

Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy

Paul Janssen,¹ Renata Coopman,¹ Geert Huys,¹ Jean Swings,¹
Marjo Bleeker,² Pieter Vos,² Marc Zabeau² and Karel Kersters¹

Author for correspondence: Paul Janssen. Tel: +32 9 264 5121. Fax: +32 9 264 5346.
e-mail: paul.janssen@rug.ac.be

¹ Laboratorium voor
Microbiologie, Universiteit
Gent, B-9000 Gent,
Belgium

² KeyGene nv, Wageningen,
The Netherlands

We investigated the usefulness of a novel DNA fingerprinting technique, AFLP, which is based on the selective amplification of genomic restriction fragments by PCR, to differentiate bacterial strains at the subgeneric level. In total, 147 bacterial strains were subjected to AFLP fingerprinting: 36 *Xanthomonas* strains, including 23 pathovars of *Xanthomonas axonopodis* and six pathovars of *Xanthomonas vasicola*, one strain of *Stenotrophomonas*, 90 genotypically characterized strains comprising all 14 hybridization groups currently described in the genus *Aeromonas*, and four strains of each of the genera *Clostridium*, *Bacillus*, *Acinetobacter*, *Pseudomonas* and *Vibrio*. Depending on the genus, total genomic DNA of each bacterium was digested with a particular combination of two restriction endonucleases and the resulting fragments were ligated to restriction halfsite-specific adaptors. These adaptors served as primer-binding sites allowing the fragments to be amplified by selective PCR primers that extend beyond the adaptor and restriction site sequences. Following electrophoretic separation on 5% (w/v) polyacrylamide/8.3 M urea, amplified products could be visualized by autoradiography because one of the selective primers was radioactively labelled. The resulting banding patterns, containing approximately 30–50 visualized PCR products in the size range 80–550 bp, were captured by a high-resolution densitoscanner and further processed for computer-assisted analysis to determine band-based similarity coefficients. This study reveals extensive evidence for the applicability of AFLP in bacterial taxonomy through comparison of the newly obtained data with results previously obtained by well-established genotypic and chemotaxonomic methods such as DNA–DNA hybridization and cellular fatty acid analysis. In addition, this study clearly demonstrates the superior discriminative power of AFLP towards the differentiation of highly related bacterial strains that belong to the same species or even biovar (i.e. to characterize strains at the infrasubspecific level), highlighting the potential of this novel fingerprinting method in epidemiological and evolutionary studies.

Keywords: DNA fingerprinting, PCR, selective amplification, AFLP

INTRODUCTION

Over the past few years, the use of modern molecular biological techniques to determine the degree of sequence

conservation between bacterial genomes has led to the development of methods based solely on the detection of naturally occurring DNA polymorphisms. These polymorphisms are a result of point mutations or rearrangements (i.e. insertions, deletions, etc.) in the DNA and can be detected by scoring band presence versus absence in banding patterns that are generated by either restriction enzyme digestion or DNA amplification procedures, or both. The underlying idea is that variations in banding

Abbreviations: FAME, fatty acid methyl ester; HG, hybridization group; LMG, Culture Collection of the Laboratorium voor Microbiologie, Universiteit Gent; S_D , Dice similarity coefficient; UPGMA, unweighted-pair group method using average linkages.

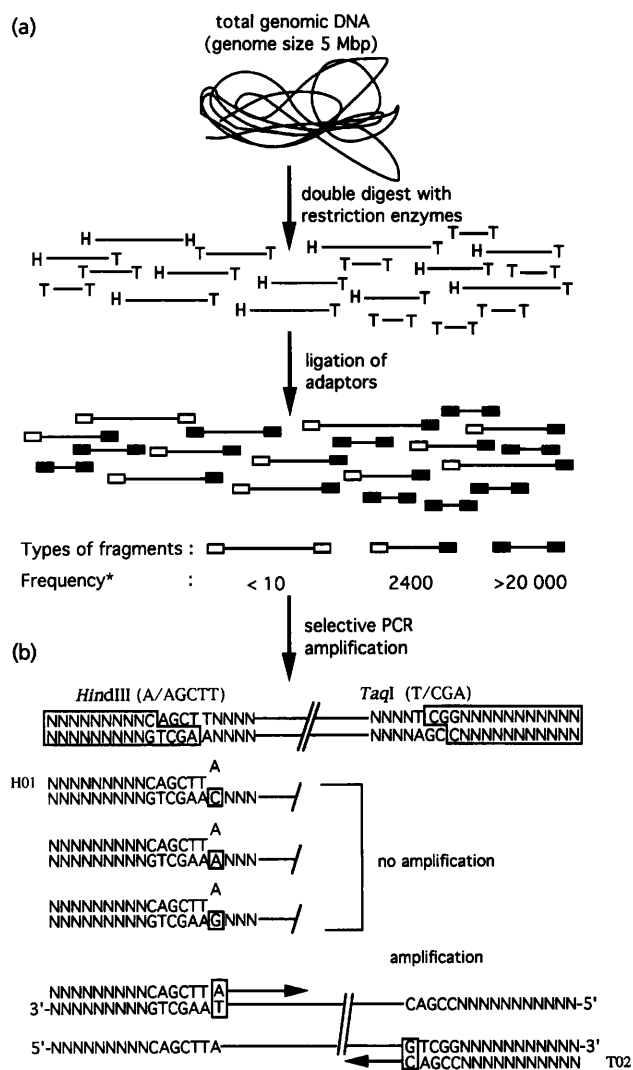


Fig. 1. Overall scheme of the AFLP technique: (a) total cellular DNA is digested with two restriction enzymes, *Hind*III (H) and *Taq*I (T), which have a 6 and 4 bp recognition sequence, respectively. This is followed by ligation of adaptors to both ends of the restriction fragments. Adaptors are restriction-halfsite-specific and have been indicated by boxes (and for H- and T-halfsites, respectively); *, the estimated frequency (number of fragments per genome) for each type of restriction fragment is based on the theoretical number of cleavage sites for 6 bp-cutters (every 4096 bp) and 4 bp-cutters (every 256 bp) assuming random base distribution and a genome size of 5 Mbp. (b) Selective amplification of restriction fragments: the adaptor sequences serve as binding sites for the PCR primers which contain selective bases at their 3'-end (primer sequences consist of the complementary sequence of the corresponding adaptors and include the restriction site sequence). Under stringent PCR conditions (see Methods), only perfectly matched primers will be elongated resulting in amplification. Sequences of oligonucleotides used in the adaptors or as PCR primers are listed in Table 1. As an example, selective amplification has been depicted with *Hind*III-primer H01 and *Taq*I-primer T02 (for alphanumeric notation of selective primers, see Methods). Because only the *Hind*III-primer is radioactively labelled, amplified *Taq*I-*Taq*I fragments will not be visualized by autoradiography. For clarity, only essential sequences are shown.

patterns are a direct reflection of the genetic relationship between the bacterial strains examined and therefore that these banding patterns can be considered as 'genomic fingerprints' allowing numerical analysis for characterization (typing) and identification purposes.

In this paper, we report the use of a novel method, AFLP, for the fingerprinting of bacterial genomes. The derivation of the name AFLP is solely based on the resemblance of the AFLP technique with RFLP (restriction fragment length polymorphism) analysis, and AFLP should therefore not be used as an acronym (Vos *et al.*, 1995). The AFLP concept, originally developed as a universal DNA fingerprinting method (Zabeau & Vos, 1993), basically consists of three steps (Fig. 1): (1) digestion of total cellular DNA with two restriction enzymes and ligation of restriction halfsite-specific adaptors to all restriction fragments; (2) selective amplification of these fragments with two PCR primers that have corresponding adaptor- and restriction-site-sequences as their target sites; and (3) electrophoretic separation of the PCR products on a 5% (w/v) polyacrylamide/8.3 M urea gel. Only a subset of fragments will be amplified because primers contain at their 3'-end one or more bases (the so-called selective bases) which are complementary to nucleotides flanking the restriction sites (Fig. 1b). Since the reaction conditions are such that only perfectly matched primers will initiate DNA synthesis, approximately one out of four restriction fragments will be amplified for each selective base used (depending on the base distribution within the analysed genome) (Vos *et al.*, 1995).

To evaluate the potential of AFLP in bacterial taxonomy, we used a large number of strains belonging to the genera *Xanthomonas* and *Aeromonas*. The taxonomies of both these bacterial genera have been studied extensively in our laboratory as well as by various other research groups using a variety of methods. This allowed us to make a reliable comparison between the previously obtained data and the results obtained by the AFLP technique. In addition, both genera include species that have been further subdivided in either pathogens (for *Xanthomonas*) or hybridization groups (HGs) (genomic species) (for *Aeromonas*). It is this complexity that prompted us to investigate the applicability of the AFLP technique for the characterization of bacterial strains at the infrasubspecific level. In addition, we explored whether AFLP can be used as a general tool in bacterial taxonomy and have produced characteristic AFLP fingerprints from other bacteria whose genomes vary in size and base composition, including members belonging to the genera *Clostridium*, *Acinetobacter*, *Bacillus*, *Pseudomonas* and *Vibrio*.

METHODS

Materials. All general chemicals were of analytical grade. Polynucleotide kinase, T4 ligase and restriction endonucleases were supplied by Pharmacia, except for the restriction enzyme *Mse*I which was purchased from New England Biolabs. EuroTaq DNA polymerase (Eurogentec) was used for all amplification procedures except for AFLP reactions on *Eco*RI-*Mse*I templates which were performed at KeyGene,

Table 1. Adaptors and PCR primers used in this study

The sequences of *ApaI*-, *EcoRI*- and *HindIII*- adaptors differ substantially from the sequences of *TaqI*- and *MseI*-adaptors to allow adaptor-specific primer binding. Primers for AFLP consist of a core region and a 3'-extension (E) of 0, 1 or 2 nt (extensions not shown, see Methods).

<i>ApaI</i>-adaptor:	5'-TCGTAGACTGCGTACAGGCC-3'
<i>ApaI</i> (GGGCC/C)	3'-CATCTGACGCATGT-5'
A-primer core sequence:	5'-GACTGCGTACAGGCCCE-3'
<i>EcoRI</i>-adaptor:	5'-CTCGTAGACTGCGTACC-3'
<i>EcoRI</i> (G/AATTC)	3'-CTGACGCATGGTTAA-5'
E-primer core sequence:	5'-GACTGCGTACCAATTCE-3'
<i>HindIII</i>-adaptor:	5'-CTCGTAGACTGCGTACC-3'
<i>HindIII</i> (A/AGCTT)	3'-CTGACGCATGGTCTGA-5'
H-primer core sequence:	5'-GACTGCGTACCAGCTTE-3'
<i>TaqI</i>-adaptor:	5'-GACGATGAGTCCTGAC-3'
<i>TaqI</i> (T/CGA)	3'-TACTCAGGACTGGC-5'
T-primer core sequence:	5'-CGATGAGTCCTGACCGAE-3'
<i>MseI</i>-adaptor:	5'-GACGATGAGTCCTGAG-3'
<i>MseI</i> (T/TAA)	3'-CTACTCAGGACTCAT-5'
M-primer core sequence:	5'-GATGAGTCCTGAGTAAE-3'

Wageningen, using AmpliTaq DNA polymerase (Perkin Elmer). [γ - 32 P]ATP (111 TBq mmol⁻¹) or [γ - 33 P]ATP (37–111 TBq mmol⁻¹) were purchased from Amersham. Milli-Q water (Millipore) was used for all DNA preparations, enzymic reactions and amplification procedures. Oligonucleotides used in adaptors or as PCR primers (purified and unpurified, respectively) were obtained from Eurogentec. Acrylamide gel matrix was obtained as pre-mixed SequiGel solutions (National Diagnostics) to which ammonium persulfate (Pharmacia) and N,N,N',N'-tetramethylethylenediamine (TEMED; Bio-Rad) was added according to the manufacturers' instructions.

Bacterial strains and cultures. All bacterial cultures were obtained from the Culture Collection of the Laboratorium voor Microbiologie (LMG), Universiteit Gent, Belgium, the Culture Collection of the University of Göteborg, Göteborg, Sweden, and the Institute of Medical Microbiology, University of Zürich, Zürich, Switzerland. Details concerning the *Xanthomonas* strains can be found in Vauterin *et al.* (1991a, b, 1992) and Yang *et al.* (1993). *Aeromonas* strains used in this study have been described previously by Huys *et al.* (1994) and were grown aerobically on trypticase soy agar [TSA; 3%, w/v, trypticase soy broth (BBL) supplemented with 1.5%, w/v, Bacto agar (Difco)] at 28 °C. *Acinetobacter* and *Pseudomonas* strains were grown aerobically on nutrient agar [nutrient broth (Gibco BRL) supplemented with 1.5% Bacto agar] at 28 °C. All strains were maintained as culture stocks in 15% (w/v) glycerol at -70 °C.

Isolation and purification of bacterial chromosomal DNA. *Xanthomonas* genomic DNAs were obtained from L. Vauterin and B. Hoste (Universiteit Gent, Belgium) and were prepared according to the method of Marmur (1961). *Clostridium* genomic DNAs were a kind gift from P. Dürre (Georg-August-Universität Göttingen, Germany) and were prepared according to Bertram & Dürre (1989). *Bacillus* and *Vibrio* genomic DNAs were obtained from M. Heyndrickx & L. Verdonck (Universiteit Gent, Belgium), respectively, and were prepared by the method outlined below. *Acinetobacter*, *Pseudomonas* and *Aeromonas* genomic DNAs were prepared using the same method as follows: a flat loop of cells (approx. 30 mg wet wt) was scraped off from TSA plates and washed once with 500 μ l

saline (150 mM NaCl, 10 mM EDTA, pH 8.0) in a 1.5 ml Eppendorf centrifuge tube. After centrifugation for 45 s at 15000 g (Heræus Biofuge 13), cells were resuspended in 100 μ l TE buffer (10 mM Tris/HCl, 1 mM EDTA, pH 7.6) and kept on ice. Cells were lysed and DNA was extracted with sarkosyl/guanidium thiocyanate (Sigma) essentially as described by Pitcher *et al.* (1989). The DNA was finally dissolved in 100 μ l TE buffer. To ensure that the DNA was completely dissolved, DNA preparations were usually kept at this stage at 4 °C overnight. Total RNA which was co-purified with the DNA was degraded by the addition of 25 μ l DNase-free RNase (0.25 mg ml⁻¹) and incubation for 1.5 h at 37 °C. The concentration of all DNA solutions was routinely determined on a Uvikon 940 spectrophotometer (Kontron Instruments) by measuring A_{260} (1 absorbance unit = 50 μ g ml⁻¹). The integrity of the DNA was checked by electrophoresis of aliquots (approx. 2 μ g) in submerged horizontal agarose gels (0.8–1.0%, w/v) using TAE buffer (40 mM Tris/acetate, 1 mM EDTA, pH 8.0) containing ethidium bromide (0.5 μ g ml⁻¹) as the electrophoresis buffer. Final DNA preparations were stored at -20 °C.

Adaptors. The adaptors were designed in such a way that the original restriction site would not be restored after ligation of the adaptor to a restriction fragment (Zabeau & Vos, 1993) (Fig. 1; Table 1). This has the advantage that restriction enzymes can be present during ligation to prevent fragment-to-fragment ligation. In addition, since adaptors consist of non-phosphorylated oligonucleotides, adaptor-to-adaptor ligation is avoided. The sequences of *ApaI*-, *EcoRI*- and *HindIII*-adaptors differ substantially from the sequences of *TaqI*- and *MseI*-adaptors to allow adaptor-specific primer binding (Table 1).

Preparation of template DNA. All procedures were performed essentially as described by Zabeau & Vos (1993) and Vos *et al.* (1995), with slight modifications. In general, 1 μ g DNA was digested with 10 U each restriction enzyme (i.e. *ApaI* and *TaqI*) in a final volume of 30 μ l according to the method of Vos *et al.* (1995). Following digestion, adaptors were added to a final concentration of 0.04 μ M (for *ApaI*-, *EcoRI*- and *HindIII*-adaptors) and 0.4 μ M (for *TaqI*- and *MseI*-adaptors) and ligated to the restriction fragments (Vos *et al.*, 1995). The DNA was

subsequently precipitated in the presence of 1.25 M ammonium acetate and 50 % (v/v) 2-propanol. The DNA pellet was washed once with 70 % (v/v) ethanol and finally dissolved in 100 µl T0.1E buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0). Template DNAs were stored at -20 °C.

PCR primers. Core sequences of PCR primers are listed in Table 1. In addition to these core sequences, selective primers contain at their 3'-end one or more bases. Throughout the text, primers are indicated with an alphanumeric notation corresponding to the type of core sequence (A for *ApaI*, E for *EcoRI*, H for *HindIII*, T for *TaqI* and M for *MseI*-primers) and 3'-moiety (0 for no selective base, 1 for an adenine, 2 for a cytosine, 3 for a guanine, 4 for a thymine, 5 for two adenines); for example, primer M02 is an *MseI*-primer with a 3'-terminal cytosine. When necessary, oligonucleotides were checked for length and purity by TLC using SureCheck (USB) or were further purified using a Sephadex NAP-10 column (Pharmacia). Working solutions of 50 ng ml⁻¹ were prepared in Milli-Q water.

AFLP reactions. Primers were labelled with either [γ -³²P]ATP or [γ -³³P]-ATP using bacteriophage T4 polynucleotide kinase (Vos *et al.*, 1995). Typically, a 30–40 % incorporation of radioactivity was obtained, yielding ³²P- or ³³P-labelled oligonucleotides with a specific radioactivity of approximately 26–33 or 15–26 TBq mmol⁻¹, respectively. Selective amplification was done as described previously (Vos *et al.*, 1995) except that 2 µl (20 ng) of the resuspended template DNA, 6.25 ng labelled primer (for *ApaI*-, *EcoRI*- or *HindIII*-primers), 30 ng unlabelled primer (for *TaqI*- and *MseI*-primers), and 0.6 U *Thermus aquaticus* DNA polymerase were used. The reactions were performed on a Perkin Elmer 9600 thermal cycler using the following cycle profile. Cycle 1: 60 s at 94 °C, 30 s at 65 °C, 60 s at 72 °C; cycles 2–12: 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 13–24: 30 s at 94 °C, 30 s at 56 °C, 60 s at 72 °C. After completion of the cycle programme, an equal volume (20 µl) of 98 % (w/v) formamide, 10 mM EDTA, 0.1 % (w/v) bromophenol blue, 0.1 % (w/v) xylene cyanol was added to the reaction mixtures. Prior to gel loading and electrophoresis, mixtures were heated for 3 min at 95 °C and then rapidly cooled on ice to prevent nucleic acid secondary structures from reannealing.

Electrophoresis of PCR products. The amplified fragments were analysed on 5 % denaturing polyacrylamide gels containing 4.75 % (w/v) acrylamide, 0.25 % (w/v) methylene bisacrylamide, 8.3 M urea, 100 mM Tris, 100 mM boric acid, 2 mM EDTA, pH 8.3 (obtained as pre-mixed SequaGel solutions from National Diagnostics). For each gel, 100 ml casting solution was prepared and mixed with 0.8 ml 10 % (w/v) ammonium persulfate and 40 µl TEMED. The gel was immediately poured into 38 × 50 cm Sequi-Gen cassettes (Bio-Rad) and left overnight at room temperature for polymerization. Spacers and combs (Gibco BRL) were 0.4 mm thick. TBE buffer (100 mM Tris, 100 mM boric acid, 2 mM EDTA, pH 8.3) was used as electrophoresis buffer. Usually, a 1.8 µl sample of each reaction mixture was loaded on the gel. Gels were run using a Sequi-Gen sequencing unit (Bio-Rad) at a constant power of 110 W to maintain an optimal gel temperature (near 50 °C) during electrophoresis. Gels were routinely run for about 150 min until the bromophenol blue front passed the lower cooling rib of the Sequi-Gen Integral Plate/Chamber by approximately 2 cm (approx. 12 cm from the bottom of the gel).

Autoradiography and data acquisition. After electrophoresis, the gel was unmounted and transferred to a 3 MM Whatman chromatography paper (a 5–6 cm strip of the upper part of the

gel was removed so that the dimensions of the gel were kept at approx. 35 × 43 cm; this was necessary since standard Röntgen films and exposure cassettes were used) and vacuum dried on a model 583 gel dryer (Bio-Rad) for 50 min at 80 °C. The dried gel was placed in a 36.6 × 43.2 cm Mavig cassette (Siemens) and exposed to a 35 × 43 cm sheet of Hyperfilm-MP (Amersham). Exposure times depended on how much radioactivity could be measured by a hand-held Geiger-Müller counter. Typically, exposure times of 16–22 h were used for activities of 600–120 c.p.m. cm⁻². If necessary, a Cronex Lightning Plus intensifying screen (Dupont) was used with exposure at -70 °C. After exposure, films were developed using a Fuji RGII X-ray film processor. Resulting autoradiograms were then scanned by a large-format (max. scan area 36 × 91 cm) RayVen RSU1 densitoscanner (X-Ray Scanner Corp.). Images were maintained as TIFF files which were used to measure the optical density along each lane. Densitometric values were further processed using GelCompar 3.1 software (Applied Maths bvba, Kortrijk, Belgium).

Numerical analysis of banding patterns. Digital images were normalized by the alignment of AFLP patterns obtained on multiple specimens of a reference strain included at regular intervals in each gel. Background was subtracted using the curve fitting algorithm which forms part of the GelCompar software package. Digital images could be combined by assigning one particular reference track as a standard and alignment of all other reference tracks versus this standard. A similarity matrix was created using the band-based Dice similarity coefficient, S_D , equal to the ratio of twice the number of common bands in two compared patterns to the sum of all bands in both patterns (Sneath & Sokal, 1973). Bands were assigned to each track automatically by GelCompar using a band search filter in which bands with an area smaller than 0.5 % of the total area of the pattern were disregarded. In addition, for each track, only bands with a minimal profiling value of 5 % (e.g. a band area of at least 5 % of that of the heaviest band in the same track) were selected. For band comparison, a band position tolerance value of 0.8 % (expressed as a percentage of the total length of the pattern) was allowed to compensate for misalignment of homologous bands due to technical imperfections. Using these parameters, reference tracks of the same gel grouped together with an S_D of 95–100 %, whereas reference tracks derived from different gels grouped together at a mean linkage level of 90 % or more. The unweighted-pair group method using average linkages (UPGMA) (Sneath & Sokal, 1973) was used to cluster the patterns (Vauterin & Vauterin, 1992). Using GelCompar, a standard deviation for every branch of the dendrogram was calculated. These standard deviations are an indication of the significance and stability of the formed clusters and are given as \pm values when discussing linkage levels in the text or are represented by error flags at the major branching point of each cluster of the dendrogram.

RESULTS

Reproducibility of AFLP using bacterial genomic DNA

Although the reproducibility of the AFLP technique is reportedly very high (Zabeau & Vos, 1993; Vos *et al.*, 1995), we undertook some simple tests to assess possible variations in AFLP patterns (experimental results not shown). For six *Aeromonas* strains, total DNA was prepared on six separate occasions following subculture at intervals of 1–3 d. From the resulting 36 DNA preparations, template was prepared (at one time) which then was

Table 2. Enzyme combinations (ECs) and primer combinations (PCs) used for the generation of AFLP patterns from bacterial strains*

Genus†	G + C (mol %)‡	EC	PC§	Selection
<i>Clostridium</i>	24–54	–	–	–
<i>beijerinckii</i>	26–28	<i>EcoRI</i> / <i>MseI</i>	E02/M02	C/C
<i>acetobutylicum</i>	28–29	<i>EcoRI</i> / <i>MseI</i>	E02/M02	C/C
<i>Bacillus</i>	32–69	–	–	–
<i>thuringiensis</i> ^a	34	<i>EcoRI</i> / <i>MseI</i>	E02/M01	C/A
<i>cereus</i> ^a	36	<i>EcoRI</i> / <i>MseI</i>	E02/M01	C/A
<i>larvae</i> ^b	42	<i>EcoRI</i> / <i>MseI</i>	E02/M01	C/A
<i>Acinetobacter</i>	38–47	<i>EcoRI</i> / <i>MseI</i>	E02/M04	C/T
		<i>HindIII</i> / <i>TaqI</i>	H01/T05	A/AA
<i>Vibrio</i>	40–50	<i>HindIII</i> / <i>TaqI</i>	H01/T01	A/A
		<i>ApaI</i> / <i>TaqI</i>	A01/T01	A/A
<i>Aeromonas</i>	57–63	<i>ApaI</i> / <i>TaqI</i>	A01/T01	A/A
		<i>EcoRI</i> / <i>MseI</i>	E01/M04	A/T
<i>Pseudomonas</i>	58–68	<i>ApaI</i> / <i>TaqI</i>	A01/T02	A/C
<i>Xanthomonas</i>	63–71	<i>ApaI</i> / <i>TaqI</i>	A03/T03	G/G
		<i>EcoRI</i> / <i>MseI</i>	E00/M02	–/C

* The use of ECs and PCs listed here resulted in banding patterns containing about 30–60 bands.

† Due to the wide range in G + C content for the genera *Clostridium* and *Bacillus*, we have indicated the particular species with corresponding mol % G + C: (a) type strain; (b) strain LMG 9820 (M. Heyndrickx & P. De Vos, unpublished results).

‡ Values for *Clostridium* and *Bacillus* from Sneath *et al.* (1986); values for *Acinetobacter*, *Vibrio*, *Aeromonas*, *Pseudomonas* and *Xanthomonas* from Krieg & Holt (1984).

§ See Table 1 and Methods for sequences.

|| Selective base(s) at the 3'-end of the PCR primers.

used for selective amplification. For each strain, the AFLP patterns obtained with the six separate DNA preparations were indistinguishable. Also, for four of the same six *Aeromonas* strains, template was prepared from each of the original DNA preparations and subjected to PCR. Again, no differences in banding patterns could be detected.

Choice of restriction enzymes to generate DNA templates for AFLP

The digestion of representative bacterial DNAs was tested using a wide variety of restriction enzymes (results not shown). The following general observations could be made. As expected, A + T-rich recognition sequences such as those of *EcoRI* (G/AATTC) or *MseI* (/TTAA) were more frequent in G + C-poor genomes, and conversely G + C-rich recognition sequences, such as that of *ApaI* (GGGCC/C), were more frequent in G + C-rich genomes. *HindIII* (A/AGCTT) appeared to be most suited to digest genomic DNA with a G + C content of 40–50 mol %. Digestion with the 'frequent-cutter' *MseI* was insignificant for bacterial DNAs with a high G + C

content (> 65 mol %), intermediate for DNAs with a G + C content between 50–65 mol % and optimal for DNAs with a G + C content below 50 mol %. A moderately high frequency of cleavage was obtained with *TaqI* (T/CGA) for all DNAs, irrespective of their G + C content, except for DNAs obtained from *Clostridium acetobutylicum* (26–28 mol %) or *Bacillus thuringiensis* (34 mol %) which were digested intermediately. As a consequence, we chose template preparations using the combined digest products of *EcoRI*/*MseI*, *HindIII*/*TaqI* or *ApaI*/*TaqI* (Table 2).

Differentiation of *Xanthomonas* species and pathovars

ApaI–*TaqI* DNA templates were made from a total of 36 *Xanthomonas* strains, including 23 pathovars of *X. axonopodis*, six pathovars of *X. vasicola*, two pathovars of *X. translucens*, and the type strains of *X. campestris*, *X. populi*, *X. fragariae*, *X. oryzae* and *X. albilineans*, and one strain of *Stenotrophomonas*. AFLP reactions on these template DNAs were performed using primers A03 and T03 (see Methods). The resulting PCR products were separated on the same polyacrylamide gel and the banding patterns were visualized by autoradiography (a detail of typical banding patterns has been given in Fig. 2). The autoradiogram was scanned and the data processed using GelCompar, resulting in the dendrogram depicted in Fig. 2. Strains of *X. axonopodis* grouped in two clusters, I and II, at linkage levels of 61.5 ± 4.8 and 61.8 ± 3.2 %, respectively. Cluster I represents mainly strains that cause bacterial canker disease of citrus (formerly known as *X. campestris* pv. *citri* strains) and cluster II contains strains that cause diseases in grasses (Poaceae). A third cluster (III), composed of six strains belonging to two different pathovars of *X. vasicola*, was found at a linkage level of 79.5 ± 2.7 %. Within cluster I, strains belonging to *X. axonopodis* pv. *citri*, *X. axonopodis* pv. *aurantifolii*, and *X. axonopodis* pv. *citrumelo* grouped in three subdivisions (α , β and γ) with linkage levels of 84.9 ± 1.0 , 78.9 ± 2.3 and 71.0 ± 3.7 %, respectively. In addition, *X. axonopodis* pv. *alfalfae* strain LMG 497 and *X. axonopodis* pv. *coracanae* strain LMG 686 clearly grouped in the γ subdivision. An important observation was that individual strains within these subdivisions of *X. axonopodis* still could be differentiated from each other based on their AFLP banding patterns (banding patterns not shown), demonstrating the high discriminatory power of the AFLP technique.

Influence of primer and template use on the numerical analysis of AFLP patterns

ApaI–*TaqI* template DNA obtained from 16 *Xanthomonas* strains was used to produce AFLP banding patterns with an additional four sets of primer combinations (not all results are shown). As was the case for the set A03/T03 (G/G) (selective bases in parentheses) used above (typical banding patterns are shown in Fig. 3a), the use of set A03/T02 (G/C) (data not shown) resulted in banding

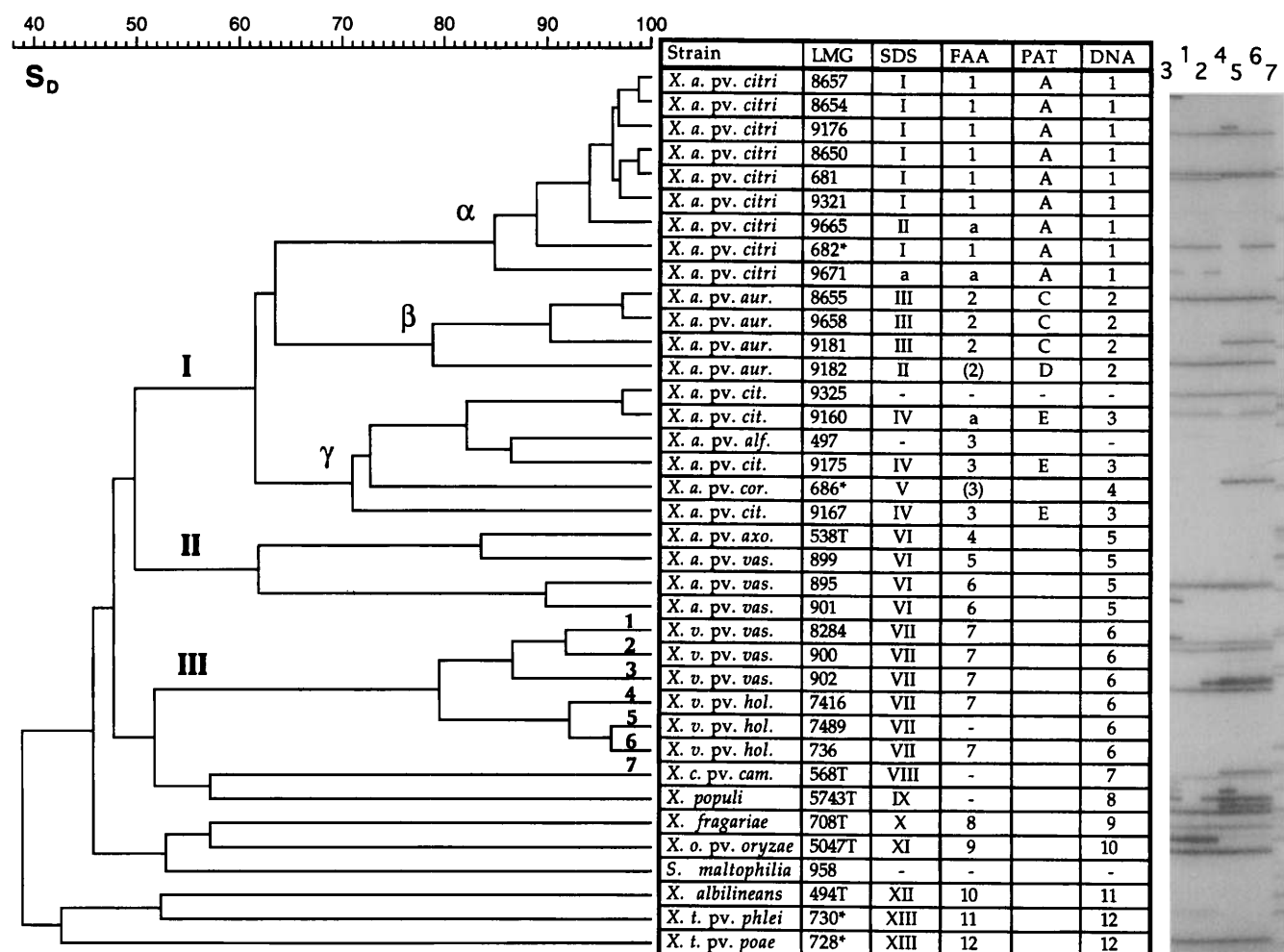


Fig. 2. Numerical analysis of AFLP patterns from 37 *Xanthomonas* strains and comparison of AFLP-based clustering with classification according to other methods. AFLP patterns, obtained from *Apal*-*TaqI* template DNAs using the selective PCR primers A03 and T03, are derived from the same gel. To the right, typical AFLP banding patterns are partly shown (size range approx. 80–350 bp) for the six *X. vasicola* strains (lanes 1–6) and one *X. campestris* pv. *campestris* strain (lane 7) numbered in the dendrogram. Full gel length, corresponding to a size range of approximately 80–550 bp, was used for numerical analysis. Species names are according to Vauterin et al. (1995). Abbreviations are as follows: *X.*, *Xanthomonas*; *S.*, *Stenotrophomonas*; *v.*, *vasicola*; *o.*, *oryzae*; *t.*, *translucens*; *c.* or *cam.*, *campestris*; *a.* or *axo.*, *axonopodis*; *alf.*, *alfalfae*; *aur.*, *aurantifolia*; *cor.*, *coracanae*; *vas.*, *vasculorum*; *hol.*, *holcicola*; *cit.*, *citrumelo*. Group numbers used here do not necessarily match with group numbers used in the literature but should be regarded as a compilation of existing data (Vauterin et al., 1991a, b, 1992, 1995; Yang et al., 1993). LMG, strain number: type strains and pathovar reference strains are indicated by T and *, respectively; SDS, grouping by SDS-PAGE of whole-cell proteins: 'a', atypical strains; FAA, grouping by fatty acid analysis: 'a', atypical strains, (), related strains; PAT, *X. axonopodis* grouping based on pathogenicity; DNA, grouping based on DNA–DNA hybridization studies; -, data were unavailable. Clusters I–III of *Xanthomonas* strains and subdivisions α , β and γ within cluster I are indicated. S_D , band-based Dice similarity coefficient (%). The dendrogram was constructed using UPGMA.

patterns with a mean of 35–50 bands and changes in the linkage levels of the individual clusters were minimal. In contrast, using the sets A04/T03 (T/G) (Fig. 3b), A01/A03 (A/G) (data not shown), and A01/T01 (A/A) (data not shown), banding patterns contained on average only between 15–25 bands. The decrease in the number of bands did not seem to affect the numerical analysis to a great extent but the linkage levels of some clusters dropped significantly (i.e. from 79.5 ± 2.7 to 60.2 ± 4.0 %; Fig. 3a versus b). On the other hand, *Xanthomonas* strains LMG 895 and 901, which displayed identical AFLP

banding patterns when using primer sets A01/T01 (A/A) and A01/T03 (A/G), could be differentiated from each other using the primer sets A04/T03 (T/G) or A03/T02 (G/C) (results not shown), indicating that highly related bacterial strains may still be distinguished with AFLP using an appropriate set of selective primers.

EcoRI–*MseI* templates were prepared from the same 37 strains listed in Fig. 2 and banding patterns were produced using the primer combination E00(–)/M02(C) (Fig. 3c; not all results are shown). Numerical analysis on these

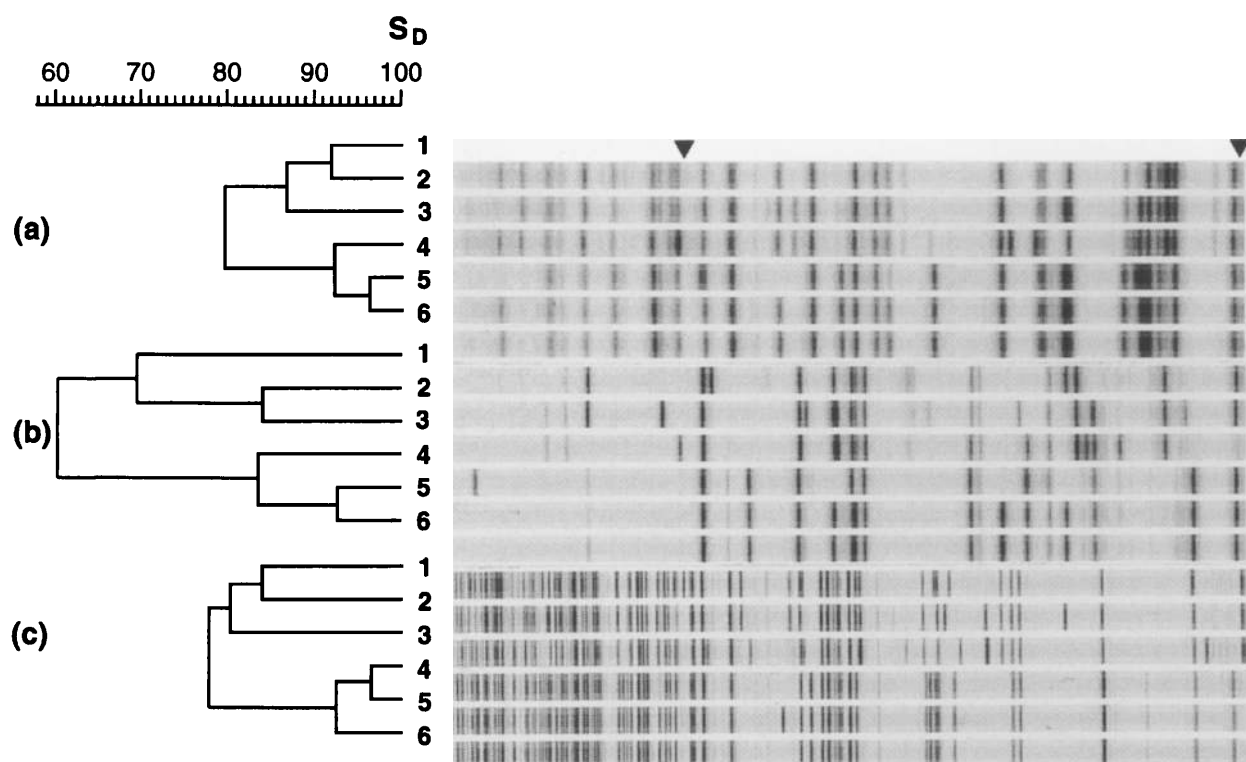


Fig. 3. Digitized AFLP patterns obtained from the genomic DNA of six *X. vasicola* strains (corresponding to cluster III in Fig. 2). (a) and (b) AFLP reactions performed on *ApaI*-*TaqI* templates using the primer sets A03/T03 and A04/T03, respectively; (c) AFLP reactions performed on *EcoRI*-*MseI* templates using the primer set E00/M02 (for sequences of primers see Table 1 and Methods). Each set of six patterns was derived from one single gel. S_D , band-based Dice similarity coefficient (%). The dendrograms were constructed using UPGMA. The various linkage levels (including standard deviations) of cluster formation are discussed in the text. Strain numbers 1–6 correspond to the numerical notation in the large dendrogram of Fig. 2. Digitized AFLP patterns in panel (a) partly correspond with AFLP patterns shown in Fig. 2 (corresponding part of the patterns indicated by arrowheads).

banding patterns correlated very well with results obtained on *ApaI*-*TaqI* templates although linkage levels between the various groups as well as between the individual strains were 5–10% lower (Fig. 3; not all results shown). *X. axonopodis* strains again clearly grouped together although the originally obtained subdivisions (α , β and γ) this time grouped slightly differently with respect to each other (results not shown). However, in view of the very close branching points of these subdivisions (Fig. 2) this must be seen as only a minor change. As expected, strains LMG 497 and 686 again grouped with subdivision γ confirming the results on *ApaI*-*TaqI* templates (data not shown). Grouping of the three *X. axonopodis* pv. *holcicola* strains LMG 895, 899 and 901, the single *X. axonopodis* pv. *axonopodis* strain LMG 538T, the three *X. vasicola* pv. *vasculorum* strains LMG 8284, 900 and 902, and the three *X. vasicola* pv. *holcicola* strains LMG 7489, 736 and 7416, was also in perfect concordance with the AFLP results on *ApaI*-*TaqI* templates depicted in Fig. 2 (data not shown). Finally, similar to the use of other PCR primers on the same DNA template (see above), the use of another DNA template allowed the discrimination of highly related strains. For instance, although *Xanthomonas* strains LMG 7489 and 736 displayed identical AFLP patterns when using *ApaI*-*TaqI* templates (Figs 3a, b, lanes 5 and 6), these strains could still be differentiated from each other

when *EcoRI*-*MseI* templates were used (Fig. 3c, lanes 5 and 6).

Classification of genotypically characterized *Aeromonas* strains

We prepared total cellular DNA from 90 strains encompassing 14 currently described hybridization groups in the genus *Aeromonas* and applied the AFLP technique to *ApaI*-*TaqI* templates using selective primers A01 and T01 (see Methods). Numerical analysis of the resulting AFLP patterns showed that AFLP-based grouping of these strains (Fig. 4) correlated very well with DNA homology data (Altwegg, 1990; Altwegg *et al.*, 1990). Strains belonging to the same HG (genomic species) clearly clustered together although, within a given HG, most strains could be easily differentiated from each other (banding patterns not shown). Overall, an $S_D \geq 64\%$ separated the AFLP clusters at (genomic) species level.

AFLP fingerprints from bacterial genomes with different G+C content

As an example of the general applicability of the AFLP technique to fingerprint bacterial genomes, we produced AFLP banding patterns for bacteria belonging to the genera *Clostridium*, *Bacillus*, *Acinetobacter*, *Vibrio* and

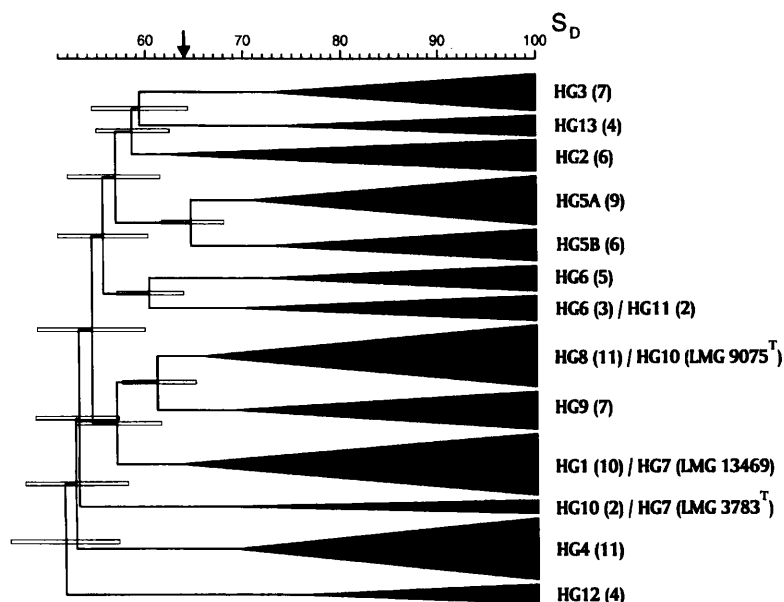


Fig. 4. Cluster analysis on 90 genotypically characterized *Aeromonas* strains using *ApaI*-*TaqI* templates. AFLP patterns used for numerical analysis were derived from three different gels. Amplification reactions were performed using primers A01 and T01 (Table 1; see Methods). Genomic species are represented as HGs (the DNA relatedness between *Aeromonas* strains belonging to the same HG is at least 70% at 60 °C and/or above 60% at 75 °C with a maximal divergence of 5%; Altwegg, 1990). The following HGs were included: HGs 1 and 2 (*A. hydrophila*), HG3 (*A. hydrophila* and *A. salmonicida*), HGs 4 and 5A (*A. caviae*), HG5B (*A. caviae* and *A. media*), HG6 (*A. eucrenophila*), HG7 (*A. sobria*), HG8 (*A. veronii* biogroup *sobria*), HG9 (*A. jandaei*), HG10 (*A. veronii* biogroup *veronii*), HG11 (*A. veronii*-like), HG12 (*A. schubertii*) and HG13 (*A. trota*). Numbers in parentheses reflect the number of strains analysed for a particular HG. *A. sobria* strains LMG 3783^T and LMG 13469 and HG10 strain LMG 9075^T have been indicated (see text). The dendrogram was constructed using UPGMA. Standard deviations are given as error flags. S_D , band-based Dice similarity coefficient (%) and the arrow marks the cut-off S_D value at 64% to define the AFLP clusters within the dendrogram.

Pseudomonas (Fig. 5). The genomic base composition of these bacteria ranges from about 30 to 70 mol % G + C. For each genus, four strains were chosen of which three belonged to the same species or – in the case of *Bacillus cereus* and *B. thuringiensis* – were highly related based on DNA–DNA hybridization studies. Notably, in each case, the three related strains had very similar AFLP patterns, whereas the fourth strain showed a much different banding pattern with limited similarity to the first three. Also, in all cases, detection of DNA polymorphisms by AFLP clearly permitted the discrimination of the three related strains.

DISCUSSION

AFLP is a novel DNA fingerprinting method that can be used on any DNA, irrespective of its origin or complexity (Zabeau & Vos, 1993; Vos *et al.*, 1995). The general strategy of the AFLP technique is outlined in Fig. 1. Briefly, the genomic DNA is digested with two restriction enzymes, and restriction halfsite-specific double-stranded oligonucleotide linkers (adaptors) are ligated to all restriction fragments. An important feature of these adaptors is that they serve as target sites for PCR primer annealing, thereby allowing amplification of the restriction fragments. Because the AFLP primers contain at their 3'-end one or more nucleotides that extend beyond the restriction site, only a subset of restriction fragments are amplified. Obviously, a correct choice of restriction enzymes is crucial to AFLP analysis as this will predetermine the size and number of restriction fragments. Although attempts have been made to predict the cleavage frequency from the mol % G + C ratio (Nei & Li, 1979;

Owen, 1989) or by Markov chain analysis of di- and trinucleotide frequencies (Forbes *et al.*, 1991), an empirical determination of the cleavage frequency for a given restriction enzyme and a particular DNA seemed to us the best approach. Based on our observations, the combined digests *EcoRI*/*MseI* and *ApaI*/*TaqI* were most suited for bacterial genomes with low and high G + C contents, respectively. For bacteria with a G + C ratio of 40–50 mol %, *HindIII* combined with *TaqI* also gave an adequate number of suitably sized restriction fragments, and *HindIII*-*TaqI* templates were used occasionally for AFLP analysis of *Vibrio* and *Acinetobacter* strains.

The use of well-characterized *Xanthomonas* and *Aeromonas* strains allowed us to investigate the validity of AFLP for the classification and typing of bacteria. In general, grouping of *Xanthomonas* strains according to AFLP corresponded well with the classification obtained with other fingerprinting methods or based on pathogenicity and DNA homology data (Fig. 2). It was interesting that the *X. axonopodis* pv. *alfalfae* strain LMG 497 and *X. axonopodis* pv. *coracanae* strain LMG 686 both grouped convincingly in the *X. axonopodis* γ subdivision, supporting the fatty acid analysis results on these strains (Yang *et al.*, 1993). Grouping of *X. axonopodis* pv. *axonopodis* strain LMG 538 with three *X. axonopodis* pv. *vasculorum* strains correlated well with the classification of these strains demonstrated previously by SDS-PAGE of whole-cell proteins, GLC analysis of fatty acid methyl esters (FAMES) and DNA–DNA hybridization data (Fig. 2). In fact, comparison of the AFLP patterns showed that *X. axonopodis* pv. *axonopodis* is more related to strain LMG 899 than to the other two strains of *X. axonopodis* pv.

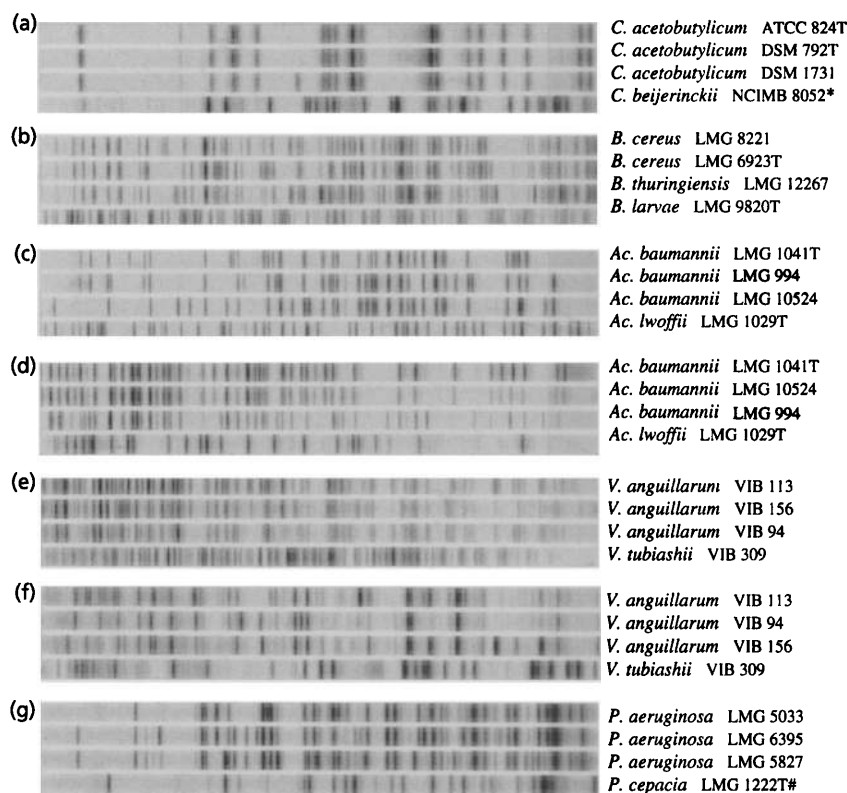


Fig. 5. Chromosomal AFLP patterns of bacteria representing a wide range in mol% G+C: (a) *Clostridium*, (b) *Bacillus*, (c) and (d) *Acinetobacter* (abbreviated as *Ac.*), (e) and (f) *Vibrio*, and (g) *Pseudomonas*. The following DNA templates were used: *EcoRI*–*MseI* (a–c), *HindIII*–*TaqI* (d–e), and *Apal*–*TaqI* (f–g). Primer sets used for selective amplification are given in Table 2. Strain notations are according to the LMG catalogue. *Vibrio* strains with the prefix VIB originate from the *Vibrio* Culture Collection which is also maintained in our laboratory. *Based on 16S rRNA sequence determination, strain NCIMB 8052, formerly known as the type strain of *C. acetobutylicum*, has been assigned to the species *C. beijerinckii* (Wilkinson *et al.*, 1995); #*Pseudomonas cepacia* has been transferred to the genus *Burkholderia* (Yabuuchi *et al.*, 1992).

vasculorum (Fig. 2, banding patterns not shown). In addition, *X. vasicola* pv. *vasculorum* strains, which are indistinguishable from *X. vasicola* pv. *bolvicola* based on SDS-PAGE of whole-cell proteins, FAME or DNA–DNA hybridization analyses (Vauterin *et al.*, 1992), could still be recognized as a separate group despite highly similar AFLP patterns (Figs 2, 3a).

A good correlation between AFLP data and existing taxonomic data was also found for the 90 *Aeromonas* strains used in this study. The AFLP clusters obtained in the dendrogram (Fig. 4) are in perfect concordance with the grouping of these strains according to DNA hybridization data (Altwegg, 1990; Altwegg *et al.*, 1990), assigning 85 strains to the correct HG (the DNA relatedness between strains belonging to the same HG is at least 70% at 60 °C and/or above 60% at 75 °C with a maximal divergence of 5%; Altwegg, 1990). In addition, AFLP-based grouping of the 90 *Aeromonas* strains basically confirmed the phenotypic analysis of Kämpfer & Altwegg (1992) and our own FAME data (Huys *et al.*, 1994). However, *A. veronii* type strain LMG 9075^T (HG10) clearly grouped with HG8 strains (Fig. 4). This may be explained by the fact that HG8 and HG10 each represent different biogroups (*A. veronii* biogroup *sobria* and *A. veronii* biogroup *veronii*, respectively) of the same nomenspecies. Indeed, both taxa are very closely related based on DNA–DNA hybridization studies (Altwegg, 1990), and at present they can only be differentiated on phenotypic (Kämpfer & Altwegg, 1992) or chemotaxonomic (Huys *et al.*, 1994) grounds. Furthermore, strains LMG 13469 and 3783^T (the sole representatives of *Aeromonas sobria* HG7), which are heterogeneous ac-

cording to their FAME profiles (Huys *et al.*, 1994), grouped with HG1 and HG10, respectively, and the only two strains allocated to HG11 grouped with HG6 (Fig. 4; banding patterns not shown). These results are partly in agreement with fatty acid analysis data (Huys *et al.*, 1994), but the inclusion of HGs which are presented by only a small number of strains may give rise to statistical variations in the performed cluster analysis.

We were interested to see whether or not the use of other restriction enzymes would influence AFLP-based grouping of bacterial strains so we prepared *EcoRI*–*MseI* templates from the 37 strains listed in Fig. 2. Numerical analysis of the AFLP fingerprints obtained with the new templates resulted in a grouping of strains that was highly similar to the grouping found with *Apal*–*TaqI* templates. However, *EcoRI* and *MseI* have a much lower cleavage frequency in G+C-rich DNA and far fewer restriction fragments were available for selective amplification, so that banding patterns obtained from *EcoRI*–*MseI* templates contained unevenly distributed bands, with most bands stacked in the upper part of the gel (Fig. 3c; all results are not shown). As a result, linkage levels between the various groups as well as between the individual strains were 5–10% lower, presumably because correlation between the banding patterns was less accurate in the high-molecular-mass zone of the gel. This was also the case for 12 strains representing HGs 1, 2 and 3 of *Aeromonas hydrophila*. These strains were analysed using *EcoRI*–*MseI* templates and AFLP patterns generated with primer set E01/M04 also contained unevenly distributed bands with most bands located in the upper third part of the gel (results not shown). Linkage levels of the three

HGs dropped considerably and HGs 1 and 2 in fact grouped together in one cluster. This clearly shows that the choice of restriction enzymes may be detrimental to the accuracy of AFLP analysis and that the best combination of endonucleases should be determined empirically for each organism.

The complexity of the final AFLP fingerprint also depends on the length and base composition of the selective nucleotides. For highly complex genomes, such as those from plants or animals, at least three selective nucleotides at the 3'-end of both primers are required to obtain suitable AFLP fingerprints (Vos *et al.*, 1995). Bacterial genomes however are relatively small and range from 0.6 to 9.4 Mbp as determined by PFGE (Cole & Saint Girons, 1994), and usually one selective nucleotide for both primers suffices. When AFLP analysis was performed on 16 *Xanthomonas* strains using the same DNA template (*ApaI*-*TaqI*) but with different primer combinations (Fig. 3a, b; all results are not shown), it was found that the base composition of the selective nucleotides clearly had an effect on the complexity of the resulting AFLP banding patterns. A logical explanation for this phenomenon is the high G+C content of the *Xanthomonas* genome (63–71 mol %) so that relatively less As or Ts are available for base-pairing with the 3'-end of the selective primers A01(A), A04(T), and T01(A). A similar bias for selective bases that are consistent with the G+C content of the analysed genome was also found when *ApaI*-*TaqI* templates were used for AFLP analysis on *Vibrio*, *Pseudomonas* and *Xanthomonas* strains: primer sets A01/T01 (A/A), A01/T02 (A/C), and A03/T03 (G/G), respectively, were required to obtain informative AFLP fingerprints that contained an adequate number of evenly distributed bands (Table 2; Figs 3a and 5f, g). In addition, since *EcoRI* and *MseI* digested the A+T-rich genomic DNAs of *Clostridium* and *Bacillus* extremely well, selective primers had to be adapted accordingly to reduce the complexity of the AFLP fingerprint (Table 2; Fig. 5a, b). Similarly, due to the high cleavage frequency of *HindIII* in *Acinetobacter* genomic DNA, a selective *Taq*-primer containing two selective bases was used for AFLP analysis on *Acinetobacter* *HindIII*-*TaqI* templates (Table 2; Fig. 5d).

The AFLP technique has specific advantages in comparison with other genomic fingerprinting methods. In contrast to ribotyping (Grimont & Grimont, 1986; Stull *et al.*, 1988) and RFLP analysis of PCR-amplified rDNA (rRNA encoding DNA) (Vanechoutte *et al.*, 1992; Jensen *et al.*, 1993), AFLP analyses the whole genome. This may be advantageous because the interpretation and utility of rDNA-based data – i.e. the use of one single gene or operon – in phylogenetic methods is sometimes questionable (Waterhouse & Glover, 1993; Sneath, 1993; Linton *et al.*, 1994; Clayton *et al.*, 1995). Attempts to use RFLP analysis of whole bacterial genomes for identification and typing purposes (Drahoš *et al.*, 1985; Lazo *et al.*, 1987) failed because RFLP fingerprints obtained on conventional agarose gels were too complex. Even the use of polyacrylamide gels for small-fragment restriction endonuclease analysis (SF-REA) only marginally impro-

ved the resolution of electrophoretic fragment separation (Haertl & Brandlow, 1990; Giovannetti *et al.*, 1990), and low-frequency restriction fragment analysis (LFRFA), although applied successfully for typing purposes and taxonomic studies (Prevost *et al.*, 1992; Rainey *et al.*, 1994), required expensive restriction enzymes, cell preparation in agarose-embedded plugs, and lengthy electrophoretic separations. Most recently, restriction fragment end-labelling (RFEL), a genomic fingerprinting technique that produces 30–50 well-separated fragments in the size range 100–400 bp, has been reported (van Steenberg *et al.*, 1995; Hermans *et al.*, 1995). Although AFLP also makes use of restriction enzymes, it largely differs from the above-mentioned RFLP-based methodology in two respects. Firstly, genomic fragments of predetermined size are tagged with adaptors that serve as primer-binding sites allowing selective amplification. This enhances the flexibility of the AFLP technique to a large extent since the complexity of AFLP fingerprints can be adjusted by simply changing the number or the base composition of the selective bases that are present at the 3'-end of the AFLP primers. Secondly, PCR products are labelled during amplification allowing AFLP fingerprints to be visualized by autoradiography in contrast to SF-REA, where banding patterns must be visualized by DNA silver staining requiring additional manipulations such as washing and soaking of gels. Also, bands in the lower molecular mass range are usually not efficiently visualized by SF-REA. In RFEL, radioactivity incorporated in the DNA must be measured for each reaction so that an equal amount can be loaded in each lane of the gel and the technique produces artefactual bands caused by the unequal mobility of the two labelled strands. These problems are avoided in AFLP because radioactive label is equally incorporated via standardized PCR conditions and only one strand of the PCR products is labelled. In addition, the technique can be easily adapted to automation by the use of fluorescent-dye-labelled primers in combination with recently developed DNA fragment separation techniques (Ansorge *et al.*, 1992; Landers, 1993).

Over the past five years, many PCR-based genomic fingerprinting methods to analyse entire genomes have been devised. These methods make use of synthetic oligonucleotides that prime specific (Welsh & McClelland, 1991; Doll *et al.*, 1993; Woods *et al.*, 1993; Haas *et al.*, 1993) or arbitrary (Williams *et al.*, 1990; Welsh & McClelland, 1990; Caetano-Anollés *et al.*, 1991; Bassam *et al.*, 1992) segments of the genomic DNA thereby producing a characteristic set or pattern of amplified fragments. Although some of these techniques have been applied extensively over the past few years for the differentiation of bacterial strains (Towner & Cockayne, 1993; van Belkum, 1994), they are prone to variations in amplification efficiency due to small temperature changes (Meunier & Grimont, 1993; Penner *et al.*, 1993; He *et al.*, 1994). In addition, several studies have showed that primer- and DNA concentration (Kernodle *et al.*, 1993; MacPherson *et al.*, 1993; Davin-Regli *et al.*, 1995), the type of DNA polymerase used (Schierwater & Ender, 1993)

and DNA template quality (Kernodle *et al.*, 1993; Micheli *et al.*, 1994) appear to be crucial factors that affect the reproducibility of arbitrary PCR-amplified fingerprinting methods. Also, it is unknown to what extent mismatches between primer and template are allowed, and the use of lower annealing temperatures and/or the use of degenerate primers may possibly cause small variations in the banding patterns.

In contrast, AFLP primers perfectly match their target site (consisting of adaptor and restriction site sequences). The only possible mismatch that may occur is between the selective base(s) at the 3'-end of the primer and the corresponding base(s) adjacent to the restriction site (Kwok *et al.*, 1990). However, according to Vos *et al.* (1995), primers with mispaired 3'-ends are not expected to participate in the amplification process, especially if only one or two selective bases are used. The reason for this lies in the fact that high-stringency conditions are used during the first 12 cycles of the AFLP reaction (see Methods; Zabeau & Vos, 1993), and initiation of DNA synthesis by imperfectly matched primers will be inefficient or non-existent. A comparative study using different types of DNA polymerases or using various PCR cycle conditions has not yet been undertaken for AFLP. However, Vos *et al.* (1995) reported that AFLP is insensitive to variations in the template DNA concentration, confirming our own results when using up to 100-fold less template DNA (unpublished results). They also showed that the complexity of the banding pattern was always reciprocal to the number of selective bases used (i.e. the number of bands is reduced approximately fourfold for each additional selective base, assuming random base distribution), and adding one selective nucleotide to one of the selective primers always resulted in a simplified fingerprint of which all bands were previously present in the original fingerprint. These are good indications that selective PCR in AFLP is highly specific and that mismatching – if it occurs at all – does not reach the detectable level. Recently, a simplified version of AFLP methodology has been used for the molecular typing of *Legionella pneumophila* isolates in a single-step reaction using only one restriction enzyme (Valsangiacomo *et al.*, 1995). These authors found that AFLP results were in good agreement with standard RFLP analysis using either an *Escherichia coli* ribosomal probe or a *L. pneumophila* genomic probe. Also, they found a direct correlation between the number and base composition of selective bases used and the complexity of the banding patterns. However, because template DNA was prepared by a single digest of genomic DNA, and unlabelled primers were used for PCR, selective amplification yielded only a limited number (5–10) of PCR products in the size range 400–2000 bp, and electrophoresis patterns had to be visualized by ethidium bromide staining. Such fingerprints may be sufficient for typing purposes but are less suited for accurate numerical analysis in taxonomic studies. In addition, the use of only one restriction enzyme constitutes a significant loss in flexibility which is one of the hallmarks of the AFLP technique.

Finally, it could be argued that the use of different enzyme

and primer combinations (depending on the G+C content of the studied genomes) may represent a drawback in the general application of AFLP for the DNA genotyping of bacteria. However, here we report the use of AFLP at the subgeneric level, i.e. the use of particular sets of enzymes and AFLP primers to characterize and type bacteria that belong to the same genus (and thus differ by less than 10% in their G+C content; Goodfellow & O'Donnell, 1993). A possible solution to generalize the application of the AFLP technique for bacterial taxonomy and identification would lie in the use of a standard set of enzymes and primers. The simultaneous use of three or four well-chosen restriction enzymes would perhaps circumvent the problem of G+C contents. For instance, it is conceivable that a quadruple-digest with the enzymes *EcoRI*, *ApaI*, *MseI* and *TaqI* would result in an adequate number of suitably sized fragments, irrespective of the G+C content of the digested DNA. Ideally, selective amplification with a fixed set of AFLP primers would then provide banding patterns with limited complexity and good band distribution for any given genome. Such an approach would greatly stimulate the general use of AFLP for bacterial DNA genotyping and would enable the establishment of a uniform database.

AFLP as a taxonomic tool – concluding remarks

AFLP methodology shares many characteristics with other genomic fingerprinting methods but in essence combines the power of RFLP analysis with the flexibility of PCR-based technology. Obvious technical advantages are its reproducibility and high resolution due to the use of stringent PCR conditions and deployment of polyacrylamide gels. The technique does not require prior knowledge of genomic sequences, produces data that are in perfect concordance with existing taxonomic data, and allows differentiation of highly related bacterial strains. Possible applications such as the use of AFLP patterns for the authentication and legal protection of patent strains or the AFLP-based detection of DNA polymorphisms in genome evolution studies should be investigated.

ACKNOWLEDGMENTS

We are grateful to G. Brown for careful reading of the manuscript. This research was carried out in the framework of contract GOA 91/96-2 of the Ministerie van de Vlaamse Gemeenschap, Bestuur Wetenschappelijk Onderzoek (Belgium). G.H. acknowledges the support received from the Vlaams Instituut voor de bevordering van Wetenschappelijk-Technologisch onderzoek in de industrie (IWT, Belgium) in the form of a Bursary for Advanced Study. P.J. is indebted to the European Environmental Research Organisation (The Netherlands) for a short-term fellowship.

REFERENCES

- Altwegg, M. (1990). *Taxonomy and epidemiology of Aeromonas spp.: the value of new typing methods*. Habilitationsschrift, Universität Zürich.
- Altwegg, M., Steigerwalt, A., Altwegg-Bissig, R., Lüthy-Hottenstein, J. & Brenner, D. J. (1990). Biochemical identification

of *Aeromonas* genospecies isolated from humans. *J Clin Microbiol* **28**, 258–264.

Ansorge, W., Voss, H., Wiemann, S., Schwager, C., Sproat, B., Zimmerman, J., Stegemann, J., Erfle, H., Hewitt, N. & Rupp, T. (1992). High-throughput automated DNA sequencing facility with fluorescent labels at the European Molecular Biology Laboratory. *Electrophoresis* **13**, 616–619.

Bassam, B. J., Caetano-Anollés, G. & Gresshoff, P. M. (1992). DNA amplification fingerprinting of bacteria. *Appl Microbiol Biotech* **38**, 70–76.

van Belkum, A. (1994). DNA fingerprinting of medically important microorganisms by use of PCR. *Clin Microbiol Rev* **7**, 174–184.

Bertram, J. & Dürre, P. (1989). Conjugal transfer and expression of streptococcal transposons in *Clostridium acetobutylicum*. *Arch Microbiol* **151**, 551–557.

Caetano-Anollés, G., Bassam, B. & Gresshoff, P. M. (1991). DNA amplification using very short arbitrary oligonucleotide primers. *Bio/Technology* **9**, 553–557.

Clayton, R. A., Sutton, G., Hinkle, P. S., Jr, Bult, C. & Fields, C. (1995). Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. *Int J Syst Bacteriol* **45**, 595–599.

Cole, S. T. & Saint Girons, I. (1994). Bacterial genomics. *FEMS Microbiol Rev* **14**, 139–160.

Davin-Regli, A., Abed, Y., Charrel, R. N., Bollet, C. & de Micco, P. (1995). Variations in DNA concentrations significantly affect the reproducibility of RAPD fingerprint patterns. *Res Microbiol* **146**, 561–568.

Doll, L., Moshitch, S. & Frankel, G. (1993). Poly(GTG)₅-associated profiles of *Salmonella* and *Shigella* genomic DNA. *Res Microbiol* **144**, 17–24.

Drahos, D., Brackin, J. & Barry, G. (1985). Bacterial strain identification by comparative analysis of chromosomal DNA restriction patterns. *Phytopathology* **75**, 1381.

Forbes, K. J., Bruce, K. D., Jordens, J. Z., Ball, A. & Pennington, T. H. (1991). Rapid methods in bacterial DNA fingerprinting. *J Gen Microbiol* **137**, 2051–2058.

Giovannetti, L., Ventura, S., Bazzicalupo, M., Fani, R. & Materassi, R. (1990). DNA restriction fingerprint analysis of the soil bacterium *Azospirillum*. *J Gen Microbiol* **136**, 1161–1166.

Goodfellow, M. & O'Donnell, A. G. (1993). Roots of bacterial systematics. In *Handbook of New Bacterial Systematics*, pp. 3–54. Edited by M. Goodfellow & A. G. O'Donnell. London: Academic Press.

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

Haas, W. H., Butler, W. R., Woodley, C. L. & Crawford, J. T. (1993). Mixed-linker polymerase chain reaction: a new method for rapid fingerprinting of isolates of the *Mycobacterium tuberculosis* complex. *J Clin Microbiol* **31**, 1293–1298.

Haertl, R. & Bandlow, G. (1990). Application of small fragment restriction endonuclease analysis (SF-REA) to the epidemiological fingerprinting of *Staphylococcus aureus*. *J Med Microbiol* **33**, 91–96.

He, Q., Viljanen, M. & Mertsola, J. (1994). Effects of thermocyclers and primers on the reproducibility of banding patterns in randomly amplified polymorphic DNA analysis. *Mol Cell Probes* **8**, 155–160.

Hermans, P. W., Sluijter, M., Hoogenboezem, T., Heersma, H., van Belkum, A. & de Groot, R. (1995). Comparative study of five different DNA fingerprint techniques for molecular typing of *Streptococcus pneumoniae* strains. *J Clin Microbiol* **33**, 1606–1612.

Huys, G., Vancanneyt, M., Coopman, R., Janssen, P., Falsen, E., Altwegg, M. & Kersters, K. (1994). Cellular fatty acid composition as a chemotaxonomic marker for the differentiation of phenospecies and hybridization groups in the genus *Aeromonas*. *Int J Syst Bacteriol* **44**, 651–658.

Jensen, M. A., Webster, J. A. & Straus, N. (1993). Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphisms. *Appl Environ Microbiol* **59**, 945–952.

Kämpfer, P. & Altwegg, M. (1992). Numerical classification and identification of *Aeromonas* genospecies. *J Appl Bacteriol* **72**, 341–351.

Kernodle, S. P., Cannon, R. E. & Scandalios, J. G. (1993). Concentration of primer and template qualitatively affects products in random-amplified polymorphic DNA PCR. *Biotechniques* **14**, 362–364.

Krieg, N. R. & Holt, J. G. (editors) (1984). *Bergey's Manual of Systematic Bacteriology*, vol. 1. Baltimore & London: Williams & Wilkins.

Kwok, S., Kellog, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C. & Sninsky, J. J. (1990). Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus 1 model studies. *Nucleic Acids Res* **18**, 999–1005.

Landers, J. P. (1993). Capillary electrophoresis: pioneering new approaches for biomolecular analysis. *Trends Biochem Sci* **18**, 409–414.

Lazo, G. R., Roffey, R. & Gabriel, D. W. (1987). Pathovars of *Xanthomonas campestris* are distinguishable by restriction fragment length polymorphism. *Int J Syst Bacteriol* **37**, 214–221.

Linton, D., Clewley, J. P., Burnens, A., Owen, R. J. & Stanley, J. (1994). An intervening sequence (IVS) in the 16S rRNA gene of the eubacterium *Helicobacter canis*. *Nucleic Acids Res* **22**, 1954–1958.

Louws, F. J., Fulbright, D. W., Taylor Stephens, C. & de Bruijn, F. (1994). Specific genomic fingerprints of phytopathogenic *Xanthomonas* and *Pseudomonas* pathovars and strains generated with repetitive sequences and PCR. *Appl Environ Microbiol* **60**, 2286–2295.

MacPherson, J. M., Eckstein, P. E., Scoles, G. J. & Gajadhar, A. A. (1993). Variability of the random amplified polymorphic DNA assay among thermal cyclers, and effects of primer and DNA concentration. *Mol Cell Probes* **7**, 293–299.

Marmur, J. (1961). A procedure for the isolation of deoxyribonucleic acid from microorganisms. *J Mol Biol* **3**, 208–218.

Meunier, J.-R. & Grimont, P. A. D. (1993). Factors affecting reproducibility of random amplified polymorphic DNA fingerprinting. *Res Microbiol* **144**, 373–379.

Micheli, M. R., Bova, R., Pascale, E. & D'Ambrosio, E. (1994). Reproducible DNA fingerprinting with the random amplified polymorphic DNA (RAPD) method. *Nucleic Acids Res* **22**, 1921–1922.

Nei, M. & Li, W.-H. (1979). Mathematical model for studying genetic variations in terms of restriction endonucleases. *Proc Natl Acad Sci USA* **76**, 5269–5273.

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.

Penner, G. A., Bush, A., Wise, R., Kim, W., Domier, L., Kasha, K., Laroche, A., Scoles, G., Molnar, S. J. & Fedak, G. (1993). Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. *PCR Methods Applic* **2**, 341–345.

- 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.
- Prevost, G., Jaulhac, B. & Piemont, Y. (1992). DNA fingerprinting by pulsed-field gel electrophoresis is more effective in distinguishing among methicillin-resistant *Staphylococcus aureus* isolates. *J Clin Microbiol* 30, 967–973.
- Rainey, P. B., Bailey, M. J. & Thompson, I. P. (1994). Phenotypic and genotypic diversity of fluorescent pseudomonads isolated from field-grown sugar beet. *Microbiology* 140, 2315–2331.
- Schierwater, B. & Ender, A. (1993). Different DNA polymerases may amplify different RAPD products. *Nucleic Acids Res* 21, 4647–4648.
- Sneath, P. H. A. (1993). Evidence from *Aeromonas* for genetic crossing-over in ribosomal sequences. *Int J Syst Bacteriol* 43, 626–629.
- Sneath, P. H. A. & Sokal, R. R. (1973). *Numerical Taxonomy: the Principles and Practice of Numerical Classification*. San Francisco: W. H. Freeman.
- Sneath, P. H. A., Mair, N. S., Sharpe, M. E. & Holt, J. G. (editors) (1986). *Bergey's Manual of Systematic Bacteriology*, vol. 2. Baltimore & London: Williams & Wilkins.
- van Steenberg, T. J. M., Colloms, S. D., Hermans, P. W. M., de Graaff, J. & Plasterk, H. A. (1995). Genomic fingerprinting by restriction fragment end labeling. *Proc Natl Acad Sci USA* 92, 5572–5576.
- 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.
- Towner, K. J. & Cockayne, A. (1993). *Molecular Methods for Microbial Identification and Typing*. London: Chapman & Hall.
- Valsangiocomo, C., Baggi, F., Gaia, V., Balmelli, T., Peduzzi, R. & Piffaretti, J.-C. (1995). Use of amplified fragment length polymorphism in molecular typing of *Legionella pneumophila* and application to epidemiological studies. *J Clin Microbiol* 33, 1716–1719.
- Vaneeschoutte, M., Rossau, R., De Vos, P., Gillis, M., Janssens, D., Paepe, N., De Rouck, A., Fiers, T., Claeys, G. & Kersters, K. (1992). Rapid identification of bacteria of the Comamonadaceae with amplified ribosomal DNA-restriction analysis (ARDRA). *FEMS Microbiol Lett* 93, 227–234.
- Vauterin, L. & Vauterin, P. (1992). Computer-aided objective comparison of electrophoresis patterns for grouping and identification of microorganisms. *Eur Microbiol* 1, 37–41.
- Vauterin, L., Swings, J. & Kersters, K. (1991a). Grouping of *Xanthomonas campestris* pathovars by SDS-PAGE of proteins. *J Gen Microbiol* 137, 1677–1687.
- Vauterin, L., Yang, P., Hoste, B., Vancanneyt, M., Civerolo, E. L., Swings, J. & Kersters, K. (1991b). Differentiation of *Xanthomonas campestris* pv. *citri* strains by sodium dodecyl sulfate-polyacrylamide gel electrophoresis of proteins, fatty acid analysis, and DNA–DNA hybridization. *Int J Syst Bacteriol* 41, 535–542.
- Vauterin, L., Yang, P., Hoste, B., Pot, B., Swings, J. & Kersters, K. (1992). Taxonomy of xanthomonads from cereals and grasses based on SDS-PAGE of proteins, fatty acid analysis and DNA hybridization. *J Gen Microbiol* 138, 1467–1477.
- Vauterin, L., Hoste, B., Kersters, K. & Swings, J. (1995). Reclassification of *Xanthomonas*. *Int J Syst Bacteriol* 45, 472–489.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Freijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: a new concept for DNA fingerprinting. *Nucleic Acids Res* 21, 4407–4414.
- Waterhouse, R. N. & Glover, L. A. (1993). Differences in the hybridization pattern of *Bacillus subtilis* genes coding for rRNA depend on the method of DNA preparation. *Appl Environ Microbiol* 59, 919–921.
- Welsh, J. & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* 18, 7213–7218.
- Welsh, J. & McClelland, M. (1991). Genomic fingerprints produced by PCR with consensus tRNA gene primers. *Nucleic Acids Res* 19, 861–866.
- Wilkinson, S. R., Young, M., Goodacre, R., Morris, J. G., Farrow, J. A. E. & Collins, M. D. (1995). Phenotypic and genotypic differences between certain strains of *Clostridium acetobutylicum*. *FEMS Microbiol Lett* 125, 199–204.
- Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18, 6531–6535.
- Woods, C. R., Versalovic, J., Koeuth, T. & Lupski, J. R. (1993). Whole-cell repetitive element sequence-based polymerase chain reaction allows rapid assessment of clonal relationships of bacterial isolates. *J Clin Microbiol* 31, 1927–1931.
- Yabuuchi, E., Kosako, Y., Oyaizu, H., Yano, I., Hotta, H., Hashimoto, Y., Ezaki, T. & Arakawa, M. (1992). Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. *Microbiol Immunol* 36, 1251–1275.
- Yang, P., Vauterin, L., Vancanneyt, M., Swings, J. & Kersters, K. (1993). Application of fatty acid methyl esters for the taxonomic analysis of the genus *Xanthomonas*. *Syst Appl Microbiol* 16, 47–71.
- Zabeau, M. & Vos, P. (1993). Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Office, publication 0 534 858 A1.

Received 21 November 1995; revised 23 February 1996; accepted 18 March 1996.