMA) seeks clusters in the data, thereby allowing the investigator to group objects together on the basis of their perceived closeness; this process requires in-depth knowledge of how to interpret dendrograms and may be subjective if the tree is complicated and includes polytomies (unresolved nodes). It is therefore desirable to exploit "supervised" learning methods (24), construct identification databases containing known bacterial taxa, and effect probabilistic matching against these.

The method that we have used is based on Bayes' theorem for the identification of bacteria against an identification matrix (9, 20, 39). Forty-five of the FAFLP patterns were chosen randomly (Table 1 shows the details) and used to construct (i.e., supervise) two probabilistic identification matrices; the first contained 8 taxa, 1 for each genus, and the second comprised 10 taxa, 1 for each species. These identification matrices comprised 291 columns, 1 for each possible FAFLP fragment from 60 to 350 bp, and each cell was the arithmetic mean for each taxon for each of the possible 291 fragments, scaled between 0.01 and 0.99. The 24 FAFLP profiles that were not used for construction of these identification matrices (Table 1) were employed to test whether this method could be used to identify an unknown hospital UTI isolate. Probabilistic identification was performed as detailed above, and likelihood scores for each taxon were computed. If the likelihood was > 0.999 for one taxon, this was taken as its identity. Each of the 24 unknown isolates was correctly identified at the genus and species levels with a score of 1.0000 (as were the 45 isolates used to construct the databases). It was therefore evident that this approach does indeed work for an independent test set and that it produced results that no longer needed the interpretation of a dendrogram.

In cases of UTI relapse, it is necessary to identify the causal organism to the subspecies level. It is known that AFLP has sufficient resolving power to allow this (6, 8, 10, 12, 17, 27, 32); therefore, the question arises as to whether the use of an FAFLP probabilistic database could identify the infecting organism to the subspecies level. Hence, we constructed a database of only the first FAFLP replicate set from the 25 *E. coli* isolates and calibrated it to identify the 11 different *E. coli* FAFLP types (A10a to -k) observed in the dendrogram at 0.05 genetic difference (Fig. 1; detailed in Table 1). The second and third replicate FAFLP patterns were subsequently matched against this, and all 50 isolates were correctly identified with scores of 1.0000. This suggests that the approach has sufficient resolving power to identify UTI relapse to the subspecies level.

In summary, this study clearly demonstrates that FAFLP can be used to identify the causal agents of UTI. A single primer combination was discovered that gave information-rich AFLP profiles from gram-negative and gram-positive bacteria encompassing eight genera. Furthermore, the identification of bacteria against an FAFLP identification matrix using Bayes' theorem was illustrated with unknown UTI isolates that were not used in the construction of the database. In conclusion, we believe that the approach we have developed using an AFLP probabilistic database constructed with FAFLP patterns from the common causal agents of UTI offers great potential as an objective UTI identification procedure in the clinical laboratory.

ACKNOWLEDGMENTS

Y.K. thanks the Overseas Research Students Awards Scheme (ORS) for his studentship, and R.G. is indebted to the United Kingdom BBSRC Engineering and Biological Systems Committees for financial support.

REFERENCES

1. **Antonishyn, N. A., R. R. McDonald, E. L. Chan, G. Horsman, C. E. Woodmansee, P. S. Falk, and C. G. Mayhall.** 2000. Evaluation of fluorescence-based amplified fragment length polymorphism analysis for molecular typing in hospital epidemiology: comparison with pulse-field gel electrophoresis for typing strains of vancomycin-resistant *Enterococcus faecium*. J. Clin. Microbiol. **38:**4058–4065.
2. **Arnold, C., L. Metherell, G. Willshaw, A. Maggs, and J. Stanley.** 1999. Predictive fluorescent amplified-fragment length polymorphism analysis of *Escherichia coli*: high-resolution typing method with phylogenetic significance. J. Clin. Microbiol. **37:**1274–1279.
3. **Bouza, E., R. San Juan, P. Munoz, A. Voss, and J. Kluytmans.** 2001. A European perspective on nosocomial urinary tract infections. II. Report on incidence, clinical characteristics and outcome (ESGNI-004 study). Clin. Microbiol. Infect. **7:**532–542.
4. **Bryant, T. N.** 1994. A program for analyzing microbial binary characteristics. Binary Comput. Microbiol. **6:**97–100.
5. **Bryant, T. N., A. G. Capey, and R. C. W. Berkeley.** 1985. Microcomputer assisted identification of *Bacillus* species. Comput. Appl. Biosci. **1:**23–27.

6. **Craig, J. C., L. M. Irwig, P. Sureshkumar, and L. P. Roy.** 1998. Symptomatic urinary tract infection in preschool Australian children. J. Paediatr. Child Health **34:**154–159.
7. **Desai, M., E. J. Threlfall, and J. Stanley.** 2001. Fluorescent amplified-fragment length polymorphism subtyping of the *Salmonella enterica* serovar Enteritidis phage type 4 clone complex. J. Clin. Microbiol. **39:**201–206.
8. **De Zoysa, A., and A. Efstratiou.** 2000. Use of amplified fragment length polymorphisms for typing *Corynebacterium diphtheriae*. J. Clin. Microbiol. **38:**3843–3845.
9. **Dybowski, W., and D. A. Franklin.** 1968. Conditional probability and the identification of bacteria: a pilot study. J. Gen. Microbiol. **54:**215–229.
10. **Foxman, B., R. Barlow, H. D'Arcy, B. Gillespie, and J. D. Sobel.** 2000. Urinary tract infection: self reported incidence and associated costs. Ann. Epidemiol. **10:**509–515.
11. **Goodacre, R., É. M. Timmins, R. Burton, N. Kaderbhai, A. M. Woodward, D. B. Kell, and P. J. Rooney.** 1998. Rapid identification of urinary tract infection bacteria using hyperspectral, whole organism fingerprinting and artificial neural networks. Microbiology **144:**1157–1170.
12. **Goulding, J. N., J. V. Hookey, J. Stanley, W. Olver, K. R. Neal, D. A. A. Ala'Aldeen, and C. Arnold.** 2000. Fluorescent amplified fragment length polymorphism genotyping of *Neisseria meningitidis* identifies clones associated with invasive disease. J. Clin. Microbiol. **38:**4580–4585.
13. **Goulding, J. N., J. Stanley, N. Saunders, and C. Arnold.** 2000. Genome sequence-based fluorescent amplified-fragment length polymorphism analysis of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **38:**1121–1126.
14. **Grady, R., D. Blanc, P. Hauser, and J. Stanley.** 2001. Genotyping of European isolates of methicillin-resistant *Staphylococcus aureus* by fluorescent amplified-fragment length polymorphism analysis (FAFLP) and pulsed-field gel electrophoresis (PFGE) typing. J. Med. Microbiol. **50:**588–593.
15. **Gruneberg, R. N.** 1994. Changes in urinary pathogens and their antibiotic sensitivities 1971–1992. J. Antimicrob. Chemother. **33**(Suppl. A)**:**1–8.
16. **Janssen, P., R. Coopman, G. Huys, J. Swings, M. Bleeker, P. Vos, M. Zabeau, and K. Kersters.** 1996. Evaluation of the DNA fingerprinting method AFLP as a tool in bacterial taxonomy. Microbiology **142:**1881–1893.
17. **Jiang, S. C., M. Matte, G. Matte, A. Huq, and R. Colwell.** 2000. Genetic diversity of clinical and environmental isolates of *Vibrio cholerae* determined by amplified fragment length polymorphism fingerprinting. Appl. Environ. Microbiol. **66:**148–153.
18. **Jones, C. J., K. J. Edwards, S. Castaglione, M. O. Winfield, F. Sala, C. vandeWiel, G. Bredemeijer, B. Vosman, M. Matthes, A. Daly, R. Brettschneider, P. Bettini, M. Buiatti, E. Maestri, A. Malcevschi, N. Marmiroli, R. Aert, G. Volckaert, J. Rueda, R. Linacero, A. Vazquez, and A. Karp.** 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol. Breeding **3:**381–390.
19. **Keim, P., A. Kalif, J. Schupp, K. Hill, S. E. Travis, K. Richmond, D. M. Adair, M. Hugh-Jones, C. R. Kuske, and P. Jackson.** 1997. Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers. J. Bacteriol. **179:**818–824.
20. **Lapage, S. P., S. Bascomb, W. R. Willcox, and M. A. Curtis.** 1973. Identification of bacteria by computer: general aspects and prospects. J. Gen. Microbiol. **77:**273–290.
21. **Lewis, D. A.** 1989. Bacteriology of urine, p. 1–19. *In* P. M. Hawkey and D. A. Lewis (ed.), Medical bacteriology: a practical approach. IRL Press, Oxford, United Kingdom.
22. **Lin, J. J., J. Kuo, and J. Ma.** 1996. A PCR based DNA fingerprinting technique—AFLP for molecular typing of bacteria. Nucleic Acids Res. **24:** 3649–3650.
23. **Lyoda, S., A. Wada, J. Weller, S. J. A. Flood, E. Schreiber, B. Tucker, and H. Watanabe.** 1999. Evaluation of AFLP, a high-resolution DNA fingerprinting method, as a tool for molecular subtyping of enterohemorrhagic *Escherichia coli* O157:H7 isolates. Microbiol. Immunol. **43:**803–806.
24. **Martens, H., and T. Næs.** 1989. Multivariate calibration. John Wiley, Chichester, United Kingdom.
25. **Morgan, M. G., and H. McKenzie.** 1993. Controversies in the laboratory diagnosis of community acquired urinary tract infection. Eur. J. Clin. Microbiol. Infect. Dis. **12:**491–504.
26. **Panaretto, K. S., J. C. Craig, J. F. Knight, R. Howman-Giles, P. Sueresh-kumar, and L. P. Roy.** 1999. Risk factors for recurrent urinary tract infection in preschool children. J. Paediatr. Child Health **35:**454–459.
27. **Pappas, P. G.** 1991. Laboratory in the diagnosis and management of urinary tract infections. Med. Clin. N. Am. **75:**313–325.
28. **Savelkoul, P. H. M., H. J. M. Aarts, J. deHaas, L. Dijkshoorn, B. Duim, M. Otsen, J. L. W. Rademaker, L. Schouls, and J. A. Lenstra.** 1999. Amplified-fragment length polymorphism analysis: the state of an art. J. Clin. Microbiol. **37:**3083–3091.
29. **Slack, R. C. B.** 1995. Urinary infections, p. 243–250. *In* D. Greenwood (ed.), Antimicrobial chemotherapy. Oxford University Press, Oxford, United Kingdom.
30. **Smith, D., G. Willshaw, J. Stanley, and C. Arnold.** 2000. Genotyping of verocytotoxin-producing *Escherichia coli* O157: comparison of isolates of a prevalent phage type by fluorescent amplified-fragment length polymorphism and pulsed-field gel electrophoresis analyses. J. Clin. Microbiol. **38:** 4616–4620.
31. **Sneath, P. H. A., and R. R. Sokal.** 1973. Numerical taxonomy. W. H. Freeman, San Francisco, Calif.
32. **Vandamme, P., B. Pot, M. Gillis, P. DeVos, K. Kersters, and J. Swings.** 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol. Rev. **60:**407–440.
33. **Van de Peer, Y., and R. de Wachter.** 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comp. Appl. Biosci. **10:**569–570.
34. **Van der Zwet, W. C., G. A. Parlevliet, P. H. M. Savelkoul, J. Stoof, A. M. Kaiser, J. G. M. Koeleman, and C. M. J. E. Vandenbroucke Grauls.** 1999. Nosocomial outbreak of gentamicin-resistant *Klebsiella pneumoniae* in a neonatal intensive care unit controlled by a change in antibiotic policy. J. Hosp. Infect. **42:**295–302.
35. **Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijter, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau.** 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23:**4407–4414.
36. **Wagenaar, J. A., M. A. P. van Bergen, D. G. Newell, R. Grogono-Thomas, and B. Duim.** 2001. Comparative study using amplified fragment length polymorphism fingerprinting, PCR genotyping, and phenotyping to differentiate *Campylobacter fetus* strains isolated from animals. J. Clin. Microbiol. **39:**2283–2286.
37. **Wilkie, M. E., M. K. Almond, and F. P. Marsh.** 1992. Diagnosis and management of urinary tract infection in adults. Br. Med. J. **303:**1137–1141.
38. **Willcox, W. R., S. P. Lapage, S. Bascomb, and M. A. Curtis.** 1973. Identification of bacteria by computer: theory and programming. J. Gen. Microbiol. **77:**317–330.
39. **Willcox, W. R., S. P. Lapage, and B. Holmes.** 1980. A review of numerical methods in bacterial identification. Antonie Leeuwenhoek **46:**233–299.
40. **Zhao, S. H., S. E. Mitchell, J. H. Meng, S. Kresovich, M. P. Doyle, R. E. Dean, A. M. Casa, and J. W. Weller.** 2000. Genomic typing of *Escherichia coli* O157:H7 by semiautomated fluorescent AFLP analysis. Microbes Infect. **2:**107–113.